Received: 26 April 2017 | Revised: 3 February 2019 | Accepted: 12 February 2019 | Published on: 2 April 2019

DOI: 10.1002/smj.3015

STRATEGIC MANAGEMENT JOURNAL WILEY

RESEARCH ARTICLE

Transparency and replicability in qualitative research: The case of interviews with elite informants

Herman Aguinis[1] | Angelo M. Solarino[2]

[1]School of Business, The George Washington University, Washington, District of Columbia

[2]Leeds University Business School, The University of Leeds, Leeds, UK

Correspondence
Herman Aguinis, School of Business, The George Washington University, 2201 G Street, NW, Washington, DC 20052.
Email: haguinis@gwu.edu

***Research Summary*:** We used interviews with elite informants as a case study to illustrate the need to expand the discussion of transparency and replicability to qualitative methodology. An analysis of 52 articles published in *Strategic Management Journal* revealed that none of them were sufficiently transparent to allow for exact replication, empirical replication, or conceptual replication. We offer 12 transparency criteria, and behaviorally-anchored ratings scales to measure them, that can be used by authors as they plan and conduct qualitative research as well as by journal reviewers and editors when they evaluate the transparency of submitted manuscripts. We hope our article will serve as a catalyst for improving the degree of transparency and replicability of future qualitative research.

***Managerial Summary*:** If organizations implement practices based on published research, will they produce results consistent with those reported in the articles? To answer this question, it is critical that published articles be transparent in terms of what has been done, why, and how. We investigated 52 articles published in *Strategic Management Journal* that reported interviewing elite informants (e.g., members of the top management team) and found that none of the articles were sufficiently transparent. These results lead to thorny questions about the trustworthiness of published research, but also important opportunities for future improvements about research transparency and replicability. We offer recommendations on 12 transparency criteria, and how to measure them, that can be used to evaluate past as well as future research using qualitative methods.

KEYWORDS
credibility of science, methodology, repeatability, research transparency, research trustworthiness

1 | INTRODUCTION

Strategic management studies and many other fields are currently immersed in an important discussion regarding the transparency and replicability of research (e.g., Aguinis, Cascio, & Ramani, 2017; Aguinis, Ramani, & Alabduljader, 2018). However, to date, this stream of research has focused mainly on quantitative research (e.g., Bergh, Sharp, Aguinis, & Li, 2017; Bettis, Ethiraj, Gambardella, Helfat, & Mitchell, 2016; Bettis, Gambardella, Helfat, & Mitchell, 2014; Bettis, Helfat, & Shaver, 2016). In contrast, our focus is on transparency and replicability in qualitative research. Specifically, our goal is to empirically investigate the extent to which replicating qualitative studies that have already been published is possible given the information that is usually available.

Before describing our study, we clarify two important issues: Our ontological perspective and the desirability of replicability in qualitative research. First, our ontological and epistemological perspective is "qualitative positivism" as used by Eisenhardt (1989) and described in detail by Yin (2014), which is similar to what Miles and Huberman (1994) labeled "transcendental realism." Specifically, this means that "social phenomena exist not only in the mind but also in the objective world—and that some lawful and reasonably stable relationships are to be found among them… Our aim is to register and 'transcend' these processes by building theories that account for a real world that is both bounded and perceptually laden" (Miles & Huberman, 1994, p. 4). This ontological perspective is dominant in strategy (as was also evidenced by our own study) and also shared by Bettis, Helfat, and Shaver (2016) and Bettis, Ethiraj, et al. (2016) in that a key goal is to produce replicable and cumulative knowledge. As such, our article describes criteria that can be used to evaluate the extent to which qualitative research is transparent because if replication is a desirable goal, then transparency is a required step (Aguinis et al., 2018).

Second, regarding the desirability of replicability, in contrast to quantitative research, this is a potentially contentious issue in qualitative research. For example, in ethnography replicability is not necessarily meaningful because the researcher takes on the role of the research instrument (Welch & Piekkari, 2017). But, most qualitative researchers would not argue against the need for transparency. Therefore, we focus on transparency criteria and the extent to which transparency is necessary for three different types of replication studies: (a) *exact replication* (i.e., a previous study is replicated using the same population and the same procedures), (b) *empirical replication* (i.e., a previous study is replicated using the same procedures but a different population), and (b) *conceptual replication* (i.e., a previous study is replicated using the same population but different procedures). In the case of qualitative researchers who are not necessarily interested in empirical or conceptual replicability, or believe that these two types are not necessary or even appropriate based on their ontological perspective, there is still an interest in transparency and perhaps in exact replication—which is about finding possible errors and the falsifiability of the knowledge produced. Moreover, there is also a need to understand the trustworthiness, meaning, and implications of a study's results for theory and practice (Denzin & Lincoln, 1994) and this can also be achieved more easily with a greater degree of transparency.

In sum, our article makes a unique value-added contribution by expanding the discussion of transparency and replicability to the domain of qualitative methods. We use the particular qualitative technique of interviews with elite informants as a case study or empirical boundary. As a result of our analysis, we offer best-practice recommendations about 12 transparency criteria that can be used by authors in conducting their work and also by journal editors and reviewers when evaluating manuscripts that adopt qualitative methodological approaches.

Next, we describe the particular qualitative methodological technique of interviews with elite informants (IEIs), which we use as an illustration and case study regarding the degree of transparency in qualitative research. Then, we describe a study in which we assessed the degree of transparency of *Strategic Management Journal* (SMJ) articles that used IEIs. Finally, based on our results, we describe implications for future qualitative research replicability and also offer recommendations based on 12 criteria for authors as well as journal editors and reviewers on how to improve transparency in future qualitative research.

2 | INTERVIEWS WITH ELITE INFORMANTS (IEIS)

For approximately 50 years, interviews with elite informants (i.e., those in the upper echelon of organizations) have been a staple in the methodological toolkit of strategic management researchers (Kincaid & Bright, 1957). In fact, IEIs were used in an article published in the very first volume of SMJ (Ajami, 1980). The reason for the use and important role of IEIs is rather obvious: for certain research questions and foci, the input provided by elite informants is critical for building and testing theories in strategic management research (Basu & Palazzo, 2008; Hambrick & Mason, 1984). For instance, IEIs are useful in exploring the role of organizational narratives in enhancing or constraining new CEO or board-member decisions, the micro-foundations of performance differences between performers at the top and bottom in an industry, and how those differences are interpreted and addressed by the executives. In other words, IEIs offer a unique opportunity to explore the micro-foundations of firms' strategies (Felin, Foss, & Ployhart, 2015; Foss & Pedersen, 2016). Also, they offer insights into how the highest level of the organization shapes the lower levels (Aguinis & Molina-Azorin, 2015) because they allow researchers to assess how cognitions and intra-psychological processes residing in members of the organization's upper echelons can be related to organization-wide processes, policies, and actions (including strategic decisions that are usually the purview of upper management). Accordingly, an examination of IEIs is appropriate as a case study of the degree of transparency in qualitative research.

There are several definitions of elite informants, ranging from top-ranking executives (Giddens, 1972; Kincaid & Bright, 1957) to highly skilled professionals (McDowell, 1998) to people with substantial expertise not possessed by others (Richards, 1996; Vaughan, 2013). Stephens (2007) noted that the term "elite" is used in a relative or ipsative sense in these contexts, defining such individuals in terms of their social position compared to the average person. Based on this literature, we offer the following unified definition: Elite informants are key decision makers who have extensive and exclusive information and the ability to influence important firm outcomes, either alone or jointly with others (e.g., on a board of directors).

Conducting qualitative research using IEIs includes features that are also present in many other types of qualitative methodologies and these include the roles of the investigator and informant as well as the relationship between the two (Ostrander, 1993; Thomas, 1993). Also, as in other types of qualitative research, the study takes place in a particular research setting, researchers make choices about sampling procedures, they have a specific position along the insider-outsider continuum, and make decisions about the saturation point as well as data coding and analysis. Again, these features point to

the appropriateness of using the IEI literature as a case study. Next, we describe an empirical study in which we assessed the degree of transparency in articles published in SMJ that used IEIs.

3 | METHOD

3.1 | Transparency criteria in qualitative research

We consulted articles and books to develop transparency criteria for qualitative research. First, we conducted an extensive literature review that included both substantive and methodological journals from management, business, sociology, psychology (i.e., general psychology, applied psychology, organizational psychology), education, nursing studies, and geography. Our search included the terms "quality," "transparency," "reproducibility," "trustworthiness," and "rigor." We conducted a search using Web of Science, PsycINFO, and Google Scholar. This first step in the process led to 127 articles and the list is in Appendix S1A (Supporting Information). Second, we used the same keywords and databases but now focusing on books. This search process resulted in 14 books and the list is also available in Appendix S1A (Supporting Information). Although Appendix S1A shows that this list is quite extensive, it may be possible that we did not include every single relevant source. But, we doubt the addition of more articles or books would produce a substantive change in the major transparency criteria we identified, as described next.

Third, once we identified the potential transparency criteria sources, we adopted an inclusive and open-coding approach (Strauss & Corbin, 1998) to extract an initial list of unique criteria covering the various phases of the research process: design, measurement, analysis, reporting of results, and data availability. We continued the data collection until no new codes were added to the code book suggesting that we had reached theoretical saturation (Locke, 2001), defined as the moment when "new information produces little or no change to the codebook" (Guest, Bunce, & Johnson, 2006, p. 65). This process led the identification of 40 uniquely identifiable criteria and they are included in Appendix S1B (Supporting Information).

Fourth, we used theoretical coding (Charmaz, 2006; Strauss & Corbin, 1998) to identify transparency criteria. In an effort to be as inclusive as possible, we grouped the initial set of 40 criteria into 12 different transparency criteria. The process was interactive and both authors agreed fully on the final list of criteria. Both authors also agreed that seven criteria (out of a total of 40) did not apply specifically to transparency and we therefore excluded them.[1] The 12 criteria are defined and described in Table 1.

The transparency criteria included in Table 1 cover the sequential aspects of the qualitative research process and include research design (i.e., kind of qualitative method, research setting, position of researcher along the insider-outsider continuum, sampling procedures, relative importance of the participants/cases), measurement (documenting interactions with participants; saturation point; unexpected opportunities, challenges, and other events; management of power imbalance), data

[1]Both authors agreed that the following seven criteria (out of a total of 40) did not apply specifically to transparency and, instead, referred to the appropriate or inappropriate use of various types of methods and techniques (Welch & Piekkari, 2017): Quality of fact checking, appropriate cutoff for inter-rater reliability, approval by institutional review board (i.e., for conducting research with human subjects), appropriate channels for contacting participants, appropriate authors' credentials, use of appropriate software, and whether authors engaged in an appropriate level of reflexivity. Certainly, some of these are indirectly related to transparency. But, overall, these are about appropriateness, quality, and rigor rather than transparency per se. For example, regarding the criterion "fact-checking," the literature refers to recommendations regarding the relative appropriateness of various choices such as what is the appropriate number of facts that should be checked and how specific facts can be attributed to some sources and not others. As a second example, the literature offers recommendations regarding what is an appropriate level of inter-coder reliability.

TABLE 1 Transparency criteria in qualitative research and their relevance for exact, empirical, and conceptual replication

ID	Transparency criterion	Definition	Criterion is necessary for replicability because...	Exact replication	Empirical replication	Conceptual replication
1	*Kind of qualitative method*	The particular qualitative methodology used in the study (e.g., action research, case study, grounded theory) (Creswell, 2007; Flick, 2014; Patton, 2002)	… a method's assumptions, beliefs, and values affect theory, design, measurement, analysis, and reporting choices, as well as the interpretation of results	✓	✓	
2	*Research setting*	The physical, social, and cultural milieu of the study (e.g., firm conditions, industry, participants' social status) (Bhattacharya, 2008; Patton, 2002)	...it clarifies the structure, the sources and the strength of the pre-existing conditions in the research setting	✓		✓
3	*Position of researcher along the insider-outsider continuum*	The researcher's relationship with the organization and study participants; the closer the relationship, the more the researcher is an insider rather than an outsider (Evered & Louis, 1981; Griffith, 1998)	... it allows for an understanding of the researcher's relationship with the organization and participants, which can alter accessibility of data, what participants disclose, and how the collected information is interpreted	✓		✓
4	*Sampling procedures*	The procedures used to select participants or cases for the study (e.g., convenience, purposive, theoretical) (Patton, 2002; Teddlie & Yu, 2007)	... given that samples are not probabilistic, it clarifies what kind of variability the researcher is seeking (and along which specific dimensions), and the presence of possible biases in the sampling procedure	✓		✓
5	*Relative importance of the participants/cases*	The study's sample and the relative importance of each participant or case (Aguinis, Gottfredson, & Joo, 2013; Dexter, 1970)	.. it allows for the identification of participants and cases with similar characteristics as in the original study	✓		✓
6	*Documenting interactions with participants*	The documentation and transcription of the interviews and all other forms of observations (e.g., audio, video, notations) (Kowal & O'Connell, 2014)	...different means of documenting interactions may alter the willingness of participants to share information and therefore affect the type of information gathered	✓	✓	
7	*Saturation point*	It occurs when there are no new insights or themes in the process of collecting data and drawing conclusions (Bowen, 2008; Strauss & Corbin, 1998)	…identifying the saturation point can include judgment calls on the part of the researcher (e.g., when a researcher believes that additional information will not result in new discoveries or that new information will not add new categories to the coding scheme)	✓	✓	
8	*Unexpected opportunities, challenges, and other events*	Unexpected opportunities (e.g., access to additional sources of data), challenges (e.g., a firm's unit declines to participate in the last data collection stage and is replaced by a different one), and events (e.g., internal and	… the way in which researchers react and actions they take in response to these unexpected events affect data collection and subsequent conclusions	✓	✓	

TABLE 1 (Continued)

ID	Transparency criterion	Definition	Criterion is necessary for replicability because...	Exact replication	Empirical replication	Conceptual replication
		external changes such as a new CEO or changes in market conditions during the study) that occur during all stages of the research process (Dexter, 1970; Harvey, 2010; Ostrander, 1993)				
9	*Management of power imbalance*	The differential exercise of control, authority, or influence during the research process (Ostrander, 1993; Thomas, 1993)	… it allows other researchers to adopt similar strategies (e.g., endorsement from a prestigious institution, self-acquaintance, asking sensitive questions) that affect the type of information gathered as well as a study's conclusions	✓	✓	
10	*Data coding and first-order codes*	The process through which data are categorized to facilitate subsequent analysis (e.g., structural coding, descriptive coding, narrative coding) (Maxwell & Chmiel, 2014; Saldana, 2009; Strauss & Corbin, 1998; Taylor, Bogdan, & DeVault, 2016)	... it allows other researchers to follow similar procedures and obtain similar conclusions	✓	✓	
11	*Data analysis and second- and higher-order codes*	The classification and interpretation of linguistic or visual material to make statements about implicit and explicit dimensions and structures (Flick, 2014) and it is generally done by identifying key relationships that tie the first order codes together into a narrative or sequence (e.g., pattern coding, focused coding, axial coding) (Saldana, 2009; Taylor et al., 2016)	... it allows other researchers to use a similar analytical approach and obtain similar conclusions	✓	✓	
12	*Data disclosure*	Raw material includes any information collected by the researcher before any manipulation (i.e., analysis) (e.g., transcripts, video recordings) (Ryan & Bernard, 2000; Schreiber, 2008)	… others can reuse the original material and attempt to obtain the same results and reach the same conclusions	✓		

Note. In the case of qualitative researchers who are not necessarily interested in empirical or conceptual replication, or believe that these two types are not necessary or even appropriate based on their ontological perspective, there is still an interest in transparency and perhaps in exact replication—which is about finding possible errors and the falsifiability of the knowledge produced.

analysis (i.e., data coding and first-order codes; data analysis and second- and higher-order codes), and data disclosure (i.e., raw material availability).

An important characteristic of the transparency criteria is that they are not mutually exclusive and, rather, they have a cumulative effect on the trustworthiness and replicability of knowledge. In

other words, transparency is a continuous variable and a matter of degree. So, the larger the number of criteria that are fully met, the better.

Finally, the 12 criteria are applicable and sufficiently broad so they can be used to assess transparency in many types of qualitative methods and across substantive domains. However, we readily acknowledge that additional criteria could be added to the list.

3.2 | Transparency criteria and three types of replicability

As shown in Table 1, there are specific reasons why each transparency criterion is relevant for replicability. However, not all criteria are necessarily relevant for all three types of replicability. This is an important issue in light of our previous discussion that not all types of replicability are always necessary—or even desirable—across ontological perspectives.

All of the criteria are relevant for *exact replication* where a previous study is replicated using the same population and the same procedures. In an exact replication study, the goal is to assess whether the findings of a past study are reproducible (Bergh et al., 2017; Tsang & Kwan, 1999). Thus, the goal is to remain as close as possible to the original study in terms of methodological approach, population and sampling criteria, data coding, analysis, and all other procedures.

In the case of *empirical replication*, a previous study is replicated using the same procedures but a different population. The purpose is to assess the extent to which results are generalizable to another population. In this second type of replication, the goal is to remain as close as possible to the original study in terms of methodological procedures but not in terms of study participants. Accordingly, transparency criteria related to methodological procedures, but not necessarily about characteristics of the sample and population, are most relevant (i.e., criteria 1, and 6–11).

Finally, in a *conceptual replication* a previous study is replicated using the same population but different procedures. The purpose of this third kind of replication is to assess whether findings, in terms of constructs and relationships among constructs, can be replicated using different methodological procedures and instruments. Because this type of replication study is based on the same theory as the original study (Tsang & Kwan, 1999), transparency criteria related to characteristics of the population, but not necessarily methodological procedures, are most relevant (i.e., criteria 2–5).

3.3 | Article selection

We manually searched all articles published in SMJ since the year 2000 and also included in press articles as of November 2017. We included studies for which IEIs served as key input for substantive results and conclusions. This led to the exclusion of articles that mentioned the use of elite informants but did not clarify the specific role of the interviews regarding substantive conclusions (e.g., Shipilov, Godart, & Clement, 2017), and those studies where the contribution of the elite informant was limited to a subsequent survey development procedure (e.g., Baum & Wally, 2003; Reuer, Klijn, & Lioukas, 2014). On the other hand, we did include studies that identified elite informant interviewing as a post-hoc analysis (e.g., Zhou & Wan, 2017). The final sample included 52 articles (but 53 independent samples because one article included two studies) and they are listed in Appendix S1C (Supporting Information).

We make three clarifications about our sample of 52 SMJ articles. First, none of the articles included aimed specifically at conducting replications. Rather, it was our own goal to empirically investigate the extent to which replicating studies that have already been published is possible given the information that is usually available.

Second, we included some studies that may be perceived as being quantitative but used IEIs in a clearly qualitative manner (e.g., they included questions across the units of analysis outside of pre-existing "if-then" branching, they included questions in fundamentally different ways contingent on circumstances, there was substantial variation in number of informants per unit of analysis based on what was learned after the research has begun). To examine this issue more precisely, we used two coders to classify each of the 52 articles into one of the following categories (inter-rater reliability was .92): (1) case study (i.e., a study in which an issue was investigated through one or more cases within a bounded system); (2) grounded theory (i.e., a study in which the researcher generated a theory of a process, action or interaction shaped by the view of a large number of participants); (3) mixed methods (i.e., a study combined qualitative and quantitative data collection and data analysis within a single study); (4) other qualitative approaches; or (5) mainly quantitative (i.e., the study included a qualitative component but it was used in post-hoc analyses or to confirm or explain the quantitative analysis). Results revealed that purely qualitative and mixed methods articles (which have a clear qualitative component) accounted for 85% of the sample. Specifically, 27 (52%) are purely qualitative, 17 (33%) used mixed methods, and 8 (15%) are mainly quantitative. Appendix S1D (Supporting Information) includes the categorization of each of the 52 articles. In sum, the vast majority of the articles included in our study was qualitative in nature (85%) and espoused a (post) positivist approach (94%).[2]

Third, we assessed whether studies in our sample were in the (post) positivist mode because if they are not, the use of our measures to assess transparency (described in the next section) would not be appropriate. So, we coded the articles using Lincoln, Lynham, and Guba's (2011) taxonomy, which includes (1) positivism and post positivism, (2) critical theory, (3) participatory, and (4) constructionist/interpretivist ontologies. Because of their commonality and clear differentiation from the other types, we treated positivism and post positivism as one category. We used two independent raters to classify each of the 52 articles (inter-rater reliability was .96). Results showed that 94.23% (n = 49) espoused a (post) positivist ontology and 5.77% (n = 3) adopted a constructivist/interpretivist approach. Based on these results, the articles we analyzed are particularly suitable for our purposes. Appendix S1D (Supporting Information) includes the ontological classification of each of the 52 articles.

3.4 | Measures and data collection

We developed behaviorally-anchored rating scales (BARS) to measure the extent to which the 52 SMJ articles met each of the 12 transparency criteria in Table 1. The use of BARS as a measurement instrument has been used extensively in human resource management and organizational behavior (HRM&OB) (e.g., Aguinis, 2019; Cascio & Aguinis, 2019; Hauenstein, Brown, & Sinclair, 2010; Maurer, 2002). The use of BARS is particularly suited for our study because it includes anchors along an evaluative continuum with behavioral examples exemplifying outcomes at different levels of that continuum rather than unspecific and generic anchors such as "agree" and "disagree." In our study, BARS aim to reduce rater errors due to differing interpretation of scales by defining transparency in behavioral terms and offering concrete, specific examples of actions that exemplify transparency at different levels. Table 2 includes the BARS we used in our study.

We followed a best-practice deductive approach in developing our BARS (Guion, 2011). First, we identified the domain of each transparency criterion and then gathered critical incidents

[2]We classified case and grounded theory studies based on what the authors themselves wrote (i.e., most researchers use grounded theory in a (post) positivistic way).

TABLE 2 Behaviorally-anchored ratings scales (BARS) to measure transparency in qualitative research

ID	Transparency criterion	1 **Criterion not mentioned** ***complete absence of information on the specific criterion making replication not possible***	2 **Criterion mentioned but not elaborated** ***criterion is mentioned but no additional information is offered making replication highly unlikely***	3 **Criterion partially met** ***some elements are present but information is missing or incomplete making replication unlikely***	4 **Criterion is met** ***detailed or full disclosure of information on the specific criterion making replication possible***
1	*Kind of qualitative method*	The authors do not describe the type of qualitative research approach they adopted in their study	The authors mention the use of a particular qualitative research approach but do not describe it	The authors describe the key elements of their qualitative research approach but fail to identify it by name	The authors clearly identify the type of qualitative research approach they adopted
2	*Research setting*	The authors do not describe the research setting of the study	The authors identify the setting without describing the pre-existing conditions that make the setting appropriate for the study	The authors describe only the key pre-existing conditions in the research setting that make it appropriate for the study	The authors offer a detailed and rich description of the research setting that goes beyond the description of the key pre-existing conditions (e.g., chronic excess capacity in a small competitive industry)
3	*Position of researcher along the insider-outsider continuum*	The authors do not disclose their position along the insider-outsider continuum	The authors mention but do not describe the existence of a relationship between them and the organization or the participants	The authors describe the type of relationship with the organization and participants	The authors clearly position themselves on the insider-outsider continuum
4	*Sampling procedures*	The authors do not describe the sampling procedures	The authors describe the sampling procedure (e.g., snowball sampling, international sampling)	The authors describe the kind of variability sought through their sampling procedure	The authors describe the kind of variability they seek and how they identified the participants or cases
5	*Relative importance of the participants/cases*	The authors do not describe the final sample or the importance of specific types of participants	The authors describe the final sample	The authors describe the final sample and identify the key participants	The authors describe how each participant was instrumental to developing one or more themes
6	*Documenting interactions with participants*	The authors do not describe how the interactions with participants were documented	The authors describe how some of the interactions with participants were documented	The authors describe how each interaction was documented	The authors describe how each interaction was documented and the associated content
7	*Saturation point*	The authors do not describe when theoretical saturation was reached	The authors report whether they reached theoretical saturation or not	The authors describe how they reached theoretical saturation	The authors describe the precise criteria used to conclude that they have reached theoretical saturation
8	*Unexpected opportunities, challenges, and other events*	The authors do not describe whether any unexpected opportunities, challenges, and other events occurred	The authors report whether any unexpected opportunities, challenges, and other events occurred	The authors describe any unexpected opportunities, challenges, and other events that occurred and how they handled them	The authors describe any unexpected opportunities, challenges, and other events, how they were handled, and their

TABLE 2 (Continued)

ID	Transparency criterion	1 **Criterion not mentioned** ***complete absence of information on the specific criterion making replication not possible***	2 **Criterion mentioned but not elaborated** ***criterion is mentioned but no additional information is offered making replication highly unlikely***	3 **Criterion partially met** ***some elements are present but information is missing or incomplete making replication unlikely***	4 **Criterion is met** ***detailed or full disclosure of information on the specific criterion making replication possible***
		during the research process			impact on substantive conclusions
9	*Management of power imbalance*	The authors do not describe how they addressed the power imbalance between them and the participants	The authors report whether there was any power imbalance with the participants	The authors describe the strategies used to address a general power imbalance with participants	The authors describe specific strategies used to address power imbalance with specific participants
10	*Data coding and first-order codes*	The authors do not describe how they performed the first-order coding of the data nor disclose the first-order codes	The authors offer a general statement about how they conducted the first-order coding, but do not specify a particular approach to doing so	The authors describe the first-order coding methodology (e.g., in vivo coding) and present the first-order code list	The authors describe the first- order coding methodology and present the full code list
11	*Data analysis and second- or higher-order codes*	The authors do not disclose how they performed the data-analysis nor disclose the second-order codes	The authors describe how they approached the identification of key themes in generic terms	The authors describe the second-order coding methodology (e.g., axial coding) and present the second-order code list	The authors describe the second-order coding methodology and present the full code list
12	*Data disclosure*	The authors do not disclose the raw materials (e.g., transcripts, video recordings) gathered and examined during the study	The authors identify the typology of sources gathered and examined during the study	The authors list or identify all the sources gathered and examined during the study	The authors disclose the raw materials gathered and examined during study

Note. These BARS are based on our ontological perspective based on qualitative positivism/transcendental realism. This perspective seems to be dominant in strategy and it was evidenced in our results given that the majority of articles included in our study also follow this ontological perspective.

(Flanagan, 1954) with the goal of defining those domains concretely (Kell et al., 2017). In the domain of HRM&OB, critical incidents consist of reports by knowledgeable observers of things employees did that were especially effective or ineffective in accomplishing parts of their jobs (Aguinis, 2019; Cascio & Aguinis, 2019). Thus, they provide a behavioral base for appraising performance. Similarly, our process of gathering critical incidents involved searching for qualitative studies not only in articles published in SMJ, but also in *Academy of Management Journal* (AMJ), *Administrative Science Quarterly*, and *Organization Science*. The two authors discussed each example until agreement was reached and then we pilot-tested the scale. First, we checked that the full range of possible "transparency behaviors" was represented in BARS in a sample of articles not included in our study and did not identify additional behaviors. Second, we tested the clarity of the BARS during a departmental research seminar at one of the author's university and no adjustments were required.

Two coders used the BARS included in Table 2 to assess each of the 52 articles. First, the two coders proceeded to independently code 10 randomly selected articles. The inter-rater reliability after the first batch of 10 articles was .95 across the 12 criteria. The two coders discussed the few areas of minor disagreement until they reached full consensus. Then, they coded the remaining 42 articles also independently. The inter-rater reliability for this second set of 42 articles was .98 across the 12 criteria. Again, the two coders discussed the few areas of minor disagreement until consensus was reached.

4 | RESULTS

Results of our analysis are summarized in Figure 1, which shows the percent of articles falling into each of the four BARS anchors (i.e., 1: criterion not met, 2: criterion mentioned but not elaborated, 3: criterion partially met, and 4: criterion met) for each of the 12 transparency criteria. As shown in this figure, the vast majority of articles were not sufficiently transparent to allow for replication. Overall, and across the 12 criteria, none of the 52 articles were sufficiently transparent to allow for exact replication, empirical replication, or conceptual replication. But, of the three types, conceptual replication is relatively more likely.

We calculated a *transparency score* to uncover how many articles could be exactly, empirically, or conceptually replicated. Results indicated that, on the four-point scales shown in Table 2, the mean (Mdn) empirical and conceptual replication scores are 1.7 (1.8) and 1.6 (1.4) respectively. Only three articles received transparency scores of at least 2.5 for empirical replication (i.e., Guo, Huy, & Xiao, 2017; Shaffer & Hillman, 2000; Szulanski & Jensen, 2006) and none for conceptual replication. Regarding exact replication, the mean (Mdn) transparency scores across the 52 articles are 1.7 (1.6) and only one article received a score greater than 2.5 (i.e., Shaffer & Hillman, 2000).[3]

Results also showed that 34% of articles explicitly mentioned the kind of qualitative method used (e.g., case study methodology, grounded theory). The second most transparent criterion is the description of the research setting: 25% of the studies clearly described why the specific research setting was chosen and was appropriate, offered a description of the key characteristics of the setting itself, and enough information to identify alternative and similar settings. Another 25% of the articles limited the description of the research setting to some specific characteristics, making conceptual replication more difficult. Half of the studies did not offer enough information to identify a similar setting for a conceptual replication.

Regarding the position of the researcher along the insider-outsider continuum, none of the 52 articles provided explicit information on this issue, but 6% of authors offered some type of information about the relationship existing between them and the target organization. Similarly, we found an insufficient level of transparency regarding sampling criteria. This transparency criterion was fully

[3]As suggested by an anonymous reviewer, we conducted a subsample analysis by comparing transparency scores for the purely qualitative ($N = 27$), mainly quantitative ($N = 8$), and mixed methods ($N = 17$) studies. We compared thee three subgroups regarding their exact replication, empirical replication, and conceptual replication mean scores. All three subgroups received average scores between 1.0 (i.e., complete absence of information on the specific criterion making replication not possible) and 2.0 (i.e., criterion is mentioned but no additional information is offered making replication highly unlikely) for exact replication, empirical replication, and conceptual replication. Results also showed that scores were higher for the purely qualitative studies for just three of the 12 criteria: kind of qualitative method, research setting, and data analysis and second- or higher-order coding. Additional details regarding these analyses and results are available in Appendix E (online supplement).

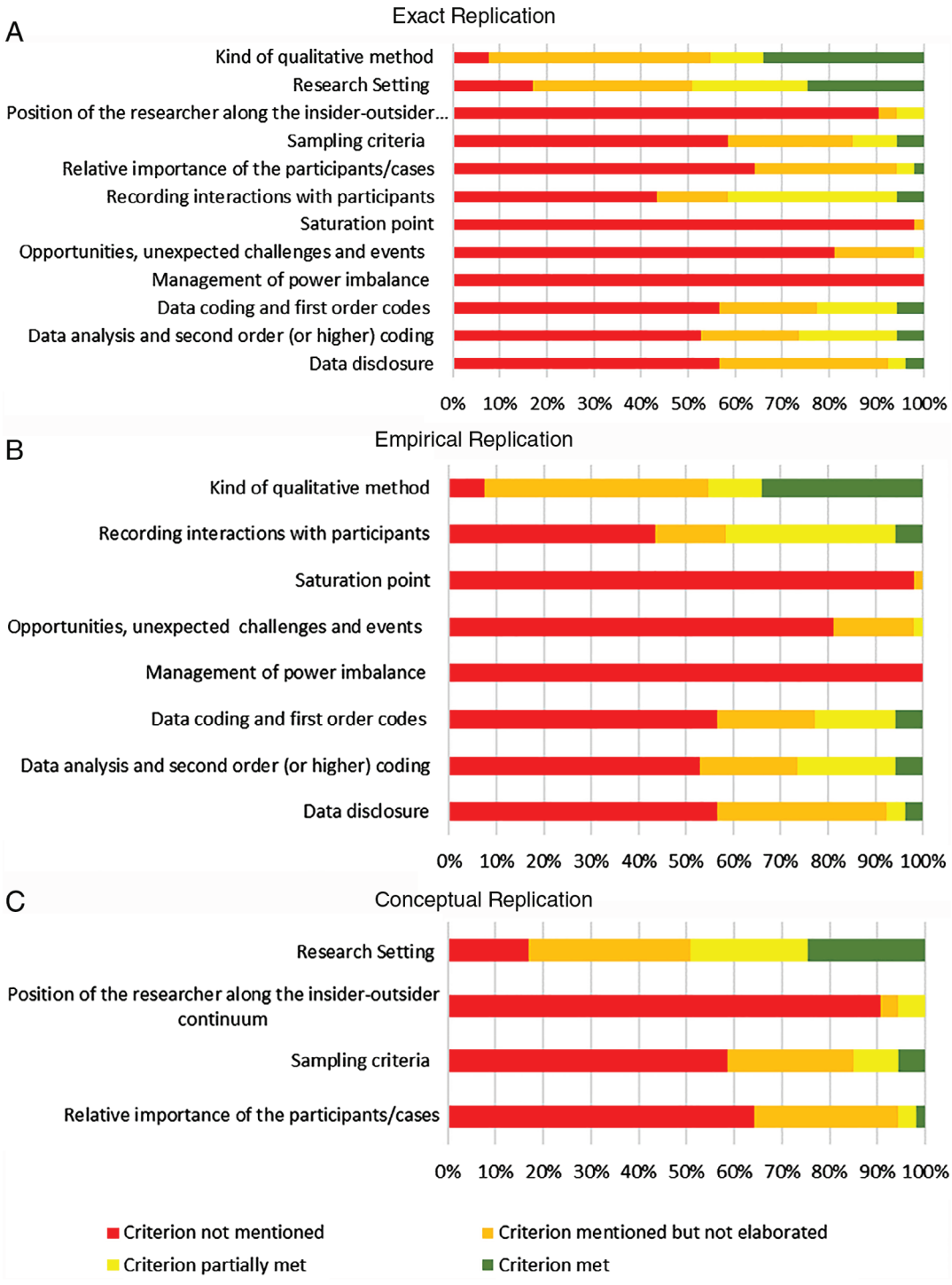

FIGURE 1 Transparency scores for exact, empirical, and conceptual replication using behaviorally-anchored rating scales in Table 2 (based on 52 *Strategic Management Journal* articles that used interviews with elite informants)

met in only 6% of the studies. Lack of transparency emerged also with regard to how authors reported how their interactions with the participants. While 36% of the articles described to what extent each interview was recorded, only 6% described the content of these interviews.

Reaching theoretical saturation is an issue mentioned prominently in most qualitative research textbooks. But, as results showed, most researchers were not sufficiently transparent on whether they reached theoretical saturation and how exactly it was defined and operationalized.

Articles also lacked sufficient transparency with regard to whether there were any opportunities, unexpected challenges, and other events, with only 19% reporting whether something did not go according to plan and 2% describing how they handled such changes. This was also surprising to us given that unexpected events are common in qualitative research in strategic management studies and many other fields. Similarly surprising was the lack of information with regards to the management of power imbalance, given that the literature has addressed not only how to mitigate the power of elite informants (e.g., Dexter, 1970; Ostrander, 1993; Welch, Marschan-Piekkari, Penttinen, & Tahvanainen, 2002), but also how to mitigate the power differences with non-elite informants (e.g., Gubrium & Holstein, 2002) and the risks associated with poor power imbalance management.

A criterion that we also expected would have a higher level of transparency involves data handling. The transparency criterion was met only in 6% of the studies regarding both data coding and analysis. First, only 17% of the articles described the methodology used for developing the codes. Second, only 21% reported how they identified the key themes. In sum, the majority of the studies did not offer sufficient information on how the data were analyzed.

Finally, few authors offered their raw materials (e.g., transcripts) or data (4%). Equally few studies clearly listed all the sources used, while one third of the total sample identified the nature of the sources. Even fewer mentioned the reasons for not sharing their raw materials.

4.1 | Relationship among transparency criteria

We also investigated whether there is a consistently low, moderate, or high degree of transparency across the 12 criteria. Specifically, we expected that if a study is transparent regarding some of the criteria, it would also be transparent regarding others. This expectation is based on a "researcher effect" in that the use of particular methodological procedures (i.e., level of transparency) should be consistent within research teams.

To address this issue, we calculated correlations among the transparency scores. In other words, we examined whether articles that scored high on one transparency criterion also scored high on others. As shown in Figure 1, the distributions of transparency scores are heavily skewed because the majority of articles received a low score on transparency. So, although Pearson's r is the most frequently used correlational test, results can be biased when variables are from distributions with heavy tails (Bishara & Hittner, 2012; de Winter, Gosling, & Potter, 2016). Accordingly, we used Spearman's ρ rather than Pearson's rs and results are included in Table 3. Spearman's ρ is interpreted in the same way as a Pearson's r (Aguinis, Ramani, Alabduljader, Bailey, & Lee, 2019). So, for example, if Spearman's ρ = .40, it means that there is .40 * .40 variance overlap (or 16%) between the two criteria.

Based on results in Table 3, the mean correlations among all 12 transparency criteria is only 0.25 (Mdn = 0.26). So, only 6% (i.e., 0.25 * 0.25) of variance in a given criterion is explained by any other criterion (on average). In other words, the fact that an article is transparent regarding some of the criteria does not necessarily mean that it is transparent regarding others. Accordingly, these results do not support the notion of a "researcher effect" in terms of a set of authors being consistently transparent (or not) across several criteria in a given article.

Table 3 also shows an exception to this overall result: Transparency regarding data coding and data analysis share about 76% of variance. Also, transparency regarding sampling criteria shares about 34% of variance with transparency regarding recording interactions with participants, data coding, and data analysis. All other pairs of variables share trivial amounts of variance.

TABLE 3 Descriptive statistics and Spearman correlations among transparency scores

	Mean	Median	*SD*	1	2	3	4	5	6	7	8	9	10	11	12
1. Kind of qualitative method	2.72	2.00	1.03	—											
2. Research setting	2.57	2.00	1.05	0.32*	—										
3. Position of the researcher along the insider-outsider continuum	1.15	1.00	0.50	0.14	0.13	—									
4. Sampling criteria	1.62	1.00	0.88	0.48**	0.04	0.10	—								
5. Relative importance of the participants/cases	1.43	1.00	0.67	0.27	0.32*	0.31*	0.38**	—							
6. Recording interactions with participants	2.04	2.00	1.02	0.43**	0.21	0.06	0.58**	0.39**	—						
7. Saturation point	1.02	1.00	0.14	0.17	−0.08	0.40**	0.12	0.17	0.14	—					
8. Opportunities, unexpected challenges and events	1.21	1.00	0.45	0.13	0.30*	−0.16	0.19	−0.08	0.32*	−0.07	—				
9. Management of power imbalance	1.00	1.00	—	—	—	—	—	—	—	—	—	—			
10. Data coding and first-order codes	1.72	1.00	0.95	0.39**	0.19	0.07	0.56**	0.28*	0.33*	0.09	0.24	—	—		
11. Data analysis and second- or higher-order codes	1.79	1.00	0.97	0.52**	0.32*	0.08	0.60**	0.29*	0.39**	0.18	0.32*	—	0.87**	—	
12. Data disclosure	1.55	1.00	0.75	0.41**	0.19	0.08	0.34*	0.47**	0.41**	0.13	−0.07	—	0.22	0.28*	—

Note. $N = 53$, but the sample included 52 articles because Paroutis and Heracleous (2013) included two independent studies. "Management of power invariance" has zero variance and, therefore, correlations between this criterion and all others are zero.
*$p < .05$; **$p < .01$.

4.2 | Robustness checks

As a check of the robustness of our results and to assess whether our sample of SMJ articles may be biased, we gathered the 18 AMJ articles with an IEI focus published between 2010 and 2013 and coded them regarding their ontology, methodology, and degree of transparency (we considered that a 4-year window was sufficient for our purposes). The list of these AMJ articles is included in Appendix S1F (Supporting Information). Results based on 12 independent-samples *t* tests (i.e., one for each transparency criterion) comparing SMJ and AMJ articles showed that only one was statistically significant at the .05 level and *d* values (i.e., standardized mean difference for transparency scores between the journals) were small (i.e., across the 12 criteria, $\bar{d} = 0.31$). So, we did not find substantive differences between AMJ and SMJ articles regarding their degree of transparency, which suggests that our results based on SMJ articles are likely generalizable to other journals. Detailed information on these robustness check procedures and results are in Appendix S1G (Supporting Information).

5 | DISCUSSION

We used interviews with elite informants to expand the discussion of transparency and replicability to the domain of qualitative methods. In using IEIs as a case study, we uncovered that insufficient transparency is as pervasive in qualitative research as it has been documented in quantitative research (Aguinis et al., 2018; Bergh et al., 2017; Bettis et al., 2014; Bettis, Ethiraj, et al., 2016; Bettis, Helfat, & Shaver, 2016).

Our results have several implications for replication studies using qualitative research, as well as for authors and journal editors and reviewers in terms of how to enhance the transparency of qualitative research in the future. Also, improved replicability is likely to lead to improvements in quality because manuscripts that are more transparent allow for a more trustworthy assessment of a study's contributions for theory and practice (Brutus, Aguinis, & Wassmer, 2013). Specifically regarding managerial practices, increased transparency means that if organizations implement policies and actions closely based on detailed and open information available in published research, they are more likely to produce results consistent with those reported in the articles.

5.1 | Implications for future replication studies using qualitative research

Our results uncovered the need for more transparency in qualitative research. More specifically, the form of replication that suffers the most given the current low level of transparency is exact replication (i.e., a previous study is replicated using the same population and the same procedures). None of the 52 studies we examined were sufficiently transparent to allow for exact replication.

But, exact replication is not likely to be the most interesting form of replication in strategic management research. Indeed, Ethiraj, Gambardella, and Helfat (2016, p. 2191) stated that SMJ is mostly interested in "quasi-replications," which address the "robustness of prior studies to different empirical approaches" or the "generalizability of prior studies results to new contexts." Using the more specific terminology in our article, quasi-replications include a combination of conceptual (i.e., robustness: same population and different procedures) and empirical (i.e., generalizability: same procedures and a different population) replications. Our results uncovered substantial barriers for these types of replication studies which are those most sought after by SMJ.

Conceptual replication requires a high level of transparency regarding four criteria including the research setting and sampling procedures. Our results revealed that 25% of the studies met the research setting criterion, but only 6% met the sampling procedures criterion. And, the correlation between these criteria is only 0.04 (see Table 3), which means that it is unlikely that if a study met the research setting transparency criterion, it also met the sampling procedures criterion.

TABLE 4 Summary of recommendations for enhancing transparency and replicability in qualitative research for authors and exemplars of the implementation of each recommendation

Transparency criterion	Authors should…	Exemplar
1. *Kind of qualitative method*	… be explicit about what specific kind of qualitative method has been implemented (e.g., narrative research, grounded theory, ethnography, case study, phenomenological research)	Monteiro and Birkinshaw (2017), Ma and Seidl (2018)
2. *Research setting*	… provide detailed information regarding contextual issues regarding the research setting (e.g., power structure, norms, heuristics, culture, economic conditions)	Wood (2009)
3. *Position of researcher along the insider-outsider continuum*	… provide detailed information regarding the researcher's position along the insider-outsider continuum (e.g., existence of a pre-existing relationship with study participants, the development of close relationships during the course of data collection)	Gioia and Chittipeddi (1991)
4. *Sampling procedures*	… be explicit about the sampling procedures (e.g., theoretical sample, purposive sample, snowballing sample, stratified sample)	Ferlie, Fitzgerald, Wood, and Hawkins (2005)
5. *Relative importance of the participants/cases*	… be explicit about the contribution that key informants made to the study	Shaffer and Hillman (2000)
6. *Documenting interactions with participants*	… document interactions with participants (e.g., specify which types of interactions led to the development of a theme).	Bruton, Ahlstrom, and Wan (2003); Shaffer and Hillman (2000)
7. *Saturation point*	… identify the theoretical saturation point and describe the judgment calls the researcher made in defining and measuring it	Guest et al. (2006)
8. *Unexpected opportunities, challenges, and other events*	… report what unexpected opportunities, challenges, and other events occurred during the study, how they were handled (e.g., participants dropped out of the study, a new theoretical framework was necessary), and implications	Michel (2014)
9. *Management of power imbalance*	… report and describe whether power imbalance exits between the researcher and the participants and how it was addressed (e.g., endorsement from a prestigious institution, self-acquaintance, asking sensitive questions)	Yeung (1995); Thomas, (1993); Richards (1996); Stephens (2007)
10. *Data coding and first-order codes*	… be clear about the type of coding strategies adopted (e.g., structural, in vivo, open/initial, emotional, vs.)	Dacin, Munir, and Tracey (2010)
11. *Data analysis and second—or higher-order codes*	… how the data were analyzed (e.g., focused, axial, theoretical, elaborative, longitudinal)	Klingebiel and Joseph (2016)
12. *Data disclosure*	… make raw materials available (e.g., transcripts, video recordings)	Gao, Zuzul, Jones, and Khanna (2017)

Note. These criteria should not be applied rigidly to all qualitative research because, although they are broad in nature, not all of them apply to every situation and type of qualitative study. Overall, our view is that when a larger number of criteria are met, there will be a greater degree of transparency. But this does not mean that the absence of any particular item has veto power over a manuscript's publication deservingness.

Empirical replication requires that even more (i.e., seven) transparency criteria be met. The criteria that were met most frequently were kind of qualitative method (34%), followed by data coding (6%) and data analysis (6%). Not a single one of the 52 studies was sufficiently transparent regarding other criteria needed for empirical replication such as saturation point and the management of power imbalance.

Based on our results, future conceptual replication is relatively more possible than empirical replication, but both types of replication studies are highly unlikely to succeed unless transparency is improved in the future. To facilitate greater transparency and replicability in future qualitative research, next we offer recommendations for authors as well as journal editors and reviewers.

5.2 | Implications for enhancing qualitative research transparency: recommendations for authors

As a preview of this section, Table 4 summarizes recommendations for authors regarding each of the 12 transparency criteria. We reiterate that transparency should be understood as a continuum. The higher the transparency level across the 12 criteria, the more the study becomes trustworthy and reproducible, and the higher the likelihood that future replication will be possible. In other words, the transparency criteria have a cumulative effect in terms of the trustworthiness and replicability of results. Also, we offer recommendations on what features or information to include. Moreover, to show that our recommendations are practical and actionable, and not just wishful thinking, we offer examples of published articles for which a particular criterion was fully met.

Kind of qualitative method. Future research should be explicit about what specific kind of qualitative method has been implemented (e.g., narrative research, grounded theory, ethnography, case study, phenomenological research). This is an important issue regarding transparency because different methodologies in qualitative studies have different goals, objectives, and implications for how the study is executed and how results are interpreted (Creswell, 2007). Moreover, making an explicit statement about the kind of qualitative method used also clarifies researchers' assumptions, beliefs, and values. For example, Monteiro and Birkinshaw (2017) clearly stated that they adopted a grounded theory approach and Ma and Seidl (2018) noted that they implemented a longitudinal multiple-case study approach.

Research setting. Future qualitative research should provide detailed information regarding contextual issues regarding the research setting (e.g., power structure, norms, heuristics, culture, economic conditions). In qualitative research, the research setting is a bundle of pre-existing conditions that alters how the data are collected and interpreted (Van Bavel, Mende-Siedlecki, Brady, & Reinero, 2016). Transparency about the research setting is particularly important in contemporary strategic management studies given the increased use of less conventional settings (e.g., Bamberger & Pratt, 2010; Boyd & Solarino, 2016). An exemplar of a highly transparent research setting is Wood (2009). The author conducted a longitudinal case study on the British brick manufacturing industry describing how the industry evolved, the economic drivers of the industry, the presence of rivalry among manufacturers, the role of the institutions, and several other contextual issues.

Position of researcher along the insider-outsider continuum. Future qualitative research should provide detailed information regarding the researcher's position along the insider-outsider continuum (e.g., existence of a pre-existing relationship with study participants, the development of close relationships during the course of data collection). The presence or absence of these relationship can alter accessibility of data, what participants disclose, and how the collected information is interpreted

(Berger, 2015). As an exemplar of high degree of transparency regarding this criterion, Gioia and Chittipeddi (1991, p. 436) noted that they "employed both an 'insider' and an 'outsider' researcher" perspective and then described the role of each of them.

Sampling procedures. Future qualitative research should be explicit about sampling procedures (e.g., theoretical sample, purposive sample, snowballing sample, stratified sample). This is particularly relevant for qualitative research because samples are often non probabilistic. An examplar is Ferlie et al. (2005, p. 119), who wrote the following: "We constructed a two-by-two cell design to explore effects of stronger/weaker scientific evidence and the degree of innovation complexity on spread pathways… We undertook theoretical rather than random sampling, choosing a pair of innovations in all four cells, giving us a total of eight cases."

Relative importance of the participants/cases. Future qualitative research should be explicit about the contribution that key informants made to the study. In qualitative research not all cases are equally informative. There are circumstances in which some participants are more informative than others because they are those who know and can better articulate how things are actually done (Aguinis et al., 2013). For instance, Shaffer and Hillman (2000, p. 180) identified one of their key informants stating that "the primary interview subject was… who had experience in both state and federal government relations."

Documenting interactions with participants. Future qualitative research should document interactions with participants (e.g., specify which types of interactions led to the development of a theme). This issue is important because different means of documenting interactions (e.g., audio, video, and notations) capture different types of information and alter the willingness of participants to share information (Opdenakker, 2006). A good example of transparency about documenting interactions with participants is Bruton et al. (2003). The authors described how each interview was documented and how the interviews with specific informants were instrumental in understanding how private and state-owned firms are managed in China.

Saturation point. Future qualitative research should identify the saturation point and describe the judgment calls the researcher made in defining and measuring it. The saturation point occurs when there are no new insights or themes in the process of collecting data and drawing conclusions (Bowen, 2008; Strauss & Corbin, 1998). Authors should therefore report how they defined the saturation point and how they decided that it was reached. As an illustration, Guest et al. (2006) described how adding interviews resulted in novel codes and how they decided that theoretical saturation was reached (i.e., the codes generated after the 12th interview were variations of already existing themes).

Unexpected opportunities, challenges, and other events. Future qualitative research should report what unexpected opportunities, challenges, and other events occurred during the study and how they were handled (e.g., participants dropped out of the study, a new theoretical framework was necessary). Because these unexpected events may affect data accessibility and substantive conclusions, researchers should report and describe any unexpected events and highlight whether they had an impact on the data collection and data analysis. For instance, Michel (2014, p. 1086) described how she took advantage of an unexpected request from her informants to "ask questions that would have been inappropriate [otherwise]."

Management of power imbalance. Future qualitative research should report and describe whether power imbalance exists between the researcher and the participants and how it has been addressed (e.g., endorsement from a prestigious institution, self-acquaintance, asking sensitive questions). This issue is important because it allows for similar strategies to be adopted in future replication studies. Yeung (1995), for instance, used the exchange of business cards with the informants to reduce the

power differential, and Stephens (2007) used phone interviews to not reveal the age difference between him and his informants.

Data coding and first-order codes. Future qualitative research should be clear about the type of coding strategies adopted (e.g., structural, in vivo, open/initial, emotional, and vs.). This is an important issue because different types of coding procedures have different goals. A good example of transparency regarding data coding is Dacin et al. (2010). The authors clearly stated that they used in vivo coding to develop the first-order codes and then reported the full list of the codes in the paper.

Data analysis and second- or higher-order codes. Future qualitative research should be clear about how the data were analyzed (e.g., focused, axial, theoretical, elaborative, and longitudinal). As an exemplar of transparency, Klingebiel and Joseph (2016) identified the methodologies adopted in data analysis (axial, and selective coding) and reported the final higher order codes along with the first-order codes that generated them.

Data disclosure. Future qualitative research should make raw materials available (e.g., transcripts, video recordings). While this criterion is necessary only for exact replication, the disclosure of the raw material is useful for error checking. Authors could make the data available to others researchers directly, in data repositories, or by request. An example is Gao et al. (2017), whose data are available for downloading from the Business History Initiative website.

As an additional issue, the aforementioned recommendations can also be beneficial for future research adopting a mixed-methods approach. Mixed methods research combines qualitative and quantitative procedures and data and, therefore, all of the recommendations described above are applicable. The application of these recommendations to mixed-methods research can be particularly beneficial for two reasons. First, the 165 mixed-methods articles published in SMJ from 1980 to 2006 have had more influence on subsequent research compared to articles adopting a single-method approach based on the average number of citations they received (i.e., the mean citation count of 59.13 for mixed-methods articles and 37.08 for single-methods articles; Molina-Azorin, 2012). Second, issues of transparency in mixed-methods research require immediate attention, as highlighted in the introductory article of a special issue on mixed-methods in *Organizational Research Methods.* Specifically, Molina-Azorin, Bergh, Corley, and Ketchen (2017, p. 186) offered recommendations for moving mixed methods forward and wrote that "to maximize the potential for subsequent research to be able to replicate a mixed methods study, researchers need to be as transparent as possible in reporting their methodological decisions and the rationale behind those choices... future researchers should be able to understand how … data were collected, analyzed, and integrated such that similar methodological efforts could be recreated in different contexts collection and data analysis within a single study."

5.3 | Implications for enhancing qualitative research transparency: recommendations for journal editors and reviewers

Journal articles published in strategic management studies and many other fields go through several revisions until they are eventually published. But, readers only have access to the final and published version. Thus, they do not know whether authors may have been more transparent in earlier versions of their manuscript but, ironically, the level of transparency decreased as a result of the review process. For example, editors and reviewers may have asked that authors remove information on some aspects of the study to comply with a journal's word or page limit. Also, information directly relevant to some of the 12 transparency criteria may have been included in the authors' responses to the review team, but not in the actual manuscript. So, the overall low level of transparency uncovered by our

study is likely the result of a complex review process involving not only authors but also a journal's review team. Moreover, the fact that the mean correlation among the 12 criterion is only 0.25 suggests that the relative degree of transparency across the criteria is quite idiosyncratic. Some reviewers may have suggested that information on specific transparency criteria be omitted (or added), whereas other reviewers may have made suggestions about different transparency criteria. Most authors are quite familiar with the many differences between a manuscript's first submission and what is ultimately published.

So, to supplement recommendations for authors in the previous section, we offer the following recommendations for journal editors and reviewers, which can also inform journal submission and review policies. Specifically, the BARS that we developed for our study, which are included in Table 2, can be used by editors and reviewers to enhance transparency of manuscripts *before* they are published. In particular, reviewers can use the BARS to make a judgment on the relative level of transparency regarding the 12 transparency criteria. Moreover, the resulting evaluation can be used for making a recommendation regarding the suitability of a particular manuscript in terms of their ultimate publication and also in terms of offering authors developmental feedback on which particular criteria should be improved in a revision. In other words, the BARS can be used for evaluative purposes, but also as a developmental tool that would allow the review team to offer concrete advice on actions authors can take to improve transparency regarding specific issues.

From a practical and implementation standpoint, an important advantage of using the BARS in Table 2 is that they can be added to existing reviewer evaluation forms without much cost or effort. Also, implementing recommendations about transparency provided by reviewers is now greatly facilitated by the availability of Supporting Information, as is done customarily in journals such as *Nature* and *Science*. In these journals, articles are usually very short compared to those in strategic management studies. But the Supporting Information are much longer and include details about research design, measurement, data collection, data analysis, and data availability.

Finally, an anonymous reviewer raised the concern that the widespread use of the BARS in Table 2 to assess the transparency criteria in the journal review process may reduce alternative perspectives such as practice theory, interpretive or narrative approaches, and ethnographies. It is certainly not our intention to imply that these alternative perspectives should be limited. In fact, we see the 12 transparency criteria as an extension to what Lincoln and Guba (1985) referred to as "thick description" in qualitative research. In other words, it is important that qualitative researchers provide a thick description of all procedures and choices so that readers and other consumers of research (i.e., practitioners) are able to interpret results and conclusions correctly. In addition, we do not believe that increased transparency regarding the 12 criteria will serve as a "screening out" device for manuscripts adopting alternative perspectives. For example, manuscripts adopting practice theory, interpretive or narrative approaches, and ethnographies would still benefit from increased transparency regarding the description of the research setting (criterion #2) and documenting interactions with participants (criterion #6), just as much as would manuscripts adopting a post positivist approach.

5.4 | Limitations and suggestions for future research

Our study as well as recommendations for authors and journal editors and reviewers is based on the 12 specific transparency criteria that we developed (which we also applied to our own study). We developed these criteria based on an extensive literature review of 127 articles and 14 books and believe they are widely applicable and sufficiently broad so they can be used to assess transparency in many types of qualitative methods and across substantive domains. Moreover, taken together, the 12 criteria cover the usual phases of the research process: design, measurement, analysis, and data

availability. In spite of these strengths and advantages, we readily acknowledge that additional criteria could be added to the list. Future research could compare how various levels of transparency based on our 12 criteria may correlate with the level of transparency of additional criteria. If the degree of overlap is low, this would suggest the value-added of including additional criteria to the list. On the other hand, a high degree of overlap would suggest redundancy with existing criteria.

Second, also related to the transparency criteria, they should not be applied rigidly to all qualitative research. The reason is that, although they are broad in nature, not all of them apply to every situation and type of qualitative study. For example, a manuscript may describe a qualitative study that is only a small portion of a larger effort. Relatedly, a manuscript may adopt a qualitative approach that is more positivistic in nature such as the use of computer-aided text analysis and some of the criteria (e.g., theoretical saturation) may not apply. Overall, our view is that when a larger number of criteria are met, there will be a greater degree of transparency. But this does not mean that the absence of any particular item has veto power over a manuscript's publication deservingness. This is a decision that reviewers and action editors will have to weigh.

Third, we once again openly acknowledge our ontological perspective based on qualitative positivism/transcendental realism. This perspective seems to be dominant in strategy and it was evidenced in our results given that the majority of articles included in our study also follow this ontological perspective. Clearly, other perspectives exist, including in the particular domain of IEIs (e.g., Empson, 2018).

Fourth, another suggestion regarding future research is the assessment of transparency in the use of other qualitative methodologies. For example, is the degree of transparency in research using interviews with non-elite informants more transparent than that using elites? Also, contemporary qualitative research in strategic management studies is making increased use of Big Data and, overall, the Internet and other technology-based enablements to collect and analyze information. Although the term Big Data is typically used for quantitative research, the availability of large amounts of texts, videos, and other non-numerical information (e.g., photos) posted online opens up a host of interesting possibilities (e.g., Christianson, 2018). The use of computer-based text analysis (CATA) is another example (e.g., McKenny, Aguinis, Short, & Anglin, 2018).[4] So, future research could examine whether the use of these innovations is resulting in higher or lower levels of transparency.

Finally, there are initiatives to enhance transparency that have been implemented in other fields that could be implemented in strategy to assess the extent to which they are beneficial. For example, these include the use of pre-registered reports, which are manuscripts submitted for possible journal publication describing a study's method and proposed analyses, but not the results. Other initiatives include open source collaboration and sharing tools like Slack. Future efforts could examine benefits, and also potential pitfalls, regarding these and other initiatives such as posting datasets online and various changes in policies by journals, professional organizations, and funding agencies aimed at reducing impediments to transparency.

6 | CONCLUDING REMARKS

Similar to Bettis et al. (2016, p. 257), our goal was to "to promote discussions and educational efforts among Ph.D. students, scholars, referees, and editors in strategic management regarding

[4]CATA is "a form of content analysis that enables the measurement of constructs by processing text into quantitative data based on the frequency of words" (McKenny et al., 2018, p. 2910). As such, it includes a combination of qualitative and quantitative procedures.

the replicability and cumulativeness of our… research knowledge." But, because these issues have to date been limited to quantitative research (Aguinis et al., 2017; Bergh et al., 2017) and "statistical results" (Bettis, Ethiraj, et al., 2016, p. 2194), we expanded the discussion of transparency and replicability to the domain of qualitative methods. We used the particular qualitative technique of interviews with elite informants and results of our study of articles published in SMJ revealed insufficient transparency. Overall, and across the 12 criteria, none of the 52 articles were sufficiently transparent to allow for exact replication, empirical replication, or conceptual replication. We hope our recommendations for authors as well as journal editors and reviewers will serve as a catalyst for improving the degree of transparency and replicability of future qualitative research.

ACKNOWLEDGEMENTS

Both authors contributed equally to this research. We thank Stefano Brusoni and three *Strategic Management Journal* anonymous reviewers for extremely constructive and detailed feedback that allowed us to improve our article substantially.

ORCID

Herman Aguinis https://orcid.org/0000-0002-3485-9484
Angelo M. Solarino https://orcid.org/0000-0001-5030-7375

REFERENCES

Aguinis, H. (2019). *Performance management* (4th ed.). Chicago, IL: Chicago Business Press.

Aguinis, H., Cascio, W. F., & Ramani, R. S. (2017). Science's reproducibility and replicability crisis: International business is not immune. *Journal of International Business Studies*, *48*(6), 653–663.

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best–practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, *16*(2), 270–301.

Aguinis, H., & Molina-Azorin, J. F. (2015). Using multilevel modeling and mixed methods to make theoretical progress in microfoundations for strategy research. *Strategic Organization*, *13*(4), 353–364.

Aguinis, H., Ramani, R. S., & Alabduljader, N. (2018). What you see is what you get? Enhancing methodological transparency in management research. *Academy of Management Annals*, *12*(1), 83–110.

Aguinis, H., Ramani, R. S., Alabduljader, N., Bailey, J. R., & Lee, J. (2019). A pluralist conceptualization of scholarly impact in management education: Students as stakeholders. *Academy of Management Learning and Education*, *18*(1), 11–42.

Aguinis, H. (2019). *Performance management* (4th ed.). Chicago, IL: Chicago Business Press.

Ajami, R. A. (1980). Determinants of attitudes toward multinational firms: Host Arab society elites. *Strategic Management Journal*, *1*(3), 193–208.

Bamberger, P. A., & Pratt, M. G. (2010). Moving forward by looking back: Reclaiming unconventional research contexts and samples in organizational scholarship. *Academy of Management Journal*, *53*(4), 665–671.

Basu, K., & Palazzo, G. (2008). Corporate social responsibility: A process model of sensemaking. *Academy of Management Review*, *33*(1), 122–136.

Baum, R. J., & Wally, S. (2003). Strategic decision speed and firm performance. *Strategic Management Journal*, *24*(11), 1107–1129.

Berger, R. (2015). Now I see it, now I don't: Researcher's position and reflexivity in qualitative research. *Qualitative Research*, *15*(2), 219–234.

Bergh, D. D., Sharp, B. M., Aguinis, H., & Li, M. (2017). Is there a credibility crisis in strategic management research? Evidence on the reproducibility of study findings. *Strategic Organization*, *15*(3), 423–436.

Bettis, R. A., Ethiraj, S., Gambardella, A., Helfat, C., & Mitchell, W. (2016). Creating repeatable cumulative knowledge in strategic management. *Strategic Management Journal*, *37*(2), 257–261.

Bettis, R. A., Gambardella, A., Helfat, C., & Mitchell, W. (2014). Quantitative empirical analysis in strategic management. *Strategic Management Journal*, *35*(7), 949–953.

Bettis, R. A., Helfat, C. E., & Shaver, J. M. (2016). The necessity, logic, and forms of replication. *Strategic Management Journal*, *37*(11), 2193–2203.

Bhattacharya, H. (2008). Research setting. In L. M. Given (Ed.), *The Sage encyclopedia of qualitative research methods* (pp. 787–788). Thousand Oaks, CA: Sage.

Bishara, A. J., & Hittner, J. B. (2012). Testing the significance of a correlation with nonnormal data: Comparison of Pearson, Spearman, transformation, and resampling approaches. *Psychological Methods*, *17*(3), 399–417.

Bowen, G. A. (2008). Naturalistic inquiry and the saturation concept: A research note. *Qualitative Research*, *8*(1), 137–152.

Boyd, B. K., & Solarino, A. M. (2016). Ownership of corporations a review, synthesis, and research agenda. *Journal of Management*, *42*(5), 1282–1314.

Bruton, G. D., Ahlstrom, D., & Wan, J. C. (2003). Turnaround in east Asian firms: Evidence from ethnic overseas Chinese communities. *Strategic Management Journal*, *24*(6), 519–540.

Brutus, S., Aguinis, H., & Wassmer, U. (2013). Self-reported limitations and future directions in scholarly reports: Analysis and recommendations. *Journal of Management*, *39*(1), 48–75.

Cascio, W. F., & Aguinis, H. (2019). *Applied psychology in talent management* (8th ed.). Thousand Oaks, CA: Sage.

Charmaz, K. (2006). *Constructing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.

Christianson, M. K. (2018). Mapping the terrain: The use of video-based research in top-tier organizational journals. *Organizational Research Methods*, *21*(2), 261–287.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks, CA: Sage.

Dacin, M. T., Munir, K., & Tracey, P. (2010). Formal dining at Cambridge colleges: Linking ritual performance and institutional maintenance. *Academy of Management Journal*, *53*(6), 1393–1418.

de Winter, J. C., Gosling, S. D., & Potter, J. (2016). Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological Methods*, *21*(3), 273–290.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (1994). *Handbook of qualitative research*. Thousand Oaks, CA: Sage.

Dexter, L. A. (1970). *Elite and specialized interviewing*. Evanston, IL: Northwestern University Press.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, *14*(4), 532–550.

Empson, L. (2018). Elite interviewing in professional organizations. *Journal of Professions and Organization*, *5*(1), 58–69.

Ethiraj, S. K., Gambardella, A., & Helfat, C. E. (2016). Replication in strategic management. *Strategic Management Journal*, *37*(11), 2191–2192.

Evered, R., & Louis, M. R. (1981). Alternative perspectives in the organizational sciences: "Inquiry from the inside" and "inquiry from the outside". *Academy of Management Review*, *6*(3), 385–395.

Felin, T., Foss, N. J., & Ployhart, R. E. (2015). The microfoundations movement in strategy and organization theory. *The Academy of Management Annals*, *9*(1), 575–632.

Ferlie, E., Fitzgerald, L., Wood, M., & Hawkins, C. (2005). The nonspread of innovations: The mediating role of professionals. *Academy of Management Journal*, *48*(1), 117–134.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, *51*(4), 327–358.

Flick, U. (Ed.). (2014). *The Sage handbook of qualitative data analysis*. Thousand Oaks, CA: Sage.

Foss, N. J., & Pedersen, T. (2016). Microfoundations in strategy research. *Strategic Management Journal.*, *37*(13), E22–E34.

Gao, C., Zuzul, T., Jones, G., & Khanna, T. (2017). Overcoming institutional voids: A reputation-based view of long-run survival. *Strategic Management Journal*, *38*(11), 2147–2167.

Giddens, A. (1972). Elites in the British class structure. *The Sociological Review*, *20*(3), 345–372.

Gioia, D. A., & Chittipeddi, K. (1991). Sensemaking and sensegiving in strategic change initiation. *Strategic Management Journal*, *12*(6), 433–448.

Griffith, A. I. (1998). Insider/outsider: Epistemological privilege and mothering work. *Human Studies*, *21*(4), 361–376.

Gubrium, J. F., & Holstein, J. A. (Eds.). (2002). *Handbook of interview research: Context and method*. London, UK: Sage.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, *18*(1), 59–82.

Guion, R. M. (2011). *Assessment, measurement, and prediction for personnel decisions*. New York, NY: Routledge.

Guo, Y., Huy, Q. N., & Xiao, Z. (2017). How middle managers manage the political environment to achieve market goals: Insights from China's state-owned enterprises. *Strategic Management Journal*, *38*(3), 676–696.

Hambrick, D. C., & Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. *Academy of Management Review*, *9*(2), 193–206.

Harvey, W. S. (2010). Methodological approaches for interviewing elites. *Geography Compass*, *4*(3), 193–205.

Hauenstein, N., Brown, R., & Sinclair, A. (2010). BARS and those mysterious, missing middle anchors. *Journal of Business and Psychology*, *25*(4), 663–672.

Kell, H. J., Martin-Raugh, M. P., Carney, L. M., Inglese, P. A., Chen, L., & Feng, G. (2017). Exploring methods for developing behaviorally anchored rating scales for evaluating structured interview performance. *ETS Research Report Series.*

Kincaid, H. V., & Bright, M. (1957). Interviewing the business elite. *American Journal of Sociology*, *63*(3), 304–311.

Klingebiel, R., & Joseph, J. (2016). Entry timing and innovation strategy in feature phones. *Strategic Management Journal*, *37*(6), 1002–1020.

Kowal, S., & O'Connell, D. C. (2014). Transcription as a crucial step of data analysis. In U. Flick (Ed.), *The Sage handbook of qualitative data analysis* (pp. 64–79). Thousand Oaks, CA: Sage.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Lincoln, Y. S., Lynham, S. A., & Guba, E. G. (2011). Paradigmatic controversies, contradictions, and emerging confluences, revisited. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (pp. 97–128). Beverly Hills, CA: Sage.

Locke, K. (2001). *Grounded theory in management research*. Newbury Park, CA: Sage.

Ma, S., & Seidl, D. (2018). New CEOs and their collaborators: Divergence and convergence between the strategic leadership constellation and the top management team. *Strategic Management Journal*, *39*(3), 606–638.

Maurer, S. D. (2002). A practitioner–based analysis of interviewer job expertise and scale for–mat as contextual factors in situational interviews. *Personnel Psychology*, *55*(2), 307–327.

Maxwell, J. A., & Chmiel, M. (2014). Generalization in and from qualitative analysis. In U. Flick (Ed.), *The Sage handbook of qualitative data analysis* (pp. 540–553). Thousand Oaks, CA: Sage.

McDowell, L. (1998). Elites in the city of London: Some methodological considerations. *Environment and Planning A*, *30*(12), 2133–2146.

McKenny, A. F., Aguinis, H., Short, J. C., & Anglin, A. H. (2018). What doesn't get measured does exist: Improving the accuracy of computer–aided text analysis. *Journal of Management*, *44*(7), 2909–2933.

Michel, A. (2014). The mutual constitution of persons and organizations: An ontological perspective on organizational change. *Organization Science*, *25*(4), 1082–1110.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Newbury Park, CA: Sage.

Molina-Azorin, J. F. (2012). Mixed methods research in strategic management: Impact and applications. *Organizational Research Methods*, *15*(1), 33–56.

Molina-Azorin, J. F., Bergh, D. D., Corley, K. G., & Ketchen, D. J. (2017). Mixed methods in the organizational sciences: Taking stock and moving forward. *Organizational Research Methods*, *20*(2), 179–192.

Monteiro, F., & Birkinshaw, J. (2017). The external knowledge sourcing process in multinational corporations. *Strategic Management Journal*, *38*(2), 342–362.

Opdenakker, R. (2006). Advantages and disadvantages of four interview techniques in qualitative research. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, *7*(4) http://www.qualitative–research.net/index.php/fqs/article/view/175/391

Ostrander, S. A. (1993). "surely you're not in this just to be helpful": Access, rapport, and interviews in three studies of elites. *Journal of Contemporary Ethnography*, *22*(1), 7–27.

Paroutis, S., & Heracleous, L. (2013). Discourse revisited: Dimensions and employment of first-order strategy discourse during institutional adoption. *Strategic Management Journal*, *34*(8), 935–956.

Patton, M. Q. (2002). *Qualitative evaluations and research methods*. Newbury Park, CA: Sage.

Reuer, J. J., Klijn, E., & Lioukas, C. S. (2014). Board involvement in international joint ventures. *Strategic Management Journal*, *35*(11), 1626–1644.

Richards, D. (1996). Elite interviewing: Approaches and pitfalls. *Politics*, *16*(3), 199–204.

Ryan, G. W., & Bernard, H. R. (2000). Data management and analysis methods. In N. Densin & Y. Lincol (Eds.), *Handbook of qualitative research* (pp. 769–802). Thousand Oaks, CA: Sage.

Saldana, J. (2009). *The coding manual for qualitative researchers*. Los Angeles, CA: Sage.

Schreiber, J. B. (2008). Data. In L. M. Given (Ed.), *The Sage encyclopedia of qualitative research methods* (pp. 185–186). Thousand Oaks, CA: Sage.

Shaffer, B., & Hillman, A. J. (2000). The development of business–government strategies by diversified firms. *Strategic Management Journal*, *21*(2), 175–190.

Shipilov, A., Godart, F. C., & Clement, J. (2017). Which boundaries? How mobility networks across countries and status groups affect the creative performance of organizations. *Strategic Management Journal*, *38*(6), 1232–1252.

Stephens, N. (2007). Collecting data from elites and ultra elites: Telephone and face–to–face interviews with macroeconomists. *Qualitative Research*, *7*(2), 203–216.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research*. Newbury Park, CA: Sage.

Szulanski, G., & Jensen, R. J. (2006). Presumptive adaptation and the effectiveness of knowledge transfer. *Strategic Management Journal*, *27*(10), 937–957.

Taylor, S. J., Bogdan, R., & DeVault, M. L. (2016). *Introduction to qualitative research methods*. Hoboken, NJ: John Wiley & Sons.

Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research*, *1*(1), 77–100.

Thomas, R. J. (1993). Interviewing important people in big companies. *Journal of Contemporary Ethnography*, *22*(1), 80–96.

Tsang, E. W., & Kwan, K. M. (1999). Replication and theory development in organizational science: A critical realist perspective. *Academy of Management Review*, *24*(4), 759–780.

Van Bavel, J. J., Mende-Siedlecki, P., Brady, W., & Reinero, D. A. (2016). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences*, *113*(23), 6454–6459.

Vaughan, S. (2013). Elite and elite-lite interviewing: Managing our industrial legacy. In A. Franklin & P. Blyton (Eds.), *Researching sustainability: A guide to social science methods, practice and engagement* (pp. 105–119). London, U.K.: Routledge.

Welch, C., Marschan-Piekkari, R., Penttinen, H., & Tahvanainen, M. (2002). Corporate elites as informants in qualitative international business research. *International Business Review*, *11*(5), 611–628.

Welch, C., & Piekkari, R. (2017). How should we (not) judge the "quality" of qualitative research? A re–assessment of current evaluative criteria in international business. *Journal of World Business*, *52*(5), 714–725.

Wood, A. (2009). Capacity rationalization and exit strategies. *Strategic Management Journal*, *30*(1), 25–44.

Yeung, H. W. C. (1995). Qualitative personal interviews in international business research: Some lessons from a study of Hong Kong transnational corporations. *International Business Review*, *4*(3), 313–339.

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Thousand Oaks, CA: Sage.

Zhou, Y. M., & Wan, X. (2017). Product variety and vertical integration. *Strategic Management Journal*, *38*(5), 1134–1150.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Aguinis H, Solarino AM. Transparency and replicability in qualitative research: The case of interviews with elite informants. *Strat. Mgmt. J*. 2019;40: 1291–1315. https://doi.org/10.1002/smj.3015