



BSc Creative Technology

Thesis paper

**Analyzing the Role of
Team Effectiveness in
Educational Outcomes: A
Bibliometric Analysis**

Daffa Paramitha

Supervisor: Faiza Bukhsh, Yeray Barrios-Fleitas

July, 2024

Department of Creative Technology

Faculty of Electrical Engineering,

Mathematics and Computer Science,

University of Twente

Contents

1	Introduction	1
2	Methods	5
2.1	Identification	5
2.2	Screening	6
2.3	Selection of analysis techniques mapped to the research questions . .	8
2.4	Selection of Large Language Model (LLM) for qualitative data analysis	10
2.5	Protocol for Generative AI use in analysis	12
3	Results	17
3.1	Descriptive analysis	18
3.2	Co-authorship analysis	21
3.3	Landmark analysis	24
4	Conclusion	27
5	Discussion	29
A	Team effectiveness definition per author	35

B Team effectiveness characteristics studied per article	42
C Acknowledgements	50

Abstract

Team effectiveness is a highly researched topic, with many use cases. This paper focussed on the educational context and answered the question of how team effectiveness is defined and what team effectiveness characteristics are studied within the educational context. A novel method of analysis was used and developed for this research; a semi-automatic extraction of data where the efforts of a human researcher and an LLM are combined. A protocol was developed and validated for this purpose. The conclusions obtained from this research also infer about the use of an LLM for this type of research.

Keywords: bibliometric, analysis, education

Chapter 1

Introduction

Team effectiveness in educational settings is pivotal as it significantly impacts critical educational outcomes such as collaborative learning, student satisfaction, and overall academic performance. In the modern educational panorama, where teamwork and collaborative skills are increasingly valued by employers, effective team dynamics in educational settings not only enhance learning but also prepare students for future professional environments. This context underscores the necessity to cultivate and enhance team effectiveness to foster environments where collaborative skills and cooperative learning are optimized.

Achieving optimal team effectiveness in educational settings is challenged by a multitude of factors. Firstly, there is no universally accepted definition of team effectiveness, which complicates efforts to evaluate and measure it consistently. This lack of consensus often leads to ambiguity in assessment methodologies and the criteria used to determine what makes a team effective in varying educational contexts. Another

significant challenge is the heterogeneity in student populations. Team effectiveness is inherently multidimensional, influenced by a myriad of factors related to team members' characteristics. These include demographic attributes, personality traits, conflict resolution styles, and preferred team roles, but also emergent states and mediators such as trust, cohesion, and collective efficacy dynamically influence team effectiveness throughout the team's life cycle. The unique combination of individual characteristics among team members shapes the overall diversity of the team.

However, while research on team diversity is extensive, its findings are inconclusive, showing positive, negative, and mixed outcomes. The positive effects are based on the premise that diversity brings a greater variety of resources, perspectives, and knowledge to the team, facilitating problem-solving, information sharing, creativity, and innovation. On the contrary, the adverse effects are based on the premise that diversity generates a greater social and psychological distance between team members, which hinders communication, cohesion, trust, and team performance. When forming student teams, teachers must decide which characteristics to consider in order to create more heterogeneous or homogeneous groups, depending on the specific objectives they aim to achieve.

Over time, theories of team effectiveness in education have evolved from simple structures to complex models that consider continuous feedback, diversity, and interdependence. Initially, research in team dynamics focused on Input-Process-Output (IPO) models, which described teams in terms of input variables such as member skills, team processes such as communication, and outputs such as team performance [1] [2] [3]. This model established a foundation for understanding how

different factors could influence a team's effectiveness. Over time, it was recognized that teams are not static and that their states and processes evolve and feed each other. Input-Mediator-Output-Input (IMOI) models emerged to capture this continuous dynamic, highlighting the role of emergent states such as trust and cohesion, which are vital in educational settings to foster effective collaborative learning [4]. Van Knippenberg's team diversity theory [5] and Marks and Mathieu's team self-regulation theory [6] offered new perspectives on how to manage team diversity and adaptability. Van Knippenberg suggested that diversity could be both a source of creative richness and a challenge to cohesion. On the other hand, Marks and Mathieu focused on how teams adapt and adjust their strategies to achieve goals, a crucial concept for student teams facing varied academic tasks. Cohen's cohesion theory [7] and Johnson and Johnson's task interdependence theory [8] delved into how task cohesion and structure can facilitate or impede team effectiveness. In parallel, Salas and Burke's "BIG FIVE"-model highlighted the importance of leadership, mutual support and adaptability in team dynamics, essential components for collaborative education [9]. Tuckman's stages of team development [10] [11] and Hackman's team dysfunctionalities [12] offered frameworks for understanding and diagnosing the challenges of teams throughout their life cycle. These theories are especially applicable in education, where the teams formed can last six months or annually.

Finally, Biggs' SOLO taxonomy [13] is not a theory of team dynamics per se, but its focus on assessing the complexity of understanding can be invaluable for designing tasks that promote deep, collaborative learning in student teams. Given the

challenges of the lack of a consensual definition of team effectiveness and the complexity inherent in the multidimensional and diverse nature of teams, and following the example of Barrios et al.'s analysis [14], which examined 6051 articles on team effectiveness in various contexts, this research aims to use the same database to specifically focus on studies related to the context of education. A new bibliometric analysis will be carried out to identify the most influential publications and classify the articles according to their topic. Artificial intelligence techniques will then be used to process and synthesize large volumes of academic texts to semi-automatically extract answers to the following crucial research questions:

1. How is team effectiveness defined in the context of education?
2. What team effectiveness characteristics are studied in the context of education?

The structure of the paper is as follows: Section 2 describes the methodology, focusing on how the interventions were tailored to diverse educational contexts and team compositions. Section 3 presents detailed results from both quantitative and qualitative analyses. Section 4 concludes with a summary of the key insights and discusses the contributions to educational practice and theory. Finally, Section 5 discusses these findings within the broader theoretical context and explores their implications for enhancing team effectiveness in education, along with outlining future research directions.

Chapter 2

Methods

This section outlines the systematic and rigorous methodology employed for the collection, analysis, and interpretation of data from existing literature. It details the origin and structure of the data, along with the inclusion and exclusion criteria used to select pertinent publications. Furthermore, we specify the combination of bibliometric and artificial intelligence techniques chosen to answer the research questions of this study.

2.1 Identification

This study utilized the database of publications on team effectiveness compiled in the bibliometric systematic review by Barrios et al. [14]. This comprehensive database includes a total of 6,051 publications related to team effectiveness across various contexts, covering the period from 1992 to 2022. The extensive scope and depth of this database enabled us to conduct a second phase of coding to specifically extract

publications framed within the educational context.

2.2 Screening

The eligibility phase of this study was conducted in accordance with the PRISMA guidelines, as detailed in the work of Barrios et al. [14]. During the screening phase, the primary focus was on identifying papers from the original database that were framed within the educational context. Titles and abstracts were reviewed to determine the context, categorizing them into organizational/business, health-care/emergency, educational, and sports contexts. The inclusion criteria for identifying literature relevant to the educational context were as follows:

1. Research papers explicitly stating in the title or abstract that their research was conducted in, or for, educational context were always included.
2. Research papers using scholar-based demographic elements as dataset, including gender, cultural, or age diversity in (large) classrooms or courses.
3. Research papers discussing multiple ad niche contexts, such as "sports education" or "nursing education", were included only if they generalize the team effectiveness elements to broader situations. For example, a study on a nursing school discussing how different nurse personality types work together in a team was included, whereas a study on the use of medical tools in different nursing teams was not.
4. Author's keywords were considered. Papers that remained ambiguous in context after reading the title and the abstract were included only if the author's

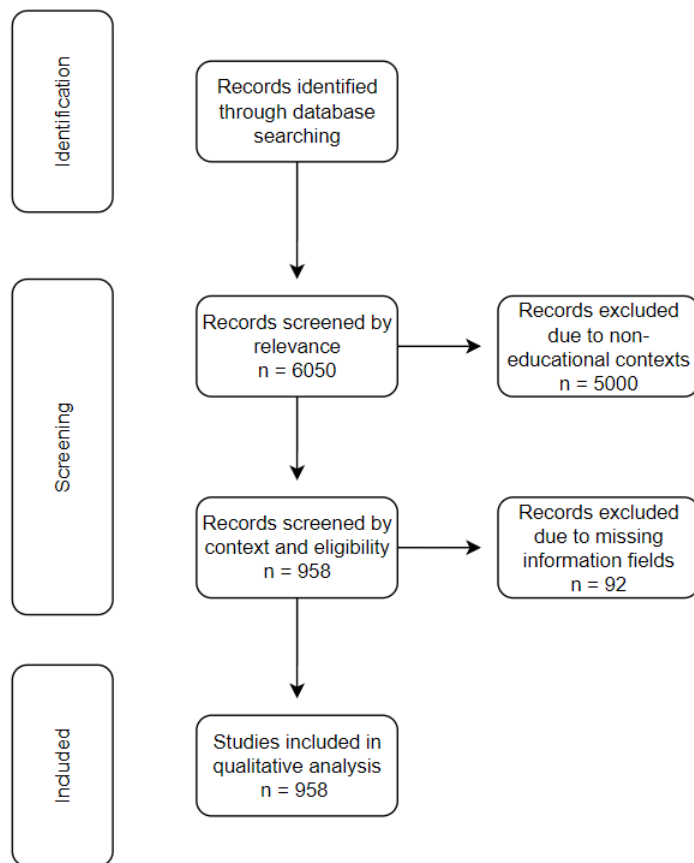


FIGURE 2.1: PRISMA diagram of the coding process

keywords included terms such as, but not limited to "education", "classroom", or "students".

If any doubt remained after applying all context clues, the paper was coded as "Not sure". Following a third round of screening, all papers coded as "Not sure" were re-evaluated using the same context clues and inclusion criteria. Those still doubtful in context were read in full, specifically looking for the context clues required. After all screening processes, this stage in produced $n = 958$ papers included within the database for analysis. A PRISMA diagram of the coding processes can be found in Figure 2.1.

2.3 Selection of analysis techniques mapped to the research questions

The different research questions required different analysis techniques to be performed to come to an answer. The reasoning and steps taken will be described in the following section.

For RQ1: *How is team effectiveness defined in the context of education?*, several assumptions were made in order to retrieve the required data from the database. First, it was assumed that those who are experts in the field of team effectiveness would be the ones to turn to when looking for a proper definition of the term. The second assumption was that those who have written several articles on the subject would be considered experts. On the basis of these two assumptions, a Lotka's Law analysis was done in Bibliometrix [15], of which the result is a Lotka's Law graph. This graph can be seen in Figure 3.4 in Section 3.

Lotka's Law shows what percentage of authors have written a certain number of papers. The top 1% of authors within the team effectiveness field were then identified, which are the presumed experts, after which their most recent papers were used for analysis in a Large Language Model (LLM). The number of papers to analyze was chosen to be 3 per author, because the Lotka's Law graph showed that the top 1% of authors have written at least 3 papers in their career, thus there would be an equal number of papers used for each author. The decision to only use the most recent papers published by the expert authors was based on a third assumption; these experts have presumably been involved in the field for a longer period of time,

so the definition they used for their first publication may have changed significantly compared to their latest publication. The variation and evolution of the definition of team effectiveness used by the expert authors is expected to be less drastic within their latest publications. Considering all these elements, the set of papers required to answer RQ1 was determined.

For RQ2: *What team effectiveness characteristics are studied in the context of education?*, different assumptions were made to retrieve the necessary data. This time, it was assumed that the most productive papers indicated the amount of interest in the characteristics studied within those papers. The productivity was calculated using the following formula:

$$\text{productivity} = \frac{\text{total number of citations}}{\text{amount of years in active publication}}$$

For RQ2, the distinction between positive and negative citations was disregarded, and the citations were taken as a whole. As the question looks to answer the most studied characteristics, even the negative citations can be used as an indicator of interest in the article, and consequentially, the characteristics studied within the article.

After ranking the database entries by highest productivity, the top 5% was taken from the list. The decision to only take the top 5% was motivated by the time constraints of this project. The papers that were obtained by this method were then used for the semi-automatic analysis by an LLM. A visual representation of the steps taken to answer RQ1 and RQ2 can be seen in Figure 2.2 below.

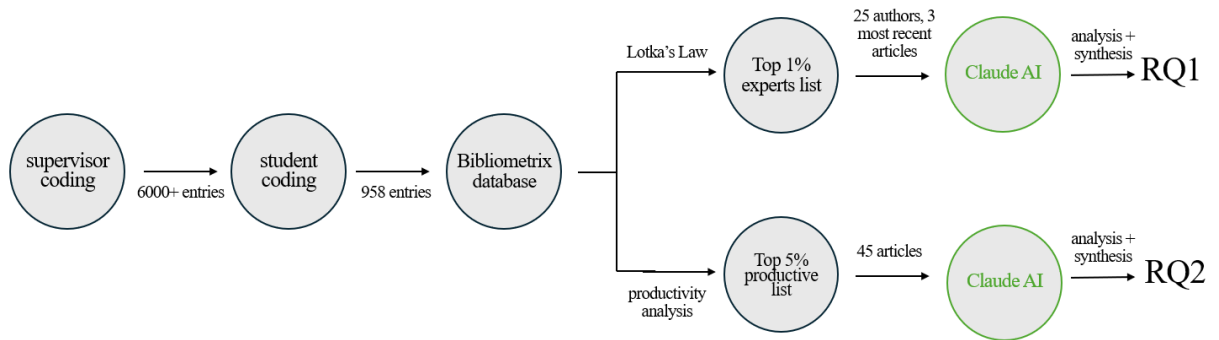


FIGURE 2.2: Diagram of the steps taken to answer each research question

2.4 Selection of Large Language Model (LLM) for qualitative data analysis

A Large Language Model (LLM) is a type of artificial intelligence (AI) that is capable of, among others, recognizing and generating texts. These LLMs are built on machine learning and trained using massive amounts of data [16]. LLMs are able to recognize and interpret human language because of the amount of data they have been trained on; they are able to recognize patterns in human language and use those identified patterns to generate a response. The unique responses that the LLMs are able to generate also makes them known by another name; Generative AI. There are many LLMs for public use today, such as OpenAI's household name ChatGPT, or the newly launched Mixtral [17], all of which have different strengths and features that can be useful in various use-cases. Because of this diverse choice of tools, choosing the right LLM for this analysis was an important step to ensure valid research. Based on current popular use of LLMs, ChatGPT, Gemini and Claude were the three proposed as starting points for this project.

To start the decision process, 11 foundational papers from the database were read

and analyzed by a human researcher. These papers were obtained through a co-citation analysis in Bibliometrix [15]. These analyses were then used as checks for the validity of the answers given by the LLMs in the exploratory stage. The three LLMs were given similar prompts, along with 3 papers for analysis. This step was performed several times for all three LLMs, after which the answers given were compared and these considered when making the decision on which LLM to continue this research with.

First, ChatGPT only allows for PDF upload on the pro version of the AI. This already formed a constraint as the text within the PDFs had to be extracted and fed manually. Some texts also had to be divided into several messages to comply with the character limit in the ChatGPT chat. ChatGPT “hallucinated” and gave answers that could not be traced back to any of the papers it was given. It did this sporadically, so some parts of the answer would be accurate, but others would be completely wrong. Furthermore, it struggled with “remembering” that it had been given several papers to analyse, since it focused on a single paper instead, and did this for several iterations.

For Gemini, the most time-consuming task was the copy and pasting of the texts within the obtained PDFs. Gemini does not allow for PDF uploads in either free or pro version, thus the text within those files had to be extracted and fed manually. Similar to ChatGPT, this formed a constraint considering the number of papers that needed to be analyzed. As for the synthesis of answers, Gemini was able to give somewhat accurate answers. However, because the texts had to be divided into several messages because of the character limit, Gemini would sometimes “forget” a

previous section of the paper that was given. This made it difficult to analyze texts in one go, and because it would “forget” about sections, some of the answers were not completely accurate.

Claude seemed the most promising from the start, as it was designed to read and analyze large amounts of texts, and this was clear in its answers. Claude, unlike ChatGPT and Gemini, allowed for PDF uploads regardless of version. At the start of this research, Claude only allowed for a maximum of 5mb file size uploads, but towards the end of the research, this amount was updated to 30mb. The ability to upload PDFs was immediately a winning characteristic, which probably influenced the level of accuracy of the answers the AI was able to give. Claude was able to accurately analyze the papers given, even when the conversation was longer, and it did not “hallucinate” answers or references, unlike the other LLMs. Sometimes, Claude would even suggest that the researcher put in a prompt that could enable more in-depth analysis, something that the other LLMs did not do. Considering all the elements, it was an obvious choice to continue with Claude for the semi-automatic analysis.

2.5 Protocol for Generative AI use in analysis

The aim of using LLMs in this study was to support researchers to identify the key information across a significant amount of papers necessary to provide an answer to the research questions proposer. Given the novelty of this approach, a rigorous protocol was proposed to ensure the validity of the findings. This protocol contemplates the design and refinement of search prompts adjusted to the purposes of

each research question, as well as the manipulation of Claude answers to successfully conclude with a reliable answer. The following subsections delineates the systematic approach adopted for employing an LLM in a semi-automated analysis, structured around specific research questions.

For the first research question, *How is team effectiveness defined in the context of education?*, Claude was needed to be able to identify/interpret the way in which different authors define team effectiveness in their most recent papers. As a first approach, a highly specific and comprehensive prompt was designed that was meant to search for the required information in certain parts of the text, giving rise to the following query:

"Please analyze the 3 research papers I will provide, which are all authored by the same individual and focus on team effectiveness in the context of education, published around 5 years ago. Specifically, I need you to extract the definition of team effectiveness according to the author. This definition is likely to be found in one of the following sections: introduction, literature review, background, state of the art, or methodology. Once you have analyzed all the papers, please provide me with the most comprehensive and accurate definition of team effectiveness as stated by the author. Additionally, please identify and list the most common keywords or phrases related to team effectiveness that appear across all the papers."

The specification of the sections to search, as well as other explanatory components of the prompt such as that the authorship of the papers comes from the same

author resulted in Claude often answering with artificially generated results when the answer was not specifically found in the indicated sections. Several iterations later, the prompt was simplified by removing all the unnecessary context, resulting in more consistent and accurate results:

"Based on the provided research papers, interpret the definition of team effectiveness according to the author. Also provide your sources, tell me the following things: -which paper did you base your analysis on? -What page did you find that information?"

To mitigate the inherent risk of high variability associated with this type of technology [18], it was decided to execute the prompt five times for each participant. The responses generated by Claude, focusing on authors ranked in the top 1%, were recorded in a spreadsheet. Random inspections were conducted to manually verify the accuracy of Claude's responses against those extracted by a human researcher. These inspections revealed that Claude consistently extracted the required information with high precision and minimal variability across iterations. Finally, a synthesis analysis was conducted to develop a consistent definition of team effectiveness based on the extracted information. For the full table of definitions, please see B.

For the second research question, *What team effectiveness characteristics are studied in the context of education?*, the starting point was the prompt used in RQ1. That prompt was modified to fit the needs of RQ2, which read:

"I am providing you with a research paper. Based on the provided re-

search paper, please tell me what team effectiveness characteristics (variables) are studied. Also provide your sources, tell me the following things: -Which paper did you base your analysis on? -What page did you find that information?"

With the prompt described above, it was noted that Claude was not able to differentiate between a passing mention of a characteristic and an actual study of the characteristic. An example of this is a paper titled *A functional model of hierarchy: Why, how, and when vertical differentiation enhances group performance* [19], where the authors wrote about the different ways in which hierarchy can affect team effectiveness. In essence, the main team effectiveness characteristic studied in that paper was hierarchy, but Claude only picked up on the team effectiveness characteristics that hierarchy had an influence on, such as team performance or team satisfaction. Thus, it became obvious that the prompt had to be refined to find the primary characteristic studied, and not the secondary characteristics that can be influenced by the primary characteristic. To do this, Claude itself was asked to design the prompt, which was as follows:

"I am trying to analyze a paper to extract the primary team effectiveness characteristics that are studied within the paper. I've used the following prompt before: "I am providing you with a research paper. Based on the provided research paper, please tell me what team effectiveness characteristics (variables) are examined. Also provide your sources, tell me the following things: -which paper did you base your analysis on? -what page did you find that information?" Could you please improve my prompt so

that I can more accurately extract the information that I need?"

Using this prompt, Claude answered:

"Please analyze the provided research paper and identify the primary team effectiveness characteristics (variables) that are directly examined or manipulated in the study. Focus on the key variables that the researchers are testing or investigating, rather than secondary characteristics that may be influenced by these primary variables. In your response, please provide the following information:

- 1. A list of the primary team effectiveness characteristics (variables) examined in the study.*
- 2. The title of the paper you based your analysis on.*
- 3. The specific section(s) and page number(s) where you found information about the primary team effectiveness characteristics. If the paper does not explicitly examine any team effectiveness characteristics, please state this in your response."*

The prompt was then tested on several papers; the answers given by Claude were then checked and validated. Once done, it was concluded that this is the prompt that would most accurately provide the information needed to answer RQ2.

Chapter 3

Results

Following the different methods previously described, this chapter will outline the results that were extracted from the techniques performed. First, a general overview is described. Second, the results of the analysis for RQ1 will be provided. Lastly, the results of the analysis done for RQ2 will be put forth as well.

The results from the analysis as described in Section 2 are categorized by three different sections, these being:

1. Descriptive analysis: General figures of the field of team effectiveness in the educational context
2. Co-authorship analysis: Identification of expert authors and semi-automated extraction of their definitions of team effectiveness
3. Landmark analysis: Identification of key publications and semi-automated extraction of the most researched team characteristics

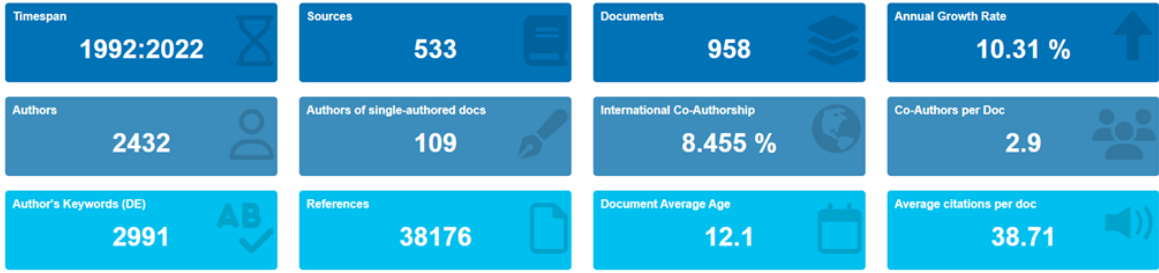


FIGURE 3.1: Overview of the data corpus

3.1 Descriptive analysis

The bibliometric analysis of team effectiveness in education encompasses a comprehensive dataset of 958 documents sourced from 533 different journals, books, and other type of publications, as seen in Figure 3.1. From the overview, it is evident that the corpus of data is extensive, with a wide range of sources contributing to the scholarly dialogue on team effectiveness in education. The dataset spans three decades, reflecting both the historical context and the modern developments in this field. The main types of documents within the database include articles, conference papers, reviews, and book chapters. Articles form the majority, accounting for 67.1% of the total, indicating a strong preference for journal publications. Conference papers represent a significant portion, 27.7%, highlighting active participation in academic conferences and the dissemination of research findings through these forums. Reviews and book chapters are relatively less common, making up 4.6% and 0.6% of the total respectively, suggesting areas where more comprehensive studies and contributions could be valuable.

The annual scientific production has shown varied but generally increasing trends over the analysed period, as seen in Figure 3.2. There is a clear upward trajectory

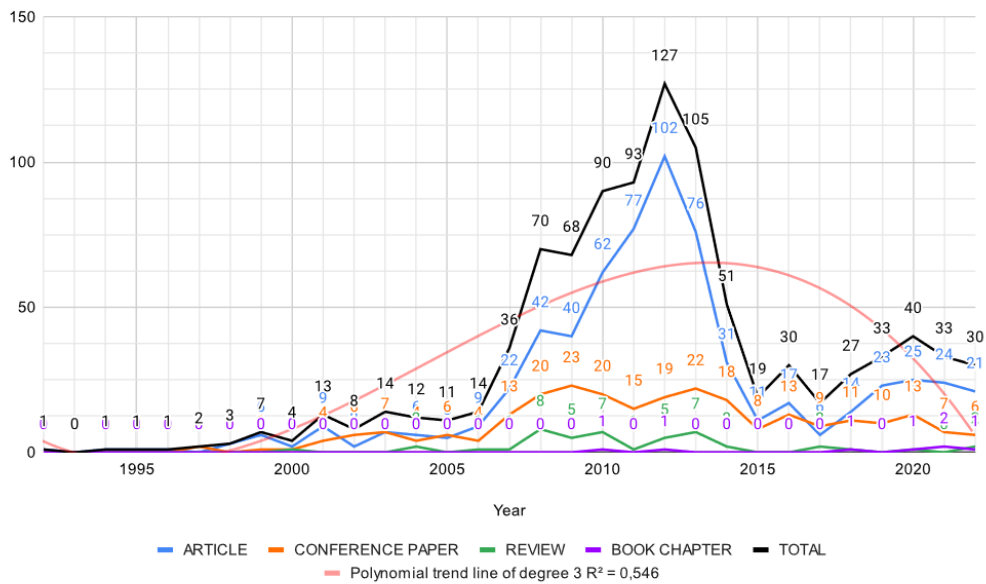


FIGURE 3.2: Scientific production per document type and global trend

in the number of publications, particularly in the years following 2010. This growth indicates a burgeoning interest and intensified research activity in the field of team effectiveness within educational settings. The early years (1992-2000) saw a steady but slow increase in publications, while more recent years have seen a significant surge, underscoring the field's growing importance and relevance.

Applying Bradford's Law to this dataset categorizes the sources into different zones based on the frequency of documents [20]. The analysis identifies the most prolific sources as seen in Figure 3.3, such as "Small Group Research," which emerges as the leading publication with the highest frequency of documents. Conference proceedings, including those from the ASEE Annual Conference and the Frontiers in Education Conference, also feature prominently, reflecting their critical role in disseminating research findings. Journals like the "Journal of Applied Psychology" contribute significantly, albeit with fewer documents compared to conference pro-

R	Source	h_index	TC	NP	PY_start
1	SMALL GROUP RESEARCH	21	1216	39	1992
2	JOURNAL OF APPLIED PSYCHOLOGY	14	3643	15	1994
3	ORGANIZATIONAL BEHAVIOR & HUMAN DECISION PROCESSES	12	1081	12	2001
4	GROUP DYNAMICS	10	334	12	1999
5	COMPUTERS IN HUMAN BEHAVIOR	9	462	9	2007
6	GROUP AND ORGANIZATION MANAGEMENT	9	688	11	1998
7	INTL.JOURNAL OF ENGINEERING EDUCATION	9	214	14	2001
8	TEAM PERFORMANCE MANAGEMENT	9	244	15	2001
9	ACADEMY OF MANAGEMENT JOURNAL	8	1401	8	2008
10	ACADEMY OF MANAGEMENT REVIEW	8	1299	8	2007
11	JOURNAL OF ORGANIZATIONAL BEHAVIOR	8	397	8	2006
12	PERSONNEL PSYCHOLOGY	8	1832	8	1996
13	JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY	7	444	7	2000
14	LEADERSHIP QUARTERLY	7	655	7	2009
15	HUMAN FACTORS	6	1034	6	2008
16	PROCEEDINGS - FRONTIERS IN EDUCATION CONFERENCE, FIE	6	125	17	1997
17	ASIA PACIFIC JOURNAL OF MANAGEMENT	5	214	5	2002
18	INFORMATION AND SOFTWARE TECHNOLOGY	5	119	5	2010
19	INTL.JOURNAL OF CONFLICT MANAGEMENT	5	112	6	2008
20	INTL. JOURNAL OF HUMAN RESOURCE MANAGEMENT	5	115	5	2011

Abbreviations: R = Ranking; TC = Total citations; NP = Number of publications; PY_start = Year of first publication published.

TABLE 3.1: Top 20 most influential sources in team effectiveness in the educational context

3.2 Co-authorship analysis

The results of the Lotka's Law as seen in Figure 3.4, showed that the top 1.6% of authors published at least 3 papers. Following the overview seen in Figure 3.1, the top 1% of authors came to an amount of 24.32, which was then rounded up to 25 authors. The 3 most recent papers written by these authors were then taken from the database and fed through Claude through 5 iterations.

The characteristics that the authors used to define team effectiveness were synthesized based on Claude's answers, and then also tallied, to show similar definitions of team effectiveness across the top 25 authors. A portion of these results can be seen in Table 3.2, to see the full table, please see [A](#).

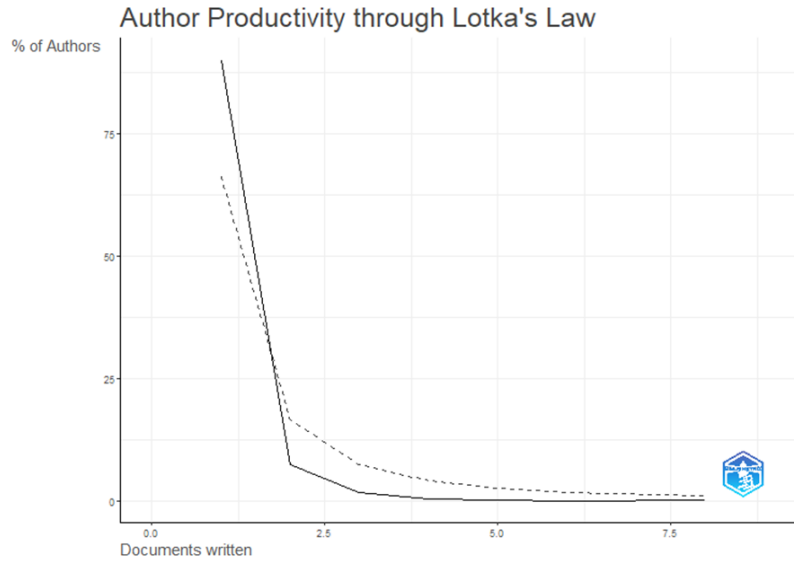


FIGURE 3.4: Lotka's Law graph

The conclusions drawn from this analysis are twofold. First, there was no two of the same answers in any of the iterations of Claude's analysis. Even so, it was easy to pick out the most important points, as Claude mentioned these in every iteration. This was almost like asking 5 people to read the same 3 papers and then ask them to analyze the contents; no two people would give the exact same answer, but the general premise of the texts would be easy to extract based on the common elements in each answer. Checking Claude's references for each iteration also ensured that the information given was valid.

Second, even though different authors define team effectiveness in slightly different ways, 18 authors agree that team performance measures are the most important indicators of team effectiveness. In different settings this could be grades, speed of task completion, or goal accomplishments. Authors also agree that proper communication and knowledge sharing are important factors in team effectiveness, with 10 and 12 authors mentioning this in their definitions, respectively.

Author	Definition	18	3	10	2	12	6	2	3	1	9	6	4	9	7	4	4	2	3
		Performance measures	Exploiting experts' knowledge	Communication	Conflict management	Sharing knowledge	Role and responsibility division	Quality of decisions	Diversity	Diversity in diversity mindset	Team member satisfaction	Adaptability to tasks given	Transactive memory system	Level of collaboration	Team engagement	Team viability	Equal participation	Creativity	Motivational states
Van Knippenberg D	Team performance is used as the key measure of team effectiveness; highly effective teams are those that achieve high levels of performance on their tasks and objectives	x							x										
Salas E	Team effectiveness is measured in the way team members collaborate, communicate and share knowledge with one another, emphasizing that a well-trained team is highly effective.			x		x								x					
Bonner BJ	Effective teams are those who recognize and exploit each member's expertise, developing transactive memory systems by transferring relevant knowledge to each other. Coordinating this, along with <u>taking into account</u> the longevity of the team, makes for a highly effective team as members can coordinate learning and predict task difficulty easily.		x	x		x						x	x			x			
Cheng A	Team effectiveness is highly impacted by the defined roles and the responsibilities thereof within the team. Effective teams comprise of those who know what their role is in the team and takes responsibility for the tasks that comes along with that role. Situational awareness and strong communication are also a <u>big influencing factors</u> in how effective a team is.			x		x	x					x							

TABLE 3.2: Table of team effectiveness definition as defined by author. For full table, see Appendix A

3.3 Landmark analysis

To answer RQ2, a landmark analysis were done. Table 3.3 shows the some of most productive papers. For the full table, please see [B](#). Initially, the top 5% were taken, which amounted to 45 articles. However, after analyzing these through Claude, this amount had to be reduced to 23. An explanation to this phenomenon will be given shortly.

From this data, it can be concluded that, first of all, Claude struggled with analyzing non-empirical papers, such as literature reviews or methodology overviews. This was apparent through the answers that Claude gave, and even through several iterations of prompts, it could not give accurate answers to the questions that were posed. This matter was detected early, and although a manual reading of the non-empirical papers averted some of the issues caused, it was then concluded that Claude indeed had good reason to struggle with these non-empirical papers. Some of the articles did not do team effectiveness studies at all, yet were included in the most productive papers because of a paragraph or two that mentioned something important about team effectiveness, but the general topic of the article would not be related to the topic at all. Further, in the non-empirical reviews, it was hard to determine how and which characteristics can be classified as being "studied", as a literature review would not necessarily study specific characteristics, but would only mention them in passing as an element of team effectiveness. Thus, it was then decided to cut the list of papers from 45 to 23, as to give the most accurate results as possible about the characteristics studied.

Second, the most studied team effectiveness characteristics in the educational context are team performance and team satisfaction, with 12 and 8 papers studying these characteristics respectively. Though that may not seem like a large number, compared to the many characteristics that were only studied in only one paper, this is quite a feat. Moreover, the number of characteristics studied within the educational context are vast. Within the top 23 most productive papers, a total of 26 characteristics were studied. Since these papers are the most productive, this means that the characteristics studied within it are of most interest within the educational context.

Author	Paper title	12	8	3	3	3	3	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
		Team performance	Team satisfaction	Creativity	Information elaboration/sharing	Conflict	Leadership	Hierarchy	Cohesion	Transactive memory systems	Communication	Closure of teamwork	Centralization of teamwork	Subgrouping of teamwork	Team viability	Teamwork processes	Team outcomes	Social loafing	Cultural diversity	Goal orientation	Trust	Quality of group decisions	Team diversity	Openness to experience	Psychological safety	Epistemic motivation	Social motivation
Crawford ER	A Configural Theory Of Team Processes: Accounting For The Structure Of Taskwork And Teamwork											x	x	x													
Halevy N	A Functional Model Of Hierarchy: Why, How, And When Vertical Differentiation Enhances Group Performance							x																			

TABLE 3.3: Team characteristics studied per article. Please see B for full table.

Chapter 4

Conclusion

To conclude, team effectiveness has been researched and studied extensively. Though the term “team effectiveness” still has an ambiguous meaning, this paper has shown that many authors do agree in many cases that team performance and team satisfaction, communication, and knowledge sharing are of utmost importance in defining the term. Moreover, the most studied team effectiveness characteristics, team performance and team satisfaction, support that these characteristics are important in defining and measuring the effectivity of a team. As for a conclusion about Claude, though it can be argued that using such AI tools have no place in research, it should also be noted that by using this tool, it has enabled a new type of method within the bibliometric analysis field to retrieve large quantities of data that go beyond the meta-data. The protocol made for the use of a Generative AI in this project could form a starting point for further research within the bibliometric field. Thus, the contribution that this paper carries is twofold; one side contributes to the bibliometric analysis of team effectiveness in educational context, the other side if

the contribution to the bibliometric analysis field by introducing a new method of semi-automated data extraction.

Chapter 5

Discussion

As this project came to an end, the researchers were suggested to use a “RAG” system, which was something unfamiliar to the project. RAG is an abbreviation for Retrieval-Augmented Generation, where an LLM would be attached to an authorized database on top of its trained knowledge to ensure correct answers and eliminate the possibility of the AI hallucinating. This is something that could have increased the validity and accuracy of the results of this research. Thus, this is something that should be considered should further research be carried out in this domain. Further, though the protocol for the use of LLM was thoroughly validated, this process was still very prone to human error. After all, if ten people would be tasked with analyzing the same article, they would all give a slightly different analysis, depending on what is asked of them to analyze and how they interpret that task. The same can be said about the LLM and is in fact even more heightened due to its precision with the definitions of words. This was more obvious while answering RQ2 than RQ1, but it shows that though the prompts that were used were validated, they

still could have been ever so slightly imperfect and changing one word could have made a large difference to the answers that were retrieved. This, on top of human error while validating and reading dozens of papers could have created a stacked error effect that could have resulted in inaccurate results. Thus, to avoid these two pitfalls in the future, an even larger number of iterations should be completed to diminish bias both from the LLM and the human. On top of that, it would be suggested that a collection of human researchers read some of the papers manually before even feeding it to the LLM. This could then be used as a more accurate way of validating the LLMs answers, instead of one human researcher going back to re-read the papers, based on the answers that the LLM provided.

References

- [1] J. McGrath, *Groups: Interaction and Performance*. Prentice-Hall, 1984, ISBN: 9780133657005. [Online]. Available: <https://books.google.nl/books?id=4pzZAAAAAAAJ>.
- [2] J. R. Hackman, “The design of work teams,” in *Handbook of organizational behavior*, J. Lorsch, Ed. Englewood Cliffs, NJ: Prentice-Hall, 1987.
- [3] S. G. Cohen and D. E. Bailey, “What makes teams work: Group effectiveness research from the shop floor to the executive suite,” *Journal of Management*, vol. 23, no. 3, pp. 239–290, 1997. DOI: [10.1177/014920639702300303](https://doi.org/10.1177/014920639702300303). eprint: <https://doi.org/10.1177/014920639702300303>. [Online]. Available: <https://doi.org/10.1177/014920639702300303>.
- [4] D. R. Ilgen, J. R. Hollenbeck, M. Johnson, and D. Jundt, “Teams in organizations: From input-process-output models to imoi models,” *Annual Review of Psychology*, vol. 56, no. Volume 56, 2005, pp. 517–543, 2005, ISSN: 1545-2085. DOI: <https://doi.org/10.1146/annurev.psych.56.091103.070250>. [Online]. Available: <https://www.annualreviews.org/content/journals/10.1146/annurev.psych.56.091103.070250>.

- [5] D. van Knippenberg, C. De Dreu, and A. Homan, “Work group diversity and group performance: An integrative model and research agenda.,” English, *Journal of Applied Psychology*, vol. 89, no. 6, pp. 1008–1022, 2004, ISSN: 0021-9010. DOI: [10.1037/0021-9010.89.6.1008](https://doi.org/10.1037/0021-9010.89.6.1008).
- [6] M. A. Marks, J. E. Mathieu, and S. J. Zaccaro, “A temporally based framework and taxonomy of team processes,” *The Academy of Management Review*, vol. 26, no. 3, pp. 356–376, 2001, ISSN: 03637425. [Online]. Available: <http://www.jstor.org/stable/259182> (visited on 04/15/2024).
- [7] S. G. Cohen and D. E. Bailey, “What makes teams work: Group effectiveness research from the shop floor to the executive suite,” *Journal of management*, vol. 23, no. 3, pp. 239–290, 1997.
- [8] D. W. Johnson and F. P. Johnson, *Joining together: Group theory and group skills*. Prentice-Hall, Inc, 1991.
- [9] E. Salas, D. E. Sims, and C. S. Burke, “Is there a “big five” in teamwork?” *Small Group Research*, vol. 36, no. 5, pp. 555–599, 2005. DOI: [10.1177/1046496405277134](https://doi.org/10.1177/1046496405277134). eprint: <https://doi.org/10.1177/1046496405277134>. [Online]. Available: <https://doi.org/10.1177/1046496405277134>.
- [10] B. W. Tuckman, “Developmental sequence in small groups.,” *Psychological Bulletin*, vol. 63, no. 6, pp. 384–399, 1965. DOI: [10.1037/h0022100](https://doi.org/10.1037/h0022100). [Online]. Available: <https://doi.org/10.1037/h0022100>.
- [11] B. W. Tuckman and M. A. C. Jensen, “Stages of small-group development revisited,” *Group & Organization Studies*, vol. 2, no. 4, pp. 419–427, 1977. DOI: [10.1177/105960117700200404](https://doi.org/10.1177/105960117700200404). eprint: <https://doi.org/10.1177/105960117700200404>.

105960117700200404. [Online]. Available: <https://doi.org/10.1177/105960117700200404>.

- [12] J. R. Hackman, “Why teams don’t work.,” in (Social psychological applications to social issues, Vol. 4.), Social psychological applications to social issues, Vol. 4. New York, NY, US: Plenum Press, 1998, pp. 245–267, ISBN: 0-306-45679-6 (Hardcover).
- [13] J. Biggs and K. Collis, “Towards a model of school-based curriculum development and assessment using the solo taxonomy,” *Australian Journal of Education*, vol. 33, no. 2, pp. 151–163, 1989. DOI: [10.1177/168781408903300205](https://doi.org/10.1177/168781408903300205). eprint: <https://doi.org/10.1177/168781408903300205>. [Online]. Available: <https://doi.org/10.1177/168781408903300205>.
- [14] Y. Barrios-Fleitas, “What’s the difference? diversity constructs as separation, variety, or disparity in organizations,” *Academy of Management Review*, vol. 32, Oct. 2024. DOI: [10.5465/AMR.2007.26586096](https://doi.org/10.5465/AMR.2007.26586096).
- [15] M. Aria and C. Cuccurullo, “Bibliometrix: An r-tool for comprehensive science mapping analysis,” *Journal of Informetrics*, vol. 11, no. 4, pp. 959–975, 2017. [Online]. Available: <https://doi.org/10.1016/j.joi.2017.08.007>.
- [16] Unknown, *What is a large language model (llm)?* Available at <https://www.cloudflare.com/learning/ai/what-is-large-language-model/>.
- [17] A. Hern, *Ai race heats up as openai, google and mistral release new models*, Available at <https://www.theguardian.com/technology/2024/apr/10/ai-race-heats-up-as-openai-google-and-mistral-release-new-models>.

- [18] J. Liu, C. S. Xia, Y. Wang, and L. ZHANG, “Is your code generated by chatgpt really correct? rigorous evaluation of large language models for code generation,” in *Advances in Neural Information Processing Systems*, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, Eds., vol. 36, Curran Associates, Inc., 2023, pp. 21 558–21 572. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2023/file/43e9d647ccd3e4b7b5baab53f03Paper-Conference.pdf.
- [19] N. Halevy, E. Chou, and A. Galinsky, “A functional model of hierarchy: Why, how, and when vertical differentiation enhances group performance,” *Organizational Psychology Review*, vol. 1, pp. 32–52, Feb. 2011. DOI: [10.1177/2041386610380991](https://doi.org/10.1177/2041386610380991).
- [20] G. Alabi, “Bradford’s law and its application,” *International Library Review*, vol. 11, no. 1, pp. 151–158, 1979. DOI: [10.1016/0020-7837\(79\)90044-X](https://doi.org/10.1016/0020-7837(79)90044-X). eprint: [https://doi.org/10.1016/0020-7837\(79\)90044-X](https://doi.org/10.1016/0020-7837(79)90044-X). [Online]. Available: [https://doi.org/10.1016/0020-7837\(79\)90044-X](https://doi.org/10.1016/0020-7837(79)90044-X).

Appendix A

Team effectiveness definition per author

Please rotate your page 90 degrees to view the table.

Author	Definition	18	3	10	2	12	6	2	3	1	9	6	4	9	7	4	4	2	3
		Performance measures	Exploiting experts' knowledge	Communication	Conflict management	Sharing knowledge	Role and responsibility division	Quality of decisions	Diversity	Diversity in diversity mindset	Team member satisfaction	Adaptability to tasks given	Transactive memory system	Level of collaboration	Team engagement	Team viability	Equal participation	Creativity	Motivational states
Van Knippenberg D	Team performance is used as the key measure of team effectiveness; highly effective teams are those that achieve high levels of performance on their tasks and objectives	X							X										
Salas E	Team effectiveness is measured in the way team members collaborate, communicate and share knowledge with one another, emphasizing that a well-trained team is highly effective.			X		X								X					
Bonner BJ	Effective teams are those who recognize and exploit each member's expertise, developing transactive memory systems by transferring relevant knowledge to each other. Coordinating this, along with <u>taking into account</u> the longevity of the team, makes for a highly effective team as members can coordinate learning and predict task difficulty easily.		X	X		X						X	X			X			
Cheng A	Team effectiveness is highly impacted by the defined roles and the responsibilities thereof within the team. Effective teams comprise of those who know what their role is in the team and takes responsibility for the tasks that comes along with that role. Situational awareness and strong communication are also a <u>big influencing factors</u> in how effective a team is.			X		X	X					X							

Tschan F	In medical teams, diagnostic accuracy is used as a measure of team effectiveness. Factors like explicit reasoning and communicating thoughts out loud contribute to raising the team effectiveness during medical emergency simulations.			X		X							X					
Van Ginkel Wp	Team effectiveness is measured by the quality of performance outcomes and the quality of decisions made within the team. Diversity can raise team effectiveness by supplying multiple perspectives, enhancing the scope of discussions related to the task. However, diversity can also lower team effectiveness when it impacts intra-team relations.	X						X	X		X							
Argote L	Team effectiveness is measured through a multi-dimensional set of factors that include both team results and team satisfaction. The author identifies three factors that influence effectivity: 1)Team performance behaviors (actions relevant to achieving team goals such as team learning, creativity, and reflectivity), 2)Team performance outcomes (the results of performance behaviors such as effectiveness and efficiency), 3)Member affective outcomes (member commitment and satisfaction). All of these combined, and facilitated with a transactive memory system, support high team effectivity.	X									X		X		X			
Baumann Mr	Team effectiveness is highly affected by the distribution of knowledge among team members, especially those who are considered experts within the team. Task distribution should consider and utilize members' expertise to accomplish goals more efficiently. Further, members should base their problem-solving strategy based on the task at hand. Sharing knowledge within the group and developing a transactive memory system are key to ensuring a highly effective team.		X	X		X							X	X				

Hollenbeck Jr	Team effectiveness is measured by how well a team can adapt to changing demands, while still optimizing the preferred performance outcomes. Further, the level of collaboration within the team also impacts the effectivity, especially in changing situations.	x										x		x				
Kirschner Pa	Team effectiveness requires both social and cognitive successes to ensure high-quality output. Engagement within the team is an indicator of team effectiveness, as higher engagement provides more in-depth discussions where multiple points of view are recognized. Further, a team's output is dependent on the effort of all team members, so " <u>free-riders</u> " negatively impact the effectivity of a team.	x		x		x	x							x	x			x
Okudan Ge	The effectivity of a (engineering design) team is based on the grades the group receives for their group output. Team quizzes, peer design evaluations, and design reports formed the basis of the grade the students received, thus reflecting their team's effectiveness.	x										x		x				
Pazos P	Team effectiveness is measured through team performance outcomes and teamwork competencies. Teamwork skills are measured through a validated instrument that has been widely used for competency assessment.	x		x		x								x				
Tekleab Ag	Team effectiveness is described in three <u>criteria</u> : <u>perceived</u> team performance, team satisfaction, and team viability. Team viability refers to the willingness of team members to remain in the team for longer periods of time. Further, the author also explores the relationship between team cohesion and team performance as an indicator of team effectiveness.	x										x						x

Zheng L	In collaborative learning contexts, collaborative knowledge building, group performance, and socially shared metacognitive regulation and interaction are all factors that contribute to team effectiveness. The author also proposes to use AI to introduce a personalized intervention approach to collaborative learning.	x		x		x							x	x				
Chen G	Team effectiveness is defined by a 6-item scale: - Team members work <u>effectively</u> .- Team members put considerable effort into their jobs. - Team members are concerned about the quality of their <u>work</u> .- Team members meet or exceed their productivity requirements.- Team members are committed to producing quality work. - Team members do their part to ensure that their products will be delivered on time.	x				x								x		x		
Cummings Jn	Team effectiveness is defined by the team's performance and ability to deliver valuable work, according to executives. There are also processes that can contribute or detract from team effectiveness, such as team learning, team cohesion, conflict, and <u>creativity</u> . The authors conclude that team effectiveness can be measured by the perspective of executives and/or team members themselves, which makes for a complicated construct.	x								x							x	
Curseu Pl	Team effectiveness is measured by the level of group performance, team member satisfaction, and team viability. The authors use a scoring system to measure these <u>factors</u> , of which the average is the general score for group effectiveness. The author uses Hackman's measurement model for the evaluation.	x								x						x		
De Dreu Ckw	The author defines team effectiveness as the quality of decisions and judgments reached by the team, which is determined by the team's motivated information processing based on their social and epistemic motivations.						x							x	x			x

Jehn KA	Team effectiveness is defined in terms of group performance outcomes, creativity, and team member satisfaction. The impact of group conflict and expertise perceptions on team effectiveness are also studied extensively. A combination of all these factors also contribute to the team viability.	x	x		x	x					x				x		x	
Johns on Dwyer	Role and responsibility division and equal participation promote team member satisfaction and interaction, which in turn enhance team effectivity.					x				x		x	x		x			x

Appendix B

Team effectiveness characteristics studied per article

Please rotate your page 90 degrees to view the table.

Author	Paper title	12	8	3	3	3	3	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
		Team performance	Team satisfaction	Creativity	Information elaboration/sharing	Conflict	Leadership	Hierarchy	Cohesion	Transactive memory systems	Communication	Closure of teamwork	Centralization of teamwork	Subgrouping of teamwork	Team viability	Teamwork processes	Team outcomes	Social loafing	Cultural diversity	Goal orientation	Trust	Quality of group decisions	Team diversity	Openness to experience	Psychological safety	Epistemic motivation	Social motivation
Crawford ER	A Configural Theory Of Team Processes: Accounting For The Structure Of Taskwork And Teamwork											x	x	x													
Halevy N	A Functional Model Of Hierarchy: Why, How, And When Vertical Differentiation Enhances Group Performance							x																			

TABLE B.1: Team characteristics studied per article

Tekleab AG	A Longitudinal Study Of Team Conflict, Conflict Management , Cohesion, And Team Effectiveness	x	x													x											
Lepine JA	A Meta-Analysis Of Teamwork Processes: Tests Of A Multidimensional Model And Relationships With Team Effectiveness Criteria	x	x					x								x											
Carton AM	A Theory Of Subgroups In Work Teams																x										
Peus C	Authentic Leadership: An Empirical Test Of Its Antecedents, Consequences, And Mediating Mechanisms	x																									

Appendix C

Acknowledgements

This page is dedicated to those who have helped and supported me to push through to complete my Bachelor's degree.

The supervisors, Mr. Yeray and Ms. Faiza Thank you for your patience, support, and wisdom throughout this process. I could not have wished for more kind and expert supervisors. I am truly grateful for you!

Jasper Thank you for your unwavering support. All the nights you stayed up with me to accompany me working and all the tears you talked me through will be forever etched in my memory.

Pablo Thank you for all the encouraging words and radiating positive energy you gave me. Talking to you always reminded me that I can chill out and things will work out!

Una Thank you for staying by my side even through the all-nighters. You always reminded me to take breaks and take care of myself by refusing to let me work when

it was past my bedtime.

To the reader who reached this point You made it! Thank you for reading my work.