The Possibilities and Limitations of Using Google Books Ngram Viewer in Research on Management Fashions

Dag Øivind Madsen * and Kåre Slåtten

Department of Business, Marketing and Law, USN School of Business, University of South-Eastern Norway, 3511 Hønefoss, Norway
* Correspondence: dag.oivind.madsen@usn.no

Abstract: Google Books Ngram Viewer (GNV) is an analytical tool that uses quantitative methods to analyze digitized text. This paper looks at the possibilities and limitations of using GNV in management fashion research, an area of management research that examines the lifecycle and evolution of management concepts and ideas. GNV provides a historical big picture of the lifecycle and popularity of specific terms and phrases in books. It is argued that this tool could have a natural application in the study of management fashions, since books are a medium through which popular management concepts and ideas have traditionally been diffused. The paper provides several illustrative examples of how GNV can be applied to study management fashions and identifies the tool’s main possibilities and limitations. Although GNV has obvious advantages such as accessibility and user-friendliness, researchers should exercise caution, as it only provides a partial picture of the impact of management fashions.

Keywords: Google Books Ngram Viewer; management fashions; management concepts; management books; evolution; digital humanities; culturomics

1. Introduction

Google Books Ngram Viewer (https://books.google.com/ngrams (25 August 2022)) is a browser-based analytical tool that allows researchers to examine historical trends in the use of specific terms and concepts in books [1].

This paper aims to provide a preliminary discussion of the possibilities and limitations of applying Google Books Ngram Viewer (GNV) as a research tool and data source in research on management fashions. Generally, management fashion research seeks to understand the lifecycle and trajectory of management concepts and ideas such as Lean Management, Agile Management, Six Sigma, and more recently Industry 4.0/5.0 [2–7]. In empirical research on management fashions, a wide range of methods can be used to study the impact of these concepts, from bibliometric methods to qualitative interviews and observations [8–10]. However, to date, this literature has only discussed to a limited extent the applicability of new analytical tools (such as GNV or Google Trends) for capturing the impact of management fashions.

Although data from GNV have been used in recent research on management fashions [11–13], these papers have used these data for mostly illustrative purposes and have not focused on in-depth discussions of the possibilities and limitations of using GNV in research on management fashions.

Instead, the current paper takes an approach similar to Madsen [14], who discussed how Google Trends, (a related Google-based tool) could potentially be applied in management fashion research. However, it should be noted that these two Google-based analytical tools have different purposes and applications for the study of management fashions. While Google Trends provides insight into the demand-side of management fashions (i.e., interest measured by the most frequent search queries), GNV provides data about activity on the
supply side of management fashions (i.e., dissemination of specific concepts and terms via books). GNV provides specific insights into one of the main channels in the diffusion of management fashions, namely books. Røvik [15], pp. 117–118 has suggested that “management literature is probably the single most important medium for the diffusion of organizational recipes.” Similarly, several other researchers have written about how management books are influential in the spread of management concepts and ideas [15,16]. Additionally, while some of the discourse about management fashion has moved to digital and social media [17,18], books remain an important channel through which management concepts and ideas are communicated. For example, in the last 15–20 years the books about the fashionable concept Blue Ocean Strategy have sold several million copies [19]. Therefore, GNV can be useful in management fashion research since it provides specific insight into books. In addition, GNV fits well with the focus of management fashion research, as it allows researchers to examine the big historical picture that crystallizes the lifecycle and stages of fashions.

The rest of the paper proceeds as follows. Section 2 provides a brief review of the use of GNV in other research fields. Section 3 provides a brief overview of management fashion research, while Section 4 provides examples of how GNV can be applied in management fashion research. In Section 5 the main benefits and shortcomings of GNV are discussed. Finally, Section 6 rounds off the paper by offering concluding comments and a discussion of limitations and ideas for future work.

2. Google Books Ngram Viewer (GNV)

2.1. Definition and Features

GNV is a browser-based analytical tool that draws on the Google Books database containing millions of scanned books published between 1500 and 2019 [20,21]. It allows researchers to examine historical trends and evolution in the use of specific terms and concepts in books. It visualizes data since it “displays a graph showing how these phrases have occurred in a corpus of books over the selected years” [1].

Some of the advantages of GNV is that it is easy to use as basically anyone with access to a computer and basic computer knowledge will be able to use it [22]. Furthermore, it allows researchers to search for specific words or terms, and the tool will make a graphic illustration of the usage over time [23].

The data in GNV are normalized, which means that it considers how many books were published at the time. After all, there has been a growth in the number of books published per year, and particularly in recent decades this growth has accelerated, so normalization is necessary to compare frequencies over time [24]. Therefore, GNV displays the percentage of total occurrences in a given year.

2.2. Digital Humanities and Culturomics

GNV uses quantitative methods to analyze digitized texts. Therefore, GNV can be seen as part of the digital humanities movement, where “big data” approaches are increasingly being used, and that has unleashed “new potential for humanities research” [24]. GNV is also closely associated with the emerging area of “culturomics”. Song and Xu [25], p. 509 define culturomics as “an emerging inter-discipline that studies human’s specific ethnic behavior and cultural trends through the quantitative analysis of digitized texts based on big data corpus.” They note that the term was first introduced in the paper by Michel, Shen, Aiden, Veres, Gray, Pickett, Hoiberg, Clancy, Norvig and Orwant [1] with the title “Quantitative Analysis of Culture Using Millions of Digitized Books”. Culturomics can be highly useful for etymology (understanding the origins of specific terms or concepts) and trend studies [25]. One key advantage of digital humanities is that it is less subject to limitations and biases of individual researchers [25].
2.3. Use of GNV in Academic Research

As noted in the previous section, GNV has several features that make it an attractive option for researchers. Younes and Reips [20] note that GNV has been used in academic research on various disciplines in the social and natural sciences, such as sociology, psychology, education, nature, and astrology, and these studies have shown that it can be used to better understand cultural changes in a long-term historical perspective.

Greenfield [26] utilized GNV to study changes in the psychology of culture in the period 1800–2000. The prediction was that the urbanization of society during this period would lead to more individualistic and materialistic values, and the changes in word frequencies from both the United States and the United Kingdom supported that individualistic values increased in importance while collectivistic values declined. In a related study, Zeng and Greenfield [27] used GNV to examine cultural evolution in China during the period 1970–2008. They found that words associated with individualistic values increased in frequency, while words associated with collectivistic values increased at a slower rate or even decreased in frequency.

Sparavigna and Marazzato [28] suggested that GNV can be used to shed light on the history of science. Another example is the study by O’Sullivan et al. [29] in which the authors applied GNV to shed light on the history of psychiatry. The authors looked at the evolution of the use of the term “psychiatry” between 1890 and 1984.

Some scattered studies have used GNV in the context of business and management studies. In a very early study, Kumar and Sahu [22] used GNV to examine the historical evolution of marketing. Kumar and Sahu [22] looked at the applicability of GNV for studying the evolution of marketing approaches and generally found that the data reveal a pattern that has been found in earlier historical examinations of the stages in the evolution of the marketing discipline (e.g., production-oriented, sales-oriented). In a similar type of study, Antoniuk et al. [30] looked at how frequently new keywords (e.g., “digital economy”) that could indicate a shift to a new socio-economic system were used. In a study of the construction industry, Daniel and Oshodi [31] used GNV to look at the rise in the popularity of the terms “Lean construction” and “Off-site” in the period 1997–2021.

As noted in the introduction, there have been some attempts to draw on GNV data to illustrate the evolution of management concepts and ideas [11–13]. For example, Salles-Djelic [11] looked at the spread of the word «management» in books published in English, French, German, and Italian. The word was shown to first have expanded in American English during the 1920s but it did not become widely used in other languages until several decades later. The study by Berg and Madsen [13] used GNV to examine the popularity of terms related to activity-based thinking in the field of management accounting, while Madsen [12] applied the same procedure in the context of Total Quality Management (TQM).

2.4. Evaluation of Possibilities and Limitations

Previous research has suggested that there are several pros and cons associated with the use of GNV in research [24,32]. Among the most apparent benefits is that it is a highly accessible and user-friendly tool that allows researchers to conduct “quick and dirty” analyses [33].

Another strength of GNV is that it provides a macro-level historical view of the evolution and popularity of terms and phrases. Ophir [24], p. 1 argues that it is “an extraordinary tool for tracking long-term usage of terms” that allows for a “very high level trend analysis” that can “shed light on hidden relations that can be discovered only at the macro resolution level.” In a similar vein, Jurić [34], p. 1 argues that GNV can be a useful tool for historical analyses, noting that it “can discover hidden patterns in history and, therefore, can be used as a window into history.”

Although the tool has several attractive features, it is important to be aware of its limitations. Younes and Reips [20] note that there has been an ongoing debate about the pros and cons of using GNV for research purposes. Younes and Reips [20] point out that even though GNV is increasingly accepted as a research tool, “critics justifiably do not
grow tired in highlighting potential problems that may challenge the reliability of existing results” (p. 2). In particular, Younes and Reips [20] note that there have been concerns about the accuracy of the results from GNV.

Another concerning issue is more fundamental and relates to the premise that “books are a tangible and public representation of culture” [27]. According to Greenfield [26], p. 1722, GNV “allows researchers to quantify culture across centuries.” However, Younes and Reips [20], p. 14 point out that the assumption that “print culture represents culture as a whole, is certainly an undifferentiated view that may not always hold.”

There are also some potential issues related to the fact that GNV does not take into account variations in the popularity and readership of different books since the “appearance of a word in a text that was read by 10 people is given the same weight as its appearance in a best-selling book that was read by millions” [29], p. 5. Related to this, it has been pointed out that context is generally lacking. In a study of the evolution of marketing, Kumar and Sahu [22] noted that a potential weakness is that words are used in a context that is different from what is being researched.

Younes and Reips [20] point out that another critique is related to the dominance of scientific literature in the GNV data set. Ignatovich [21] notes that there could be other factors that influence the quality and usability of the data, such as variations in the quality of scanning, and a dominance of English-language sources.

3. Management Fashion Research

This section provides a brief overview and introduction to the field of management fashion research. Providing a brief background is needed to contextualize the use of GNV in this research area.

The study of “management fads and fashions” emerged in the early 1990s, spearheaded by researchers such as Eric Abrahamson and Alfred Kieser [35–37]. In the 30 years since then, it has grown into an important research area in management and organizational behavior [38–43].

Management fashion researchers tend to view the diffusion and popularization of management concepts and ideas as occurring within a management fashion market shaped by supply and demand side forces [37]. The supply side of management fashions consists of different types of fashion-setting actors (e.g., consultants, gurus, software vendors, and business school professors) who propose and propagate new concepts and ideas, while the demand side of management fashions consists of organizations and managers that are consuming and applying these concepts and ideas in practice.

A key focus of management fashion research is mapping trends in the popularity of different management concepts and ideas. In management fashion research, the lifecycle of concepts and ideas is typically illustrated using a bell-shaped curve [35,37,39]. The assumption is that fashions are fleeting and transient phenomena, which after a take-off and growth phase, become widely popular and fashionable, before going out of fashion and waning away.

Over the last decades, many management concepts and ideas have been introduced, illustrated by several guides providing overviews to practitioners [44–48]. Well-known examples of (once) fashionable management concepts and ideas include Knowledge Management [49,50], Balanced Scorecard [51], and Lean [4,52]. However, it should be noted that not all new management concepts and ideas succeed in attracting attention and becoming widely adopted and diffused. In many cases, they instead fade away rather quickly, leaving few traces behind [53]. Therefore, Jung and Kieser [54], p. 329 have proposed a definition of management fashions as those “management concepts that relatively speedily gain large shares in the public management discourse.”

Management fashion research has been criticized in different ways. One often-heard criticism is that most studies focus heavily on the supply side of fashions, and largely take supply side discourse (e.g., citations in newspapers and professional magazines) as a proxy for actual take up of the fashion by demand side organizations [55,56]. Research has shown
that there is not always a “co-evolution” of supply side activity and demand-side use [57]. For example, there might be much talk about new concepts and ideas without them being widely adopted by the demand side. Therefore, it can be difficult to study the actual impact of management fashions [57]. Data about the supply side of management fashions, such as how frequently fashionable concepts are mentioned in books, journals, or magazines do not provide a complete picture of their impact [10].

Therefore, studying the impact of management fashion often requires a combination of data from both the supply and demand sides of the management fashion market. It has been suggested that management fashion researchers should combine different research methods and data sources to provide a more complete and balanced picture of the impact of management fashions [8–10].

As noted in the introduction, this paper provides a discussion and illustration of how GNV can be applied in the study of management fashions. To date, the tool has only been used for illustrative purposes in a few dispersed studies, and the possibilities and implications of using GNV have not been thoroughly addressed in the management fashion literature. As will be discussed in greater detail in the following, GNV could be used as one of the colors on a palette used to “paint an overall picture” [57], p. 638 of the impact of a particular management fashion.

4. GNV in Management Fashion Research

This section provides some examples of how researchers could use GNV to study the evolution of management concepts and ideas. Examples of management concepts and ideas used in this section (for example, Big Data, or TQM) are often mentioned in the Bain & Company biannual survey of popular management tools and trends [58] or Gartner’s annual hype cycle report for emerging technologies [59].

GNV allows researchers to search for terms such as “Big Data” or “Balanced Scorecard”. In the following, several examples of how GNV can be applied in management fashion research are presented.

Figure 1 shows a screenshot of a graph containing the evolution of the term Big Data over the period 2008–2019. Big Data is a term that has received a lot of attention in the business and organizational world during the last 10–15 years [60,61]. The Figure 1 shows that there was little mention of Big Data in books before the late 2000s, which corresponds to findings in other studies of the Big Data concept [61].


Similarly, Figure 2 shows a screenshot of a graph of the term “budgetary control” over the period 1910–2019. It shows that the term started taking off around 1920, with the influential book by James McKinsey published in 1922 [62]. The use of the term peaked during the 1950s, before starting to decline.
In the last couple of decades, budgetary control has been debated and criticized by the “Beyond Budgeting” movement [63,64]. Figure 3 displays a comparison of the use of the two terms “budgetary control” and “beyond budgeting” over the period 1997–2019. While budgetary control is still more frequently used, the two terms have opposite trajectories, with Beyond Budgeting slowly becoming more frequently used in books. Still, the Figure 3 shows that budgetary control is still a much more frequently used term than its newer alternative.

GNV also allows for comparisons of the evolution of several terms over a specific time period. This allows for a comparative view of the evolution of management fashions. Figure 4 shows the curve for three of the most well-known management fashions of the 1980s and 1990s: Total Quality Management (TQM), Benchmarking, and Business Process Reengineering (BPR). Of the three, only TQM’s GNV curve resembles a typical bell-shaped curve depicted in management fashion theory [37], but the curve also shows that it has not completely disappeared from books.

Although benchmarking’s popularity has flattened out and declined a bit, it remains at a much higher level than is the case for the other two terms. This pattern resembles what was found in the study by Madsen et al. [65] and Bain & Company’s study which shows that benchmarking continues to be one of the most widely used management tools in organizations worldwide [58]. The most surprising observation in Figure 4 is the BPR concept’s evolution. Although BPR was one of the most important management fashions of the 1990s [66,67], it did not reach nearly as high a level of dissemination in books as the other two concepts.

The last example will look at how GNV can be used to study the evolution in the relative popularity of different variations of the same management concept. In the literature on management fashions, several authors have noted that concepts become adapted and "translated" by supply side actors [68]. For example, there are multiple variations of the original Lean Production concept including, but not limited to, Lean Construction and Lean Software Development (see, e.g., [69]). A similar development can be seen in the case of Industry 4.0, where researchers have documented a multitude of variations of the original concept [3,70].

Figure 5 shows the evolution in the occurrences of four terms related to Lean: "lean production", "lean manufacturing", "lean construction" and "lean software development". As can be seen, the production and manufacturing variants stand out in terms of their use in books. However, it is also interesting to note that the lifecycles of these Lean-related terms diverge. Although Lean Production was dominant during the early to mid–1990s, Lean Manufacturing has become relatively more influential since the turn of the millennium, and the two curves have converged to a large extent.

5. Discussion

In this section, the possibilities and limitations of applying GNV in management fashion research will be discussed in more detail.

5.1. Possibilities

One of the most obvious advantages of GNV is that it is user-friendly, readily available, and free to use. Cost-effectiveness is particularly important, since collecting data on management concepts and ideas can be difficult and costly in terms of time and resources.
Another advantage of using GNV is that it is unobtrusive and objective, and not subject to the biases and limitations of individual researchers [25].

GNV can also be used to help in quantifying historical trends. It provides a visualization of data about the supply side of management fashions, specifically books. Furthermore, it provides a historical view that allows for graphic illustrations of how specific terms and content have emerged, risen, and fallen over specific periods.

Researchers have noted that GNV can be considered a “very high level trend analysis” [24]. Although this could also be viewed as a disadvantage, the bird’s eye perspective provides useful information in the context of management fashion research, particularly to understand the historical evolution and changes in the popularity of management concepts and ideas.

5.2. Limitations

GNV has several shortcomings that it is important to be mindful of. As pointed out earlier, GNV is highly accessible and easy to use. However, some might argue that the analyses are rather “quick and dirty” [33]. As noted earlier, the focus is on high level historical trends [24] and hence it sheds no insight into how these terms and concepts are applied at the micro level.

In relation to management fashion research, a key shortcoming of GNV is that it only provides data on management fashions from a supply side perspective. It does not provide any insight into the extent to which these books have been read and applied by actors on the demand-side of the market.

As several commentators have pointed out in the management fashion literature, most empirical studies of management fashions tend to draw mostly on data from the supply side [55,56]. Print-media indicators are commonly used to get insight into how the popularity of management concepts and ideas evolves [10,71]. However, this type of research has drawbacks [9,10,72], the most important being that such supply side data “must not be conflated with the actual use of concepts by organizations” [10], p. 827. Since GNV draws on supply side data, the tool should be used in conjunction with other research methods and data sources [8,9].

Furthermore, the relative importance of books in the dissemination and diffusion of management fashions has changed over time. While books were perhaps the most important channel for introducing and disseminating new concepts and ideas decades ago [15], this is not necessarily true in the age of digital and social media, since these new platforms have changed the landscape of management fashions [17,18,38]. Social media platforms (e.g., LinkedIn, Twitter) have in recent years become more influential relative to traditional management books [17].

Therefore, GNV is less applicable for examining “young” management concepts and ideas (i.e., early in the lifecycle), such as Industry 5.0 at the present time [73]. At the same time, GNV can still be useful for examining the evolution of mature management concepts and ideas that have been around for decades, such as TQM and Economic Value Added [12,74].

5.3. Evaluation

This brief discussion has shown that GNV has both possibilities and limitations when applied in the context of management fashions. GNV only captures one aspect of the supply side activity around management fashions, that is, the extent to which concepts are mentioned in books. Therefore, the tool only offers a partial picture of the impact of a fashionable management concept. To use the painting metaphor, GNV data are only one color in the palette used to paint a picture of the impact of management fashions.
6. Conclusions

6.1. Concluding Comments

The aim of this paper has been to explore the possibilities and limitations of using the tool GNV in management fashion research. We have provided some illustrations of how GNV can be used in research on specific management concepts and ideas such as Big Data, Beyond Budgeting, and Lean. In general, the paper has shown that GNV has both strengths and weaknesses when used in the context of management fashions.

The main conclusion is that management fashion researchers can apply GNV data as one of the colors in the palette used to paint a picture of the impact of a management fashion. Specifically, GNV can provide useful insight into macro-level historical trends about management fashions. More generally, it provides a “window into history” [34] that can be useful for retrospective analyses. It provides a big historical picture that crystallizes the phases and stages in the lifecycles of management concepts and ideas.

6.2. Limitations and Future Research

Like any piece of research, the current paper has limitations. This paper should be seen as an early attempt to discuss the usefulness of GNV as a research tool and data source in management fashion research. Being conceptual and speculative, the limitations of the paper should be carefully considered, and the conclusions should be viewed as tentative.

It is especially important to stress that the insight that can obtained from GNV is limited, and this type of data should only be one piece of the puzzle when researching management fashions. Other research methods and data sources should also be considered [8,9], and in the future, researchers should explore how data from GNV could be used in conjunction with different types of quantitative and qualitative data on the supply and demand for management fashions. Comparing the insights from GNV with other types of data about management fashions could also corroborate the conclusions about the use of this tool.

Another related issue is the decreasing importance of books as a channel for the diffusion of management fashions in the age of digital and social media [17]. Therefore, in the future, researchers need to be aware of these changes in the management fashion landscape. Although GNV might be highly useful for retrospective historical analyses of classic management fashions, in particular how certain terms or phrases originate and expand [11,22], it could perhaps be less useful for shedding the light on future management fashions that possibly will be launched and diffused mainly via digital and social media [17].

As the review in this paper has shown, despite its shortcomings, GNV has several applications for understanding the historical evolution of management fashions. Therefore, this paper should be seen as an invitation to further explorations of the applicability of GNV in the context of research on management fashions.

Author Contributions: Conceptualization, D.Ø.M.; methodology, D.Ø.M.; formal analysis, D.Ø.M.; investigation, D.Ø.M. and K.S.; writing—original draft preparation, D.Ø.M. and K.S.; writing—review and editing, D.Ø.M. and K.S.; visualization, D.Ø.M.; project administration, D.Ø.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.
32. Pechenick, E.A.; Danforth, C.M.; Dodds, P.S. Characterizing the Google Books corpus: Strong limits to inferences of socio-cultural and linguistic evolution. PloS ONE 2015, 10, e0137041. [CrossRef]
61. Madsen, D.Ø.; Stenheim, T. Big Data viewed through the lens of management fashion theory. Cogent Bus. Manag. 2016, 3, 1165072. [CrossRef]


73. Madsen, D.Ø.; Berg, T. An Exploratory Bibliometric Analysis of the Birth and Emergence of Industry 5.0. Appl. Syst. Innov. 2021, 4, 87. [CrossRef]