Edinburgh Research Explorer

One finding is no finding

Citation for published version:
https://doi.org/10.1016/j.jfbs.2022.100521

Digital Object Identifier (DOI):
10.1016/j.jfbs.2022.100521

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published in:
Journal of Family Business Strategy

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 15. Sep. 2023
One finding is no finding: Toward a replication culture in family business research

Jasper Brinkerink a,∗, Alfredo De Massis b,d,e,f, Franz Kellermanns c,g

a Business School, University of Edinburgh, Edinburgh, UK
b Faculty of Economics and Management, Free University of Bozen-Bolzano, Bolzano, Italy
c Belk College of Business, University of North Carolina – Charlotte, USA
d Lancaster University Management School, Lancaster, UK
e IMD Business School, Lausanne, Switzerland
f Zhejiang University, Zhejiang, China
g WHU–Otto Beisheim School of Management, Vallendar, Germany

ARTICLE INFO

Keywords: Replication Validation Family business Theory Theory testing

ABSTRACT

Our goal is to foster the development of a healthy replication culture in family business research. Replication, which advances theory by confronting existing understanding with new evidence, is of paramount importance in creating a meaningful cumulative knowledge base. In the family business field, however, as in many other fields within the broader management literature, dedicated replications are largely absent. After a brief analysis of the likely causes and consequences of our collective avoidance of replication studies, we examine four types of replication of particular importance to the field and provide guidelines and recommendations for family business scholars interested in conducting such research. We invite journals and their editors to reflect on the role they can play in changing the incentive structures to conduct and submit useful replication studies and provide actionable suggestions for improvement. We illustrate contemporary examples of family business knowledge advancement through replication research.

1. Introduction

A number of high-profile cases of academic fraud and misconduct have spurred reflection on the credibility of research efforts across social science disciplines (Bergh, Sharp, Aguinis, & Li, 2017; Byington & Felps, 2017; Earp & Trafimow, 2015; Nelson, Simmons, & Simonsohn, 2018; Tourish, 2019). Among other outcomes of such collective soul searching, large-scale metascientific evidence highlights the poor replicability—that is, the extent to which consistent results are obtained when a study is repeated—of even the most carefully designed studies in disciplines such as economics (Camerer et al., 2016), psychology (Open Science Collaboration, 2015), and the wider social sciences (Camerer et al., 2018). These efforts empirically substantiate an uncomfortable notion that, until recently, has been ignored for lack of evidence: “published and true are not synonyms” (Nosek, Spies, & Motyl, 2012, p. 616);

results of many individual research studies will be incorrect (Ioannidis, 2005) and/or severely inflated (Ioannidis, 2008).

The overabundance of false positives in academic research may be due, in part, to the behavioral consequences of bias on the part of authors (Gelman & Loken, 2014; Simmons, Nelson, & Simonsohn, 2011), or editors and reviewers (Franco, Malhotra & Simonovits, 2014; Gerber & Malhotra, 2008) yet may also result from the many sources of endogeneity, sampling and measurement error, and other forms of “noise” that are often hard to circumvent in empirical research (Loken & Gelman, 2017; Zhang, Fang, Duf, & Chrisman, 2021).

Whatever the root cause of what has become popularly known as science’s “replication crisis”, a solution can only be realized through a collective commitment to increasing the systematic empirical scrutiny to which we subject potentially influential research findings and theoretical arguments before accepting them as common understanding.

In Remembrance of Mike Wright. We were saddened when one of the scholars with whom we started this journey, Mike Wright, passed away on November 25, 2019. Mike was a pioneering scholar whose research and service had a profound impact on our profession and influenced several generations of scholars and students. The esteem in which he was held by colleagues throughout the world was reflected by his several important roles as an editor and a leading scholar. We dedicate this article to our mentor and dear friend Mike Wright and his legacy—Thanks for being such a supporting and inspiring person Mike!

E-mail addresses: jasper.brinkerink@ed.ac.uk (J. Brinkerink), alfredo.demassis@unibz.it (A. De Massis), kellermanns@uncc.edu (F. Kellermanns).

https://doi.org/10.1016/j.jfbs.2022.100521

Available online 28 September 2022
1877-8585/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Table 1

<table>
<thead>
<tr>
<th>Replication type</th>
<th>Reproduction</th>
<th>Direct replication</th>
<th>Conceptual replication</th>
<th>Empirical validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose</td>
<td>Assess truthfulness of findings and detect potential errors in original study.</td>
<td>Assess generalizability of significance and magnitude of findings across repeated tests. Impossible.</td>
<td>Assess sensitivity of significance and magnitude of findings to methods/operationalization. Possible.</td>
<td>Assess empirical validity of conceptual predictions. Reversed (i.e., theory → empirical manifestation)</td>
</tr>
<tr>
<td>Theoretical abstraction</td>
<td>(Potential) high-impact study. We believe however that verifying the reproducibility of novel findings should be a structural prerequisite to the publication responsibility of the journals. The challenge resides rather than in the original authors’ academic peers after publication.</td>
<td>(Potential) high-impact study. Especially those based on small / idiosyncratic samples and with complicated hypotheses.</td>
<td>(Potential) high-impact study.</td>
<td>(Potential) high-impact conceptual or empirical theory-building study with un(der)tested propositions or theoretical claims.</td>
</tr>
<tr>
<td>Best targeted original study</td>
<td>Same sample, same methods and measurements</td>
<td>Different sample, same methods / operationalizations</td>
<td>Same or different sample, different methods / operationalizations</td>
<td>Correspondence with original authors before collecting and analyzing data to ensure conceptual validity of intended methods could be helpful. Be mindful though of risk of being ‘influenced away’ from result neutrality.</td>
</tr>
<tr>
<td>Data and methods w.r.t. original study authors</td>
<td>In pre-publication reproducibility checks by journals or third-party representatives of the journals, original authors supply their materials with to those performing the check. In post-publication reproduction attempts of peer researchers, original authors can be helpful by sharing materials and/or reconcile potential differences/uncertainties in communication with reproducing researchers.</td>
<td>Perhaps supply code and other materials. If original authors are unresponsive or replicating authors decide it is better to not contact original authors, replicating authors may instead follow procedures reported in original study as closely as possible.</td>
<td>Not necessary to correspond, unless replicating authors want to use the same data/sample and it is not publicly available.</td>
<td>Conceptually valid sample, methods, and operationalizations</td>
</tr>
<tr>
<td>Role of original study authors</td>
<td>Required</td>
<td>Strongly recommended</td>
<td>Strongly recommended</td>
<td>Strongly recommended</td>
</tr>
<tr>
<td>Independence of replicating authors w.r.t. original study</td>
<td>Not recommended (sensitive) Reproduction work is generally not done for publication purposes.</td>
<td>Recommended Paper can be a short research note, where replicating authors: 1. Introduce the original study, explain its influence, and the reason for doing a direct replication. 2. Briefly restate the original study’s hypotheses/ reasoning. 3. Explain sample and methods in detail, highlighting especially how, where, and why the replication deviates from the original study. 4. Report analyses and findings. 5. Provide a comparison of the replication’s and the original findings, and provide nuanced interpretations of (proposed) reasons for potential differences.</td>
<td>Strongly recommended Paper can be a short research note, where replicating authors: 1. Introduce the original study, explain its influence, and the reason for doing a conceptual replication. 2. Briefly restate the original study’s hypotheses/ reasoning. 3. Explain sample and methods in detail, highlighting especially how, where, and why the replication deviates from the original study. 4. Report analyses and findings. 5. Provide a comparison of the replication’s and the original findings, and provide nuanced interpretations of (proposed) reasons for potential differences.</td>
<td>Strongly recommended Paper can be a short research note, where validating authors: 1. Introduce the original study, explain its influence, and the reason for doing the validation. 2. Briefly describe the original study’s conceptual argument and propositions/predictions. 3. Explain sample and methods in detail, highlighting especially the conceptual/construct validity vis-à-vis the constructs and relationships theorized on in the original study. 4. Report analyses and findings. 5. Provide a comparison of the validation’s findings and the original predictions, and provide nuanced interpretations of (proposed) reasons for potential differences.</td>
</tr>
</tbody>
</table>

(Coffman & Niederle, 2015; Munafo et al., 2017). Recent literature surveys, however, document that in the wider management and economics disciplines, replication is still neither a priority for researchers (Block, Fisch, Kanwal, Lorenzen, & Schulze, 2022; Mueller-Langer, Fecher, Harhoff, & Wagner, 2019; Ryan & Tipu 2022) nor high on the agenda of journals and their editors (Easley, Madden, & Gray, 2013; Hensel, 2019). Moreover, published replication studies are rarely independent – that is, conducted by researchers not related to the original work and its authors – and, therefore, not likely to have been result-neutral ex ante (Ryan & Tipu, 2022).

The family business field is no exception. At the time we started developing this article, we were able to identify only three dedicated independent replications (Hirshich & Cahill, 1995; López-Delgado & Dieguez-Soto, 2015; Weismeier-Sammer, 2011) in a literature that now easily spans thousands of publications (Rovelli, Ferasso, De Massis, & Kraus, 2021) and also is not immune to biased research practices (Brinkerink, 2021). The lack of replication culture may be particularly worrisome for family business research. Not only do we, like researchers in other management fields, risk basing future theories on insufficiently tested underlying assumptions (Brinkerink, 2022; Tsang, 2021), our research is embedded in an applied empirical context that has a clear nonacademic target audience: family businesses and related practitioners. Hence, persistent false positives in the family business literature could be disproportionately damaging, as managerial and policy recommendations based on nonreplicable findings may enter the practical arena directly and thereby encourage real-world interventions based on erroneous or highly idiosyncratic evidence.

Our collective failure to replicate has a quite simple explanation: there is no incentive to do so. Generally speaking, journals and their editors expect authors to submit findings that provide new insights and...
an original contribution to theory (Hambrick, 2007; Neubaum & Mie- lotta, 2021). Replication, which advances theory by confronting existing understanding with new evidence (Nosek & Errington, 2020), is therefore at odds with the prevailing editorial demands for a priori theoretical novelty. From an opportunity cost perspective, pursuing replication studies may thus even be disincentivized as it siphons time and resources away from other (original) research that is easier to publish.

It is in this spirit that we followed recent examples set by colleagues in fields such as strategic management (Bettis, Helfat, & Shaver, 2016), marketing (Babin, Ortinau, Herrmann, & Lopez, 2021), and entrepreneurship (Van Witteloostuijn, Dejardin, & Pollack, 2021) to provide an incentive for the replication of prior influential findings and well-cited yet untested conceptual arguments in the family business literature (De Massis, Kellermanns, Wright, & Brinkerink, 2020). Besides briefly introducing some contemporary examples of family business replication studies that were conducted in response to our call, in what follows we present guidelines and considerations that will help authors develop useful and high-quality replication studies. We also provide recommendations to journals and their editors, who hold the key to changing the incentive structures currently preventing the establishment of a healthy replication culture in family business research. Our hope is that this article will serve as both a benchmark and an inspiration for members of the wider family business research community willing to give replication a more prominent role in the field.

2. Guidelines and recommendations for authors

Given the idiosyncrasies of research practices and processes applied, common modes and objectives of replication vary widely across scientific disciplines, Block and Kuckertz (2018) highlight that in the natural sciences potentially groundbreaking findings are only accepted as scientific truths when independent research teams using the exact same experimental conditions repeatedly obtain the exact same results. Such exact reproduction of results across repeated tests, however, is not feasible in social science fields such as family business. Unable to predict human (and, by extension, organizational) behavior with absolute accuracy, we need to rely on replications of results across samples and methods rather than exact reproductions to gradually establish “quasi-laws on how human beings in organizations or organizations themselves behave” (Block & Kuckertz, 2018, p. 356). Without (cl)aiming to be exhaustive, and after reflecting on considerations relevant to all forms of replication, we consider four types of replication studies that may be especially helpful given the central weaknesses of the family business literature to date: reproduction (i.e., as a means of detecting errors and assessing a study’s truthfulness), direct replication, conceptual replication, and empirical validation. Table 1 offers a side-by-side comparison of these four replication types and describes the characteristics of each.

2.1. Considerations across replication types

2.1.1. Result neutrality

The ideal of the value-neutral scientist is, especially in the social sciences, neither feasible nor desirable per se (Tsui, 2016). Also in replication work, researchers must judge which original studies are, and which are not, worthy of being replicated. However, researchers do need to maintain absolute neutrality toward the results of their replication studies. Regarding the overarching purpose of replication – advancing theory by confronting existing understanding with new evidence (Nosek & Errington, 2020) – the only thing worse than a family business literature without replications would be one full of biased replications. It should not be underestimated how difficult it is to remain result neutral throughout the replication process. At a basic level, researchers must check their own (very human) desire to find at least “something” (Simmons et al., 2011). Indeed, one reason why null (i.e., nonsignificant) results are not published is researchers’ own unwillingness to submit them for publication to begin with (Franco et al., 2014). Other stakeholders may, either actively or in spirit, exert their own pressures on the replicating researchers’ efforts to remain objective. Journal editors or reviewers may enforce HARKing (hypothesizing after results are known) a submitted replication’s findings if these are significant yet in the opposite direction of those in the original study (Kerr, 1998; Murphy & Aguinis, 2019), and may be more interested in replication findings that overturn prior results in general (Gertler, Galiani, & Romero, 2018). On the other hand, the original study’s authors will generally not enjoy seeing their findings questioned and/or overturned, anticipated consequences of which may invite a certain confirmation bias in a replication project. Hence, it is important to be mindful of such forces irrespective of the type of replication undertaken. Generally speaking, we therefore advise against replicating one’s own prior work, or that of one’s current or former supervisors or frequent co-authors.

2.1.2. Preventing arbitrary and ad-hoc decisions

Researchers can maintain result neutrality by preventing arbitrary and ad-hoc decisions throughout the replication process. One way to do so is by publicly preregistering a replication protocol prior to data collection and analysis, which details the hypotheses to be tested, measurements to be collected, the targeted sample size in relation to desired statistical power, analytical procedures, and other elements of the replication plan (Nosek, Ebersole, DeHaven, & Mellor, 2018). Researchers can then hold themselves accountable by adhering to the preregistered protocol irrespective of the findings their replication generates, making note in the submitted article of any unforeseen deviations. Having an ex ante preregistered protocol may also aid in resisting reviewers’ and/or editorial pressures for post-analysis changes in theory or hypotheses. We also encourage researchers to follow a replication template such as that of Brandt et al. (2014), which, although targeted at experimental psychologists, provides excellent detailed guidance on how to mindfully conduct and report a meaningful replication study.

2.1.3. Finding an original study to replicate

In essence, every scientific claim should be subject to collective scrutiny. Nevertheless, we mirror Crawford et al.’s (2022) “suggestion to concentrate on those studies that have a marked influence on a domain’s thinking” (p. 3), of which the core findings or theoretical propositions have been subject to scant subsequent empirical scrutiny. Although we agree with Block and Kuckertz (2018, p. 356) that “such studies are most likely (but not necessarily) published in top tier management [and, in our literature, also entrepreneurship (cf. Chrisman, Chua, Kellermanns, Matherne, & Debicki 2008)] journals,” we also recognize that the quality of studies and empirical rigor in those journals varies substantially (Anderson, Wennberg, & McMullen 2019; Maula & Stam, 2020; cf. Starbuck, 2005). Potentially, the most interesting studies to target are those highly cited top-tier journal publications with apparent empirical limitations (e.g., underpowered, weak measurements, not addressing obvious potential sources of endogeneity) or those published in our most

1 But see how one can transparently (as opposed to secretly) HAR King in a way that may actually be meaningful to scientific progress in certain circumstances (Hollenbeck & Wright, 2017).

2 Internal replications – i.e., replicating one’s own findings in sequential studies presented within the same paper (Lamberger, 2019) – may be an exception to this general rule. Researchers have to be mindful, though, of the risk of selective reporting involved in these internal replications and internal meta-analyses (Ueno, Fastrich, & Murayama, 2016; Vogserau, Simonson, Nelson, & Simmons, 2019).

3 Preregistrations can be uploaded on online registries such as those hosted by the Center for Open Science (see https://www.cos.io/initiatives/prereg) and the American Economic Association (https://www.socialscienceregistry.org/). A particularly straightforward and user-friendly format is available at https://as predicted.org/.
2.2. Four types of replication

2.2.1. Reproduction

A reproduction tries to obtain an original study’s exact same findings, normally using the same data, code, and other materials to verify the study’s truthfulness and freedom from error (Peng, 2011). In an ideal world, researchers would not (intentionally or unintentionally) misrepresent study findings and/or procedures, only submitting work for publication that can be exactly reproduced by others. Evidence suggests, however, that a substantial proportion of published findings cannot be reproduced when using the exact same materials (Bergh et al., 2017; McCulloough, McGearry, & Harrison, 2006), even when communicating with the original authors (Crawford et al., 2022). These empirical observations call for better checks and quality assurance mechanisms to prevent the publication of nonreproducible research, perhaps by increasing the anticipated probability that submitted findings will eventually undergo some form of reproduction. While reproduction can of course be performed by independent peer researchers after an original study’s publication (cf. Crawford et al., 2022), we advise researchers to first carefully consider the risks involved.

Leaving verification of the reproducibility of published work up to the original authors’ academic peers may be somewhat undesirable, whether findings eventually turn out to be reproducible or not. Most obviously, an exact (or near-exact) reproduction of prior findings is unpublishable on its own, thereby skewing ex ante incentives toward finding irregularities. Even when reproducing researchers do their best to retain result-neutrality however, if findings cannot be (completely) reproduced it is not always obvious how to proceed. One possibility is to contact the original study’s authors to reconcile any differences between the reproduction procedures and those underlying the reported findings in the original study. If resolution is achieved, one is again left with an unpublishable product. Even if resolution is not reached it may often however not be in the best interest of both the original study’s authors (obviously) and the reproduction study’s authors to publish the alternative findings.

Unlike direct replication, strongly diverging reproduction findings cannot be attributed to the inherent randomness of sampling-based research, but necessarily result from differences between the attempted reproduction and the original authors’ approach. Pursuing publication of such observed differences between the original study and the reproduction would involve a substantial degree of reputational risk, not only for the original study’s authors whose published findings are called into question, but also for the reproduction study’s authors. Risk on both sides may be further exacerbated based on the impact of the original study; not surprisingly, high-impact studies are likely to be authored by scholars with a central role—and thus quite a bit of influence—in the family business research community.

We do not wish to discourage family business scholars from taking a healthily critical approach to each other’s work. Taking the above considerations into account however, we believe that the onus of securing the exact reproducibility of empirical findings should be on the journal that decides to publish them, as we will discuss later in more detail.

2.2.2. Direct replication

The purpose of a direct replication is to assess the generalizability of a specific statistical association in terms of its significance, directionality, and magnitude. By the very nature of sampling-based statistical research, estimates of population parameters will differ across even the most randomly drawn samples; therefore, it is important to gauge the
consistency of findings across repeated samples (Coffman & Niederle, 2015). A pure direct replication tests the exact same hypothesis of the original study, using the same methodology and measurements, but in a different sample drawn from the same population; that is, the theoretical population of interest (e.g., privately owned family firms). A different sample from that same theoretical population may come from a different geographical or temporal context, as long as that context does not conceptually define the population of interest. While we do not have a preference for specific family business topics or studies on which direct replication efforts should be focused, it may be particularly fruitful to target those high-impact original studies based on small and/or highly idiosyncratic samples. Largely because of current academic incentive structures (Higginson & Munafò, 2016), the family business literature suffers from an overabundance of underpowered studies, which are not only unlikely to capture true statistical associations (Type II errors; Aguinis, 1995), but are also prone to capturing spurious significant results (Type I errors); this risk increases the more improbable the tested hypotheses are (Christley, 2010; Murphy & Aguinis, 2019). Much could thus be gained by replicating such underpowered original studies testing seemingly counterintuitive or provocative hypotheses, by using samples large enough in relation to the complexity of the model, intended methodology, and anticipated effect sizes.

Conducting a direct replication of an influential family business study is not likely to be as sensitive an issue as reproduction, since a failure to replicate can always be attributed to the randomness of statistical sampling rather than errors or opportunism of either the replication’s or the original study’s authors. Nevertheless, we advise authors of replication studies to be mindful in deciding whether and how to involve the original study’s authors in the replication process. On the one hand, it could be beneficial to obtain the exact materials used (syntax, measurement scales, etc.) prior to conducting the replication study to ensure complete comparability of findings across both samples. On the other hand, especially when the original study’s authors are significantly more established in the field than the replication study’s authors, it may be better to delay communication – if desired at all – until after the replication study has been accepted for publication to avoid its authors being influenced away from result neutrality.

Similar to reproductions (successful) direct replications do not allow for further theoretical abstraction of the relationship under study (Nosek et al., 2012), as they are likely to suffer from the same kinds of measurement error and construct validity problems as the original study. This, however, is where conceptual replications come in.

### 2.2.3. Conceptual replication

The purpose of a conceptual replication is to assess the potential for “abstraction of the explanation for an effect [or other form of statistical association] from the particulars of a given operationalization to the theoretical variable that operationalizations attempt to manipulate or assess” (Nosek et al., 2012, p. 619) by testing the same idea using an alternative approach (Uhlmann et al., 2019); for instance, changing how independent constructs (e.g., manifestations of family involvement or influence) or dependent constructs are measured (Nosek et al., 2012). It is important to refrain from making compromises to the quality of the data and measures (Block & Kuckertz, 2018), which includes attending to the statistical power considerations discussed before. As far as data collection methodologies allow, it may be advisable to capture both the intended (improved) alternative measure and the measure used in the original study to explicitly address the sensitivity of a conceptual replication’s findings to the operationalizations used. Researchers should be mindful, however, that measuring and testing multiple operationalizations implicitly binds them to report all of these tests, irrespective of convenient or inconvenient outcomes. Again, preregistration of replication plans may help authors prevent being held hostage by their data (Wagenmakers & Dutilh, 2016).

Besides variation in operationalizations, conceptual replications may also demonstrate improvements in study design and methodology relative to the original study. Many impactful findings in the family business literature stem from cross-sectional, single-informant survey studies, which provide low-hanging fruit in terms of findings to be replicated with improved longitudinal and/or multi-informant designs.

### Table 2

Replication climate at ten journals most frequently publishing family business research during 2011–2020.

<table>
<thead>
<tr>
<th>Journal</th>
<th># FB studies 2011–2020&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Explicitly invites replications?</th>
<th>Published replications?</th>
<th>Data sharing condition&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>J. Fam. Bus. Strategy</td>
<td>206</td>
<td>No</td>
<td>Yes (two)</td>
<td>No</td>
<td>First family business journal to commission a replication Special Issue</td>
</tr>
<tr>
<td>Fam. Bus. Research</td>
<td>169</td>
<td>No</td>
<td>Yes (one)</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Ent. Theory &amp; Practice</td>
<td>128</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>J. Fam. Bus. Manag.</td>
<td>114</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>J. Bus. Research</td>
<td>65</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>J. Bus. Ethics</td>
<td>63</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>J. Small Bus. Manag.</td>
<td>60</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>J. Corporate Finance</td>
<td>58</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Small Bus. Economics</td>
<td>57</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Corporate Governance</td>
<td>52</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

Note. <sup>a</sup> Only journals included that published more than 50 family business studies in the period 2011–2020 according to Aparicio, Ramos, Casillas, and Iturralde (2021). <sup>b</sup> Submission guidelines of several of these journals include standard publisher-issued statements encouraging data sharing, yet do not require it nor provide any incentive for doing so.

—

<sup>7</sup> While, in essence, every replication will be informative as to the true size and/or significance of a statistical effect or association, simulations by Coffman and Niederle (2015) suggest that a small number of (unbiased!) direct replications (e.g., 3–5) should suffice to be confident about the generalizability of the original finding (i.e., in terms of associations between specific variables, not in terms of theoretical abstraction).

<sup>8</sup> Not many scholars seem to be aware that, for instance, including interaction terms in a model substantially increases the sample size required to reliably pick up even relatively large true interaction effects (Gelman, Hill & Vehtari, 2020; see Gelman’s, 2018 blogpost for an interactive discussion of the matter).
<table>
<thead>
<tr>
<th>Replicating author(s)</th>
<th>Original study</th>
<th>Type of replication</th>
<th>Purpose of replication</th>
<th>Core constructs/variables</th>
<th>Summary of key findings in relation to original study</th>
<th>Replicating authors’ explanations for differences in relation to original study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dou, Wu &amp; Fang</td>
<td>Chrisman et al. (2012)</td>
<td>Combination of direct replication and conceptual replication.</td>
<td>Addressing the generalizability and external and internal boundary conditions of original findings (US sample) to a different geographical context (China).</td>
<td>DV(s): Family-centered non-economic (FCNE) goals; Family-centered economic goals IV(s): Family involvement Mediator(s): Family essence Moderator(s): Market environment; Firm age; Firm size</td>
<td>Results partially corroborate the original findings as only family management has significant total effects on family centered goals. The relationships between FE and FCNE goals are moderated by various contingencies.</td>
<td>Importance of recognizing cultural, economic, and institutional contexts and organizational characteristics.</td>
</tr>
<tr>
<td>Gerken, Hülsbeck, Ostermann &amp; Hack</td>
<td>Hauck et al. (2016)</td>
<td>Combination of direct and conceptual replication.</td>
<td>Addressing the sensitivity and generalizability of the original findings with larger sample size and using additional methods.</td>
<td>F I B E R DV (for testing predictive validity): business model innovation</td>
<td>Results for F dimension replicated. Extension of R E I scale by integrating the F dimension, resulting in a 17-item I B E R scale. Validating a 12-item short scale F: does not seem to be a dimension of SEW B: it seems that the items that define B in the original scale should be modified for a better measurement. R also presents some problems although, it was included in our scale.</td>
<td>Larger, more heterogeneous sample and additional methods allow for a more psychometric soundness of the scale.</td>
</tr>
<tr>
<td>Herrero &amp; Gomez-Mejia</td>
<td>Berrone, Cruz, and Gomez-Mejia (2012); Hauck et al. (2016); Filser et al. (2018)</td>
<td>Direct replication.</td>
<td>The original scale of Berrone et al. was developed conceptually based on literature review. A few subsequent validations showed problems with some of its dimensions.</td>
<td>F I B E R</td>
<td>Results differ to the original scale and are similar to those of Hauck et al. (2016) in Austria/Germany and those of Filser, De Massis, Gast, Kraus, and Niemand (2018) in Finland even with some Spanish specificities, suggesting a certain homogeneity among European countries vs USA. A potential reason could simply be that the original scale was set at the theoretical level. Future research should compare these scales in different cultural contexts to corroborate findings.</td>
<td>Direct measurement of socioemotional wealth instead of a proxy measure; fsQCA methodology.</td>
</tr>
<tr>
<td>Kosmidou &amp; Holt</td>
<td>Sciarca et al. (2014)</td>
<td>Conceptual replication.</td>
<td>Addressing the sensitivity of original findings to 1) the use of a different method (fuzzy set qualitative comparative analysis) and 2) the use of direct measures of socioemotional wealth.</td>
<td>DV(s): Firm Performance IV(s): Socioemotional wealth; Family management; Generational stage</td>
<td>IV – DV results replicate, when socioemotional wealth is measured directly via socioemotional wealth importance. Results do not replicate when using other measures of socioemotional wealth. Results partially corroborate original findings and extend to include State-owned enterprises as separate and distinct form of non-family firm.</td>
<td>Results for F dimension replicated, Extension of R E I scale by integrating the P dimension, resulting in a 17-item I B E R scale. Validating a 12-item short scale F: does not seem to be a dimension of SEW B: it seems that the items that define B in the original scale should be modified for a better measurement. R also presents some problems although, it was included in our scale.</td>
</tr>
<tr>
<td>Markin, Skorodizyevskiy, Zhu, Chrisman &amp; Fang</td>
<td>Miller et al. (2007)</td>
<td>Combination of direct replication and conceptual replication.</td>
<td>Addressing the generalizability of original findings (US sample) to a different geographic context (China).</td>
<td>DV(s): Firm performance IV(s): Family firms; Lone-founder firms; Non-family firms; State-owned enterprises (new); First- and second-generation family firms</td>
<td>Results partially corroborate original findings and extend to include State-owned enterprises as separate and distinct form of non-family firm.</td>
<td>Larger sample, new IV, and expanded set of controls variables. Cultural differences between original study (US) and replication (China) show that the findings of original study are generalizable in a new context.</td>
</tr>
<tr>
<td>Stutz, Schell &amp; Hack</td>
<td>Hueh (2018)</td>
<td>Combination of direct and conceptual replication.</td>
<td>Firstly, retesting original findings in a different context and with a different subject pool. Secondly, enhancing experimental methodology by conducting a manipulation check. Thirdly, improving understanding of relationships by integrating mediators.</td>
<td>DV(s): Information credibility; Stakeholders’ intentions to buy, invest and to apply for a job IV(s): Family firm vs. non-family firm Mediator(s): Benevolence perception</td>
<td>The original study’s findings cannot be replicated. Extensions of the original study reveal opposite effect.</td>
<td>Exclusion of subjects who wrongly classified firms as family or non-family (manipulation check); Inclusion of an omitted key variable (benevolence perception); Assigning subjects to specific stakeholder roles (customers, investors, and job seekers).</td>
</tr>
<tr>
<td>Szewczyk, Kurzhals, Graf &amp; Vlachy, König et al. (2013)</td>
<td>Empirical validation.</td>
<td>Operationalizing and empirically examining central elements of conceptual work.</td>
<td>DV(s): Discontinuous technology adoption speed; discontinuous technology adoption</td>
<td></td>
<td></td>
<td>Potential theoretical explanations – e.g., different operationalizations of family influence might (continued on next page)</td>
</tr>
</tbody>
</table>
Barring certain exceptions (Lude & Prügl, 2021), the theoretical relationships examined in family business research are hard to capture using experimental methods. As most family business studies are inherently correlational, this leaves open many different paths for endogeneity issues. As such, we mirror Shaver’s (2020) concern that too “many replications focus on the generalizability of existing tests rather than advancing better-identified tests” (p. 1254) and encourage scholars to consider designing conceptual replication studies allowing for better nonexperimental identification of causal effects (Zhang et al., 2021).

2.2.4. Empirical validation

Finally, empirical validations put previously published abstract theoretical or conceptual ideas and propositions under empirical scrutiny. Anecdotally, we observe that both purely conceptual works and inductively derived theoretical models often remain largely untested, yet are considered part of the common understanding about the nature of family firms. As a result, also “our field has an absurdly high ratio of ideas to tests of ideas, [such that we] suppose much more than we know” (Hambrick, 2007; p. 1350). We therefore urge both quantitative scholars and qualitative researchers comfortable with theory-testing⁸ (cf. Bitkine, 2008) to consider the myriad of well-cited, yet un(der)tested conceptual and theoretical models available in the extant family business literature when searching for research ideas.

On the one hand, the prospect of validating already published conceptual models may not be attractive ex ante, as the validation study’s authors will not receive credit for the groundbreaking ideas they are testing. On the other hand, the fact that the conceptual arguments to be tested already carry the stamp of approval provided by peer review, and possibly a favorable reception in the literature, should be an advantage when dealing with reviewers and editors provided they, in turn, appreciate the necessity of cumulative evidence-based science (cf. Hambrick, 2007; more on this later).

We encourage researchers, prior to commencing a validation study, to consider whether or not to communicate their plans to the authors of the selected original study. This is a sensitive issue, as scholars can be protective of their conceptual brainchild and may dissuade attempts at empirical scrutiny of their study’s principal merit. On the other hand, limited coordination with the original authors in the design of the validation project may aid in ensuring maximum ex ante conceptual validity of the validation’s intended methods and measurements. In that regard, we cannot overemphasize the benefits of preregistration, as the preregistration timestamp would form a logical barrier separating the phases of the validation project in which coordination with the original authors may be acceptable and productive (i.e., before preregistering the final research design) versus somewhat questionable and counterproductive (i.e., after preregistration).

2.2.5. Combining replication types

All four of these replications are “ideal” types. In practice, useful replications may combine elements of more than one type. Consider, for example, the recent preregistered meta-analytical replication of the Duran, Kammerlander, Van Essen, and Zellweger (2016) “doing more with less” study by Block, Hansen, and Steinmetz (2022). Block et al. (2022) provide a (near) exact reproduction of Duran et al. (2016), and proceed to present a direct replication by appending the meta-analytic sample with family business innovation studies that appeared since the publication of the original meta-analysis. They extend the replication conceptually in a meaningful way by demonstrating the sensitivity of their findings to the inclusion of a crucial control variable. Despite the potential for combining elements of multiple forms of replication, we encourage authors to explicitly define the overarching purpose of their replication project ex ante, in which the ideal types can be used as flexible benchmarks.

3. Guidelines and recommendations for journals and journal editors

The replication crisis essentially boils down to a social dilemma (Everett & Earp, 2015). Even though the family business field would benefit from a better balance between original and replication studies, scholars focus most of their efforts on the former at the expense of the latter, given the current publication system. As it is unreasonable to expect that many individual scholars will autonomously pursue replication, thereby acting against their self-interest, we strongly encourage those members of the academic family business research community with the most direct influence on publication standards and incentives to address both current problems and potential solutions. Journals publishing family business research, and the editors serving on those journals, can do much to make most types of replication – direct, conceptual, and validation – easier to conduct, worthwhile, and useful for family business scholars. We however first contend that securing the reproducibility (i.e., truthfulness and freedom from error) of new, potentially impactful findings should, ideally, be the responsibility of the journals publishing them.

3.1. Taking responsibility for reproducibility

We argued that verification of the reproducibility of empirical findings should perhaps not be left to the original authors’ academic peers, who may have reasonable objections to engaging in such sensitive work and lack publication incentives for doing so. Reproducibility should be considered a joint responsibility of the original authors and

---

⁸ Where, obviously, testing should not be interpreted in a statistical sense.
the journals publishing their findings. We encourage journals to standardize reproduction checks after conditional acceptance, yet before publication. Once a manuscript has been conditionally accepted, journals could commission independent experts to verify that the authors’ reported results can be reproduced using the data and code shared with those experts and that reported procedures match those actually performed. If not, the authors may be asked to address identified errors or inconsistencies; if necessary, editors may return the paper to reviewers.

While these additional steps may initially delay the publication process (Jacoby et al., 2017), such deficiencies should largely disappear over time as authors will become less likely to submit erroneous or misreported findings (Gertler et al., 2018). Such independent reproduction checks, then, should be both unbiased, as editors generally will not want conditionally accepted results overturned, and time- and labor-efficient, as authors have strong incentives to cooperate with journals during the final pre-publication stage (Gertler et al., 2018).

Implementing such procedures will take time, resources, and, above all, a learning curve. We suggest that journals could initially conduct reproducibility checks on only a fraction of accepted manuscripts, that are picked at random and without editorial interference such that a manuscript’s chances of being selected for verification of reproducibility are unrelated to characteristics of its authors. Additionally, we point journals and editors to already available tools that allow for identifying potentially problematic statistical findings using only information reported in a manuscript (e.g., correlation tables and descriptive statistics; see Bergh et al., 2017).

3.2. Making replication easier

Journals can facilitate easier replication of their own published studies. For instance, researchers can more effectively design and conduct replication studies if they can access the materials used to produce the results of the original study without having to personally contact the original authors (Block & Kuckertz, 2018): the data, the questionnaires or experimental materials used for data collection, as well as the software, code, and/or algorithms used to analyze the data. While strictly enforcing open sharing of data and other materials as a condition for publication may currently be a bridge too far given the comparatively underdeveloped open science culture in the broader management discipline (Aguinis, Ramani, & Alabduljader 2018; Barc zak et al. 2022; Castille, Kreamer, Albritton, Banks & Rogelberg, 2022), journals could better facilitate and encourage the voluntary post-acceptance publication of study materials.

Individual researchers’ philosophical convictions aside, it generally is not (yet) in authors’ best self-interest to share study materials. Scholars can be particularly protective of their collected data. Authors’ reluctance to publish their data alongside an accepted study often results from the desire to pursue additional research ideas based on the same data set, and thus the reasonable fear that other researchers would “beat them to the mark” with their own data (e.g., Laine, 2017). There are, however, open access models and policies available that offer authors some protection in this regard, by including a data embargo option in which their data are made available to the public only after a certain number of months/years, thereby giving the creators of a data set the first opportunity to exploit its further publication potential (Roche et al., 2014).

Besides lowering (perceived) barriers, journals can also make such open practices attractive per se (Heutkoop et al., 2018). A recent initiative of the Center for Open Science enables journals to incentivize authors to follow open-science practices using badges (see Open Science Framework, 2016). For example, published articles can be stamped with an Open Data Badge when researchers make the materials needed to reproduce their findings publicly available. Such relatively simple initiatives would not unnecessarily punish authors who are not able or willing to share their materials, yet would allow scholars to credibly signal their transparency, thereby offering journals an easily implemented incentive toward more open scientific practices.¹²

3.3. Making replication worthwhile

Journals publishing family business research hold substantial power when it comes to setting field-wide incentives for replication studies. The highest-impact journals may consider replication research as unoriginal, and thereby a matter for lower-ranking journals. Nevertheless, the leading publishing institutions in (family) business research must become agents of change, as their reputational values will be needed to convince authors to undertake replication studies.¹³ Journals can both remove disincentives and improve or create incentives for replication.

First, journals and their editors should stop implicitly disincentivizing replication studies and recognize their fundamental importance in theory development (Nosek & Errington, 2020) rather than lacking an original contribution to theory (cf. Neubaum & Micloti, 2021). Moreover, editors and reviewers should subject authors of original empirical research to greater scrutiny in terms of, for instance, claims of generalizability, causality, and unfounded implications. Presenting the results and implications of the first published test of a hypothesis in an overly bold and conclusive fashion narrows the space for replications. Forcing authors to instead consider the implications of their original study in a more realistic way in relation to the study’s limitations and as part of a necessarily cumulative body of evidence (cf. Aguinis et al., 2021; Shaver, 2020) should naturally create more room for replication studies.

Second, submitting replication studies should be made attractive and explicitly incentivized (Koole & Lakens, 2012), perhaps by creating a separate short “research note” submission format with specific submission guidelines and the explicit promise that reviewers and editors will not judge the replication for its lack of theoretical novelty.¹⁴ Authors may feel discouraged to conduct replications if they do not anticipate many future citations of such work. Suggestions have been made to bind authors citing original studies that have undergone some form of replication to simultaneously co-cite the replication attempts (e.g.,

¹⁰ These experts would, ideally, be third parties without affiliations to the journal (editors), the authors, and even the field at large to ensure the absence of any conflicts of interest (see Jacoby, Lafferty-Hess and Christian (2017) for experiences at the American Journal of Political Science), although other seemingly effective strategies involve journal-appointed data- or method editors, such as at the American Economic Review (Vilhuber, Turrito, & Welch, 2020).

¹¹ Note that the sharing of data and other materials (either at submission or after publication) is generally normative in the more reputable empirical economics and finance journals and obligatory in most of the leading general science journals (a.o., all Science, Nature, and Royal Society Journals). It is therefore not unreasonable to assume that the management discipline will eventually move in a similar direction. We understand, however, that given current prevailing practices, management (and thus family business) journals may envision a substantial decrease in submissions if they were to be (one of) the first to make sharing of data and materials a condition for publication.

¹² While awarding badges for desirable behavior may seem like a childish strategy, Kidwell et al. (2016) document a sharp increase in open data reporting in a top-tier psychology journal after it implemented the badges initiative, while no such increases were observed in comparison journals over the same period, suggestive of the initiative’s effectiveness.

¹³ Howard Aldrich, for instance, has been quoted (rightly) wondering why “… whereas there’s a Review and Annals and Perspectives and so on, there’s no Academy of Management Replications” (Berman, 2014). A new journal hosted by the Southern Management Association – JoM Scientific Reports (https://smgmt.org/jomsr/) – bears some promise in this regard.

¹⁴ Some journals relevant to family business have recently adopted similar replication research note submission formats (e.g., ETP, see https://journals.sagepub.com/author-instructions/ETP). Replication studies tend to need less space given the absence of substantial original theory building.
Koole & Lakens, 2012). As for replication studies published in the same journal (or even with the same publisher) as the original study, journals could facilitate this co-citation practice by internally linking the original and the replication study, akin to how retracted papers and retraction notices are currently linked on most journal webpages.

3.4. Making replication useful

An increase in published replication studies will benefit the field only if these studies are conducted in a useful way. Usefulness should be interpreted as the degree to which a replication could (ex-ante) contribute to advancing theory through confronting existing understanding with new evidence (cf. Nosek & Errington, 2020). Accordingly, for the replication to be of maximum use it should address an underlying theoretical question with substantial (potential) implications for developing understanding, yet maintain objectivity toward the evidence that understanding is confronted with; that is, be indifferent to whether the replication confirms or opposes the original study’s findings and whether its findings are statistically significant or null (cf. Landis, James, Lance, Pierce, & Rogelberg, 2014).

Journals can be proactive in ensuring that replication studies address the most important questions in the field. For instance, recent meta-scientific work in psychology demonstrates the power of crowdsourcing hypothesis tests as a means to rapidly collect replicates of new and potentially groundbreaking findings (e.g., Landy et al., 2020; Uhlmann et al., 2019). Journal editors may set up a “tender for each of the...most cited articles per volume two or three years post-publication, and crowdsourced independent teams of researchers to submit detailed research proposals for a replication of one of those [editor-selected] studies” (Brinkerink, 2022, p. 3). These proposals can be reviewed and revised based on expert reviewers’ feedback until rejection or in-principle acceptance of the proposed replication is reached. In-principle accepted replication proposals can then be executed as agreed upon by the proposing authors and, once received, the written up replication can be linked to the original study on the journal’s webpage.

This review model, the “registered report” format (Chambers, 2013), should also secure the replication’s objectivity. Conducting the review (and potential revisions) of hypotheses and proposed methods prior to data collection and analyses (i.e., registered reports), as opposed to afterwards (i.e., most current peer review processes), should protect authors against their own urges to search for significant results because their analytical procedures have been fixed ex ante. Authors should no longer be subject to any publication biases or preferences from editors and reviewers, as the only remaining condition for acceptance should be that the replication study has been conducted as promised in the already accepted proposal, and not whether or not its findings are statistically significant and in line with or contrary to the original study (Nosek & Lakens, 2014).

A slightly less demanding alternative approach is for example practiced by Entrepreneurship Theory & Practice (ETP), which requires every submitted replication to be based on a proposal preregistered on an independent registry. This approach is designed to aid authors in critically evaluating their replication design prior to collecting data and subsequently holding themselves to their research plan. Compared to registered reports however, preregistrations do not yet entail a commitment of the journal and its editors and reviewers to treating the replication with neutrality toward its findings. Nevertheless, we generally commend ETP for its leadership in normalizing replication work in entrepreneurship and family business studies. Besides considering innovations in terms of submission formats and review processes, however, journals and editors should be especially mindful in selecting reviewers when it comes to replication studies given the sensitivities discussed in the previous section. At a minimum, editors should select reviewers who are independent from both the authors of the replication and the original study, are aware of the purpose of and need for high-quality replication studies, and will judge the submitted replication study (or proposal) with this in mind rather than on its relative lack of independent theoretical innovation.

A quick survey of ten journals regularly publishing family business studies suggests that most of these outlets can do much more to create a favorable replication climate with respect to the issues discussed (cf. Table 2). We hope our suggestions will be supportive in moving ahead. In closing, we highlight the family business replication studies that are published in this issue of JFBS.

4. Advancing family business knowledge through replication research

Table 3 provides an overview of each of the papers published in the JFBS replication Special Issue. It lists the original study, the type of replication performed and its purpose, the core constructs central to original and replication study, and a summary of the replication’s findings in relation to the original study, accompanied by the replication study’s authors’ tentative explanations for any differences.

These articles collectively demonstrate how replication research can advance knowledge. They represent the different replication types described in this article. For instance, Kosmidou and Holt’s (2022) replication of Sciascia, Mazzola, and Kellermans (2014) is primarily conceptual in nature; Gómez-Mejía and Herrero (2022) provide a direct replication of Hauck, Suess-Reyes, Beck, Prügl, and Frank (2016) using a sample from a different national context; Szewczyk, Kurzhals, Graf-Vlachy, Kammerlander and König (2022) conduct a “pure” empirical validation of König, Kammerlander, and Enders (2013). However, most of the included replications combine elements from different ideal types. Elements of both direct and conceptual replication appear in Dou, Wu, and Fang (2022) replication of Chrisman, Chua, Pearson, and Barnett (2012); Gerken, Hülsbeck, Ostermann, and Hack (2022) replication of Hauck et al. (2016); Markin, Skorodziyevskiy, Zhu, and Chrisman (2021) replication of Miller, Le Breton-Miller, Lester, and Cannella Jr (2007); and Stutz, Schell, and Hack (2022) replication of Hsu (2018).

5. Conclusion

Replication studies deserve a more prominent place in family business research. A stronger replication culture will lead to more robust and reliable knowledge about family business phenomena. We have discussed the importance of replication studies for the advancement of family business research, examined some key issues preventing the development of a stronger replication culture, identified different replication types and examined their characteristics, and offered guidelines and recommendations for authors, journals, and journal

---

15 For instance, going forward one could cite Duran et al. (2016) as (Duran et al., 2016; replication attempt by Block et al., 2022).

16 Reviewers should be independent from the authors of the original study, obviously. An exception here may be empirical validation, where the review of validation proposals may be the optimal stage in which authors of the original study could secure conceptual validity of the proposed approach; perhaps not in a blinded way as to add transparency.

17 The registered report format should improve the quality and objectivity of all forms of theory-testing in family business research – both original and replication studies – as discussed in more detail in Brinkerink (2021). This broader discussion, although much needed, goes beyond the scope of the current article.

18 Additionally, ETP offers a separate registered report format not (explicitly) targeting replication studies, such that family business scholars could presumably still consider sending in their replication plan as a registered report format. See https://journals.sagepub.com/author-instructions/etp#ArticleTypes.