

Block, Jörn H.; Fisch, Christian; Kanwal, Narmeen; Lorenzen, Solvej; Schulze, Anna

**Article — Published Version**

## Replication studies in top management journals: An empirical investigation of prevalence, types, outcomes, and impact

Management Review Quarterly

**Provided in Cooperation with:**

Springer Nature

*Suggested Citation:* Block, Jörn H.; Fisch, Christian; Kanwal, Narmeen; Lorenzen, Solvej; Schulze, Anna (2022) : Replication studies in top management journals: An empirical investigation of prevalence, types, outcomes, and impact, Management Review Quarterly, ISSN 2198-1639, Springer International Publishing, Cham, Vol. 73, Iss. 3, pp. 1109-1134, <https://doi.org/10.1007/s11301-022-00269-6>

This Version is available at:

<https://hdl.handle.net/10419/307499>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>



# Replication studies in top management journals: An empirical investigation of prevalence, types, outcomes, and impact

Jörn H. Block<sup>1,2,3</sup> · Christian Fisch<sup>4</sup> · Narmeen Kanwal<sup>1</sup> · Solvej Lorenzen<sup>3</sup> · Anna Schulze<sup>1</sup>

Received: 8 February 2022 / Accepted: 3 April 2022 / Published online: 23 June 2022  
© The Author(s) 2022

## Abstract

Replication studies are important for the empirical research process. Yet, while there is an increased awareness of the need for replication in management research, it appears that such studies are rarely published in leading management journals. Importantly, we lack a comprehensive overview of replication studies in the top management journals that spans all sub-disciplines. Our systematic review closes this gap and provides an overview of the prevalence, types, outcomes, and impact of replication studies in management journals. We find that differences in the prevalence of replications between sub-disciplines exist and that most replications are wide replications. With regard to the replication outcome, our review shows that the share of non-confirming replications is low. Moreover, such replications are cited less often than confirming replications pointing towards a confirmation bias in management research. We discuss the implications of our results for authors, reviewers, and editors of management journals.

**Keywords** Replication · Independent replication · Systematic review · Impact · Empirical research

## 1 Introduction

Replication is an essential part of the empirical research process. Replication studies help to build and establish knowledge about a particular phenomenon or relationship of interest. Recently, the need for replication studies has been stressed in many fields and sub-disciplines of management research, including strategic management

---

Note to reviewers We uploaded all data and analyses presented in this manuscript anonymously to the Open Science Foundation. The files and the online appendix can be accessed here: <https://osf.io/ejcaz/>. Extended author information available on the last page of the article

(e.g., Bettis et al., 2016), organization studies (Wright and Sweeney, 2015), international business (e.g., Aguinis et al., 2017), family business (e.g., De Massis et al., 2020), and entrepreneurship (e.g., Maula and Stam, 2020). Replication studies are also an important part of the aims and scope of Management Review Quarterly (MRQ) (Block and Kuckertz, 2018), which publishes replication studies of various types and from various disciplines (e.g., Dettori and Floris, 2022; Szumal et al., 2021; Van Scotter, 2020; Yuan et al., 2020).

Yet, the current status quo of replication studies is opaque because we lack an overview of how many studies are actually published in the leading management journals. Coherently, we lack detailed information about these replication studies. This includes information on the types of replications studies, information on how often the findings of the replicated studies (hereafter: original studies) are (partially) supported or not, and their impact on the scientific community. Without such an overview, the discussion about the need and value of replication studies as well as the existence of a replication crisis in management research remains superficial and anecdotal. Currently, proponents and skeptics on the necessity of replication studies in management research exchange arguments without truly knowing the current status quo.

Our study addresses this research gap. We conduct a comprehensive and systematic review of 56 leading management journals that encompasses all sub-disciplines of management research. Our focus is on *independent* replication studies that seek to replicate prior published peer-reviewed journal articles. That is, we focus on those replications, where one author (team) tries to replicate the peer-reviewed empirical article of another author (team). We also categorize replications based on how closely they resemble the original study.

Our review is guided by the following four questions: (1) How prevalent are replication studies in the top management journals overall and across the different sub-disciplines? (2) What types of replication studies do appear in these journals? (3) What are the replication outcomes? (4) What is their impact on the scientific field in terms of citations and how does this impact differ by replication outcome?

We systematically scan 56 journals for empirical studies that use the term “replication” (or a related term) in the title, abstract or main text. We do not impose time constraints on our search process and consider all articles published in these journals that can be searched electronically. Our comprehensive identification strategy results in a sample of 240 independent replication studies. Our review shows that independent replications are rarely published in the leading management journals. The sub-disciplines with the highest prevalence of independent replications are (organizational) psychology (82 replications in 13 journals), general management, ethics, and social responsibility (73 replications in 12 journals), and strategy (28 replications in 4 journals). A strong time trend towards an increasing number of replication studies cannot be observed. The overall prevalence was and continues to be low.

In addition to these general insights, our overview provides a more nuanced picture of the replication studies that exist to date. First, our findings show that some differences in the prevalence of replications exist between sub-disciplines. This could point towards disciplinary differences in the value or legitimacy of replications within the field of management research. Alternatively, it could simply be coincidental or

driven by differences between individual editors. Second, our overview shows that so-called quasirandom replications (replications that differ from the original study without clearly enhancing it, Köhler and Cortina, 2021), constitute a large part of replication studies in leading management journals (57.9% of all replication studies in our sample). Yet, it is exactly this type of replication that is of questionable value to the empirical research process (e.g., Köhler and Cortina, 2021). This is the case because quasirandom replications use a different but not a clearly superior research design compared to the original study so that their contribution for solidifying empirical knowledge about a particular relationship is limited. Third, we find the share of non-replicated results (i.e., the findings of the original study are not replicated) in our sample of replication studies is 20.4%; 79.6% of the replication studies in our sample (at least partially) replicate the original study. In contrast to other disciplines, the share of non-confirming results is comparatively low. For example, comprehensive and influential studies have estimated replication rates of 36% in psychological sciences (Open Science Collaboration, 2015), 62% in experimental social sciences (Camerer et al., 2018), and 61% in experimental economics (Camerer et al., 2016). One possible explanation would be that management research suffers from a bias against publishing non-replicated results. Fourth, our results show that replication studies which do not replicate the results of the original study are cited less often than those replication studies which do replicate the results of the original study. Similar to the low rate of non-replicated results, a confirmation bias might exist in management research and primarily those replications become visible to the community that confirm the results of the original study.

With these findings, our study contributes to an ongoing and very recent discussion about the status quo and value of replication in management research (e.g., Aguinis et al., 2017; Bergh et al., 2017; Dau et al., 2021; De Massis et al., 2020; Hensel, 2021; Köhler and Cortina, 2021; Maula and Stam, 2020; Ryan and Tipu, 2022). By showing an extremely low prevalence of replication studies in the leading management journals, our study adds to the discussion of how open the management field really is to replication studies. Our findings suggest that, despite the recent acknowledgments of the importance of replication and the many articles and editorials that encourage producing replication studies, there is still a long way to go for management research as an empirical discipline. This is particularly true for the top journals of the field. Our paper also contributes to the discussion about the interrelationship between scientific impact and replication. For the field of economics, Mueller-Langer et al., (2019) show that more impactful articles and articles from leading institutions are more likely to be replicated while replication is less likely for articles from the very top journals. Our study shows that this relationship also exists in the other direction in that the replication outcome also determines its impact, however not in the way it should be. Confirming results are cited more often than non-confirming results pointing towards a confirmation bias.

## 2 Prior literature on replication studies in the field of management

Our study ties in with and updates the small number of prior studies that seek to synthesize the state of the art of replication studies in management research. We briefly review them below.<sup>1</sup>

Hubbard and Vetter (1996) review replication studies published between 1970 and 1991 in 18 leading business journals, covering the domains of accounting, economics, finance, management, and marketing. Using a broad definition of replication studies, they find that 6.2% (266 of 4,270) of all empirical studies in that period are replications. Focusing on the field of management (i.e., the journals *AMJ*, *ASQ*, *JAP*, *Organizational Behavior and Human Decision Processes*), the replication incidence is 5.3% (65 of 1,222 empirical studies). The authors also assess the outcomes of the replication studies and highlight that replication studies in management less often conflict with the findings of the original studies, relative to other disciplines (e.g., marketing).

The second overview of replication studies is by Hubbard et al. (1998), who investigate replication studies in strategic management research, thus focusing on a sub-discipline within management research. Using a broad categorization, they determine that the replication study prevalence in strategic management research is 5.3% (37 of 701 articles) for the period 1976–95. This is in line with the findings of Hubbard and Vetter (1996). Additionally, the authors differentiate journal tiers within strategic management to explore differences in replication studies across these journal tiers. While their results do not reveal any major differences regarding the timeliness and prevalence of replications when comparing journals across tiers, Hubbard et al. (1998) find that first-tier journals are more likely to publish replications with supportive than conflicting results. The opposite is true for third-tier journals.

Published more than 20 years after the initial studies by Hubbard and Vetter (1996) and Hubbard et al. (1998), the recent work by Köhler and Cortina (2021) is the only subsequent overview on replication studies in management that we could identify. Köhler and Cortina (2021) comprehensively review and categorize replications of empirical articles published in three leading management journals (i.e., *AMJ*, *JAP*, *JoM*) between 2007 and 2017. Overall, the authors identify 79 independent replication studies (i.e., conducted by different researchers than the original study). The primary purpose of their study is a nuanced categorization of the different types of replication studies and their prevalence. For example, the authors show that quasirandom replications are the most common type of replication study (i.e., the replication repeats some procedures of the original study, while others are varied; the replication does not seek to improve the original study), followed by constructive replications (i.e., the replication seeks to improve aspects of the original study).

Our brief review provides initial insights on replications in management research and their characteristics. Important subjects addressed in previous research consider

<sup>1</sup> Other overviews on the status quo of replication studies focus on other disciplines, such as marketing (e.g., Darley, 2000; Evanschitzky et al., 2007; Hubbard and Armstrong, 1994) or economics (e.g., Mueller-Langer et al., 2019). These overviews typically indicate a low prevalence of replication studies in the respective fields, well below 5%.

replication outcomes (Hubbard and Vetter, 1996), differences across journals and sub-disciplines (Hubbard et al., 1998), and replication types (Köhler and Cortina, 2021). We revisit these initial findings via a comprehensive empirical assessment of empirical replication studies in the leading management journals that considers a broad sample and coverage of top journals, sub-disciplines, and time periods. We also investigate the (citation) impact of replication studies, a subject that has not received any attention so far. Investigating the scientific impact of replications in the management field has been overlooked so far and our study closes an important research gap in this regard. Knowing more about the impact of replications and how it depends on the replication outcome is important as it shows how well replications are received by the respective community and to what extent management research learns and updates its knowledge base.

### 3 Data

#### 3.1 Selection of journals

Our selection of studies is based on top-tier management journals included in the 2018 ABS (Association of Business Schools) list (ABS, 2018; Hubbard and Vetter, 1996; Walker et al., 2018). The ABS list is comprehensive and contains 1,582 journals classified in 22 sub-disciplines (ABS, 2018; Walker et al., 2018) on a scale from 1 (lowest rating) to 4\* (highest rating). In order to identify independent replication studies in top-tier management journals, we focused on articles published in journals that are ranked 3, 4, or 4\*. We considered the sub-disciplines of (1) entrepreneurship and small business management (ENT-SBM), (2) general management, ethics, gender and social responsibility (ETHICS-CSR-MAN), (3) international business and area studies (IB&AREA), (4) innovation (INNOV), (5) operation research and management science (OR&MANSCI), (6) organization studies (ORG STUD), (7) psychology (organizational) (PSYCH (WOP-OB)), and (8) strategy (STRAT). Furthermore, we only considered journals that publish empirical articles. Thus, we excluded theoretical and conceptual journals (e.g., *Academy of Management Review*). Moreover, we also omitted journals that exclusively focus on specific regions (e.g., *African Affairs*). Our final journal list included 56 journals.

#### 3.2 Identification of replication studies

In order to identify independent replication studies in these journals, we developed a set of keywords (or word stems) linked to replication studies. In line with prior research, we used the keywords “replicate”, “replication”, “replicating”, “revisit”, “reexamine”, “retesting” (Köhler and Cortina, 2021; Mueller-Langer et al., 2019). We entered these keywords in every journal’s search engine, which enabled a comprehensive search of all electronically available articles published until the last volume of 2020. Since we did not have access to all volumes on the websites of the respective journals, we identified all contributions available in these journals on Google Scholar

(advanced search function). Notably, we searched the keywords in the full body of text as well as in the title and abstract.

56 of the 56 journals had at least one article that matched the keywords. (Overall, the 56 journals had a total of 159,242 articles published). Out of a total of 159,242 published contributions, 24,595 articles were gathered in response to our keywords search. Subsequently, we manually assessed all 24,595 articles to identify actual independent replication studies. In order to do so, we proceeded in several steps. First, we excluded editorials, book reviews, and author guidelines. However, research notes and commentaries were included. Second, we excluded qualitative-empirical, purely theoretical, and conceptual studies. Third, we excluded studies that mentioned the search words (e.g., “replicate”) somewhere in the text or the reference list but were, in fact, not replications. Fourth, we excluded within-study replications, in which the authors simply use a multi-study design but do not replicate a prior study. Although such multi-study designs are highly desired and recommended, our research focus was on replications that aim to replicate prior published studies. Finally, we excluded replications of monographs, (dissertation) theses, conference papers, and working papers as our emphasis was on replications of peer-reviewed journal articles.

These search and exclusion steps resulted in a sample of 438 studies from 48 journals.<sup>2</sup> This sample is our broadest sample of replication studies and we refer to it as replication study sample 1 (RS 1) in Table 1. After having derived our broadest sample, we determined two further samples. The replication study sample 2 (RS 2) encompasses 348 studies and only includes independent replication studies with no overlap of the authors between replication and original study. Finally, in replication study sample 3 (RS 3), we excluded those studies seeking to replicate more than one study. Although such replications are of course highly desirable, it is difficult to determine the type of replication and outcome if multiple studies are replicated, as we will describe in the following section.

The full list of journals considered, volumes, the total population of articles, the keyword hits, and the respective number of replications identified are displayed in Table 1.

### 3.3 Coding and variables

Guided by our four overarching research questions, we developed an initial coding scheme to code the identified replication studies. Our coding efforts focus on the studies included in R3. Based on regular discussions among the team of authors, the coding scheme was revised several times in the course of the coding process. Unclear cases were discussed among the coders until an agreement was reached.

In order to determine the replication type, we compared the samples, variables, measurements, and empirical analyses of the replication study and the corresponding original study. We then coded whether the *quality of the sample* of the replica-

<sup>2</sup> We could not identify any replication studies for RS 1 in 8 of the 56 journals considered (i.e., International Small Business Journal, Business Ethics Quarterly, Gender and Society, Gender, Work and Organization, Journal of Common Market Studies, Organization, Global Strategy Journal, Long Range Planning). These journals are omitted from Table 1.

**Table 1** Overview of the replication studies included in our study

Column	(1) Journal characteristics				(2) Replication studies			(3) Replication type			(4) Replication outcome			
Sub-discipline and journal	ABS rating	First year in our sample	# Articles	RS 1	RS 2	RS 3	Literal	Con-structive	Regressive	Con-founded	Quasi-random	Not repl.	Par-tial repl.	Fully repl.
Entrepreneurship and Small Business Man. (ENT-SBM)				18	16	12	0	5	0	0	7	1	5	6
Entrepreneurship, Theory and Practice	4	1976	1,549	2	1	1	0	0	0	0	1	0	0	1
Journal of Business Venturing	4	1985	1,427	3	3	2	0	0	0	0	2	0	0	2
Strategic Entrepreneurship Journal	4	2007	389	1	1	0	0	0	0	0	0	0	0	0
Entrepreneurship and Regional Development	3	1989	978	1	1	1	0	0	0	0	1	0	1	0
Family Business Review	3	1988	655	3	3	3	0	1	0	0	2	0	2	1
International Small Business Journal	3	1982	838	0	0	0	0	0	0	0	0	0	0	0
Journal of Small Business Management	3	1963	905	3	3	2	0	2	0	0	0	0	1	1
Small Business Economics	3	1989	2,312	5	4	3	0	2	0	0	1	1	1	1
General Man., Ethics, Gender, SR (ETHICS-CSR-MAN)				125	113	83	2	34	0	3	44	11	37	35
Academy of Management Journal	4	1950	4,155	16	14	9	0	2	0	1	6	2	1	6
Administrative Science Quarterly	4	1999	1,314	4	3	2	0	0	0	0	2	0	1	1
Journal of Management	4	1979	1,979	10	10	5	0	3	0	0	2	2	3	0



Table 1 (continued)

Column	(1) Journal characteristics			(2) Replication studies			(3) Replication type			(4) Replication outcome				
	4	1990	1,312	6	6	6	0	3	0	1	2	0	5	1
British Journal of Management	4	1991	1,604	0	0	0	0	0	0	0	0	0	0	0
	4	1964	2,591	4	4	3	0	1	0	0	2	0	3	0
	3	1960	941	1	1	1	0	1	0	0	0	0	1	0
European Management Review	3	2004	537	3	2	1	0	1	0	0	0	0	0	1
Gender and Society	3	1987	920	0	0	0	0	0	0	0	0	0	0	0
Gender, Work and Organization	3	1994	1,325	0	0	0	0	0	0	0	0	0	0	0
Journal of Business Ethics	3	1982	8,850	37	34	22	1	4	0	1	16	2	9	11
Journal of Business Research	3	1973	7,840	44	39	34	1	19	0	0	14	5	14	15
International Business (IB&AREA)				31	30	22	1	8	0	0	13	7	10	5
Journal of International Business Studies	4	1970	2,796	9	8	5	1	2	0	0	2	1	3	1
Journal of World Business	4	1965	4,777	2	2	2	0	1	0	0	1	0	1	1
Asia Pacific Journal of Management	3	1984	1,297	8	8	5	0	2	0	0	3	1	2	2
International Business Review	3	1992	1,844	3	3	2	0	0	0	0	2	1	1	0
Journal of Common Market Studies	3	1962	3,660	0	0	0	0	0	0	0	0	0	0	0
Journal of International Management	3	1999	839	1	1	1	0	0	0	0	1	1	0	0
Management International Review	3	1960	9,550	8	8	7	0	3	0	0	4	3	3	1
Innovation (INNOV)				10	8	6	0	2	0	0	4	1	2	3

**Table 1** (continued)

Column	(1) Journal characteristics			(2) Replication studies			(3) Replication type			(4) Replication outcome				
Research Policy	4	1971	4,180	5	4	3	0	2	0	0	1	0	1	2
Journal of Product Innovation Management	4	1984	2,062	3	3	2	0	0	0	0	2	1	0	1
R and D Management	3	1970	2,184	1	0	0	0	0	0	0	0	0	0	0
Technovation	3	1981	3,130	1	1	1	0	0	0	0	1	0	1	0
Operations Research and Man. Science (OR&MANSCI)				10	9	6	0	2	0	0	4	1	4	1
Management Science	4	1954	9,625	10	9	6	0	2	0	0	4	1	4	1
Organization Studies (ORG STUD)				33	27	17	1	3	1	0	12	2	11	4
Organization Science	4	1990	1,956	1	1	0	0	0	0	0	0	0	0	0
Leadership Quarterly	4	1990	2,100	16	13	7	0	2	0	0	5	0	5	2
Human Relations	4	1947	8,150	10	9	7	1	0	1	0	5	1	4	2
Organization Studies	4	1980	3,635	1	1	1	0	0	0	0	1	0	1	0
Organizational Research Methods	4	1989	791	3	2	2	0	1	0	0	1	1	1	0
Group and Organization Management	3	1976	1,642	2	1	0	0	0	0	0	0	0	0	0
Organization	3	1994	1,583	0	0	0	0	0	0	0	0	0	0	0
Psychology (Organizational) (PSYCH (WOP-OB))				166	129	92	0	26	2	0	64	16	44	32
Journal of Applied Psychology	4	1917	9,963	53	46	34	0	11	1	0	22	6	16	12
Journal of Occupational and Organizational Psychology	4	1975	1,300	11	7	7	0	0	0	0	7	0	6	1
Journal of Occupational Health Psychology	4	1996	888	3	0	0	0	0	0	0	0	0	0	0

Table 1 (continued)

Column	(1) Journal characteristics			(2) Replication studies			(3) Replication type			(4) Replication outcome				
	4	1980	2,462	13	12	6	0	2	0	0	4	1	3	2
Journal of Organizational Behavior	4	1971	3,590	35	28	20	0	8	0	0	12	2	10	8
Journal of Vocational Behavior	4	1966	2,510	8	7	5	0	1	1	0	3	1	2	2
Organizational Behavior and Human Decision Processes	4	1948	7,940	17	11	10	0	0	0	0	10	3	4	3
Personnel Psychology	3	1952	2,109	6	4	3	0	1	0	0	2	0	1	2
Applied Psychology: An International Review	3	1991	1,246	3	2	2	0	1	0	0	1	1	1	0
European Journal of Work and Org. Psychology	3	1958	4,278	10	6	3	0	1	0	0	2	1	0	2
Human Factors: JHFES	3	1988	672	2	2	0	0	0	0	0	0	0	0	0
Human Performance	3	1986	1,750	2	1	1	0	0	0	0	1	0	1	0
Journal of Managerial Psychology	3	1987	1,071	3	3	1	0	1	0	0	0	1	0	0
Work and Stress	4	1980	3,264	46	42	28	0	15	0	1	12	12	9	7
Strategic (STRAT)	4	1980	3,264	45	41	27	0	15	0	1	11	12	8	7
Strategic Management Journal	3	2011	300	0	0	0	0	0	0	0	0	0	0	0
Global Strategy Journal	3	1979	7,226	0	0	0	0	0	0	0	0	0	0	0
Long Range Planning	3	2003	451	1	1	1	0	0	0	0	1	0	1	0
Strategic Organization														
Total				439	374	266	4	95	3	4	160	51	122	93

Notes: R1 (replication study sample 1): The sample includes both dependent and independent replication studies; R2: (replication study sample 2): the sample includes only independent replication studies; R3 (replication study sample 3): the sample includes only independent replication studies and excludes those studies that seek to replicate more than one study.

tion study is better, worse, same/similar, or different but neither better nor worse as compared to the sample of the original study. If the sample size of the replication study was larger and/or broader and external validity improved, we coded it as better. In turn, if the sample was smaller and/or narrower and external validity decreased, we coded it as worse. In some cases, the samples of the original and replication study were the same or almost identical. Finally, in many cases, the samples of the original and replication study were different but neither better nor worse. This was the case, for example, when the samples were from different industries, years, and countries but neither sample was clearly better than the other. We then compared the (in)dependent variables used by the original and replication study and coded whether the *quality of variables* used in the replication study was better, worse, same/similar, or different but neither better nor worse as compared to the variables of the original study. If the replication study, for instance, used a larger set of control variables than the original study, we considered the variables of the replication study as better. We further coded whether the *quality of measurement* of the central constructs was better, worse, same/similar, or different but neither better nor worse in the replication as compared to the original study. If, for example, a larger (smaller) and/or more (less) detailed set of items was used to measure a central construct in the replication as compared to the original study, we classified the measurement as better (worse). In a final step, we considered the *quality of empirical analysis* and coded whether the quality was better, worse, same/similar, or different but neither better nor worse in the replication as compared to the original study. In cases where the method used in the empirical analysis of the replication study was better able to account for endogeneity resulting from selection bias or reverse causality as compared to the method used in the original study, we thus considered the quality of the empirical analysis to be better.

### 3.4 Categorizing replication studies into different replication types

Following Köhler and Cortina (2021), we distinguish between literal, constructive, quasirandom, confounded, and regressive replication studies.

A *literal replication* describes a replication study in which the study design directly mirrors the original study (e.g., Köhler and Cortina, 2021; Lykken, 1968; Stroebe and Strack, 2014). Thus, if the quality of the sample, variables used, measurement, and empirical analysis were rated ‘same/similar’, we labeled the respective study as a literal replication. A *constructive replication* maintains the characteristics of the original study but enhances it in some way (e.g., Köhler and Cortina, 2021; Stroebe and Strack, 2014). We considered a replication study as constructive if the study exceeded the original study in external validity (i.e., quality of the sample) or internal validity (i.e., quality of the variables used, measurement, or empirical analysis), or both. If a study was rated worse in any quality dimension, we would consider it as a quasirandom replication. The term *quasirandom replication* refers to a replication that differs from the original study without clearly enhancing it (e.g., Köhler and Cortina, 2021; Tsang & Kwan, 1999). If the replication study was rated ‘different but neither better nor worse’ regarding one of the quality dimensions, we classified the replication as quasirandom. A *confounded replication* is a replication where the

external validity is lower but the internal validity has improved as compared to the original study (e.g., Köhler and Cortina, 2021). Accordingly, we classified a replication study as a confounded replication when the quality of the sample was worse but the quality of the variables used, measurement, or empirical analysis was better as compared to the original study. Finally, the term *regressive replication* refers to a replication that is similar regarding all quality dimensions of the original study except for one where it is worse (Köhler and Cortina, 2021). Hence, we coded a replication study as a regressive replication when the replication was worse regarding one of the quality aspects but ‘same/similar’ regarding the remaining aspects.

### 3.5 Measurement of replication outcomes

Regarding the replication outcome, we distinguished between several possible results. Namely, we differentiated between cases in which the replication study fully replicated the findings of the original study (i.e., all findings of the original study are replicated), only partially replicated the original findings (i.e., a subset of the findings of the original study is replicated), or not at all (i.e., none of the findings of the original study is replicated). This coding is in line with Hubbard and Vetter (1996), who were the first to explore different replication outcomes in business research, who distinguish the categories “support”, “partial support”, and “conflict”.

### 3.6 Measurement of the replication study's impact

We operationalized the impact of each replication study and the corresponding original study via the number of citations the study received on Google Scholar. Receiving a higher number of citations from subsequent studies (i.e., forward citations) indicates a higher scientific impact of an article for following research studies. Furthermore, Google Scholar provides comprehensive coverage of citation metrics and is thus especially suitable to capture citation data in social sciences (e.g., Harzing, 2013). With regard to the time frame, we considered article citations until the end of 2020 and manually collected the citation data from Google Scholar in May 2021.

### 3.7 Other study characteristics

Our regression analyses consider a broad set of further study characteristics that could explain the impact of the replication study. We distinguish between characteristics of the replication study and characteristics of the original study.

Regarding characteristics of the replication study, we considered whether the replication study was a literal, constructive, regressive, confounded, or quasirandom replication study. We constructed a dummy variable for each type. We further investigated whether the replication study replicates the findings of the original study fully, partially, or not at all. Again, we constructed a dummy variable for each outcome. We also devised a set of dummy variables that accounted for the journal's sub-discipline according to the ABS ranking (i.e., ENT-SBM, ETHICS-CSR-MAN, IB&AREA, INNOV, OR&MANSCI, ORG STUD, PSYCH (WOP-OB), or STRAT). Finally, we constructed a variable that measured the publication lag between the replication

study and the original study in years, based on the publication dates of the respective studies.

Regarding the characteristics of the original study, we took the age of the original study into account. The latter was operationalized via the number of years since the study was published as of 2021. Subsequently, we used a dummy variable in order to investigate whether the original study appeared in the same journal as the replication. Moreover, to assess the impact of the original study, we included the number of Google Scholar citations the original study received until December 2020. Due to the skewness of the citation data, the variable was included in logged form. Finally, we accounted for the number of co-authors in the original study.

Table 2 gives a detailed overview of all variables and their coding. We included a list of all replication studies as well as the corresponding original studies in our publicly available online Appendix<sup>3</sup>. The Appendix also provides the exact coding for every variable for each replication study.

## 4 Descriptive results

### 4.1 Prevalence of replication studies

Table 1; Fig. 1 show the prevalence of replication studies over time and by sub-discipline. Overall, the absolute number of independent replication studies increased from 15 in the years 1971 to 1980 to 81 from 2011 to 2020. Table 1 (Column 2), reveals the differences across sub-disciplines in management research. While we were able to identify 82 independent replication studies in Psychology (Organizational) (in 13 journals) and 73 in General Management, Ethics, Gender, and Social Responsibility (in 12 journals), we could identify only 11 in Entrepreneurship and Small Business Management (in 8 journals), 6 in Operations and Management Science (in 1 journal), and 5 in Innovation (in 4 journals). Moreover, as illustrated in Fig. 1, while replications in Psychology (Organizational) and General Management, Ethics and Social Responsibility, have a longer tradition of publishing independent replication studies, Strategy research only very recently started to publish replications.

Notes: N=240. Years before 2000 are omitted for the sake of brevity.

Table 1 also highlights the prevalence of the replication studies across the journals sampled. The journals with the highest absolute numbers of independent replication studies are the *Journal of Applied Psychology* (32 studies), *Strategic Management Journal* (27 studies), *Journal of Business Research* (30 studies), *Journal of Business Ethics* (18 studies), and *Journal of Vocational Behavior* (16 studies). 15 of the 56 journals in the list did not publish any independent replication studies at all. We can also observe strong differences within the sub-disciplines. For example, while the *Journal of Applied Psychology* is the journal with the highest absolute number of independent replication studies in our sample, *Applied Psychology: An International Review* only published 3 independent replication studies, and *Human Performance*

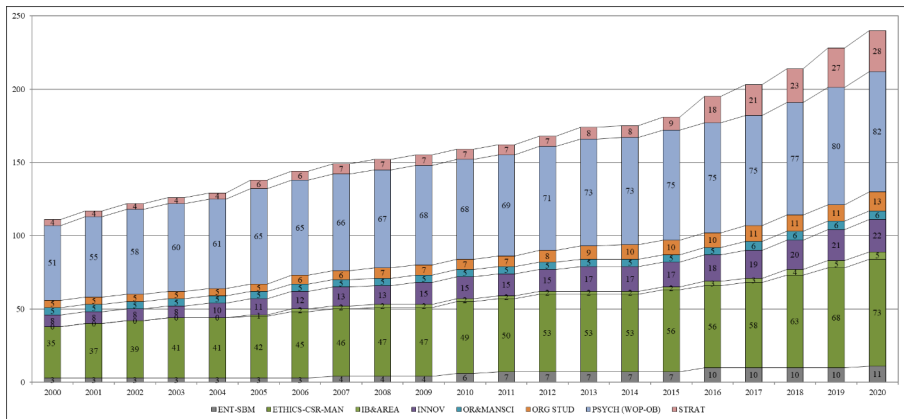
<sup>3</sup> Our appendix and the full reference list of all replication studies and original studies are available here: <https://osf.io/ejcaz/> (accessed March 15th, 2022).

**Table 2** Variables and coding

Variable	Operationalization
<i>Replication type</i>	
Type: literal	Dummy variable; replication directly mirrors the original study.
Type: constructive	Dummy variable; replication maintains the characteristics of the original study but enhances it in some way.
Type: quasirandom	Dummy variable; replication differs from the original study without clearly enhancing it.
Type: confounded	Dummy variable; replication where the external validity is lower but the internal validity has improved as compared to the original study.
Type: quasirandom	Dummy variable; replication is similar regarding all quality dimensions of the original study except for one where it is worse.
<i>Replication outcome</i>	
Outcome: not replicated	Dummy; none of the findings of the original study is replicated.
Outcome: partly repl.	Dummy; a subset of the findings of the original study are replicated.
Outcome: fully replicated	Dummy; all findings of the original study are replicated.
<i>Variables used in coding process</i>	
Quality of sample	The sample in the replication study is (1) better, (2) worse, (3) same/similar, or (4) different but neither better nor worse as compared to the original study.
Quality of variables	The (in-)dependent variables used are (1) better, (2) worse, (3) same/similar, or (4) different but neither better nor worse in the replication as compared to the original study.
Quality of measurement	The measurement of the central constructs is (1) better, (2) worse, (3) same/similar, or (4) different but neither better nor worse in the replication study as compared to the original study.
Quality of the empirical analysis	The quality of the empirical analysis in the replication study is (1) better, (2) worse, (3) same/similar, or (4) different but neither better nor worse in the replication as compared to the original study.
<i>Control variables: characteristics of the replication study</i>	
Citations replication study	Number of citations that the replication study received on Google Scholar until 2020 (collected in May 2021).
Sub-discipline: ETHICS-CSR-MAN	Dummy; sub-discipline general management, ethics, and social responsibility.
Sub-discipline: INNOV	Dummy; sub-discipline international business and studies area.
Sub-discipline: IB&AREA	Dummy; sub-discipline innovation.
Sub-discipline: OR&MANSCI	Dummy; sub-discipline operations research and management science.
Sub-discipline: ORG STUD	Dummy; sub-discipline organization studies.
Sub-discipline: PSYCH (WOP-OB)	Dummy; sub-discipline psychology (organizational).
Sub-discipline: STRAT	Dummy; sub-discipline strategy.
Sub-discipline: ENT-SBM	Dummy; sub-discipline entrepreneurship and small business management.
Publication lag	Time between the publication of the replication study and the original study in years
<i>Control variables: characteristics of the original study</i>	
Article age original study	Age of the original study in years as of 2021

**Table 2** (continued)

Variable	Operationalization
Same journal	Dummy variable that captures whether the original study was published in the same journal as the replication study
Citations original study (log.)	Number of citations that the original study received on Google Scholar until April 2021 (in logged form).
# Authors original study	Number of authors on the original study. The variable is categorical (i.e., 1, 2, 3, 4, 5, 6 or more).

**Fig. 1** Cumulative number of replication studies published per sub-discipline

published 0. In the context of Strategy research, especially when comparing *Strategic Management Journal* (27 studies) and *Long Range Planning* (0 studies), the situation is similar. Hence, apart from differences regarding the research culture across sub-disciplines, editorial policies of specific journals seem to matter as well.

## 4.2 Types of independent replication studies

How do the replication studies compare to the original studies? Following our coding described above, Column (3) of Table 1 reports the prevalence of the different replication types (i.e., literal, quasirandom, constructive, confounded, and regressive). The categorization is based on a comparison of the replication study with the original study along the four quality dimensions *sample*, *variables*, *measurement*, and *empirical analysis*. Only 4 out of 240 studies are *literal replications*. Most replications differ in one way or another from the original study. 139 out of 240 studies (57.9%) are *quasirandom*. These replications are neither better nor worse than the original study. In 91 out of 240 studies (37.9%), the replication can be classified as a *constructive replication*. In such cases, the replication is improved in at least one of the four quality dimensions over the original study but not worse in the other three dimensions. In addition, *confounded replications* account for 4 cases (1.7%) and *regressive replications* account for two cases (0.8%).



In the vast majority of cases, the replication hence differs from the original study. Yet, this difference is not always an improvement. When analyzing the differences between sub-disciplines, the field of Strategy stands out. It is the only sub-discipline where the number of constructive replications (15 studies) is higher than the number of quasirandom replications (11 studies).

### 4.3 Replication outcomes

Column (4) of Table 1 displays the results of the replication studies across the entire sample and the different sub-disciplines. 20.4% of the replication studies were able to fully replicate the findings of the original study, 47.9% could at least partially replicate the findings, and 31.7% could not replicate the findings at all. These percentages differ by sub-discipline and in this context two disciplines stand out. In the categories strategy and in international business, the number of studies where the result of the original study cannot be replicated is higher than the number of studies where the result is fully replicated. In all other sub-disciplines, this relationship is the other way round or the numbers are equal.

### 4.4 Impact of replication studies

In order to determine the impact of the replication studies, we collected the number of Google Scholar citations as of the end of 2020 for both the original study and the replication. Table 3 displays the mean and the median as well as the 0.10-, 0.25-, 0.75-, and 0.90-percentiles for the sample of original studies and the sample of replication studies.

The original studies receive substantially more citations. The mean (median) number of citations is 1,038.7 (462.5) for the original studies versus 142.6 (68.5) for the replication studies. The impact of the replication studies however varies significantly,

**Table 3** Descriptive comparison of our samples of replication studies and original studies

Statistic/Variable	Sample of replication studies (RS 3)	Sample of original studies
Total (Google Scholar) citations	34,213	249,052
Mean (SD)	142.55 (206.43)	1,037.72 (1,486.19)
Percentile: 10%	4.5	80
Percentile: 25%	22	159
Percentile: 50%	68.5	462.5
Percentile: 75%	170	1,274
Percentile: 90%	367	2,364
Percentile: 99%	1,028	7,522
Mean of article age (as of 2021)	19.36 years	27.95 years

*Notes:* N=480 (replication studies; N=240; original studies; N=240)

ranging from 4.5 at the 0.10 percentile to 367 at the 0.90 percentile. The standard deviation is 206.4.

## 5 Multivariate results on the determinants of impact

To further explore the impact of the replication studies, we assessed the determinants of the replication study's impact in a multivariate regression framework. The dependent variable was the number of citations that the replication study had received until the end of 2020. Because this variable is a count variable that only includes non-negative integers, we used a negative binomial regression as our main estimator.<sup>4</sup> As explanatory variables, we included a set of characteristics of the replication study and the original study. The results are displayed in Table 5. Model 1 considers the characteristics of the replication study, Model 2 considers the characteristics of the original study, and Model 3 considers both groups of variables together. Table 4 displays the correlation statistics for the variables used in our regression analysis.

Concerning the characteristics of the replication study, Model 3 shows that regressive replication studies receive a significantly higher amount of citations. The reference group is quasirandom replications. A similar, yet less pronounced effect emerges for literal and constructive replication studies. Further, replication studies that do not replicate the findings of the original study receive fewer citations than studies that are fully or partially replicated. The effect is highly significant. Regarding the citations received across the different sub-disciplines, the results show that replication studies in the sub-disciplines of innovations and operations research and management science receive a higher amount of citations.<sup>5</sup> Finally, a higher citation lag (i.e., a longer period elapsed between the publication of the original study and the replication study) result in significantly fewer citations. This indicates that studies published quickly after the original study are more impactful.

With regard to the characteristics of the original study, the results demonstrate that replications of older and more impactful studies (i.e., measured via the citations the original study received) receive a higher number of citations.

## 6 Interpretation of results and implications for management research

### 6.1 Summary of main results and interpretations

We provide a comprehensive overview of replication research in the management field by outlining the prevalence, types, outcomes, and impact of replication studies in the leading management journals.

<sup>4</sup> Table 6 re-estimates the same model with a OLS regression. The results are similar in magnitude and significance.

<sup>5</sup> In the OLS re-estimation (Table 6), no differences emerge across the sub-disciplines.

**Table 4** Descriptive statistics, correlations, and variance inflation factors

	SD	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	VIF
(1) Ci-ta-	0.00																						1.43
(2)	0.02	0.13	0.00	1.00	0.03																		1.07
(3)	0.38	0.49	0.00	1.00	-0.05	-0.10																	1.16
(4)	0.01	0.09	0.00	1.00	0.03	-0.01	-0.07																1.09
(5)	0.02	0.13	0.00	1.00	0.03	-0.02	-0.10	-0.01															1.07
(6)	0.58	0.50	0.00	1.00	0.03	<b>-0.15</b>	<b>-0.92</b>	-0.11	<b>-0.15</b>														-
(7)	0.20	0.40	0.00	1.00	<b>-0.16</b>	-0.07	0.03	-0.05	-0.07	0.01													1.51
Out-																							
(8)	0.48	0.50	0.00	1.00	0.04	<b>0.14</b>	-0.06	-0.09	0.01	0.04	<b>-0.49</b>												1.42
Out-																							
(9)	0.32	0.47	0.00	1.00	0.10	-0.09	0.04	<b>0.13</b>	0.05	-0.05	<b>-0.34</b>	<b>-0.65</b>											-
Out-																							
(10)	0.05	0.21	0.00	1.00	-0.04	-0.03	0.03	-0.02	-0.03	-0.01	-0.06	-0.05	0.11										-
(11)	0.30	0.46	0.00	1.00	0.05	0.06	0.04	-0.06	0.13	-0.08	-0.11	0.00	0.10	<b>-0.14</b>									5.58
(12)	0.02	0.14	0.00	1.00	0.11	-0.02	0.01	-0.01	-0.02	0.01	0.00	-0.02	0.03	-0.03	-0.1								1.50
(13)	0.09	0.29	0.00	1.00	-0.05	0.07	-0.01	-0.03	-0.04	0.01	0.09	-0.02	-0.06	-0.07	<b>-0.21</b>	-0.05							2.87
(14)	0.03	0.16	0.00	1.00	0.12	-0.02	-0.02	-0.01	-0.02	0.03	-0.01	0.06	-0.05	-0.04	-0.11	-0.02	-0.05						1.64
(15)	0.05	0.23	0.00	1.00	-0.07	0.11	-0.07	-0.02	-0.03	0.05	-0.08	<b>0.14</b>	-0.08	-0.05	<b>-0.16</b>	-0.03	-0.08	-0.04					2.20
(16)	0.34	0.48	0.00	1.00	0.00	-0.09	-0.09	<b>0.13</b>	-0.09	0.12	-0.02	0.03	-0.02	<b>-0.16</b>	<b>-0.48</b>	-0.11	<b>-0.23</b>	-0.12	<b>-0.17</b>				6.08
(17)	0.12	0.32	0.00	1.00	-0.07	-0.05	0.12	-0.03	0.05	-0.11	<b>0.20</b>	-0.11	-0.05	-0.08	<b>-0.24</b>	-0.05	-0.12	-0.06	-0.09	<b>-0.26</b>			3.36
(18)	8.60	6.39	0.00		<b>-0.14</b>	0.04	<b>0.22</b>	0.11	-0.03	<b>-0.24</b>	-0.03	0.00	0.03	-0.02	0.02	-0.05	-0.03	0.06	-0.03	<b>-0.14</b>	<b>0.23</b>		1.52
(19)		2.00			<b>0.15</b>	0.00	-0.10	0.12	0.02	0.07	0.05	-0.06	0.02	-0.12	-0.04	<b>-0.14</b>	-0.08	<b>0.17</b>	-0.04	<b>0.27</b>	<b>-0.18</b>	<b>0.20</b>	1.50
Ar-ticle age																							

**Table 4** (continued)

	SD	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	VIF			
(20)	0.37	0.48	0.00	1.00	0.11	-0.03	-0.06	0.03	0.04	0.05	-0.02	-0.04	0.06	0.08	<b>-0.16</b>	-0.05	<b>-0.15</b>	-0.01	-0.07	<b>0.16</b>	<b>0.15</b>	<b>-0.23</b>	-0.01	1.28		
(21)	6.12	1.39	1.39	9.17	<b>0.32</b>	0.03	-0.05	-0.08	0.11	0.02	-0.01	0.04	-0.04	0.05	0.10	0.00	0.09	0.05	-0.05	<b>-0.26</b>	<b>0.14</b>	<b>0.23</b>	-0.05	1.52		
Ci-																										
ta-																										
(22)	2.23	1.05	1.00		0.01	-0.09	0.10	0.02	-0.06	-0.06	0.06	0.03	-0.09	0.03	-0.03	0.00	0.00	-0.01	0.05	0.04	-0.07	<b>-0.15</b>	<b>-0.26</b>	0.05	0.05	1.18
#																										
u-																										

Notes: Variance inflation factors (VIF) calculated based on Table 5 (Model 3). Correlation coefficients highlighted in bold denote \*  $p < 0.05$ .  $N = 240$

**Table 5** Negative binomial regression analysis (dependent variable: citations of the replication study)

Model	(1)		(2)		(3)	
Variables	Coeff.	(SE)	Coeff.	(SE)	Coeff.	(SE)
<i>Characteristics of replication study</i>						
Type: literal	0.614	(0.314)*			0.751	(0.428)*
Type: constructive	0.133	(0.182)			0.219	(0.119)*
Type: regressive	0.567	(0.248)**			1.214	(0.272)***
Type: confounded	0.147	(0.440)			-0.277	(0.299)
Type: quasirandom	Ref.				Ref.	
Outcome: not replicated	-0.769	(0.248)***			-0.634	(0.190)***
Outcome: partially replicated	-0.192	(0.151)			-0.087	(0.171)
Outcome: fully replicated	Ref.				Ref.	
Sub-discipline: ETHICS-CSR-MAN	0.527	(0.391)			0.439	(0.355)
Sub-discipline: INNOV	1.032	(0.591)*			1.030	(0.675)*
Sub-discipline: IB&AREA	0.321	(0.500)			0.254	(0.373)
Sub-discipline: OR&MANSCI	1.166	(0.367)***			1.359	(0.341)****
Sub-discipline: ORG STUD	-0.083	(0.375)			0.064	(0.356)
Sub-discipline: PSYCH (WOP-OB)	0.473	(0.467)			0.330	(0.360)
Sub-discipline: STRAT	0.541	(0.405)			0.302	(0.319)
Sub-discipline: ENT-SBM	Ref.				Ref.	
Publication lag to original study	-0.043	(0.009)***			-0.080	(0.018)***
<i>Characteristics of the original study</i>						
Age of original study			0.020	(0.007)***	0.034	(0.012)***
Published in same journal			0.279	(0.166)*	0.109	(0.125)
Citations original study (log.)			0.418	(0.069)***	0.481	(0.064)***
# Authors original study			0.022	(0.093)	0.013	(0.074)
Observations	240		240		240	
Log-Likelihood	-1396.838		-1382.766		-1355.170	

Notes: Standard errors clustered at the journal level in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

First, our findings document that independent replication studies are rarely published in the leading management journals. We were only able to identify a sample of 240 independent replication studies despite engaging in a comprehensive research and coding effort in 56 management journals. The sub-disciplines with the highest prevalence of independent replications were (organizational) psychology (82 replica-

tions in 13 journals), general management, ethics, and social responsibility (73 replications in 12 journals), and strategy (28 replications in 4 journals). Despite ongoing calls for an increase in replication studies (e.g., Aguinis et al., 2017; De Massis et al., 2020; Maula and Stam, 2020), an increase in replication studies over time was not evident. Regarding the type of replication studies that are published, we found that

**Table 6** OLS regression analysis (dependent variable: log(citations of the replication study + 1))

Model	(1)		(2)		(3)	
Variables	Coeff.	(SE)	Coeff.	(SE)	Coeff.	(SE)
<i>Characteristics of replication study</i>						
Type: literal	0.521	(0.992)			0.511	(0.879)
Type: constructive	-0.246	(0.192)			0.039	(0.153)
Type: regressive	1.369	(0.342)***			1.708	(0.374)***
Type: confounded	0.652	(0.440)			-0.023	(0.369)
Type: quasirandom	Ref.				Ref.	
Outcome: not replicated	-0.563	(0.296)*			-0.658	(0.222)***
Outcome: partially replicated	-0.111	(0.184)			-0.104	(0.183)
Outcome: fully replicated	Ref.				Ref.	
Sub-discipline: ETHICS-CSR-MAN	0.496	(0.626)			0.397	(0.514)
Sub-discipline: INNOV	0.475	(0.838)			0.852	(0.642)
Sub-discipline: IB&AREA	0.040	(0.705)			-0.066	(0.552)
Sub-discipline: OR&MANSCI	1.415	(0.611)**			0.706	(0.518)
Sub-discipline: ORG STUD	-0.224	(0.659)			-0.048	(0.475)
Sub-discipline: PSYCH (WOP-OB)	0.378	(0.688)			0.294	(0.528)
Sub-discipline: STRAT	0.196	(0.623)			0.285	(0.505)
Sub-discipline: ENT-SBM	Ref.				Ref.	
Publication lag to original study	-0.054	(0.014)***			-0.097	(0.019)***
<i>Characteristics of the original study</i>						
Age of original study			0.031	(0.008)***	0.038	(0.011)***
Published in same journal			0.376	(0.191)*	0.130	(0.182)
Citations original study (log.)			0.371	(0.054)***	0.469	(0.068)***
# Authors original study			0.016	(0.159)	-0.027	(0.126)
Observations	240		240		240	
Log-Likelihood	-429.369		-422.039		-393.441	

Notes: Standard errors clustered at the journal level in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

the majority of these studies are quasirandom (57.9%). However, this is the type of replication study of questionable value (Köhler and Cortina, 2021). Concerning the outcomes of the replication studies, we found that 79.6% of the replication studies at least partially replicate the results of the original study. A speculative interpretation would be that top management journals suffer from a bias against publishing replication studies that contradict the original study. Finally, with regard to the impact of the replication studies, the studies in our sample are cited 142.6 times, on average. Regression analysis shows that the replication studies that do not confirm the original study are cited less, further suggesting a profound bias against non-replicated research.

These results contribute to the recent discussion about the status quo and value of replication in management research (e.g., Aguinis et al., 2017; Bergh et al., 2017; De Massis et al., 2020; Köhler and Cortina, 2021; Maula and Stam, 2020). Our study adds to the discussion of how open the management field is to the topic of replications by documenting the status quo of independent replication research in management. Overall, our findings suggest that the state of replication studies in management is grim and in need of attention. This is particularly true for the field's top journals.

## 6.2 Implications for management research

Our results also have implications for (potential) authors, reviewers, and editors of management research. Prior research describes that a major reason for the lack of replication studies is the strong emphasis of the field on producing novel theoretical contributions, which are often a prerequisite for publishing in the top journals (e.g., Corley and Gioia, 2011). As a consequence, replication studies may be unattractive because they can be criticized for a lack of theoretical contribution and novelty (Köhler and Cortina, 2021). The focus on theoretical contribution presents a critical barrier that management research needs to overcome. We highlight and discuss three further observations that emerge from and that our study and that have received little attention so far.

First, a lack of consensus in the field can be observed; it is not clearly defined what a good replication study constitutes and what contribution it should deliver to get published in the leading management journals. Having analyzed the 240 replications identified in our sample of replication studies, we recognize that a huge variety exists. While Köhler and Cortina (2021) list many examples of good replications and even provide templates, we miss editorial guidance from the leading journals so far. In our view, such consensus and guidance would encourage more researchers to conduct a replication of other researchers' work. Questions that need to be addressed include the following: What constitutes a sufficient theoretical or empirical contribution in the case of a replication study? Is an extension of the original study needed to achieve an acceptable level of theoretical novelty? What type of submission format is best suited for replications? Should they be submitted in a research note format? To what extent is an own theory or hypothesis section required?

Second, another important issue concerns the data availability of the original studies. Unless the authors of the original studies use archival data from publicly available or readily accessible databases, it is often impossible for authors of replication

studies to conduct narrow forms of replications (e.g., literal replication). This data unavailability may explain why literal replications are very rare in our sample and why most replications are either constructive or quasirandom. Leading journals in neighboring disciplines such as finance (e.g., *Journal of Finance*) or economics (e.g., *American Economics Review*, *Quarterly Journal of Economics*) are one step ahead of management and have made it compulsory for authors of accepted papers to upload their datasets. This makes it possible for other researchers to conduct narrow forms of replications (see Harvey, 2019 for a discussion of code and data sharing policies in finance and economics). Such open data policies have led to impactful replications that uncovered inconsistencies in original studies and resulted in their retraction. A recent example from the *Journal of Finance*, the preeminent journal in the finance field, is the literal replication by Guest (2021). Such retractions are even more common in the natural sciences, as indicated by Retraction Watch's list of the top 10 most highly cited retracted papers (Retraction Watch, 2021). We did not discover such an extreme case during our search for replication studies in management, which solidifies our view that management research may have a blind spot in this regard.

We would also like to highlight that some management journals already have similar policies in place. One of the forerunners is *Management Science*, which has required authors of accepted papers to make their data and code available for the sake of replicability since 2019 (*Management Science*, 2021). In our view, more journals should follow these examples as this also has important research implications: Future survey-based or experimental research could investigate the factors or barriers that motivate authors to engage in replications and about why they choose a particular original study for replication. For example, possible factors or barriers include data availability, the impact of the original study, the reputation of the journal or the authors of the original study, and tenure requirements. From the perspective of the journals, a better understanding of these factors would be an essential prerequisite for understanding how to stimulate replication studies. Also, such findings would connect to the initial research conducted by Mueller-Langer et al. (2019) who have used bibliographic data to understand which studies are chosen for replication in economics.

Our third discussion point concerns the confirmation bias that seems to exist in management research. Our empirical analysis indicates that replication studies that do *not* confirm the results of the original study are rarely published in the top journals and receive fewer citations. This adds to a recent discussion on the practice of vote-counting in the meta-analytical literature (Anderson & Maxwell, 2016; Maxwell et al., 2015) according to which non-significant results should not necessarily be seen as failure but as another piece of evidence needed to enhance and move the respective research field forward. An important step to rectifying this situation would be to make editors and reviewers aware of the value of such research in the review process. This creation of awareness is already in progress and has been described in several recent editorials (e.g., Aguinis et al., 2017; Bergh et al., 2017; De Massis et al., 2020). Some top management journals such as *Entrepreneurship Theory and Practice* have also begun to publish non-confirming replications (e.g., Block et al., *in press*). However, the process should not stop there. Once a negative or non-result is published, the scholarly community needs to acknowledge its existence and give it due credit.



From the perspective of a submitting author, it is often more convenient or less risky to still reference or cite the original study even though replications could not confirm their main results. Editors, publishers, and journal administrators should also learn to “celebrate” a non-result or a result that does not confirm prior research. Concerning the nature and extent of the confirmation bias, bibliographic research could analyze and interpret the forward citations of replications and original studies over time to understand what drives this confirmation bias. Similarly, survey-based research could ask authors about their citation and referencing behavior to better understand their conscious or unconscious motivations to cite a particular replication or original study. Moreover, experimental research designs in the form of scenarios or conjoint experiments constitute a promising option. Another interesting research direction would be to analyze how the confirmation bias interacts with the bias described by Mueller-Langer et al. (2019): how does the confirmation bias depend on the impact of the replicated article as well as the reputation and quality of the respective authors and journals? The question is of high importance as it shows directly how difficult or challenging it is to update the most impactful management research and to what extent dangerous path dependencies and lock-in situations exist that have already been observed in other contextual situations (Arthur, 1989).

**Funding** The authors declare that no funds, grants, or other support were received during.

**Open Access funding enabled and organized by Projekt DEAL.**

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- ABS (2018) Academic journal guide by the Chartered Association of Business Schools. Available at: <https://facultystaff.richmond.edu/~tmattson/AJG%202018%20Journal%20Guide.pdf>. Last accessed: August 27, 2021
- Aguinis H, Cascio WF, Ramani RS (2017) Science's reproducibility and replicability crisis: International business is not immune. *J Int Bus Stud* 48:653–663
- Anderson SF, Maxwell SE (2016) There's more than one way to conduct a replication study: Beyond statistical significance. *Psychol Methods* 21(1):1–12
- Arthur WB (1989) Competing technologies, increasing returns, and lock-in by historical events. *Econ J* 99(394):116–131

- Bergh DD, Sharp BM, Aguinis H, Li M (2017) Is there a credibility crisis in strategic management research? Evidence on the reproducibility of study findings. *Strategic Organ* 15(3):423–436
- Bettis RA, Helfat CE, Shaver JM (2016) The necessity, logic, and forms of replication. *Strateg Manag J* 37(11):2193–2203
- Block J, Kuckertz A (2018) Seven principles of effective replication studies: strengthening the evidence base of management research. *Manage Rev Q* 68(4):355–359
- Block J, Hansen C, Steinmetz H (in press) Are family firms doing more innovation output with less innovation input? A replication and extension. *Entrepreneurship Theory and Practice*
- Camerer CF et al (2016) Evaluating replicability of laboratory experiments in economics. *Science* 351(6280):1433–1436
- Camerer CF et al (2018) Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nat Hum Behav* 2(9):637–644
- Corley KG, Gioia DA (2011) Building theory about theory building: what constitutes a theoretical contribution? *Acad Manage Rev* 36(1):12–32
- De Massis A, Kellermanns F, Wright M, Brinkerink J (2020) Replication and validation in family business research. *J Family Bus Strategy* 11:100415
- Dettori A, Floris M (2022) Improving continuity by simplifying the structure of family firms: a replication study. *Manage Rev Q* 1–26. <https://doi.org/10.1007/s11301-021-00255-4>
- Dau LA, Santangelo GD, van Witteloostuijn A (2021) Replication studies in international business. *Journal of International Business Studies* 1–16
- Guest PM (2021) Risk Management in Financial Institutions: A Replication. *The Journal of Finance*, forthcoming
- Harvey CR (2019) Editorial: replication in financial economics. *Crit Finance Rev* 8:1–9
- Harzing AW (2013) A preliminary test of Google Scholar as a source for citation data: a longitudinal study of Nobel prize winners. *Scientometrics* 94(3):1057–1075
- Hensel PG (2021) Reproducibility and replicability crisis: How management compares to psychology and economics e A systematic review of literature. *Eur Manag J* 39(5):577–594
- Hubbard R, Vetter DE (1996) An empirical comparison of published replication research in accounting, economics, finance, management, and marketing. *J Bus Res* 35(2):153–164
- Hubbard R, Vetter DE, Little EL (1998) Replication in strategic management: scientific testing for validity, generalizability, and usefulness. *Strateg Manag J* 19(3):243–254
- Köhler T, Cortina JM (2021) Play it again, Sam! An analysis of constructive replication in the organizational sciences. *J Manag* 47(2):488–518
- Lykken DT (1968) Statistical significance in psychological research. *Psychol Bull* 70(3):151–159
- Management Science (2021) Management Science policy for data and code disclosure. Available at: <https://pubsonline.informs.org/page/mnsc/datapolicy#1>. Last accessed: August 27, 2021
- Maula M, Stam W (2020) Enhancing rigor in quantitative entrepreneurship research. *Entrepreneurship Theory and Practice* 44(6):1059–1090
- Maxwell SE, Lau MY, Howard GS (2015) Is psychology suffering from a replication crisis? What does “failure to replicate” really mean? *Am Psychol* 70(6):487–498
- Mueller-Langer F, Fecher B, Harhoff D, Wagner GG (2019) Replication studies in economics—How many and which papers are chosen for replication, and why? *Res Policy* 48(1):62–83
- Open Science Collaboration (2015) Estimating the reproducibility of psychological science. *Science* 349(6251):aac4716
- Retraction Watch (2021) Top 10 most highly cited retracted papers. Available at: <https://retractionwatch.com/the-retraction-watch-leaderboard/top-10-most-highly-cited-retracted-papers>. Last accessed: August 27, 2021
- Ryan JC, Tipu SA (2022) Business and management research: Low instances of replication studies and a lack of author independence in replications. *Res Policy* 51(1):104408
- Stroebe W, Strack F (2014) The alleged crisis and the illusion of exact replication. *Perspect Psychol Sci* 9(1):59–71
- Szumal JL, Boglarsky CA, Cooke RA (2021) Thinking and behavioral styles as described by self versus others: a replication and extension with male and female managers. *Management Review Quarterly*, 1–30
- Tsang EW, Kwan KM (1999) Replication and theory development in organizational science: a critical realist perspective. *Acad Manage Rev* 24(4):759–780
- Van Scotter JR (2020) Narcissism in CEO research: a review and replication of the archival approach. *Manage Rev Q* 70(4):629–674

- Yuan S, Rieger MO, Caliskan N (2020) Maxing out: the puzzling influence of past maximum returns on future asset prices in a cross-country analysis. *Manage Rev Q* 70(4):567–589
- Wright TA, Sweeney DA (2015) The call for an increased role of replication, extension, and mixed-methods study designs in organizational research. *J Organizational Behav* 37(3):480–486

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Jörn H. Block<sup>1,2,3</sup> · Christian Fisch<sup>4</sup> · Narmeen Kanwal<sup>1</sup> · Solvej Lorenzen<sup>3</sup> · Anna Schulze<sup>1</sup>

---

✉ Jörn H. Block  
block@uni-trier.de

Christian Fisch  
christian.fisch@uni.lu

Narmeen Kanwal  
narmeenkanwal786@yahoo.com

Solvej Lorenzen  
solvej.lorenzen@uni-wh.de

Anna Schulze  
schulzea@uni-trier.de

<sup>1</sup> Faculty of Management, Trier University, 54296 Trier, Germany

<sup>2</sup> Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

<sup>3</sup> Witten Institute for Family Business, Witten/Herdecke University, 58448 Witten, Germany

<sup>4</sup> Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, 29, Avenue John F. Kennedy, 1855 Luxembourg, Luxembourg