THE INTERNATIONAL JOURNAL FOR THE PSYCHOLOGY OF RELIGION
2019, VOL. 29, NO. 4, 278–291
https://doi.org/10.1080/10508619.2019.1687197

Routledge
Taylor & Francis Group

The Replication Crisis and Qualitative Research in the Psychology of Religion

Adam Anczyk [a], Halina Grzymała-Moszczyńska [a], Agnieszka Krzysztof-Świderska [a], and Jacek Prusak [b]

[a]Department for Psychology of Religion and Spirituality, Institute of Psychology, Jesuit University Ignatianum, Kraków, Poland; [b]Department for Psychopathology and Preventive Psychology, Jesuit University Ignatianum, Kraków, Poland

ABSTRACT

This article forms a contribution to the discussion on the "replication crisis" in psychology from the qualitative research perspective and qualitative-oriented psychology of religion. The main theme of our article is, how should we deal, as qualitative-oriented psychologists of religion, with the issue of replicability? The introductory part of the article concentrates on validity criteria within qualitative-oriented research, and why qualitative-oriented psychologists of religion should take them into consideration in their research projects. Next, a typology of approaches (Intuitive, Field, Pragmatic, Synthetic), toward replication within qualitative studies is presented. Alongside discussing the possible ways of making qualitative research more replicable, examples of good practices in that matter are discussed. Some takeaway tips for qualitative-oriented psychologists of religion, that are to be used in order to make their research more replicable, are presented in the concluding part. Promoting CAQDAS, developing clear research protocols and procedures, describing the cultural context of the research in detail and discussing methodological issues and barriers/limitations of the study in a separate section are one of the main ideas postulated in the article, that should be included in the future qualitative studies on religion (s). The article concludes with a plea of sorts, that qualitative-oriented researchers do need to pay more attention to methodological issues while designing a research project, keeping in mind that it can (and should) be replicated in the future.

"I had discovered that there simply is no repetition and had verified it by having it repeated in every possible way"
Søren Kierkegaard (1813–1855), "Repetition"

No day copies yesterday
no two nights will teach what bliss is
in precisely the same way
with precisely the same kisses
Wisława Szymborska (1923–2012), "Nothing twice"

Introduction

Qualitative research is not immune to questions of replicability and validity. On the contrary, its critics often use arguments that may cut into psychology research as a whole, thus hitting qualitative analysis even harder. "This research could not be replicated"; "it's not representative"; "it uses a limited, casually picked sample" – these are opinions that a researcher may hear whether they

CONTACT Adam Anczyk anczyk.adam@ignatianum.edu.pl Department for Psychology of Religion and Spirituality, Institute of Psychology, Jesuit University Ignatianum, Kopernika Str. 26, 31-501 Kraków, Poland.

perform quantitative, experimental, or – especially – qualitative research (Charmaz, 2007, p. 77). Qualitative researchers cannot ignore such criticism (Ali & Yusof, 2011, p. 30), nor can they only retort "my research follows a different methodology." In the face of such criticism, they must make a stand and try to implement some standards and good practices within their research projects.

R. B. Johnson presented strategies to improve validity in a qualitative research, enumerating the five most important factors qualitative researchers should follow (Johnson, 1997, p. 283). First of them is considering the degree of accuracy of reports about people's behavior and the context in which a given behavior took place (descriptive validity). The second one is an ability researcher should have acquired for understanding meanings of statements and behaviors of research participants (interpretive validity), third is interpreting collected data from a theoretical perspective (theoretical validity), fourth is a potential for finding existing causes of observed phenomena (internal validity), and fifth is a possibility of generalization of derived findings on other research samples (external validity). The notion of "derived etic" as a strategy postulated by cultural psychology (Kim, Yang, & Hwang, 2006), represents an example of achieving external validity by accumulating research findings from qualitative research conducted in many different cultures – or in other words – culturally specific samples. In conclusion, Johnson (1997) indicates requirements, which need to be fulfilled by given research reports in order to provide an opportunity for their eventual generalization: "the number and kinds of people in the study, how they were selected to be in the study, contextual information, the nature of the researcher's relationship with the participants, information about any informants who provided information, the methods of data collection used, and the data analysis techniques used" (Johnson, 1997, p. 290).

Using qualitative methods as a main research method or including qualitative tools in methodological triangulation within a given study is getting more and more popular in the psychology of religion. If the IAPR congresses can be considered as representative for the discipline's development, it can be stated that the quantitative approach still dominates the field, but qualitative methods form an important part of research within psychology of religion: approximately 20% of presentations at IAPR congress in Bari, 2011 and 15% in the case of the congress in Lausanne, 2013 were based on or included qualitative methods (Anczyk, 2014). Therefore, research teams and researchers who use qualitative methods in their work should take the challenge of creating standards and proposing good practices that will, in a way, form an answer to the "replication crisis".

Jacob Cohen (1923– 1998), a psychologist and statistician of the New York University, had concluded his methodological essay *The Earth Is Round (p < .5)* with a statement: "given the problems of statistical induction, we must finally rely, as have older sciences, on replication." (Cohen, 1994, p. 1002). It does seem that the so-called replication crisis may have an important influence on quantitative research based on statistics. This statement also shows another trend: the "older sciences" mentioned therein are the "natural" sciences. Currently, psychology is more and more oriented toward "natural" science, where replicability is one of the criteria of good experiment design. This is understandable given the history of that discipline, since psychology, including psychology of religion, had always contained that "scientific" or "natural" element, the experimental character that freed it from the confines of philosophy and theology. On the other hand, psychology also comes with a fair share of humanist thought, pointing to the subjectivity of human experience and the inherent difficulty in studying it (and psychology of religion especially, since religiosity and spirituality are here the main research topic). The subjectivity and unique character of human experience is a challenge for methodology, as stated by Carl G. Jung: "Our way of looking at things is conditioned by what we are. And since other people are differently constituted, they see things differently and express themselves differently." (Jung, 1933, p. 134).

Within the scope of the replication crisis (Open Science Collaboration, 2015) a clue statement has been made that psychology may be a "false-positive" science – any given research outcomes can be presented as significant if appropriate statistical techniques are used (Simmons, Nelson, & Simonhson, 2011). Paraphrasing this conclusion – some psychological studies are no more than a mathematical game, conducted in a spirit "by gamers for gamers". The authors conclude their

article with a statement: "Our goal as scientists is not to publish as many articles as we can, but to discover and disseminate truth." (Simmons et al., 2011, p. 1365).

Carl G. Jung was of the opinion that phenomenology may be the best method for analyzing psychological phenomena (Jung, 1969, p. 285). This approach, while present in modern psychology, has trouble defending itself from the requirements made to modern science. It is considered unmeasurable – too subjective – but also archaic, omitting the postmodern, constructionist and cultural perspective, which doubts that a single researcher can directly glean into a phenomenon. In social sciences, which are primarily empirical, there had been a schism of sorts (not completely dualistic, however) between post-positivist researchers (Phillips & Burbules, 2000) who look for measurable social or psychological "facts" and constructionist researchers, who delve into the role and function of cultural contexts and meanings. Constructionist psychologists (Madill, Jordan, & Shirley, 2000, pp. 9–10) have turned toward qualitative research as a method of obtaining relevant data, which is perhaps a natural turn.

Qualitative research often garners criticism in the fields of academia where quantitative or experimental paradigms exist concurrently, and especially if it is the dominant one – as is the case of psychology. In those disciplines, discussion of qualitative research validity appears side by side with its criticism. Perhaps this coexistence of seemingly distinct methodological paradigms within a single scientific field may be a reason why "within psychology, researchers using qualitative methods are usually looked upon with a certain degree of suspicion" (Heyink & Tymstra, 1993, p. 291). In a way, the replication crisis has added a new aspect to that suspicion: the criticism based on the lack of (or on insufficient) replicability of research – and thus of poor validity. The criticism was directed mainly toward quantitative and experimental research. Qualitative research remained aside, not concerned, due to their adherence to another "methodological order" which shielded them from those "methodological terrorists", that Susan Friske mentioned in her broadly discussed draft letter to *The Observer* (Letzer, 2016; Singal, 2016).

Replication in a qualitative-oriented study: Into possibilities

In theory, the procedure for qualitative research is simple. If the topic is considered "innovative" by the peers then we are halfway already. Assuming that we want our research to be replicated in the future by other scholars, we must apply ourselves diligently to its design. But even then, would it be replicable? The conceptual idea and the topic of a study may be replicated, the procedure and tools as well – but not necessarily the research as a whole. And yet research, even qualitative, should be replicable, because "replications are fundamental to theoretical growth" (Lucas, Morrell, & Posard, 2013, p. 213). In that process, there are some important, methodological difficulties, of which a qualitatively oriented researcher should be aware of.

The first issue is sampling. Qualitative research is based on an in-depth exploration of data gathered from interviews or by participant observation. When changing the subjects, we modify approach, methods, and the research material. Qualitative research always asks the question "who is being studied?" Onwuegbuzie and Leech (2007) present several sampling designs. Many variables and contexts are important here, with cultural context being especially important. In that respect, transferring the research into another context may also mean starting a new study, and not just replicating the previous one in different conditions. Cultures are dynamic as well, and time is another obstacle to replication. Even if we come back to the same sample (in a longitudinal study), we come to other people – older, transformed by new experiences, both internal (psychological) and external (environmental, cultural). At the same time, the cultural environment changes in the same way, with shifting mores, attitudes, and politics.

The problem of sampling in psychology is a matter of a broader discussion, as psychology often studies a certain sample, which is in fact a specific group of people, often sophomore year students (and often WEIRD, Western Educated, Industrialized, Rich, Democratic, see Henrich, Heine, & Norenzayan, 2010) who are not representative of the whole community, on which findings are often

generalized (Hanel & Vione, 2016). The lack of representativeness is even more pertinent when convenience samples are gathered via Amazon's Mechanical Turk (MTurk) or other e-tools of recruitment. Self-selected participants perform certain research tasks ordered by the researcher and their results are then generalized within a sample, or even seen as valid for the whole population (Chandler & Shapiro, 2016). When choosing a research sample we also have to deal with probability theory, which is easier to observe in qualitative research. For instance, if we study the political preferences of stay-at-home wives, we may take 40 in-depth interviews and find that most of them have radical right-wing views. And even if, among all of a given country's stay-at-home wives, some 10% have radical right-wing views, we may still happen to have them overrepresented in our sample. And since probability is random, the same thing can happen in quantitative research. Such "randomness" of sample composition may explain why some supposedly very similar studies give different results when repeated.

The second obstacle in qualitative research replicability is defining phenomena that are to be a subject of study. For example, if "religious experience" is defined by a given sample as an event of being connected to the Divine, personal force, it would be hardly replicable in a study group, in which "religious experience" is defined in terms of being one with an impersonal, absolute force or in a group in which attaining *satori*, forms the definition of a "religious experience". Wulff (1997, p. 6) wrote that terms of "religiosity" and "spirituality" were sometimes used in the psychology of religion interchangeably, therefore making replication of research even harder. If "spirituality" is to be defined as a pan-cultural psychological feature, measurable with certain tools, different from religiosity measures – and psychologists who would adhere to that definition are as numerous as the researches who oppose that statement, we can state that the qualitative research projects within psychology of religion are even more prone to the "perennialists vs constructivists" discussion.

The next question a qualitative researcher must ask is, "who does the study?" The influence of a researcher on the process is obvious in qualitative research, because it is important who interviews or observes. Theoretically, a qualitative researcher should remain neutral, which can be difficult in practice. Everyone has their bias, which can come to light in research procedure. In psychology of religion there is a recurring question of whether an atheist or a nonbeliever can understand a religious person, and vice-versa – does a religiously engaged scholar possess the qualifications to study atheists? Some scholars, especially those focused on phenomenology, may doubt the possibility of this happening in practice. However, critical self-reflection about own biases and preconceptions might help to control researcher's shortcomings (Johnson, 1997). Sharing the experience between a researcher and respondents in order to gain insight into a phenomenon being a subject of research was already stressed by Jung: "there are ways which bring us nearer to living experience, yet we should beware of calling these way 'methods'. The very word has a deadening effect. The way to experience, moreover, is anything but a clever trick; it is rather a venture which requires us to commit ourselves with our whole being." (Jung, 1969, p. 285). In this approach, the most competent researcher is the person who shares a given experience, which can also influence replicability. It should be noted that this idea, while not new, is in accord with current methods of cultural psychology. The "hard" indigenous approach in psychology underscores that a person most qualified to conduct a study in a given culture, should be its member, and every study has its context which must be included in the process of drawing conclusions. (Allwood & Berry, 2006; Sundararajan, 2015). Ideas of this kind bring important questions when designing a replicable-like qualitative research project: does research conducted on a similar sample, with the same tools (like interview scenario, for example) count as replication, or as a new study, if other researchers are conducting it?

The "human" factor is also strongly present in the next stage, i.e., qualitative data analysis, with interview transcripts themselves being already an interpretation of sorts, the first processing of "raw data." This issue is rarely touched upon (Bielecka-Prus, 2015). Even when using CAQDAS (Computer-Assisted Qualitative Data Analysis Software) we still have to deal with the question of who and how categorizes interview data, or who prepares the "code tree" – the result of an analysis that will be discussed or used to draw conclusions. This is especially problematic in group projects,

where several scholars conduct interviews or observations in the same project and then cooperate in analyzing data.

One of the remedies for this is the Inter-Code Reliability procedure (ICR), a method of comparing material encoding performed by various coders (MacPhail, Khoza, Abler, & Ranganathan, 2016). The software used in qualitative data analysis allows calculating the conformity of the coders. However, it will never eliminate the human factor, and leveling codes to a common standard (ascribing sentences in an interview to specific codes), does not guarantee that the coding procedure is accurate. Statistical coherence is not the same as correctness – similar results obtained by four unexperienced coders do not carry similar validity as similar results obtained by four highly competent ones. These kinds of procedures may, on the one hand, support qualitative research validity, but on the other, they may be a kind of "mathematical deflector" – the ICR is fine, so the coders are in accord with their data analysis, never mind whether or not it was diligent and rigorous. The important thing is that the numbers are all right and the program shows no conflict. This attitude, which may have seeped through to qualitative research from the quantitative perspective, shows an important aspect of the replication crisis – the call for validity and innovative character of "publishable" psychological research.

Those who introduce new things, take up original, not previously explored subjects, may, eventually get published. The other kinds of research – which replicate previous studies, or which yield "no satisfying result" remain unpublished, therefore representing the so-called "file drawer problem" (PsychFileDrawer, 2019). If they ever got published, many arguments need to be offered by authors why previously established results did not get confirm in their research. What is mostly overlooked, is the fact that lack of identical results could be interpreted as a potential mistake in the previous research (especially if previous results were not confirmed in the project conducted by a younger, less known researcher, coming from a less prestigious unit or university.) One of the responses to the replication crisis is a call for emancipating this research, for their "coming out" and being published. This in turn should provoke reactions (like call for papers), especially from highly indexed, important periodicals on psychology, and this special issue of the International Journal for the Psychology of Religion is an example of a good practice in that respect.

The situation has improved significantly in the last 10 years: for example, in 2015 Nosek et al., reported replications of experiments published in prestigious psychology journals in 2008 (Lindsay, 2015; Open Science Collaboration, 2015), which started a heated discussion. However, social science periodicals are not eager to publish replication studies, as they value more "innovative" research (Makel, Plucker, & Hegarty, 2012). Those who introduce new things, take up original, not previously explored topics have better chances to get published. Those, who replicate previous studies, have fewer chances.

Methodological approaches toward replication in qualitative studies: Intuitive, field, pragmatic, synthetic

The question of qualitative research validity and credibility was and is eagerly discussed among methodologists, psychologists, and sociologists (Charmaz, 2007). The stands taken by researchers on the matter of qualitative methods validity, understood, among other factors, as the ability to replicate and generalize results are many, as the researchers are These stands can be presented and summarized as Intuitive, Field, Pragmatic and Synthetic approaches (Table 1).

By the Intuitive approach, we understand drawing conclusions from qualitative research using one's own ideas and individual hermeneutics. This approach is based on the idea that researchers, when analyzing their material, might have a certain insight into the phenomena in question, thus resulting in conclusions, some of them being of the "generalizable nature". Such analysis confronts researchers with two main issues: plausibility, in the way an experience is described and necessity of safeguarding transparency in following the correct rules of logic inference during the whole process. It is present in phenomenological psychology to a certain extent (Langdridge, 2007) and

Table 1. Methodological approaches to replication in qualitative research.

	Intuitive	Field	Pragmatic	Synthetic
Main disciplines	Anthropology Religious studies Cultural studies Early psychology Sociology (theoretical)	Anthropology Ethnology and ethnography Religious studies Cultural studies Sociology History (oral)	Sociology Psychology (qualitative) Anthropology Political sciences Health sciences	Sociology Psychology Political sciences Health sciences
Research methods used	Introspection Interview Case study Text analysis	Interview Participant observation Case study	Interview Participant observation Mixed-methods Case study Text analysis	Interview Case study Research synthesis Research meta-analysis Mixed-methods
Stand on replication	*Not necessary* researchers should be competent in interpreting phenomena and drawing generalizable conclusions	*Not crucial/ Optional* each research is a report from a specific field study, replicability is not the main aim of research	*Desirable* for verifying previous research and completing data, making qualitative research more comparable	*Crucial* Replicability possibilities as indispensable condition for generalizing results

represented by the phenomenology of religion and hermeneutic approach in religious studies. This approach is rather -researcher-centred, and conclusions may be questioned by another researcher on the basis of his/her insight. The intuitive approach was not alien to classical psychologists like Freud. His " Moses and Monotheism" being an example of nothing more than a record of intuition and even the personal views of the author (Freud, 1939). This approach is not widely used in modern psychology, being limited to certain schools, such as psychoanalysis or analytical psychology.

The Field Approach within the study of religion is typical of anthropological research. There is special journal focused on this kind of studies in religion – *Fieldwork in Religion* (e.g., Irvine, 2010). In this approach, the "Field" speaks, and researchers listen, learning things about the sample, in specific circumstances and setting. Researchers limit their conclusions to their own research context, with a disclaimer that the conclusions are drawn from reflecting on a specific environment. They present and interpret their results, with little pretense to generalization, making their work more idiographic than nomothetic (Thomae, 1999). Taking this approach, a researcher does not concern him or herself with possibilities of replication, because each contact with the field is, in a way, unique – in each attempt, the researcher changes, as do the research participants and the conditions, so each field research is not a replication but a new project. Thus, it is difficult to imagine a replication of a study "Spirits and Trance in Brazil. An Anthropology of Religious Experience" (Schmidt, 2016) or many other studies conducted in different settings. However, the correct way of archiving field data from such research may, eventually (over an extended period of time) lead to some, although limited, generalization of results.

The approach we call Pragmatic differs from the Field Approach in a blurry way, as the two approaches are intertwined. The difference is of aims: while the Field approach concentrates on a given phenomenon or a research setting, concentrating on a in-depth study, the Pragmatic approaches aim at contributing to previous research by revisiting topics, fields, and methods in the spirit of a comparative perspective. It is based on the idea that qualitative research permits conclusions, but of limited scope, narrowed by the sample choice but it is still important and possible to generalize to a certain extent. Williams and Payne, sociologists, propose the term "moderatum generalizations" to describe such conclusions, which should be cautious and guarded, distinguishing between the data and the interpretation, with a detailed description of the research elements and conditions (Payne & Williams, 2005, pp. 304–305). Within this approach, replication has a certain

comparative sense – it allows to "check" whether these "moderatum generalizations" fit a similar sample and conditions, or to change conditions and see how do the conclusions change in turn.

Some recent studies on God's image performed within a psychoanalytical theoretical frame among patients with different mental conditions, conducted on different samples, in which similar research tools were used, form a good example of the Pragmatic Approach.

Table 2 presents a comparison of studies on God's image performed from the psychoanalytic perspective and their "replicability potential". Studies conducted by Rizzuto (1979), Goodman and Manierre (2008) and Krzysztof-Świderska (2017), were qualitative in nature, researchers shared the same theoretical background (object relations theory), and used various methods to investigate a given phenomenon. Rizzuto and Krzysztof-Świderska used the same projective tool, namely the God/Family Questionnaire (1979), while Goodman and Manierre's analysis (2008) was based on the observation of group processes. This is an example of a study, in which different qualitative researchers base their research on a certain theoretical background and use both similar and different tools to investigate it, which in turn may lead eventually to some "moderatum generalizations".

The Synthetic approach represents a turn toward qualitative synthesis and/or meta-analysis of research results, meaning research of research (Braver, Thoemmes, & Rosenthal, 2014), within qualitative studies. In this approach, the replicability criterion is not important, as long as a single research/study is concerned, and – on the contrary – it is vital when results of different studies form a subject of analysis. Each single study must be conducted with due diligence, gathering as rich data as possible, taking care to archive results properly so that other researchers can reuse them as well. Payne and Williams (2005, p. 298) noted that in qualitative studies "'External validity' depended on 'thick description' of the fieldwork; the richness of data collected and full reportage of the care used in its collection (…)". The idea is to concentrate on a given topic/phenomenon multiple times, from varying angles and using various methods, switching research tools and groups, or replicate research done by others on a given subject. It should also be remembered that research which is thus analyzed is not necessarily a "faithful" copy of the original – both in qualitative and quantitative research – since its object is the changeable entity, the human being: "Behavioral research studies can never be direct or exact replications of one another" (McShane & Böckenholt, 2017, p. 1049). Once there is a satisfying amount of qualitative research done on a subject, we can consider drawing conclusions. Thus, this approach suggests (1) conducting as many *emic* research as possible on a given topic, next, (2) the summary/synthesis of the acquired results from multiple research and, after that, (3) *etic* conclusions may be formulated. The Synthetic approach may aim not only at bringing together and comparing results of various research projects, but also taking qualitative research into the realm of the quantitative. So to speak, conducting many qualitative studies, and then making them a subject of a meta-analysis will eventually end with analyzing *big data*. On that basis, we may draw some conclusions, which are of course open to discussion and further research possibilities (See Utts, 1991).

Meta-analysis, as a technique of synthesizing and analyzing data from many different studies, has a strong position in contemporary research (McShane & Böckenholt, 2017), as do the synthetic reviews of articles. Nevertheless, some scholars underestimate or exclude qualitative data from their meta-analyses (e.g., Sherman et al., 2015) or syntheses (Koenig, 2012). Integrating qualitative data with quantitative studies in a meta-analysis or synthesis gives probably the fullest possible insight into a given field (Leppink, 2017). Such an approach was applied by Weaver, Flannelly, and Flannelly (2001). They performed a review study on religious and spiritual variables found in leading gerontological periodicals – *Journal of Gerontological Nursing* and *Geriatric Nursing*. The studies, being a subject of analysis, were divided into two categories: "research" and "non-research". Research articles were defined as those containing a statement of purpose, description of research methods and findings/results. In this regard, decisions were made by at least two researchers in the role of judges/raters, and a third one if there was no consensus. Then, research articles were divided into quantitative (if they included descriptive or inferential statistics) and qualitative. What is interesting is that the authors indicated that this division "caused some studies to be classified as quantitative

Table 2. Research on God's image as an example of pragmatic approach in qualitative psychology of religion.

Study	Sample	Methods of selection	Contextual information	Researcher's relationship with the participants	Information sources	Data collection	Data analysis
"The Birth of the Living God" Ana-Maria Rizzuto (1979)	20 in total; 10 men and 10 women, with different psychiatric diagnosis (one subject with borderline diagnosis)	The basis of selection was to obtain the coverage of mental diagnostic categories	Research was conducted in Boston State Hospital, US; different education and different denominations and ethnic groups	Researcher was a part of the treatment team, however mostly not a personal therapist/doctor of subjects	Medical history prepared by the hospital staff	God/Family Questionnaire, interview (case study method), drawings (image of God)	Psychoanalytic-oriented introspection and analysis, "Ten points" Rizzuto's interpretation scheme
"God from borderlands" Agnieszka Krzysztof-Świderska (2017)	16 in total; 11 women, 5 men, all diagnosed with borderline personality organization	Purposive sampling – borderline patients. borderline personality organization was diagnosed with a Personality Organization Diagnostic Form (PODF), based on Kernberg's theory	Research was conducted in outpatient clinic of the Jagiellonian University's Medical College, Krakow, Poland; all had completed secondary or higher education, were Poles and were raised as Roman Catholics	Researcher was a part of treatment team, however mostly not a personal therapist/doctor of subjects	Medical history prepared by the hospital staff	God/Family Questionnaire, interview (case studies), drawings (image of God) – supplementary method	Psychoanalytic-oriented introspection and analysis, "Ten points" Rizzuto interpretation scheme
"Representations of God Uncovered in a Spirituality Group of Borderline Inpatients", Geoff Goodman and Manierre (2008)	9 women, all diagnosed with borderline personality disorder	Purposive sampling – borderline patients; diagnosis was mainly based on Kernberg's theory	Research was conducted in New York – Presbyterian Hospital Weill Medical College of Cornell University; 9 from 25 subjects in the unit agreed to participate in the research	Researchers were group therapists of subjects	Medical history, researchers' data	Notes from observing group processes, analysis of drawings (image of God)	Interpreting group processes, interpretation of drawings, psychoanalysis of subjects and its conclusions

even though they employed ethnographic, grounded theory, and other qualitative methods" (Weaver et al., 2001, p. 49). Content analysis was used to classify articles as pertaining to their religious and spiritual content. Inter-judge agreement seemed to improve the validity of the research in question. With all its shortcomings it is surely an example of good practice applied in analyzing a cohort of both qualitative and quantitative studies.

Another example of including qualitative studies into meta-analysis in the field of the psychology of religion is a study by Mahoney, Pargament, Tarakeshwar, and Swank (2008), who performed a meta-analytic review of links between religion, marriage, and parenting. They have investigated both quantitative (78) and qualitative studies (16). Also, they used meta-analytic techniques to summarize key quantitative findings (Mahoney et al., 2008, p. 67). The presented typology of attitudes toward validity, generalizability, and replicability of qualitative research is – like any typology – arbitrary and not final. It does, however, give a certain overall idea of the theoretical perspectives or methodological meta-ideas used in empirical research if attitudes toward replicability are concerned. The boundaries between approaches presented may be blurred, and some studies may fit within a given category, as well as in the other, but they may present a picture of sorts. It shows that every approach has its merits and flows, and that a multi-faceted approach seems to be an option to follow. Tanya Luhrmann's works on hearing the voice of God, being both an anthropological study and phenomenological inquiry forms an example of a topic analyzed from different perspectives, using various methodological approaches and research tools (Luhrmann, 2005; Luhrmann, Padmavati, Tharoor, & Osei, 2015a, 2015b; Luhrmann, 2017). Focusing on a certain topic, as Luhrmann did, observing and analyzing it from different angles may be a qualitative response for the "replication crisis". Therefore, we should review procedures of qualitative research in order to fit those criteria, and ask ourselves the question: what can increase qualitative research validity and therefore make it more replicable?

Designing a replicability-oriented qualitative study (on religion)

Research validity (including replicability) in qualitative psychology of religion can be enhanced by implementing strategies and guidelines formulated by Johnson (1997). First, an attention needs to be paid to sampling issues. Purposive sampling is the obvious choice of action. The sample should reflect the research topic, as precisely as possible (Palys, 2008, p. 697), and it is a good practice applied within many qualitative studies (just to mention ones quoted in this article, Krzysztof-Świderska, 2017; Pietkiewicz & Kołodziejczyk-Skrzypek, 2016). Contextual information (social, local, cultural) is of importance for every psychological study, as no study is performed in void, but in qualitative research, those factors are more than significant. Therefore, the context of qualitative study needs to be described in full detail. Researchers need to be aware of their position in relation to respondents (outsider/vs insider perspective) and a potential impact of this relation on the course of research. That may pertain not only to their fieldwork qualifications but also to other factors, such as their gender. For instance, a research on the role of religion in coping with refugee experiences of Syrian women in Turkey can bring different results depending on the gender of the researcher. A male researcher will be able to reach only a small, a-typical group of such women, whereas a female scholar will be able to reach a much broader sample. She will be able to visit her respondents' homes, where the majority of them spend most of their time taking care for children, and where access to men other than their husband or father is strictly limited. Such a study will also differ depending on the language competences of the researcher (Levantine Arabic or Turkish), and on the possible presence/absence (and gender) of a translator (Janas, 2018).

In order to support replicability in qualitative studies, creating "step-by-step" protocols and templates for the data collection procedures that can be easily re-used in further research is recommended. A deeper insight into research results could also be achieved by including as many sources of information present in the research context as possible in the analysis. In a clinical context, it could be a supervision of investigated cases. The inter-judge coding could provide some

support in that matter (Armstrong, Gosling, Weinman, & Marteau, 1997). This procedure is not only helpful but also brings a certain rigor to qualitative research (Mays & Pope, 1995, p. 110). Another factor of enhancing replicability of qualitative research is using Computer-Assisted Qualitative Data Analysis Software (CAQDAS), which supports the accuracy of coding, data analysis, and enables building databases of qualitative data (Bryda, 2014). The beginning of using computer tools in the analysis of qualitative data in social sciences dates back to the second half of the 1960s. The term CAQDAS was initially popularized by Raymond Lee and Nigel Fielding starting from the nineties of the XXth century. In 1994 "CAQDAS Networking Project" was established. It works to this day: on the website of the University, one can find a detailed software information, practical advices, and information about upcoming courses.

An example of using CAQDAS in psychology of religion is a study by Pietkiewicz and Kołodziejczyk-Skrzypek (2016) on gay Catholics. Full transcriptions of audio recordings were analyzed using NVivo10. Both researchers listened to each recording of an interview and read its transcript carefully for a few times. They individually produced extensive notes on content and language, writing down interpretative comments using the "annotation" function in NVivo10. Then, they divided notes into new themes according to the assignment of descriptive labels (nodes). It can be noted that working with computer programs raises the researcher's analytical awareness, prompting inter-reliability and contributes to research replication (Bryda, 2014, p. 31). Establishing a special unit dedicated to qualitative studies is also a good practice, an example of which is the QualLab in the Institute of Psychology, University SWPS in Katowice, Poland (QualLab, 2019).

A clear and detailed methodological section of any empirical article is critically important for designing a replicability-oriented qualitative study. It should contain not only a description of all completed procedures but also reflections on limitations and shortcomings encountered in the research. The collected research material should be stored and indexed in databases, becoming accessible to other scholars, who can follow the reasoning of their colleagues, interpret the data in a different way, or include the previously gathered material in their own work (Elman, Kapiszewski, & Vinuela, 2010). Practices which are postulated by the Open Science movement point to one more measure which is pre-registration of the research before the start and permitting other scholars to follow the different stages of the research project (Salmi, 2015). A short summary of good practices in qualitative research is presented in Table 3.

Conclusions

In the qualitative-oriented psychology of religion, there are certain types of "to-be-replicated" studies that have been going on for years since the very inception of the discipline. Phenomena such as conversion, possession, influence of religiosity on mental health, religious coping, etc., have been of interest to many scholars, and therefore the discipline accumulated rich material for critical analysis. As we have presented, direct replication of a given study is almost impossible, due to qualitative studies' specifics. But there may be research themes, areas of study, samples, and settings worth revisiting, and there are many to choose from. Creating more strict and standardized procedures as well as implementing some of the tips we included in Table 3 may help to raise qualitative research credibility and validity, and enable more accurate comparison between studies (as qualitative syntheses or meta-analyses of single research projects focused on the same topic are an important attempt to confirm obtained results). Properly preparing and structuralizing a qualitative study enhances its potential to be replicated, that would eventually end with a synthesis/meta-analysis of research. Of course, we must remember that there always will be studies that do not adhere to such a strict protocol, the ones that are intuitive, personal, ethnographic, maybe even emotional or poetic. And it is a good thing, because without them qualitative studies on religion would lose much of its charm.

Qualitative study differs in many ways from the quantitative, they can be "messy, time-consuming, and even frustrating", but it is in the qualitative optics "the data is full of meaning, providing rich, thick descriptions of the phenomenon under investigation" (Milacci, 2011). One

Table 3. Good practices in a qualitative-oriented research: Toward replication possibilities.

Sample	Methods of selection	Contextual information	Researcher's relationship with the participants	Information source	Data collection	Data analysis	Reporting
Purposive sampling – as precise as possible for a given research's topic, n ~ 10–40 for a single qualitative study (enabling in-depth analysis, depending on how big your research team is) make your sample visible (e. g. in the title of your report)	Aim at homogeneous group or groups that are in the scope of the research's topic rather than expanding the study on different subjects try to go beyond "snowballing" in selecting samples	Describe the context (situational, cultural) in detail in order to provide the background of the study performed mind the whole of the research procedure (e. g., keep a research journal) and describe important variables that could influence its course	Provide a short description of the researcher's role and position toward subjects mind the researcher's qualities (age, gender, ethnicity, looks etc.) that could influence the research procedure	Use as much independent data sources as possible, consult – if needed – on every stage of the research with competent experts	Create "step-by-step" protocols for the data collection procedures, to be used by the research team (templates of data collecting procedures, templates concerning ethical issues) that can be re-used in further research, creating of or supporting existing databases containing qualitative data describe the tool used in detail (and if new – provide a copy in the annex to the report)	Use Computer-Assisted Qualitative Data Analysis Software (CAQDAS) for increasing comparability of research and building qualitative data databases, follow the interpretation scheme (especially if a study is performed in a team of researchers), mind that your research analysis may inspire other researchers to do a replication of your study (you can encourage them to do so, by mentioning that in your report)	Describe methodology used in detail (do not omit methodological difficulties encountered) consider adding sections on "limitations of the study" and "prospective research possibilities" in your report both providing important information for other researchers and encouraging them to do a replication of your study (e. g., on a different sample)

may wonder, considering their differences, if the quantitative rigor can really meet the charm of the qualitative? The answer is not just a simple "yes". The two approaches should meet, date, and eventually marry. A mixed-methods approach, which accounts for the pros and cons of the various methods involved, may be a place of departure for both methodological standpoints. To continue our metaphor, a mixed-methods approach is a key meeting point, a good restaurant with a Swedish smörgåsbord of methods that can be chosen and applied to a given study when needed (and accordingly to one's taste buds). Using various research tools, approaches and methods, so using triangulation, "the practice of obtaining more reliable answers to research questions through integrating results from several different approaches" (Lawlor, Tilling, & Smith, 2016), should be a good – and more often applied – practice in the psychology of religion. It is what Cresswell calls "The Pragmatic Worldview" in methodology (Creswell, 2014), or otherwise a fine dining experience, and researchers, pardon, wizards ought to be found behind big dinners, as we know from the observations made by sir Terry Pratchett.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Adam Anczyk http://orcid.org/0000-0001-6906-3104
Halina Grzymała-Moszczyńska http://orcid.org/0000-0003-2751-3204
Agnieszka Krzysztof-Świderska http://orcid.org/0000-0002-0819-0350
Jacek Prusak http://orcid.org/0000-0002-3204-2251

References

Ali, A. M., & Yusof, H. (2011). Quality in qualitative studies : The case of validity, reliability and generalizability. *Issues in Social and Environmental Accounting*, *5*(1), 25–64. doi:10.22164/isea.v5i1.59

Allwood, C. M., & Berry, J. W. (2006). Origins and development of indigenous psychologies: An international analysis. *International Journal of Psychology*, *41*(4), 243–268. doi:10.1080/00207590544000013

Anczyk, A. (2014, October 27–29). *Psychology of religion within religious studies: A methodological challenge [conference presentation]*. Kraków, Poland: III International Kraków Study of Religion Symposium, Jagiellonian University.

Armstrong, D., Gosling, A., Weinman, J., & Marteau, T. (1997). The place of inter-rater reliability in qualitative research: An empirical study. *Sociology*, *31*, 597–606. doi:10.1177/0038038597031003015

Bielecka-Prus, J. (2015). Transkrypcja jako działanie interpretacyjne. *Rocznik Lubuski*, *41*(1), 41–54.

Braver, S. L., Thoemmes, F. J., & Rosenthal, R. (2014). Continuously cumulating meta-analysis and replicability. *Perspectives on Psychological Science*, *9*(3), 333–342. doi:10.1177/1745691614529796

Bryda, G. (2014). CAQDAS a badania jako?ciowe w praktyce. *Przegl?d Socjologii Jako?ciowej, 10*(2), 12–38. Retrieved from http://przegladsocjologiijakosciowej.org/Volume26/PSJ_10_2_Bryda.pdf

Chandler, J., & Shapiro, D. (2016). Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology*, *12*, 53–58. doi:10.1146/annurev-clinpsy-021815-093623

Charmaz, K. (2007). Tensions in qualitative research. *Sociologisk Forskning*, *44*(3), 76–85.

Cohen, J. (1994). The earth is round (p < .5). *American Psychologist*, *49*(12), 997–1003. doi:10.1037/0003-066X.50.12.1103

Creswell, J. W. (2014). *Research design: Qualitative, quantitative and mixed methods approaches*. London, UK: Sage.

Elman, C., Kapiszewski, D., & Vinuela, L. (2010). Qualitative data archiving: Rewards and challenges. *PS - Political Science and Politics*, *43*(1), 23–27. doi:10.1017/S104909651099077X

Freud, S. (1939). *Moses and monotheism*. Oxford, UK: Knopf.

Goodman, G., & Manierre, A. (2008). Representations of god uncovered in a spirituality group of borderline inpatients. *International Journal of Group Psychotherapy*, *58*, 1–15. doi:10.1521/ijgp.2008.58.1.1

Hanel, P. H. P., & Vione, K. C. (2016). Do student samples provide an accurate estimate of the general public? *PLoS One*, *11*(12), 1–11. doi:10.1371/journal.pone.0168354

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, *33*(2–3), 61–83; discussion 83–135. doi:10.1017/S0140525X0999152X

Heyink, J. W., & Tymstra, T. (1993). The function of qualitative research. *Social Indicators Research*, *29*(3), 291–305. doi:10.1007/BF01079517

Irvine, R. D. G. (2010). The experience of ethnographic fieldwork in an English benedictine monastery: Or, not playing at being a monk. *Fieldwork in Religion*, *5*(2), 221–235. doi:10.1558/firn.v5i2.221

Janas, M. (2018). *The role of religion in coping with refugee experience by Syrian women* (doctoral dissertation in print). Kraków, Poland: Jagiellonian University.

Johnson, B. (1997). Examining the validity structure of qualitative research. *Education*, *118*(2), 282–292.

Jung, C. G. (1933). *Modern man in search of a soul*. London, UK: Kegan Paul, Trench, Trubner and CO LTD.

Jung, C. G. (1969). *Psychology and religion: West and East* (H. Read, M. Fordham, G. Adler, & W. McGuire, eds.). Princeton, NJ: Princeton University Press. Retrieved from https://www.holybooks.com/c-g-jung-collected-works-free-pdf/

Kim, U., Yang, K.-S., & Hwang, -K.-K. (2006). *Indigenous and cultural psychology: Understanding people in context*. Boston, MA: Springer.

Koenig, H. (2012). Religion, spirituality, and health: The research and clinical implications. *ISRN Psychiatry*, *2012*, 1–33. doi:10.5402/2012/278730

Krzysztof-Świderska, A. (2017). *Bóg z pogranicza* [God from the Borderlands]. Kraków, Poland: Wydawnictwo Ignatianum.

Langdridge, D. (2007). *Phenomenological psychology: Theory, research and method*. Harlow, Essex: Pearson Education Limited.

Lawlor, D. A., Tilling, K., & Smith, G. D. (2016). Triangulation in aetiological epidemiology. *International Journal of Epidemiology*, *45*(6), 1866–1886. doi:10.1093/ije/dyw314

Leppink, J. (2017). Revisiting the quantitative-qualitative-mixed methods labels: Research questions, developments, and the need for replication. *Journal of Taibah University Medical Sciences*, *12*(2), 97–101. doi:10.1016/j.jtumed.2016.11.008

Letzer, R. (2016). Scientists are furious after a famous psychologist accused her peers of 'methodological terrorism'. *Business Insider*. Retrieved from https://www.businessinsider.com/susan-fiske-methodological-terrorism-2016-9?IR=T

Lindsay, D. S. (2015). Replication in psychological science. *Psychological Science*, *26*(12), 1827–1832. doi:10.1177/0956797615616374

Lucas, J. W., Morrell, K., & Posard, M. (2013). Considerations on the "replication problem" in sociology. *The American Sociologist*, *44*(2), 217–232. doi:10.1007/s12108-013-9176-7

Luhrmann, T., Padmavati, R., Tharoor, H., & Osei, A. (2015a). Differences in voice-hearing experiences of people with psychosis in the U.S.A., India and Ghana: Interview-based study. *British Journal of Psychiatry*, *206*(1), 41–44. doi:10.1192/bjp.bp.113.139048

Luhrmann, T., Padmavati, R., Tharoor, H., & Osei, A. (2015b). Hearing voices in different cultures: A social kindling hypothesis. *Topics in Cognitive Sciences*, 1–18. doi:10.1111/tops.12158

Luhrmann, T. M. (2005). The art of hearing God: Absorption, dissociation and contemporary American spirituality. *Spiritus: A Journal of Christian Spirituality*, *5*, 133–157. doi:10.1353/scs.2006.0014

Luhrmann, T. M. (2017). Knowing god. *The Cambridge Journal of Anthropology*, *35*(2), 125–142. doi:10.3167/cja.2017.350210

MacPhail, C., Khoza, N., Abler, L., & Ranganathan, M. (2016). Process guidelines for establishing intercoder reliability in qualitative studies. *Qualitative Research*, *16*(2), 198–212. doi:10.1177/1468794115577012

Madill, A., Jordan, A., & Shirley, C. (2000). Objectivity and reliability in qualitative analysis: Realist, contextualist and radical constructionist epistemologies. *British Journal of Psychology*, *91*(1), 1–20. doi:10.1348/000712600161646

Mahoney, A., Pargament, K. I., Tarakeshwar, N., & Swank, A. B. (2008). Religion in the home in the 1980s and 1990s: A meta-analytic review and conceptual analysis of links between religion, marriage, and parenting. *Psychology of Religion and Spirituality*, *S*(1), 63–101. doi:10.1037/1941-1022.S.1.63

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science*, *7*(6), 537–542. doi:10.1177/1745691612460688

Mays, N., & Pope, C. (1995). Rigour and qualitative research. *British Medical Journal*, *311*, 109–112. doi:10.1136/bmj.311.6997.109

McShane, B. B., & Böckenholt, U. (2017). Single paper meta-analysis: Benefits for study summary, theory-testing, and replicability. *Journal of Consumer Research*, *43*, 1048–1063. doi:10.1093/jcr/ucw085

Milacci, F. (2011). Qualitative research. In T. Clinton & T. R. Hawkins (Eds.), *The popular encyclopedia of Christian counseling: An indispensable tool for helping people with their problems* (pp. 546–548). Eugene, OR: Harvest House Publishers.

Onwuegbuzie, A. J., & Leech, N. L. (2007). Sampling designs in qualitative research: Making the sampling process more public. *The Qualitative Report*, *12*(2), 238–254.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716–aac4716. doi:10.1126/science.aac4716

Palys, T. (2008). Purposive sampling. In L. M. Given (Ed.), *The Sage encyclopedia of qualitative research methods* (Vol. 2, pp. 697–698). Los Angeles, CA: Sage.
Payne, G., & Williams, M. (2005). Generalization in qualitative research. *Sociology*, *39*(2), 295–314. doi:10.1177/0038038505050540
Phillips, D. C., & Burbules, N. C. (2000). *Postpositivism and educational research. Philosophy, theory, and educational research*. Lanham, MD: Rowman & Littlefield.
Pietkiewicz, I. J., & Kołodziejczyk-Skrzypek, M. (2016). Living in sin? How gay Catholics manage their conflicting sexual and religious identities. *Archives of Sexual Behavior*, *45*(6), 1573–1585. doi:10.1007/s10508-016-0752-0
PsychFileDrawer. (2019). *PsychFileDrawer. Archive of replication attempts in experimental psychology*. Retrieved from http://www.psychfiledrawer.org/TheFiledrawerProblem.php
QualLab. (2019). *QualLab - pracownia badań jakościowych*. Instytut Psychologii, SWPS Uniwersytet Humanistycznospołeczny, filia w Katowicach. Retrieved from https://www.swps.pl/nauka-i-badania/jednostki-badawcze/quallab
Rizzuto, A. M. (1979). *The birth of the living God: A psychoanalytic study*. Chicago, IL: University of Chicago Press.
Salmi, J. (2015). *Study on open science, impact, implications and policy options*. Luxembourg: Publications Office of the European Union.
Schmidt, B. E. (2016). *Spirits and trance in Brazil. An anthropology of religious experience*. London, UK: Bloomsbury.
Sherman, A. C., Merluzzi, T. V., Pustejovsky, J. E., Park, C. L., George, L., Fitchett, G., & Salsman, J. M. (2015). A meta-analytic review of religious or spiritual involvement and social health among cancer patients. *Cancer*, *121*(21), 3779–3788. doi:10.1002/cncr.29352
Simmons, J. P., Nelson, L. D., & Simonhson, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. doi:10.1177/0956797611417632
Singal, J. (2016, October 12). Inside psychology's "methodological terrorism" debate. *The Cut*. Retrieved from https://www.thecut.com/2016/10/inside-psychologys-methodological-terrorism-debate.html
Sundararajan, L. (2015). Indigenous psychology: Grounding science in culture, why and how? *Journal for the Theory of Social Behaviour*, *45*(1), 64–81. doi:10.1111/jtsb.12054
Thomae, H. (1999). The nomothetic-idiographic issue: Some roots and recent trends. *International Journal of Group Tensions*, *28*(1–2), 187–215. doi:10.1023/A:1021891506378
Utts, J. (1991). Replication and meta-analysis in parapsychology. *Statistical Science*, *6*(4), 363–378. doi:10.1214/ss/1177011585
Weaver, A. J., Flannelly, L. T., & Flannelly, K. J. (2001). A review of research on religious and spiritual variables in two primary gerontological journals 1991–1997. *Journal of Gerontological Nursing*, *27*(9), 47–54. doi:10.3928/0098-9134-20010901-09
Wulff, D. M. (1997). *Psychology of religion: Classic and contemporary*. New York, NY: John Wiley & Sons.