

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Practices and attitudes toward replication in empirical translation and interpreting studies

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Olalla Soler, C. (2020). Practices and attitudes toward replication in empirical translation and interpreting studies. TARGET, 32(1), 3-36 [10.1075/target.18159.ola].

Availability:

This version is available at: <https://hdl.handle.net/11585/792625> since: 2021-01-28

Published:

DOI: <http://doi.org/10.1075/target.18159.ola>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

Practices and attitudes toward replication in empirical translation and interpreting studies

Christian Olalla-Soler

Universitat Autònoma de Barcelona

The aim of this article is to present the results of three studies on practices in and attitudes toward replication in empirical research in translation and interpreting studies. A study based on a survey approach was conducted in which 52 researchers in translation and interpreting with experience in empirical research completed a survey on their practices in and attitudes toward replication. The survey data were complemented by a bibliometric study in which the publications indexed in the Bibliography of Interpreting and Translation (BITRA) (Franco Aixelá 2001–2019) that explicitly stated in the title or abstract that they were derived from a replication were identified and analyzed. In a second bibliometric study, a conceptual replication of Yeung’s (2017) study on the acceptance of replications in neuroscience journals was conducted by analyzing 131 translation and interpreting journals. The purpose of this article is to provide evidence-based arguments for initiating a debate about the need for replication in empirical translation and interpreting studies and its implications for the development of the discipline.

Keywords: replication, empirical translation and interpreting studies, methodology, survey approach, bibliometric approach

1. Introduction

This article presents the results obtained in a series of studies aiming to describe the practices and attitudes of researchers conducting empirical studies in the field of translation and interpreting studies (TIS) regarding replication. Replication is defined as the repetition of the methods that led to a reported finding (Schmidt 2009).

Replication is a central concept in the scientific method and a key to scientific progress (Popper 1959; Kuhn 1962). The replicability of empirical research and the

amount of replication that takes place is one of the many factors that differentiate soft and hard sciences. While the opposition between soft and hard sciences is not free of criticism (e.g., Graham and Kantor 2007), soft-science disciplines are thought to obtain less replicable results (Hedges 1987; Baker and Penny 2016), and to employ a lower degree of logical and methodological rigor in testing hypotheses (Fanelli 2010). Even if replication in soft sciences is not expected to happen as frequently as in hard sciences, scholars from many soft-science disciplines such as psychology, neuroscience and sociology have expressed concern about the low replicability of their empirical research (Baker and Penny 2016); that is, the low likelihood of being able to find the same effects as an original empirical study. Furthermore, many academic journals devoted to soft-science disciplines generally do not publish articles derived from replications because they do not constitute new, original research (Ferguson and Heene 2012; Baker and Penny 2016; Yeung 2017).

In TIS, which is considered a multidisciplinary area of research (Saldanha and O’Brien 2014), replication is mainly possible in studies based on a positivistic philosophy of science, which is characterized by “a belief that the study of human behaviour should seek to identify, measure and evaluate any phenomenon and to provide rational explanation for it” (Hale and Napier 2013, 14). TIS research based on positivistic approaches is epistemologically close to soft-science disciplines. Literature on replication in TIS is notably scarce, and the few references that can be found tend to ask for more replication to take place (Alves, Pagano, and da Silva 2011; O’Brien 2011). In view of the importance given to replication in science and the scarce reflection on replication in TIS, this article aims at providing data for the following research questions:

1. Is replication needed in empirical TIS?
2. Do researchers who carry out empirical studies in translation and interpreting conduct replications? What are the characteristics of these replications?
3. If researchers in empirical TIS do conduct replications, what problems do they face when they replicate empirical studies?
4. If researchers in empirical TIS do conduct replications, how can replication be enhanced?
5. Is there a publication bias in translation and interpreting journals that discourages publishing replications?

To answer these questions, three studies were conducted. A survey approach (Ravitch and Mittenfelner Carl 2016) was used in one study, and a bibliometric approach (Rovira-Esteva and Franco Aixelá 2018) was employed in two studies. Fifty-two researchers who had conducted empirical research on translation and interpreting filled out a survey related to their practices in and attitudes toward

This is a post-print version of the following article published as “online first”:
Olalla-Soler, C. (2020). Practices and attitudes toward replication in empirical translation and interpreting studies. *Target*, 32(1). <https://doi.org/10.1075/target.18159.ola>

replication. This study gathered answers to all five research questions. The first bibliometric study aimed to complement the survey results, providing answers to research questions 2 and 5 by identifying publications derived from replications of studies on translation and interpreting and their characteristics. The purpose of the second bibliometric study was to provide complementary data to answer research question 5 by determining whether there is a publication bias in translation and interpreting journals that discourages publishing replications.

This article is structured as follows. First, the conceptual framework is discussed, with particular reference to the role of replication in soft-science disciplines and in empirical TIS. Second, the research design of the three studies is described. Third, the results of each study are presented. Fourth, the results are discussed and linked to other investigations, and finally, the most relevant conclusions are established.

2. Replication in soft-science disciplines

The repetition of the methods that led to a reported finding (Schmidt 2009), that is, replication, is a key concept in the philosophy of science. Popper (1959) considered that scientific theories encompass predictions that might turn out to be false in future observations. Since science progresses by trial and error, confirmed hypotheses need to be tested many times until the hypothesis is revealed to be false. Thus, replication plays a central role in the scientific method because it is one of the keys to scientific progress. In Popper’s (1959, 64) words, “non-reproducible single occurrences are of no significance to science.” However, as a single confirmation of a hypothesis would not be enough to state that a theory is true, a single failed replication would not be enough to state that the theory is false.

Replication also plays a major role in the progress of science according to Kuhn’s (1962) cycle of scientific revolutions. Pre-science paves the way to the development of a shared paradigm (a disciplinary matrix of symbolic generalizations, metaphysical presumptions, judgments of theories, and models). Scientists who adhere to this paradigm practice *normal science*, in which researchers develop and articulate the paradigm. In this process, researchers might conduct studies whose results contradict previously confirmed hypotheses. Hence, researchers might question the paradigm, and this might lead to a scientific crisis in which the confirmed evidence is questioned until a new paradigm emerges. Researchers who adopt the new paradigm produce *revolutionary science*.

While most philosophers of science¹ attribute a relevant role to replication in the progress of science, many scientists have engaged in debates about what, exactly, replication is. Despite the fact that Schmidt’s (2009) broad definition of replication has been generally adopted, replications may differ greatly from one another. In soft-science disciplines such as psychology, neuroscience or sociology, scientists typically identify three types of replication regardless of the methods employed (qualitative and/or quantitative), even if replication tends to be more frequent in research based on quantitative approaches (Polit and Beck 2010). The three types of replication vary in the number and types of differences in relation to the original study, and can be defined as follows:

- Exact replication: “an attempt to conduct a study [...] in a manner as close to the original as possible. An exact replicator seeks to use the same materials, the same manipulations, the same dependent variables, and the same kind of participants” (Crandall and Sherman 2016, 93).
- Constructive replication: “follow-up studies that include an exact or close replication of an original study in an exact/close replication condition, but also include new elements in a constructive replication condition. [...] Epistemologically, constructive replications thus seek not ‘only’ to provide additional evidence for or against an existing finding but also to refine or extend findings” (Hüffmeier, Mazei, and Schultze 2016, 86).
- Conceptual replication: “an attempt to test the same fundamental idea or hypothesis behind the original study, but the operationalizations of the phenomenon, the independent and dependent variables, the type and design of the study, and the participant population may all differ substantially” (Crandall and Sherman 2016, 93).

Each replication type pursues a different purpose. Exact replications assess the reliability of the effect observed in the original study (Simons 2014), conceptual replications address issues of construct validity of the original study (Fabrigar and Wegener 2016), and constructive replications assess the reliability of the effect detected in the original study by modifying a limited number of aspects of the research design.

Many scientists favor exact replications over conceptual ones, arguing that a failed conceptual replication could indicate that the original finding was false, or the new operationalization of the study might have caused the negative result. Thus, they are often considered uninformative (Simons 2014; Fabrigar and Wegener 2016). The criticism of exact replications stems from the concept of

¹ See Feyerabend (1978) for a critique of Popper’s and Kuhn’s views on the scientific method and scientific progress.

‘exact’. Many scientists argue that it is not possible to recreate the same contextual factors and participant characteristics of the original study (Schmidt 2009; Fabrigar and Wegener 2016), especially in disciplines in which sociological factors play a major role (Crandall and Sherman 2016). Thus, exact replications are more plausible in hard-science disciplines such as physics, biology or chemistry than in soft-science disciplines, given that these sociological factors have little or no influence in hard-science disciplines. Given the limitations of this type of replication in soft-science disciplines, some scientists believe that exact replications should only be carried out by the original researcher to confirm the original finding (Cesario 2014), while others argue that exact replications should be conducted by others to control the possible bias of the original researchers (Simons 2014).

2.1 Replicability in soft-science disciplines

Using simulations, Ioannidis (2005) demonstrates that for most study designs a research finding is more likely to be false than true. Moonesinghe, Khoury, and Janssens (2007) confirm this result, but they also demonstrate that replication enhances the likelihood that published findings are true. However, many scientists working in soft-science disciplines find that in replications original effects tend not to be confirmed, especially in exact replications. An example is the study carried out by Open Science Collaboration (2015), a team of psychology researchers who conducted exact replications of 100 experimental and correlational studies published in peer-reviewed journals. The replication effects were half the magnitude of the original ones, and while 87% of the original studies had statistically significant results, only 36% of the replications did.

A survey carried out by *Nature* with a sample of 1,576 scientists from both hard- and soft-science disciplines (Baker and Penny 2016) found that more than 70% of the participants had failed to replicate another scientist’s original effects, and more than 50% were unable to reproduce their own. Replicability is considered to be higher in hard sciences than in soft sciences (Baker and Penny 2016), although the values vary from one discipline to another.

Many causes for the low replicability of scientific studies have been identified. In the *Nature* survey, scientists from both hard- and soft-science disciplines referred to pressure to publish, selective reporting, low statistical power, and methodological aspects that made replication difficult as causes of the problem. In the specific case of soft-science disciplines, scientists considered that low replicability was caused by the use of questionable research practices, such as inaccurate description of the methods and procedures employed (Begley 2013; Fiedler and Schwarz 2016), inappropriate use of statistics (Begley 2013), falsification of data (Fanelli 2009; Fiedler and Schwarz 2016), stopping data collection after achieving

This is a post-print version of the following article published as “online first”:
Olalla-Soler, C. (2020). Practices and attitudes toward replication in empirical translation and interpreting studies. *Target*, 32(1). <https://doi.org/10.1075/target.18159.ola>

the desired result (Fiedler and Schwarz 2016), and deciding to exclude data after looking at the impact of doing so (Fiedler and Schwarz 2016). In a survey completed by 1,138 psychology researchers (Fiedler and Schwarz 2016), 47% of the respondents admitted to having used questionable research practices at least once.

In soft-science disciplines, scientists are not the only ones responsible for selective reporting, since academic journals may not always welcome articles derived from replications or may be reluctant to publish negative findings, both of original and replicated studies (Ferguson and Heene 2012; Baker and Penny 2016). Yeung (2017) reviewed 465 high-impact, peer-reviewed neuroscience journals and found that only 6% explicitly stated that they accepted replications, 84.7% did not state their position on replications, 8.6% implicitly discouraged replications by emphasizing the novelty of manuscripts, and 0.7% explicitly stated that they rejected replications. In psychology, the acceptance of replications was stated as 4% in a sample of 1,151 journals (Martin and Clarke 2017).

The low replicability of the research carried out in soft-science disciplines has encouraged scientists to explore mechanisms to enhance it. Some researchers propose using more robust designs, making better use of statistics, conducting meta-analyses, and collaborating with funders and publishers to increase the number of replications (Baker and Penny 2016). Wagenmakers et al. (2012) consider that pre-registering empirical studies (i.e., a public date-time stamped registration in an institutional registration system such as Open Science Framework of the hypotheses, research design, and data analysis plans before conducting the study) would be a useful procedure to separate exploratory from confirmatory science.

Not all scientists of soft-science disciplines consider replicability of science essential (e.g., Fanelli 2018). Many believe that science is self-correcting (Crandall and Sherman 2016; Jussim et al. 2016). However, there is no evidence that science is engaged in self-correcting processes because changes in a specific discipline may be due to many factors (Pashler and Harris 2012), including the satisfactory answer of research questions, the development of new techniques, or the emergence of new research areas.

2.2 Replication in translation and interpreting studies

TIS is considered an interdisciplinary area of research (Saldanha and O'Brien 2014), in which different research methods and approaches meet. In TIS, Hale and Napier (2013) distinguish between positivistic and phenomenological research philosophies. While phenomenological approaches are described as humanistic, interpretative and subjective, positivistic approaches are described as objectivist, empirical, experimental or observational (such as corpus research), and scientific (Hale and Napier 2013). According to this dichotomy, phenomenological

approaches in TIS constitute the core philosophy of all research that is not based on the scientific method and does not meet scientific criteria such as objectivity, reliability, replicability, validity and generalizability (Neunzig and Tanqueiro 2007). TIS research based on positivistic approaches is epistemologically close to soft-science disciplines, even if the degree of ‘hardness’ or ‘softness’ varies from one study to another due to methodological reasons and decisions. While some TIS research areas could be generally associated with phenomenological approaches (such as research into translation and interpreting history or translation criticism) and others with positivistic approaches (such as interpreting or translation process research), almost all research areas could be investigated from both a positivistic and a phenomenological perspective. For example, research areas such as the development of translation or interpreting competence, the use of technologies when translating or interpreting, or the reception of translations or interpreting output could be investigated from a subjective and interpretative approach, or they could be investigated using experimental or observational designs that meet scientific criteria. While replication might not be applicable (or relevant) to research based on a phenomenological philosophy, it is applicable (and possibly relevant) to methods based on a positivistic research philosophy.

In TIS, positivistic and phenomenological research have not run in parallel. During the emergence of modern translation studies in the mid-twentieth century (Snell-Hornby 2006), research based on a phenomenological philosophy of science was predominant, and theorization played a paramount role (Tymoczko 2005).² With the development of descriptive translation studies (Holmes [1972] 1988) and more specifically during the so-called empirical turn in the 1990s (Snell-Hornby 2006), the interdisciplinarity of TIS grew as methods and instruments from other disciplines were borrowed. This is especially the case for translation process research and cognitive translation studies,³ which drew from soft sciences mainly based on a positivistic philosophy of science, such as psychology, cognitive science and neuroscience (Muñoz Martín 2010; O’Brien 2013). Currently, despite the growth in research based on a positivistic approach in different fields such as interpreting (Liu 2011), cognition (Muñoz Martín 2014), specialized translation (Way 2014), machine translation (Gupta and Dhawan 2019) and audiovisual translation (Orero et al. 2018), among others, phenomenological approaches still

² In the case of interpreting studies, in the 1960s and early 1970s, experimental research into interpreting was already being conducted by a few psychologists and psycholinguists (Gile 2000).

³ While some authors employ both terms as synonyms, others consider that they cannot be considered as such (Muñoz Martín 2014).

This is a post-print version of the following article published as “online first”:
Olalla-Soler, C. (2020). Practices and attitudes toward replication in empirical translation and interpreting studies. *Target*, 32(1). <https://doi.org/10.1075/target.18159.ola>

represent a large part of the research carried out in TIS (Zanettin, Saldanha, and Harding 2015).

Replication is regarded as a necessity for the development of TIS in most of the empirical research handbooks in the discipline (Tymoczko 2002; Neunzig and Tanqueiro 2007; Hale and Napier 2013; Saldanha and O’Brien 2014). Many authors have highlighted the need to improve replicability in empirical TIS (Li 2004), while others have identified a lack of replication and called for more replication to take place (Gile 1991; Hurtado Albir and Alves 2009; Alves, Pagano, and da Silva 2011; O’Brien 2011; House 2013). The concern for replication in empirical TIS is rather recent, due to the evolution of the discipline over time, and the contradiction between the importance attached to replication in empirical research handbooks and the lack of replication in the literature will possibly disappear as the discipline evolves and if positivistic research increases.

Against the background of the growth in research based on a positivistic approach in TIS, the importance ascribed to replication in empirical research handbooks, and the lack of data on replication practices and attitudes of researchers conducting empirical studies in the field of TIS, this article aims to provide an overview of this subject to provide evidence-based arguments for initiating a debate about the need for replication in empirical translation and interpreting studies and its implications for the development of the discipline.

3. Materials and methods

3.1 Study 1: Practices in and attitudes toward replication in empirical research on translation and interpreting

We conducted an online survey to investigate practices in and attitudes toward replication in empirical research on translation and interpreting. This section presents the nature and design of the survey, describes the characteristics of the survey participants, highlights the limitations of the survey and sample size, and outlines the analysis procedure.

3.1.1. Survey characteristics and design

The online survey consisted of three parts: 1) information about the participants; 2) information about the participants’ practices regarding replication in translation and interpreting research; and 3) information about the participants’ opinions about replication in translation and interpreting research. Part 1 was preceded by a page in which the researcher introduced himself, presented the topic of the survey, and described the profile that researchers had to meet to partici-

pate. This profile was employed as sample-selection criteria. The criteria were that the participant 1) was or had been involved in empirical research in translation or interpretation; 2) had published results from at least one empirical study in a peer-reviewed journal; and 3) was affiliated with a tertiary institution. The three parts of the survey included single-answer and multiple-choice questions, Likert scale questions, and open-ended questions. The survey was offered in two languages (English and Spanish; see Supplementary Material 1 for the English version of the survey and Supplementary Material 2 for the Spanish version), and they were accessible from 17 January to 1 March 2018.

The link to the survey and information about this research were announced in three ways: 1) it was tweeted from the author’s Twitter account approximately twice a month; 2) it was distributed through the mailing lists of several university departments, the mailing list of the Translation, Research Empiricism & Cognition (TREC) network, the newsletter of the European Society for Translation Studies (EST), and the LinguistList newsletter; and 3) the Bibliography of Interpreting and Translation (BITRA) (Franco Aixelá 2001–2019) was consulted to identify researchers who had conducted empirical research in translation and interpreting, and these researchers were contacted via email and invited to fill out the survey and to forward it to their colleagues.

The first part of the questionnaire consisted of eleven questions to profile the participants. Information was gathered on the participants’ age range, academic position, experience in empirical research in translation and interpreting, experience as reviewers of empirical research in translation and interpreting, research areas in translation and interpreting, most frequently used research methods, and previous training in statistics and empirical research methods.

The second part contained thirteen questions, designed to solicit information about the participants’ replication practices. Participants were firstly asked whether they had ever been contacted by another researcher who wanted to replicate one of their empirical studies. The next question was whether they had ever tried to replicate an empirical study originally carried out by themselves or by another researcher. Depending on the answer to this question, the questions in this part varied. If the participant answered ‘no’, he/she was asked to select from a list the reasons why he/she had not replicated empirical studies. If the researcher answered ‘yes’, the following information was gathered: the number of replicated studies, the number of self-replications, the number of exact, constructive, and conceptual replications, the number of successful replications, the number of published replications, and the number of replications that were explicitly presented as such in the abstract or body of the publication. Participants were also asked whether they had had dif-

faculties publishing replicated studies and what difficulties they faced. In most of these questions, a ‘no response’ or ‘I don’t remember’ option was included.

The third part contained thirteen questions focused on the participants’ opinions and attitudes regarding replication in translation and interpreting research. Participants who had replicated empirical studies were asked whether they would recommend that other researchers replicate, and those who had not were asked whether they were interested in doing so. They were all asked to guess what ratio of all the empirical studies in TIS that they knew of had been replicated and what ratio they considered had confirmed the results of the original trial.⁴ All participants were also asked about their perception of the importance of replication in empirical research on translation and interpreting, their perception of the frequency of occurrence and impact of several questionable research practices, and their perception of the applicability of several mechanisms to enhance replicability and the likelihood of them adopting those mechanisms in their research.

The third part was followed by a page on which participants could comment about the topic of the survey. After submitting the responses, the survey displayed a reference list of the cited works therein.

A first draft of the survey was assessed by Dr Ricardo Muñoz Martín, who has extensive knowledge of empirical research in translation. His suggestions for improvement were included. The survey was proofread by a native English speaker, and after that, the online survey was created with LimeSurvey. It was then tested by four research colleagues at the Department of Translation and Interpretation and East Asian Studies of the Universitat Autònoma de Barcelona who had experience in empirical research in translation and interpreting. The survey was modified according to their suggestions and was then translated into Spanish.

3.1.2 Participants

Fifty-seven researchers completed the survey. The responses of five researchers were discarded. In four cases, the researchers had not published any empirical studies. In one case, the data provided were inconsistent: in some questions the participant mentioned that she/he had replicated one study, and in other questions she/he indicated that she/he was the author of many replications. In all, 73% of the respondents conducted research in translation, 13.5% in interpreting, and 13.5% in both. Figure 1 presents the participants’ research areas in translation, while Figure 2 shows the participants’ research areas in interpreting. Each participant could select more than one area.

The participants frequently combined quantitative and qualitative methods in mixed methods studies (24.3% of respondents). Quasi-experimental designs were

⁴ Due to space constraints, these data are not provided in this article. However, they can be found in Supplementary Material 3.

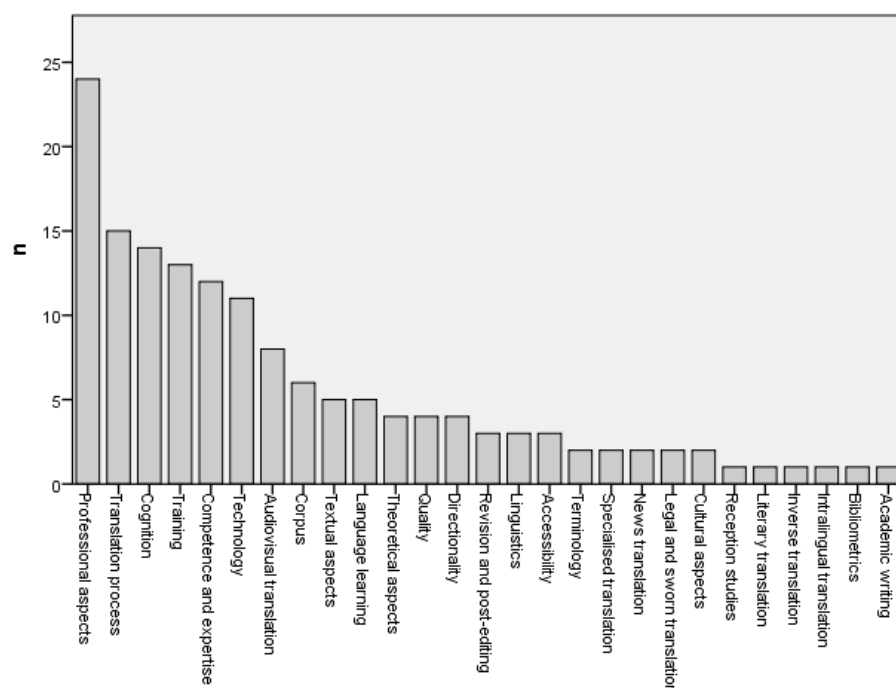


Figure 1. Participants’ research areas in translation

more frequently used than case studies (21.6% and 20.7% respectively). Observational methods were used by 17.4% of the participants. Experimental designs were the least frequently used method (16%). In addition to these methods, one respondent (1.9%) employed bibliometric analysis.

Respondents were asked whether they had received training (no training, basic training or advanced training) in statistics and in empirical research methods. While 44.2% of the respondents had received advanced training in empirical research methods, only 13.5% had received training in statistics at the same level. The percentage of respondents who had no training in statistics (21.2%) was higher than the percentage of respondents with no training in empirical research methods (13.5%).

Researchers from nineteen countries filled out the survey: Spain (32.7%); Finland (9.6%); United States (9.6%); Brazil, Poland and Switzerland (5.8% each); Belgium, Hungary, and Ireland (3.8% each); and Australia, Austria, Chile, China, Denmark, Hong Kong, Italy, Norway, South Korea, and United Kingdom (1.9% each). The age of 36.5% of participants ranged from 37 to 46 years, 30.8% from 47 to 56 years, 26.9% from 27 to 36 years, 3.8% from 57 to 66 years, and 1.9% over 67 years. In all, 46.2% of the respondents were assistant/associate professors (or equivalent in their respective countries), 36.5% were full professors, 9.6% were

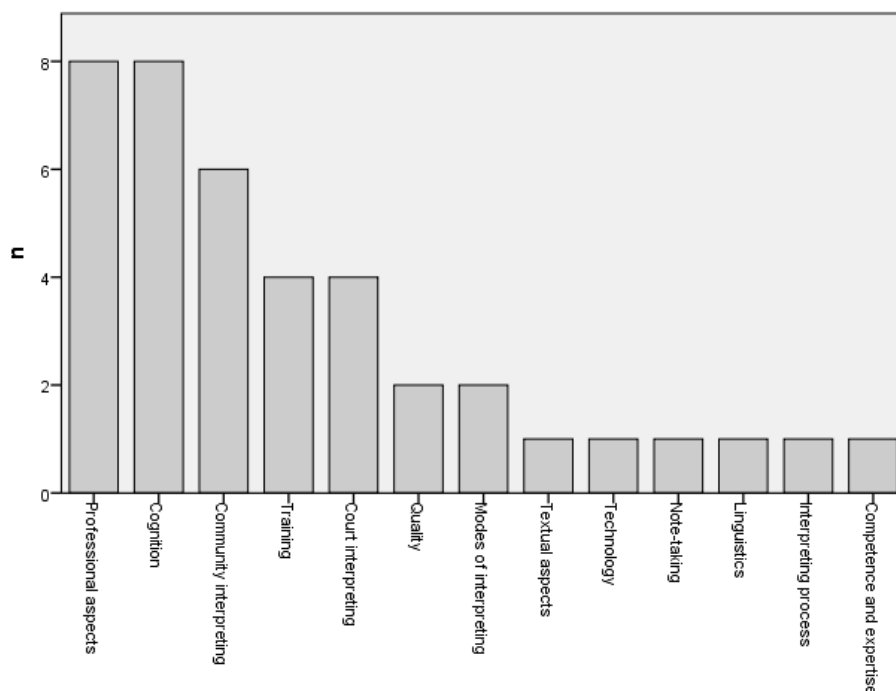


Figure 2. Participants’ research areas in interpreting

post-doctoral researchers, 5.8% were PhD candidates, and one participant (1.9%) indicated that he/she held a different position.

In all, 94.2% of the respondents were regularly involved with empirical research. For 38.5% of the respondents, papers containing results from empirical studies represented nearly all their published papers, for 28.8%, up to one quarter of their papers, for 23.1%, half of their papers, and for 9.6%, three out of four papers. Of the respondents, 78.8% had reviewed articles for peer-reviewed journals at least once. For 34.1% of them, half of the reviewed papers contained results from empirical studies, for 31.7%, nearly all the papers, for 22.0%, up to a quarter of the papers, and for 12.2%, three out of four papers. Half of the respondents were members of the editorial board of at least one peer-reviewed journal, and 51.9% had taught college-level research methods and/or statistics classes.

3.1.3 Limitations

The completion rate of the survey was 61%: of ninety-four researchers who started to fill it out, fifty-seven completed the whole survey. Despite the efforts made to disseminate the survey, a small sample of researchers participated in this research. Another limitation is the low number of participants who conduct research in interpreting. While the results are not presented separately for researchers in

translation and interpreting, these two limitations do not allow for generalizing the results obtained with this instrument to the population of researchers who have carried out empirical research in translation and interpreting (and more specifically, to those carrying out research in interpreting only).⁵

Once the survey was launched and participants had already started to fill it out, three researchers contacted the author to comment that they had experienced difficulties in answering some questions. Two of them thought the survey included too many concepts with which they were not familiar, and one considered that the survey was too oriented toward quantitative research and did not fit the reality of qualitative research.

Another factor that could have caused the low completion rate was that even though the description of the research project clearly stated that both researchers who had replicated empirical studies and those who had not were welcomed to fill out the survey, the topic of the study was in all likelihood less appealing to researchers who had not replicated studies, and only those who had some interest in replication are likely to have been willing to share their views on this matter. Hence, it is possible that a bias toward a positive view of replication in translation and interpreting research exists in the data: even if the number of participants who had not replicated empirical studies was slightly larger than the number of participants who had, the attitudes toward replication of both groups are generally positive (see Section 4.1.1).

3.1.4 Analysis

Given the low number of participants, a descriptive analysis in which frequencies and percentages were computed was considered to be a more adequate type of analysis than an inferential one. For some survey questions, percentages were calculated separately for researchers who had replicated studies and for those who had not. Chi-squared tests for homogeneity with Yates correction and Fisher’s exact tests were computed to determine whether there were any differences in the distribution of the answers between the two groups.⁶ Cramér’s *V* was computed

⁵ Although we consider this a limitation, according to BITRA (Franco Aixelá 2001–2019), research in interpreting accounts for 7,789 documents (10.4%) out of the 75,000 that are indexed. In our survey, participants who conduct research only in interpreting represent 13.5% of the sample. Even if the sample size for this study had been larger, the number of participants conducting empirical research in translation would still have been higher than that of researchers conducting research in interpreting.

⁶ While Mellinger and Hanson (2017) recommend running Fisher’s exact test instead of the chi-squared test in case of limited sample sizes, Fisher’s exact test is generally used for dichotomous categorical data (Everitt 1998; Pardo and San Martín 2012). However, both tests were run and are reported.

as an effect size measure for the chi-squared tests. However, the statistical results may not be generalizable to the whole population of researchers in empirical TIS given the limitations and the possible bias mentioned before.

Content analysis (Hsieh and Shannon 2005) was used to identify and quantify themes and patterns that emerged from the participants’ open-ended responses. In two of the survey questions, Likert scales were used, which measured the perception of occurrence and impact of questionable research practices and the perception of the applicability and likelihood of adoption of mechanisms to enhance replicability. In one of the questions, five-point Likert scales were employed (never, rarely, sometimes, often, always; and not at all negative, slightly negative, moderately negative, very negative, extremely negative). In the other question, a three-point Likert scale was used (not applicable, somewhat applicable, totally applicable), and also a four-point scale (extremely unlikely, unlikely, likely, extremely likely). Each point was converted into a consecutive ordinal value. Hence, the five-point scales were converted to 1-2-3-4-5, the three-point scale was converted to 1-2-3, and the four-point scale to 1-2-3-4. Means and standard deviations were calculated for each scale.

When a question included a ‘no response’/‘I don’t remember’/‘I don’t know’ answer, these responses were treated as missing data. The data gathered with this survey can be found in Supplementary Material 3.

3.2 Study 2: Bibliometric analysis of replications in BITRA

To complement the survey results, a bibliometric approach was used to identify replications in the TIS literature by carrying out a conceptual replication of Makel, Plucker, and Hegarty’s (2012) study on replication in psychology research. While the original study was conducted by analyzing only 100 journals with the highest 5-year impact factor in the field of psychology, in the present study, the bibliometric analysis was carried out using BITRA (Franco Aixelá 2001–2019). The database was consulted on 12 and 13 November 2018, and at that time, it contained 74,687 entries, of which 38,551 (51.6%) included the publication’s abstract. The coverage of this database of the literature produced in TIS is optimal for bibliometric analysis, both quantitatively and qualitatively. From a quantitative perspective, the more than 70,000 entries represent the majority of publications devoted to translation and interpreting from the beginning of the discipline to the present. From a qualitative perspective, the database contains not only journal articles, but also books, book chapters and PhD theses in sixty-six different languages. The ten most frequent languages are (from most to least frequent) English, Spanish, French, German, Portuguese, Italian, Chinese, Catalan, Gali-

This is a post-print version of the following article published as “online first”:
Olalla-Soler, C. (2020). Practices and attitudes toward replication in empirical translation and interpreting studies. *Target*, 32(1). <https://doi.org/10.1075/target.18159.ola>

cian and Polish. Publications in these languages represent 96.8% of the documents indexed in BITRA.

While in Makel, Plucker, and Hegarty’s (2012) study only ‘replicat*’ was used as a keyword to identify replications, we decided to use more keywords that could retrieve results in at least the most frequent languages in BITRA. Hence, to identify replications the abstracts and titles of the documents indexed in BITRA were consulted with the following keywords: ‘replica’, ‘reprodu’, ‘replizier’, ‘wiederhol’, ‘repet’, ‘riprodu’, and ‘replika’. With these keywords, results from at least nine out of the ten most frequent languages were retrieved. In the case of Chinese (which represents 1.4% of the documents indexed in BITRA), only those documents that also included the abstract or the title in English could be analyzed.

The results retrieved with each keyword were read, and the documents that were replications were identified. This was not a straightforward process, because, for example, replication was not always used in the sense of repeating an empirical study, but as a synonym of ‘translation’. Once the list of replication studies was completed, the text of each document was consulted, and the following characteristics of the replication were identified: the type of replication, the relationship of the author with the authorship of the original study, and the success of the replication of obtaining the same results as those of the original study. The list of publications that contain results from empirical studies can be found in the Supplementary Material 4.

While BITRA is an optimal database for bibliometric analysis, only half of the indexed documents include the abstract of the publication. This is an important limitation that has to be taken into account when interpreting the results. Another limitation is that we could only identify documents that explicitly stated in their titles or their abstracts that they were replications. Hence, there could be documents that were replications but were presented as such in the body of the document only, or they could have been presented as ‘based on’ someone’s study. We are aware of this limitation, and we consider the results presented in this article exploratory. Confirmatory results need a further investigation in which documents of a limited timespan would be examined one by one to identify replications.

3.3 Study 3: Analysis of the acceptance of replications in academic journals devoted to TIS

To expand and complement the results obtained with the survey and the bibliometric analysis of BITRA, a conceptual replication of Yeung’s (2017) study on the acceptance of replications in neuroscience journals was carried out by analyzing 131 translation and interpreting journals. The list of journals was extracted from

This is a post-print version of the following article published as “online first”:
Olalla-Soler, C. (2020). Practices and attitudes toward replication in empirical translation and interpreting studies. *Target*, 32(1). <https://doi.org/10.1075/target.18159.ola>

RETI (Translation and Interpreting Studies Journals) (Biblioteca d’Humanitats UAB 2013–2019), and it was limited to the journals that are also indexed in BITRA. RETI was consulted on 14 November 2018.

Following Yeung’s (2017) procedure, each journal website was accessed and the aims and scope and instructions to authors were screened to assess if they 1) explicitly stated that they accepted replications, 2) did not state their position on replications, 3) implicitly discouraged replications by emphasizing the novelty of manuscripts, or 4) explicitly stated that they rejected replications. Each journal was also classified according to its activity status (active or inactive), its access type (open or paid access), use of peer review, and indexation in the Social Science Citation Index, in Scopus or in the Scimago Journal Rank (the 2017 edition was consulted in all cases). The data can be found in Supplementary Material 5.

A limitation of this procedure is that the difference between not stating a position on replication and implicitly discouraging replications by emphasizing the novelty of manuscripts is subtle in some cases, and the category assigned to them may have been rather subjective. While categories 1 and 4 are straightforward to identify, the results for categories 2 and 3 are not to be considered completely objective.

4. Results

4.1 Study 1: Practices and attitudes toward replication in empirical research on translation and interpreting

4.1.1 *Replication practices in empirical research in translation and interpreting*

Of the fifty-two survey participants, twenty-four (46.2%) had tried to replicate an empirical study originally carried out by themselves or by somebody else. These twenty-four researchers had replicated a total of forty-seven empirical studies. The mean number of replicated studies by each researcher was 1.96 ($SD = 1.12$). The maximum number of replications carried out by a single researcher was five.

Table 1 describes the forty-seven replications carried out by the twenty-four participants in terms of authorship (the relationship of the author of the replication to the author of the original study), replication type, the results reached and the way these results were disseminated.

Table 1. Characteristics of the forty-seven replications in empirical research on translation and interpreting identified in the survey

Characteristics of the replications identified in the survey	n	%
Relationship of author of the replications to author of the original studies		
Number of self-replications	18	38.3
Number of replications of studies originally conducted by others	29	61.7
Types of replications		
Number of exact replications	4	8.5
Number of constructive replications	22	46.8
Number of conceptual replications	21	44.7
Results reached in the replications		
Number of replications in which the same results as in the original studies were obtained	11	23.4
Number of replications in which the same conclusions as in the original studies were reached	14	29.8
Dissemination of the replications		
Number of researchers who tried to publish the results of their replications	20	83.3
Number of researchers who had difficulty publishing the results because it was a replication	2	10
Number of replications presented at conferences or meetings	30	63.8
Number of replications published in peer-reviewed journals	27	57.4
Number of replications published in non-peer-reviewed journals	0	0.0
Number of replications published in books or book chapters	8	17.0
Number of non-published replications	11	23.4
Number of replications that were explicitly presented as a replication in the abstract or in the body of the article/chapter/book	19	40.4

4.1.2 Reasons not to replicate empirical research in translation and interpreting

In all, twenty-eight (53.8%) of the participants had not tried to replicate an original study. Some of them did not have any conscious reason, but some of them did. Table 2 summarizes the reasons not to replicate that these participants gave.⁷

⁷ Participants could select more than one option except for ‘I did not have any conscious reason’. When selecting this one, all other options were automatically blocked.

Table 2. Reasons not to replicate empirical studies in TIS

Reasons not to replicate in TIS	n	%
Participants who did not have any conscious reason	11	39.3
Participants who had a conscious reason	17	60.7
Replicating an empirical study will not produce the same academic impact as conducting an original empirical study	8	47.1
Editors and publishers are not interested in replicated studies	5	29.4
It is too time-consuming	4	23.5
I am concentrated on an original line of research and have no time/interest/wish to replicate others	4	23.5
It is not possible to replicate all conditions of the original study	3	17.6
Unsatisfactory replication could question my original findings	2	11.8
It is difficult to find all necessary information about the original research design	1	5.9
It is not relevant to TIS	1	5.9
Replication is not totally applicable within a qualitative framework	1	5.9

4.1.3 Attitudes toward replication

Nearly all participants who had replicated studies recommended that others replicate (Table 3). These figures are a bit different for participants who had not replicated studies: in this group, less than three out of four would be willing to replicate empirical studies. Despite this difference, the reasons why participants would recommend others replicate or would be willing to replicate are similar.⁸

Table 3. Attitudes toward replication held by the twenty-four participants who had replicated studies and the twenty-eight who had not

Attitudes of participants who had replicated (n = 24)	n	%
Researchers that would recommend others replicate	23	95.8
Their own empirical studies	0	0.0
Those of other researchers	3	13.0
Both their own and those of other researchers	20	87.0
Reasons to replicate		
To help consolidate previous results	5	21.7
To verify previous results	5	21.7
To compare results	3	13.0
To expand previous results	2	8.7
To establish substantial, scientific knowledge	2	8.7
To generalize previous results	1	4.3

⁸ Note: the reasons (not) to replicate were identified using content analysis in the participants' responses to the open-ended question.

To connect results from different studies	1	4.3
To increase the sample size of previous studies	1	4.3
To reuse previously validated instruments	1	4.3
To strengthen the discipline	1	4.3
To learn about research methods	1	4.3
Reasons not to replicate		
Reviewers negatively evaluate replications	1	100.0
Attitudes of participants who had not replicated (n = 28)	n	%
Researchers that would be willing to replicate	20	71.4
Their own empirical studies	8	40.0
Those of other researchers	0	0.0
Both their own and those of other researchers	12	60.0
Reasons to replicate		
To strengthen the discipline	5	25.0
To verify previous results	5	20.0
To expand previous results	3	15.0
To generalize previous results	1	5.0
To compare results	1	5.0
To establish substantial, scientific knowledge	1	5.0
Reasons not to replicate		
It is seen as uninteresting	7	87.5
It is not a priority at that time in TIS	4	50.0
It does not build academic reputation	2	25.0
It could question the original researcher and his/her findings	2	25.0
It is valuable but difficult to carry out	1	12.5

Both researchers who had replicated empirical studies and those who had not tend to consider replication as a necessity in the discipline (Table 4). This tendency seems to be stronger in the case of the participants who had replicated, although it is not statistically significant: $\chi^2(3) = 5.93$, $p = 0.07$, $V = 0.35$; p value of the Fisher's exact test = 0.75.

Table 4. Answers to the question ‘Do we need replication in empirical studies in translation and interpretation?’

Do we need replication in empirical TIS?	Participants who had replicated studies (n = 24)	Participants who had not replicated studies (n = 28)
Yes, absolutely	22 (91.7%)	17 (60.7%)
Yes, but it is not as important as in other disciplines	2 (8.3%)	6 (21.4%)
No, there are other aspects that need to be addressed first	0 (0.0%)	2 (7.1%)
No, not at all	0 (0.0%)	1 (3.6%)

4.1.4 Questionable research practices that hinder replication

Table 5 summarizes the questionable research practices that the participants consider occur most frequently in empirical TIS and how negative their effects are. It must be noted that the average of no responses to these questions was 19.8% ($SD = 8.6$).

Table 5. Occurrence and impact of questionable research practices in translation and interpreting research

Questionable research practices in TIS	Occurrence		Impact	
	Mean (min. = 1; max = 5)	SD	Mean (min. = 1; max = 5)	SD
Selectively reporting studies regarding a specific finding that ‘worked’	3.3	0.7	3.3	1.0
Assuming that a detected effect is broader and more generalizable than it actually is	3.3	0.7	3.4	0.8
Failing to report all dependent measurements that are relevant for a finding	3.2	0.7	3.5	0.9
Overestimating the size of the detected effect	3.2	0.7	3.5	0.9
Claiming conclusive research findings solely on the basis of exploratory studies	3.1	0.9	3.6	0.9
Failing to report all conditions that are relevant for a finding	3.1	0.8	3.7	0.9
Rounding off p values (e.g., reporting a p value of .054 as .05)	3.0	0.9	2.8	1.1
Deciding whether to exclude data after looking at the impact of doing so regarding a specific finding	2.9	0.7	3.7	0.9
Collecting more data after checking that results were non-significant	2.9	0.6	2.6	1.3
Claiming to have predicted an unexpected result	2.8	1.0	2.7	1.2
Stopping data collection after achieving the desired result concerning a specific finding	2.7	0.8	3.3	1.2
Claiming that results are unaffected by demographic variables (e.g., gender) although one is actually not sure (or knows that they are)	2.6	0.8	3.8	1.1
Detecting some empirical effect when in fact no such effect exists in reality	2.6	0.7	4.2	0.9
Falsifying data	1.9	0.5	4.8	0.6

Since not being able to replicate an empirical study may not only be due to questionable research practices, in an open-ended question in the survey, thirty-five participants commented on other factors specific to empirical research in

translation and interpreting that may make it difficult to replicate empirical studies (Table 6).⁹

Table 6. Other factors specific to empirical research in translation and interpreting that may hinder replication

Other factors specific to empirical research in translation and interpreting that may hinder replication (n = 35)	n	%
Lack of detailed information related to the research design in the original study	8	22.9
Difficulties in getting comparable samples	6	17.1
Differences related to the context in which the replication is taking place	5	14.3
Difficulties in controlling variables	5	14.3
Difficulties in getting a properly sized sample	4	11.4
Difficulties related to the complexity of the element under study	3	8.6
Difficulties in getting access to data	3	8.6
Lack of training in empirical research methods	3	8.6
Difficulties related to data-collection instruments (misuse, malfunctions or obsolescence)	3	8.6
Lack of resources	2	5.7
Difficulties in assigning subjects randomly to groups	1	2.9
Lack of validated instruments to collect data	1	2.9
Misuse of statistical procedures	1	2.9

4.1.5 Enhancing replicability in translation and interpreting research

Table 7 summarizes the participants’ responses regarding the applicability of several mechanisms to enhance replicability and the likelihood of them adopting these mechanisms. It must be noted that the average of no responses to these questions was 9.7% ($SD = 6.5$).

Table 7. Applicability and likelihood of adoption of mechanisms to enhance replicability

Mechanisms to enhance replicability	Applicability		Likelihood of adoption	
	Mean (min. = 1; max = 3)	<i>SD</i>	Mean (min. = 1; max = 4)	<i>SD</i>
Designing more robust empirical studies	2.7	0.5	3.3	0.6
Clearly differentiating between exploratory and confirmatory studies	2.7	0.5	3.2	0.6
Reporting all decisions concerning data collection and data analysis	2.7	0.5	3.2	0.8
Making better use of statistics	2.6	0.5	3.2	0.7
Training PhD students to replicate empirical studies	2.6	0.6	2.8	0.9
Conducting studies with larger samples	2.5	0.5	2.9	0.8

⁹ These categories were established using content analysis. The participants’ responses may be classified in more than a single category according to their explanations provided in the open-ended question.

Uploading open data, materials, and workflow of the empirical study to repositories	2.4	0.6	2.6	0.9
Providing additional materials to those published in academic journals	2.4	0.6	2.6	0.7
Reporting effect sizes for all statistical tests	2.4	0.7	2.9	0.8
Engaging financing bodies and editors to raise awareness of the importance of replication	2.2	0.6	2.1	0.8
Pre-registering empirical studies	1.9	0.7	2.0	0.8

4.2 Study 2: Bibliometric analysis of replications in BITRA

Of 38,551 documents in BITRA that include an abstract (out of 74,687 entries; 51.6%), twenty-three publications of empirical studies (0.06%) were found that explicitly stated in the abstract or in the title that they were replications. Table 8 summarizes the results of an in-depth analysis of the twenty-three published replications in terms of the year of publication, publication type, research area, replication type, authorship, and the results obtained in the replication.

Table 8. Analysis of the twenty-three replications indexed in BITRA (November 2018)

Replications indexed in BITRA (November 2018)	n = 23	%
Year of publication		
1993-1996	2	8.7
1997-2000	1	4.3
2001-2004	1	4.3
2005-2008	4	17.4
2009-2012	6	26.1
2013-2016	9	39.1
Publication type		
Journal article	15	65.2
Book chapter/book	4	17.4
PhD thesis	4	17.4
Research area		
Translation	12	52.2
Interpreting	11	47.8
Replication type		
Exact	2	8.7
Constructive	9	39.1
Conceptual	9	39.1
No data available	3	13.0
Authorship		

This is a post-print version of the following article published as “online first”:
 Olalla-Soler, C. (2020). Practices and attitudes toward replication in empirical translation and interpreting studies. *Target*, 32(1). <https://doi.org/10.1075/target.18159.ola>

Self-replication	7	30.4
Replication of another’s study	15	65.2
No data available	1	4.3
Replication of the original results		
Yes	13	56.5
No	2	8.7
Partially	5	21.7
No data available	3	8.7

4.3 Study 3: Analysis of the acceptance of replications in academic journals devoted to TIS

Of the 131 journals reviewed, only one (0.8%; *Journal of Interpretation*) explicitly stated that replications were accepted, 108 (82.4%) did not state their position on replications, thirteen (9.9%) implicitly discouraged replications by emphasizing the novelty of manuscripts, none (0%) explicitly stated that replications were rejected, and in nine cases (6.9%), it was not possible to access this information.

A series of comparisons was performed between journals that were grouped according to the following factors: 1) whether the journal had published at least one issue during the two previous years before this study was conducted (active journal) or if it had stopped publishing new issues during that period (inactive journal); 2) whether it was an open-access or a paid-access journal; 3) whether the journal was peer-reviewed; and 4) whether the journal was indexed in the Social Sciences Citation Index, in Scopus, or in the Scimago Journal Rank (2017 edition in all cases). Results (Table 9) show no relevant differences regarding the acceptance of replications in relation to the distribution of journals in the different categories.

Table 9. Analysis by factor of the acceptance of replications in journals (November 2018)

	Explicitly accepts replications	No position stated	Discourages replications	Rejects replications	No data available
Active/inactive journals					
Active	1 (1.1%)	95 (83.3%)	13 (11.4%)	0 (0%)	5 (4.4%)
Inactive	0 (0%)	13 (81.3%)	0 (0%)	0 (0%)	3 (18.8%)
Access type					
Open access	1 (1.1%)	80 (87%)	9 (9.8%)	0 (0%)	3 (2.2%)
Paid access	0 (0%)	28 (71.8%)	4 (10.3%)	0 (0%)	7 (17.9%)
Use of peer review					
Peer-review	1 (0.9%)	98 (83.8%)	13 (11.1%)	0 (0%)	5 (4.2%)
Non-peer-review	0 (0%)	10 (71.4%)	0 (0%)	0 (0%)	4 (28.6%)
Indexation (in Social Sciences Citation Index, in Scopus, or in the Scimago Journal Rank)					
Indexed	0 (0%)	33 (86.8%)	5 (13.2%)	0 (0%)	0 (0%)
Non-indexed	1 (1.1%)	75 (80.6%)	8 (8.6%)	0 (0%)	9 (9.7%)

5. Discussion

The discussion section is organized in sub-sections that correspond to each of the research questions posed in Section 1. When applicable, results are compared to those obtained in soft-science disciplines.

5.1 The need for replication in empirical TIS

While the survey participants who had replicated studies expressed a stronger opinion regarding the need for replication in empirical TIS than those who had not, the percentage who considered it a necessity were 100% and 92.1%, respectively. According to the exploratory bibliographic study we conducted (Table 8), the number of replications is on the increase, which could also indicate a growth in interest in replication in empirical TIS.

According to the participants, replication would help strengthen the discipline and verify and expand previous results. Thus, the benefits of replication for empirical TIS are aligned with the general benefits of replication for science (Popper 1959; Kuhn 1962). However, one out of four participants was not interested in replication since it was considered uninteresting and not a priority in empirical TIS. A possible way to increase the interest in replication among the empirical TIS community would be to give replication a major role in research methods training and to make such studies more visible. When interpreting these results, it should also be taken into account that research in TIS based on a positivistic philosophy of science does not represent the majority of the research car-

ried out within the discipline. As mentioned in Section 2.2, TIS research based on a phenomenological approach is abundant, and given the subjective and interpretative nature of these investigations, replication only makes sense in a very specific fraction of TIS research: the one based on a positivistic philosophy of science. Hence, we consider it relevant not to overestimate the relevance and the need of replication in TIS in general, but scholars carrying out research based on positivistic approaches should engage in a discussion about when replication is applicable and when it is relevant for the development of the discipline.

5.2 Characteristics of the existing replications in translation and interpreting studies

Although the results obtained in the study on the replications contained in BITRA have to be interpreted with caution given the exploratory nature of this study, the findings indicate that replications are conducted at a similar rate in both translation and interpreting research (52.2% and 47.8%). In all, 34.6% of the survey participants had been contacted by other researchers who wanted to replicate one of their studies. The figure obtained in the *Nature* survey (Baker and Penny 2016) was lower, at 20%. This may indicate a positive practice in empirical TIS, since researchers show interest in replicating others' work. The proportion of respondents who had tried to replicate at least one empirical study almost reaches 50%. However, as stated in Section 3.1.3, these results might be biased, since researchers who had not replicated studies or did not have a particular interest in replication might not have been interested in completing the survey.

While forty-seven replications were mentioned in the survey, only twenty-three were retrieved from BITRA. However, neither of these figures can be seen as definitive. In the case of the survey data, the number of survey participants was rather small. In the case of the bibliometric study, only half of the entries in BITRA include the abstract, and only those publications that explicitly stated in the title or the abstract that they were replications could be identified and included in this study. In psychology, a discipline in which replication is more common than in empirical TIS, the percentage of replications published in journals was 1.07% (Makel, Plucker, and Hegarty 2012).

The percentages of self-replications and of replications conducted by others are similar both in the survey results (38.3% and 61.7%, respectively) and in the replications retrieved from BITRA (30.4% and 65.2%, respectively). Thus, this might indicate that the most frequent practice in empirical TIS is to replicate other researchers' studies, which might be a positive procedure to reduce the possible bias of the original researcher. In soft-science disciplines like psychology, self-replication represents 52.9% of all replications (Makel, Plucker, and Hegarty

2012). However, in the *Nature* survey (Baker and Penny 2016) scientists from both hard and soft-science disciplines agreed that failure to find the original effects was more frequent when replicating other researchers' studies, compared to their own. While it was not possible to analyze this using our survey data, the results obtained from BITRA showed this trend, although the difference was rather small and not conclusive (71.4% of self-replications were successful, while 61.5% of the replications of other researchers' studies were successful).

Figures regarding the types of replication are also similar when comparing the results obtained from the survey and from BITRA: exact replications represented 8.5% and 8.7%, respectively; constructive replications 46.8% and 39.1%, respectively; and conceptual replications 44.7% and 39.1%, respectively. Thus, there seems to be a tendency to conduct constructive and conceptual replications in empirical TIS, rather than exact replications. This may be due to the difficulty of creating the same conditions as in an original study, as required in exact replications. Given that most of the replications are constructive and conceptual, it is not possible to determine whether the number of unsuccessful replications is due to the fact that the original finding was false or the new operationalization of the study caused the negative result. Since the purpose of exact replications is to assess the reliability of the findings, this type of replication should be carried out more frequently. However, the figures regarding exact replications are not very different from the ones observed in another soft-science discipline, psychology, where sociological factors also play a major role: of the articles published in psychology journals, 14% were exact replications, 81.9% were conceptual replications, and 4.1% were constructive replications (Makel, Plucker, and Hegarty 2012).

5.3 Problems encountered when replicating empirical studies in translation and interpreting

The main reason cited by respondents for not replicating studies was the low impact of replications in academia. In future investigations with larger samples, it would be important to determine whether this view is dependent on the position of the scholar. Young scholars could possibly be more interested (or urged to) produce new, original research to foster their career, whereas senior scholars with lower pressure to publish could be more open to include replications in their research agenda. Almost one out of four participants who had not replicated studies considered that editors and publishers were not interested in replications, while 90% of researchers who tried to publish the results of their replications did not have any difficulty in doing so. This could indicate that preconceived ideas about the acceptance of replications in TIS publishing exists, but this idea does not seem to match the reality.

Another frequent problem for replication indicated by the survey participants was the lack of detailed information about the research design in the original study. This could be partially solved by devoting more space to the methods section of publications derived from empirical studies and by uploading thorough descriptions of methods, instruments, methodological decisions, and sample selection criteria to public repositories and linking these supplementary materials to the publication.

Survey participants also mentioned methodological constraints inherently related to the nature of translation and interpreting: difficulties in getting comparable samples, differences related to research contexts, difficulties controlling variables, and difficulties in getting good-sized samples. These problems could partially be solved by encouraging more self-replications. Uploading supplementary materials to public repositories would also facilitate this process in the case of replicating other researchers' original studies. These methodological problems also raise several questions that the community of researchers in empirical TIS need to resolve: Is an exact replication achievable in our discipline? How exact should a replication be for the result to be considered reliable by the TIS community? Where does the TIS community want to establish the border between conceptual replications and new studies?

Questionable research practices seem to occur at a similar rate in empirical TIS and in psychology (e.g., Fiedler and Schwarz 2016). The most frequent questionable practice in empirical TIS according to the survey participants (selectively reporting studies regarding a specific finding that ‘worked’) was also identified as the most frequent practice both in the *Nature* survey (Baker and Penny 2016) and as the second most frequent one in Fiedler and Schwarz's (2016) survey in psychology. Falsifying data was the least frequent practice in empirical TIS, according to the survey participants. It was the second least frequent practice in the *Nature* survey (Baker and Penny 2016) and the least frequent one in Fiedler and Schwarz's (2016) survey. In psychology, 2% of scientists admit to having falsified data at least once (Fanelli 2009). Thus, this practice seems to be rare in soft sciences, and possibly also in empirical TIS.

5.4 Enhancing replicability in translation and interpreting studies

The survey participants considered that designing more robust empirical studies would be totally applicable to enhance replicability in empirical TIS. The same result was obtained in the *Nature* survey for both soft and hard-science disciplines (Baker and Penny 2016). While pre-registration is considered an effective way of enhancing replicability in many soft-science disciplines (Wagenmakers et al. 2012; Baker and Penny 2016), the survey participants considered it to be the least applic-

able mechanism in TIS, and stated it was unlikely that they would adopt it in their research.

While requiring pre-registration of empirical studies in empirical TIS to be able to publish results derived from these studies is a drastic mechanism that would, in all likelihood, not be widely accepted by the community at this point in time, uploading supplementary materials to accompany publications derived from empirical studies could be a first step toward higher transparency, and it would facilitate replication and the replicability of results. Of course, it would still be possible to employ questionable research practices, but it would allow for more efficient peer review of original studies and to stimulate their replication.

5.5 Publication bias against replications in translation and interpreting journals

Three out of four replications mentioned in the survey had been published, mainly in peer-reviewed journals (57.4%; Table 1). The results retrieved from BITRA (Table 8) show a similar trend in that results from replications were more frequently published as journal articles than as book chapters or PhD theses. This provides evidence against the existence of a publication bias against replications in translation and interpreting journals.

The results obtained from the analysis of the acceptance of replications in translation and interpreting journals likewise show no evidence of a bias against replications. None of the journals rejected replications. While 11.4% implicitly discouraged replications by emphasizing the novelty of manuscripts, a clear majority of the journals did not state their position on replications (82.4%). Only one journal explicitly accepted replications. Thus, there is no overt negative perception of replications among translation and interpreting journals. While the percentage of journals that do not state their position on this matter or implicitly discourage replications is similar to that in Yeung’s (2017) study of neuroscience journals, the percentage of journals that explicitly accept replications is larger in neuroscience (6%) and psychology journals (4%; Martin and Clarke 2017). It is possible that translation and interpreting journals receive such low numbers of replicated studies for review, that explicit statements about replications are not felt to be necessary. However, if the TIS community considers that replications should increase, it would be advisable to explicitly mention that replications are accepted so that researchers would feel more confident of publishing outlets interested in publishing the results derived from their replications.

A third finding that counters the existence of any bias against publishing replications is the low number of survey participants who encountered difficulties publishing results derived from replicated studies (only two respondents). How-

ever, the perception that editors and publishers would not be interested in replicated studies was a rather common reason cited for decisions not to carry out replications (29.4% of the respondents).

The results obtained can also be explained by the nature and evolution of TIS research. As discussed in Section 2.2, much of the research conducted in modern TIS is based on a phenomenological philosophy of science. Positivistic approaches appeared in TIS research only comparatively recently, and, despite its rapid growth in recent years, research based on a phenomenological approach remains abundant. Consequently, academic journals that publish research in TIS reflect this historical reality: journals that began to be published in the early development stages of the discipline (such as *Meta* or *Babel*, among others) adapted their aims and scope to the reality of the discipline at that time. In the case of journals founded in more recent developmental stages, journals have also adapted to the current characteristics of TIS research.

Irrespective of the historical development, replication is neither applicable nor relevant to much of the research published in TIS journals. Possibly, if specific academic journals dedicated to TIS research exclusively based on a positivistic approach were founded, replication would acquire much greater relevance than in other journals. For these reasons, we do not consider that there is a negative bias toward replication in academic journals devoted to TIS, but neither can we consider that the bias is positive. From the historical trajectory of the discipline and from what has been observed in the studies presented in this article, it may be concluded that there is, instead, an absence of reference to replication. In the future, if the community of researchers in empirical TIS considers replication a relevant aspect for the discipline, it would be advisable for academic journals to state explicitly whether they accept articles derived from replications.

6. Conclusions

Reaching a definitive conclusion on the relevance and need for replication in empirical TIS is not possible given the limitations of this study, although the observed trend is that in empirical TIS a certain relevance is attributed to replication and there has been an increase in replications in the last decade. Furthermore, the answer to questions regarding replication cannot (and should not) be provided by a single researcher: it is the empirical TIS community that should reach a consensus on the role replication should play in empirical TIS, and this can only be achieved by engaging the community in a debate based on evidence-based arguments. Providing some evidence to initiate this debate is the humble contribution that this article hopes to make.

This is a post-print version of the following article published as “online first”:
Olalla-Soler, C. (2020). Practices and attitudes toward replication in empirical translation and interpreting studies. *Target*, 32(1). <https://doi.org/10.1075/target.18159.ola>

Bearing in mind the limitations of the present studies, this article has shown some possible similarities between empirical TIS and other soft-science disciplines such as psychology or neuroscience (low publication rate of replications, low percentage of exact replications, similar problems encountered when replicating empirical studies, and similar occurrence of questionable research practices). While these similarities are inconclusive, there are clear suggestions of parallels between empirical TIS and other soft-science disciplines regarding attitudes toward and practices in replication.

If the community of researchers in empirical TIS consider replication to be necessary, this article could serve as a point of departure for future research devoted to practices and attitudes toward replication or to other theoretical and methodological aspects of replication. The instruments and procedures employed for the present study could be improved, modified, and expanded to obtain more refined results, and new methods and instruments could be designed to gather new and comparable data. Concretely, the study of the replications indexed in BITRA could be carried out for a shorter timespan, and a more precise procedure could be used in which each publication is individually assessed to determine whether it is a replication and, if so, what its characteristics are. If the TIS community is willing to foster a debate on the need for replication in empirical TIS based on evidence-based arguments, more research across diverse contexts and using diverse methods is needed to achieve representative results.

References

- Alves, Fabio, Adriana Pagano, and Igor da Silva. 2011. “Towards an Investigation of Reading Modalities in/for Translation: An Exploratory Study Using Eye-Tracking Data.” In *Cognitive Explorations of Translation*, edited by Sharon O’Brien, 175–196. London: Continuum.
- Baker, Monya, and Dan Penny. 2016. “Is There a Reproducibility Crisis?” *Nature* 533 (7604): 452–454.
- Begley, C. Glenn. 2013. “Six Red Flags for Suspect Work.” *Nature* 497 (7450): 433–434.
- Biblioteca d’Humanitats. 2013–2019. RETI: revistes dels estudis de Traducció i Interpretació: Indicadors de qualitat. Accessed November 14, 2018. <http://www.bib.uab.cat/human/acreditacions/planes/publicues/revistes/eti.php?area=eti&menuidioma=eng>
- Cesario, Joseph. 2014. “Priming, Replication, and the Hardest Science.” *Perspectives on Psychological Science* 9 (1): 40–48.
- Crandall, Christian S., and Jeffrey W. Sherman. 2016. “On the Scientific Superiority of Conceptual Replications for Scientific Progress.” *Journal of Experimental Social Psychology* 66: 93–99.

This is a post-print version of the following article published as “online first”:

Olalla-Soler, C. (2020). Practices and attitudes toward replication in empirical translation and interpreting studies. *Target*, 32(1). <https://doi.org/10.1075/target.18159.ola>

Everitt, Brian S. 1998. *The Cambridge Dictionary of Statistics*. Cambridge: Cambridge University Press.

Fabrigar, Leandre R., and Duane T. Wegener. 2016. “Conceptualizing and Evaluating the Replication of Research Results.” *Journal of Experimental Social Psychology* 66: 68–80.

Fanelli, Daniele. 2009. “How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data.” *PLoS ONE* 4 (5): 1–11.

Fanelli, Daniele. 2010. “‘Positive’ Results Increase Down the Hierarchy of the Sciences.” *PLoS ONE* 5 (4): e10068.

Fanelli, Daniele. 2018. “Opinion: Is Science Really Facing a Reproducibility Crisis, and Do We Need It To?” *Proceedings of the National Academy of Sciences* 115 (11): 2628–2631.

Ferguson, Christopher J., and Moritz Heene. 2012. “A Vast Graveyard of Undead Theories.” *Perspectives on Psychological Science* 7 (6): 555–561.

Feyerabend, Paul. 1978. *Against Method: Outline of an Anarchistic Theory of Knowledge*. London: Verso.

Fiedler, Klaus, and Norbert Schwarz. 2016. “Questionable Research Practices Revisited.” *Social Psychological and Personality Science* 7 (1): 45–52.

Franco Aixelá, Javier. 2001–2019. BITRA (Bibliography of Interpreting and Translation). Accessed November 13, 2018. <https://dti.ua.es/en/bitra/introduction.html>

Gile, Daniel. 1991. “Methodological Aspects of Interpretation (and Translation) Research.” *Target* 3 (2): 153–174.

Gile, Daniel. 2000. “The History of Research into Conference Interpreting.” *Target* 12 (2): 297–321.

Graham, Loren, and Jean-Michel Kantor. 2007. “Soft” Area Studies versus “Hard” Social Science: A False Opposition.” *Slavic Review* 66 (1): 1–19.

Gupta, Brij Mohan, and S. M. Dhawan. 2019. “Machine Translation Research: A Scientometric Assessment of Global Publications Output during 2007–16.” *DESIDOC Journal of Library & Information Technology* 39 (1): 31–38.

Hale, Sandra, and Jemina Napier. 2013. *Research Methods in Interpreting*. London: Bloomsbury.

Hedges, Larry V. 1987. “How Hard Is Hard Science, How Soft Is Soft Science? The Empirical Cumulativeness of Research.” *American Psychologist* 42 (5): 443–455.

Holmes, James S. (1972) 1988. “The Name and Nature of Translation Studies.” In *Translated! Papers on Literary Translation and Translation Studies*, edited by James S. Holmes, 67–80. Amsterdam: Rodopi.

House, Juliane. 2013. “Towards a New Linguistic-Cognitive Orientation in Translation Studies.” *Target* 25 (1): 46–60.

Hsieh, Hsiu-Fang, and Sarah E. Shannon. 2005. “Three Approaches to Qualitative Content Analysis.” *Qualitative Health Research* 15 (9): 1277–1288.

Hüffmeier, Joachim, Jens Mazei, and Thomas Schultze. 2016. “Reconceptualizing Replication as a Sequence of Different Studies: A Replication Typology.” *Journal of Experimental Social Psychology* 66: 81–92.

This is a post-print version of the following article published as “online first”:

Olalla-Soler, C. (2020). Practices and attitudes toward replication in empirical translation and interpreting studies. *Target*, 32(1). <https://doi.org/10.1075/target.18159.ola>

- Hurtado Albir, Amparo, and Fabio Alves. 2009. “Translation as a Cognitive Activity.” In *The Routledge Companion to Translation Studies*, edited by Jeremy Munday, 54–73. London: Routledge.
- Ioannidis, John P. A. 2005. “Why Most Published Research Findings Are False.” *PLoS Medicine* 2 (8): 696–701.
- Jussim, Lee, Jarret T. Crawford, Stephanie M. Anglin, Sean T. Stevens, and Jose L. Duarte. 2016. “Interpretations and Methods: Towards a More Effectively Self-Correcting Social Psychology.” *Journal of Experimental Social Psychology* 66: 116–133.
- Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.
- Li, Defeng. 2004. “Trustworthiness of Think-Aloud Protocols in the Study of Translation Processes.” *International Journal of Applied Linguistics* 14 (3): 301–313.
- Liu, Minhua. 2011. “Methodology in Interpreting Studies: A Methodological Review of Evidence-Based Research.” In *Advances in Interpreting Research: Inquiry in Action*, edited by Brenda Nicodemus and Lauria A. Swabey, 85–120. Amsterdam: John Benjamins.
- Makel, Matthew C., Jonathan A. Plucker, and Boyd Hegarty. 2012. “Replications in Psychology Research.” *Perspectives on Psychological Science* 7 (6): 537–542.
- Martin, G. N., and Richard M. Clarke. 2017. “Are Psychology Journals Anti-Replication? A Snapshot of Editorial Practices.” *Frontiers in Psychology* 8: 1–6.
- Mellinger, Christopher D., and Thomas A. Hanson. 2017. *Quantitative Research Methods in Translation and Interpreting Studies*. New York: Routledge.
- Moonesinghe, Ramal, Muin J. Khoury, and A. Cecile J. W. Janssens. 2007. “Most Published Research Findings Are False – But a Little Replication Goes a Long Way.” *PLoS Medicine* 4 (2): 218–221.
- Muñoz Martín, Ricardo. 2010. “Leave No Stone Unturned: On the Development of Cognitive  Translatology.” *Translation and Interpreting Studies* 5 (2): 145–162.
- Muñoz Martín, Ricardo. 2014. “A Blurred Snapshot of Advances in Translation Process Research.” *MonTI* special issue - *Minding Translation* 1: 49–84.
- Neunzig, Wilhelm, and Helena Tanqueiro. 2007. *Estudios empíricos en traducción. Enfoques y métodos*. Girona: Documenta Universitaria.
- O’Brien, Sharon, ed. 2011. *Cognitive Explorations of Translation*. London: Continuum.
- O’Brien, Sharon. 2013. “The Borrowers: Researching the Cognitive Aspects of Translation.” *Target* 25 (1): 5–17.
- Open Science Collaboration. 2015. “Estimating the Reproducibility of Psychological Science.” *Science* 349 (6251): aac4716.
- Orero, Pilar, Stephen Doherty, Jan-Louis Kruger, Anna Matamala, Jan Pedersen, Elisa Perego, Pablo Romero-Fresco, Sara Rovira-Esteva, Olga Soler-Vilageliu, and Agnieszka Szarkowska. 2018. “Conducting Experimental Research in Audiovisual Translation (AVT): A Position Paper.” *Journal of Specialised Translation* 30: 105–126.
- Pardo, Antonio, and Ricardo San Martín. 2012. *Análisis de datos en ciencias sociales y de la salud II*. Madrid: Editorial Síntesis.

This is a post-print version of the following article published as “online first”:

Olalla-Soler, C. (2020). Practices and attitudes toward replication in empirical translation and interpreting studies. *Target*, 32(1). <https://doi.org/10.1075/target.18159.ola>

- Pashler, Harold, and Christine R. Harris. 2012. “Is the Replicability Crisis Overblown? Three Arguments Examined.” *Perspectives on Psychological Science* 7 (6): 531–536.
- Polit, Denise F., and Cheryl Tatano Beck. 2010. “Generalization in Quantitative and Qualitative Research: Myths and Strategies.” *International Journal of Nursing Studies* 47 (11): 1451–1458.
- Popper, Karl. 1959. *The Logic of Scientific Discovery*. London: Hutchinson.
- Ravitch, Sharon M., and Nicole Mittenfelner Carl. 2016. *Qualitative Research: Bridging the Conceptual, Theoretical and Methodological*. Thousand Oaks, CA: SAGE.
- Rovira-Esteva, Sara, and Javier Franco Aixelá. 2018. “Bibliometric Tools.” In *A History of Modern Translation Knowledge: Sources, Concepts, Effects*, edited by Lieven D’hulst and Yves Gambier, 117–122. Amsterdam: John Benjamins.
- Saldanha, Gabriela, and Sharon O’Brien. 2014. *Research Methodologies in Translation Studies*. London: Routledge.
- Schmidt, Stefan. 2009. “Shall We Really Do It Again? The Powerful Concept of Replication Is Neglected in the Social Sciences.” *Review of General Psychology* 13 (2): 90–100.
- Simons, Daniel J. 2014. “The Value of Direct Replication.” *Perspectives on Psychological Science* 9 (1): 76–80.
- Snell-Hornby, Mary. 2006. *The Turns of Translation Studies: New Paradigms or Shifting Viewpoints?* Amsterdam: John Benjamins.
- Tymoczko, Maria. 2002. “Connecting the Two Infinite Orders: Research Methods in Translation Studies.” In *Crosscultural Transgressions: Research Models in Translation Studies II: Historical and Ideological Issues*, edited by Theo Hermans, 9–25. Manchester: St. Jerome.
- Tymoczko, Maria. 2005. “Trajectories of Research in Translation Studies.” *Meta* 50 (4): 1082–1097.
- Wagenmakers, Eric-Jan, Ruud Wetzels, Denny Borsboom, Han L. J. van der Maas, and Rogier A. Kievit. 2012. “An Agenda for Purely Confirmatory Research.” *Perspectives on Psychological Science* 7 (6): 632–638.
- Way, Catherine. 2014. “Structuring a Legal Translation Course: A Framework for Decision-Making in Legal Translator Training.” In *The Ashgate Handbook of Legal Translation*, edited by Le Cheng, King-Kui Sin, and Anne Wagner, 135–152. Aldershot: Ashgate.
- Yeung, Andy W. K. 2017. “Do Neuroscience Journals Accept Replications? A Survey of Literature.” *Frontiers in Human Neuroscience* 11: 1–6.
- Zanettin, Federico, Gabriela Saldanha, and Sue-Ann Harding. 2015. “Sketching Landscapes in Translation Studies: A Bibliographic Study.” *Perspectives: Studies in Translatology* 23 (2): 37–41.

This is a post-print version of the following article published as “online first”:
Olalla-Soler, C. (2020). Practices and attitudes toward replication in empirical translation and interpreting studies. *Target*, 32(1). <https://doi.org/10.1075/target.18159.ola>

Appendix

Supplementary Material 1. Survey in English: <https://doi.org/10.1075/target.18159.ola.additional.1>

Supplementary Material 2. Survey in Spanish: <https://doi.org/10.1075/target.18159.ola.additional.2>

Supplementary Material 3. Survey data: <https://doi.org/10.1075/target.18159.ola.additional.3>

Supplementary Material 4. Data retrieved from BITRA:
<https://doi.org/10.1075/target.18159.ola.additional.4>

Supplementary Material 5. Data from the analysis of acceptance of replication in translation and interpreting journals: <https://doi.org/10.1075/target.18159.ola.additional.5>

Address for correspondence

Christian Olalla-Soler

Department of Translation and Interpreting & East Asian Studies

Universitat Autònoma de Barcelona

Edifici K, Plaça del Coneixement

08193 BELLATERRA

Spain

Christian.Olalla@uab.cat

<https://orcid.org/0000-0001-9251-476X>

Publication history

Date received: 29 November 2018

Date accepted: 16 November 2019

Published online: 19 December 2019