Abstract

Preregistrations -- records made *a priori* about study designs and analysis plans and placed in open repositories -- are thought to strengthen the credibility and transparency of research. Different authors have put forth arguments in favor of introducing this practice in qualitative research and made suggestions for what to include in a qualitative preregistration form. The goal of this study was to gauge and understand what parts of preregistration templates qualitative researchers would find helpful and informative. We used an online Delphi study design consisting of two rounds with feedback reports in between. In total, 48 researchers participated (response rate: 16%). In round 1, panelists considered 14 proposed items relevant to include in the preregistration form, but 2 items had relevance scores below our predefined criterion (68%) with mixed argument and were put forth again. We combined items where possible, leading to 11 revised items. In round 2, panelists agreed on including the two remaining items. Panelists also converged on suggested terminology and elaborations, except for two terms for which they provided clear arguments. The result is an agreement-based form for the preregistration of qualitative studies that consists of 13 items. The form will be made available as a registration option on Open Science Framework (osf.io). We believe it is important to assure that the strength of qualitative research, which is its flexibility to adapt, adjust and respond, is not lost in preregistration. The preregistration should provide a systematic starting point.

241 words

Key words: qualitative research, preregistration, transparency

Introduction

Preregistrations -- records made *a priori* about study designs and analysis plans and placed in (open) repositories -- are thought to strengthen the credibility and transparency of research. Firstly, the openness of this information about the study encourages the researcher to carefully reflect on different study aspects and to systematically report on their design and analysis choices, including those made as the study progresses (Kern & Gleditsch, 2017; Wagenmakers & Dutilh, 2016). Secondly, the records about the study design and analysis plan help the reviewer or user of the study in understanding and assessing the study's findings, because the preregistration provides a structured insight into how the study was thought out and set up (Haven & van Grootel, 2019; Nosek, Ebersole, DeHaven, & Mellor, 2018).

The importance of preregistration is recognized by different stakeholders, such as journals, funders, and research communities. This is reflected in increasing journals and funders attention towards the topic (Arnold Foundation, 2020; Nosek et al., 2015; Open Science Framework, 2020). In addition, different research communities have advocated for preregistration (Bonniaud et al., 2018; Glasziou et al., 2014; Humphreys, Sanchez de la sierra, & Van der windt, 2013; Ioannidis, 2005; Miguel et al., 2014; Munafò et al., 2017; Nosek et al., 2018; Picciotto, 2018) and developed registries based on the type of study (e.g., clinicaltrials.gov for randomized clinical trials, PROSPERO for systematic reviews) or a field of study (e.g., EGAP for governance and politics, RIDIE for development-impact studies). Different forms are used for different types of studies and some platforms welcome preregistrations from different fields using the same form (e.g., aspredicted.org) or have multiple forms available for researchers to choose from (e.g., osf.io).

Preregistering qualitative research may seem counterintuitive as one of the functions of preregistration is to distinguish exploratory and confirmatory research, and qualitative

research is often exploratory by nature (Humphreys et al., 2013; Nosek et al., 2018). Indeed, preregistration may not be useful for all forms of qualitative research. Nonetheless, the usefulness of preregistration may extend more widely than is immediately apparent. For one thing, preregistration may be helpful for forms of qualitative research that involve an element of testing[1]. The posting of predictions in a timestamped registration allows the researcher to credibly communicate to audiences that the hypotheses being assessed were specified in advance of seeing the evidence, and to clearly identify deviations from the plan as exploratory undertakings. Yet preregistration can also bring greater clarity to research that operates on a more exploratory or iterative logic. Even when exploring, scholars usually bring prior knowledge or theoretical preconceptions to the study, and preregistration can help distinguish which aspects of a study's findings drew on -- and which diverged from -- those initial beliefs or expectations. For both sides of the spectrum, preregistration intends to make visible the connections between analytical assumptions, evidence, and decisions that form a particular interpretation of the data.

As the idea of preregistration has been extended to qualitative research, different authors have put forth suggestions for what to include in a qualitative preregistration form (Haven & van Grootel, 2019; Jacobs, 2019; Kern & Gleditsch, 2017; Piñeiro & Rosenblatt, 2016). These suggestions built on existing strategies to foster the credibility of qualitative research, such as the audit trail (Cohen & Crabtree, 2006; Lincoln & Guba, 1985; Miles & Huberman, 1994; Miller, 1997; Schwandt & Halpern, 2011), the decision trial (Koch, 2006), and reporting guidelines (Malterud, 2001; O'Brien, Harris, Beckman, Reed, & Cook, 2014; Tong, Sainsbury, & Craig, 2007).

However, the proposed items for qualitative preregistration thus far are not based on systematic empirical investigations within the qualitative research community. If the

---

[1] A recent review of qualitative articles across leading political science journals found that about 40% explicitly aimed to test hypotheses (Jacobs, 2020).

preregistration form is to benefit the qualitative community, it should reflect what they as a community consider relevant, comprehensive, and comprehensible for qualitative study preregistration (Terwee et al., 2018). Tools to foster transparent reporting in qualitative research, such as reporting guidelines (e.g., COREQ, JARS-APA), are often leveraged when the study is ready to undergo peer review. What is distinctive about preregistration is that it involves the researcher declaring plans for a study design, data collection strategy, and set of empirical expectations in advance, thus allowing for a clearer and more credible delineation between ex ante beliefs and plans, on the one hand, and actual study implementation and ex post conclusions, on the other hand. Unlike an ex post checklist, preregistration makes more transparent to readers how a study unfolded and beliefs shifted over time. In this sense, preregistration can be thought of as a systematic start of a research log (Piñeiro & Rosenblatt, 2016). As such, preregistration acts as a complement to reporting guidelines, although it is unclear what aspects of existing reporting guidelines can be amended for a preregistration form. Hence, we consider it pivotal to develop a preregistration form in close consultation with the qualitative community.

The goal of this study is to gauge and understand what parts of preregistration templates qualitative researchers would find helpful and informative. Given the diversity of epistemological starting points and methods used in qualitative research, we do not expect to be able to develop a form suited for all qualitative scholarship. We tried to accommodate this diversity by inviting qualitative researchers across a broad range of disciplines and approaches to qualitative research, ranging from social constructivism to (post) positivism. We employ the Delphi method to develop a template on which there is high agreement among these surveyed participants, while still retaining meaningful structure. At the same time, the Delphi method can only achieve agreement among participating experts; thus, we can make no claims about how generally useful the form might be to qualitative research

4

communities as a whole. The preregistration form is intended to function as a tool that will aid research planning and communication, not a prescription that hampers the flexible nature of qualitative research (Wagenmakers & Dutilh, 2016). This form will be made available with other preregistration forms available on OSF to foster the credibility and transparency of qualitative inquiry.
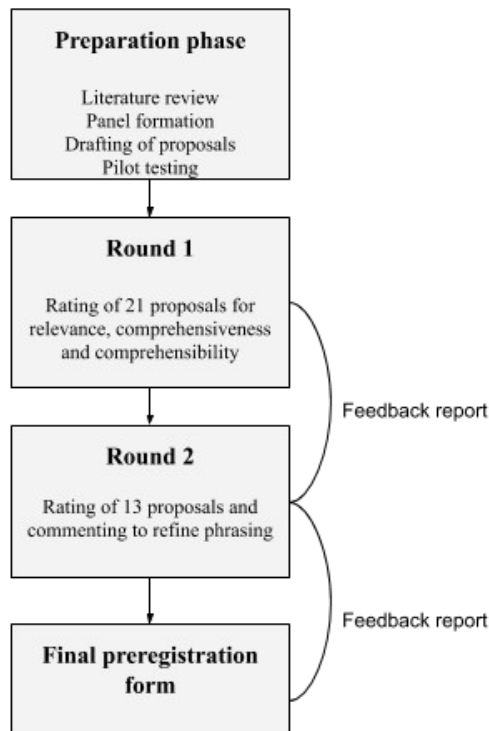
Methods

Ethics

The Institutional Review Board for the Social and Behavioral Sciences at the University of Virginia approved our protocol (# 3397).

Study design

We used an online Delphi study design. Delphi uses repeated questionnaires with feedback reports in between to attain agreement on a particular topic (Okoli & Pawlowski, 2004; Rowe, Wright, & Bolger, 1991). The Delphi design was appropriate because it provided structured group communication in order to find a set of items where there has so far been little convergence on applied principles – the Delphi terminology refers to this as "agreement" of the surveyed participants (Linstone & Turoff, 2002; Powell, 2003). In addition, Delphi's online nature allowed us to reach out to a dispersed group of international researchers and the pseudo-anonymized form precludes particular people from dominating the debate (Graham, Regehr, & Wright, 2003).

Our Delphi consisted of two rounds with feedback reports, see Figure 1. Panelists' feedback was pseudo-anonymized and panelists received no financial compensation.

Figure 1. Flowchart of study design.

5

Preparation phase

Literature review
Panel formation
Drafting of proposals
Pilot testing

Round 1

Rating of 21 proposals for
relevance, comprehensiveness
and comprehensibility

Feedback report

Round 2

Rating of 13 proposals and
commenting to refine phrasing

Feedback report

Final preregistration
form

Feedback report

Panelists

We invited 295 active qualitative researchers and qualitative methodologists to participate. As alluded to above, we wanted to represent this group as broadly as possible and to attain this, we used four different recruitment methods.

First, the steering committee and research team members suggested qualitative researchers or qualitative methodologists from their own networks ($n = 51$, 17% of invitees). Second, we approached authors of various reporting guidelines for qualitative research such as the COREQ (Tong et al., 2007) or the JARS-Qual guidelines (Levitt et al., 2018), because these guidelines inspired our proposed items ($n = 33$, 11% of invitees). Third, we searched Web of Science and Scopus for corresponding authors who published at least 5 qualitative research papers in the last 5 years, because we wanted to get an insiders' perspective from researchers active in, among other fields, biomedicine, social sciences and humanities (the exhaustive list of fields can be found here). This search yielded 213 unique email addresses that we hand-matched with names to see if the person was still active in research. Despite

this, 7 emails bounced ($n = 206$, 70% of invitees). The search block and code for unique email addresses can be found here. Finally, we asked invitees to recommend colleagues ($n = 5$, 2% of total).

Steering committee

The steering committee consisted of the day-to-day team as well as advisory committee members. The day-to-day team implemented the survey (TH with support from WM and TME) in close association with the advisory committee members (KSG, LvG, AJ, FK, BN, RP and FR). The members of the steering committee were trained in different disciplines that make use of qualitative research methodology. The steering committee guided the execution of the study. They helped in selecting potential panelists, advised on the design and content of the questionnaires, and on the analysis of the responses. They did not take part in the Delphi. However, if the panel did not converge after round 2, the steering committee made the final decision about the item.

Creation of the questionnaire

Our strategy to create the list of proposed items was fourfold. First, we integrated suggestions from existing works on preregistration of qualitative research (Haven & van Grootel, 2019; Jacobs, 2019; Kern & Gleditsch, 2017) and pre-analysis plans of qualitative research (Piñeiro & Rosenblatt, 2016). Some authors' suggestions are influenced by the logic of preregistration for quantitative work (Jacobs, 2019; Kern & Gleditsch, 2017), others are influenced by existing strategies to foster the credibility of qualitative research such as a decision trail or audit trail (Denzin & Lincoln, 1998; Koch, 2006).

Second, we searched in PubMed and PsycINFO using the terms "qualitative research" and "preregistration" at the end of January 2020. We included studies if they regarded views

7

on qualitative preregistration and excluded studies that only mentioned preregistration as an example. Articles had to be written in English. PubMed returned 88 hits. After screening the titles and abstracts, we identified one conference abstract that was relevant (Bowers, 2019). One paper discussed open science practices in the field of aging sciences and used the publication by Haven & van Grootel (Haven & van Grootel, 2019) as an example (Isaacowitz & Lind, 2019). PsycINFO returned 10 hits where after screening of the titles and abstracts, 1 paper was relevant that discussed open science in Sport Psychology (Tamminen & Poucher, 2018) focusing on qualitative inquiry.

Third, we downloaded 21 reporting guidelines that EQUATOR listed as bearing relevance to qualitative research (see here)[2]. We included these because preregistrations and reporting guidelines are complementary. Whereas preregistration is intended to help design a study in the most rigorous and transparent way, reporting checklists are intended to write up a study in the most understandable and transparent way. Hence, it is desirable that there is some overlap. Additionally, we hand-searched reporting guidelines for qualitative research in psychology and social sciences and screened the resulting Journal Article Reporting Guidelines for Qualitative research (Levitt et al., 2018) as well as guidelines for qualitative research submitted to the Journal of the Society for Social Work and Research (Wu, Wyant, & Fraser, 2016).

Finally, the above strategies resulted in a document with 32 proposed items that was put forth to the steering committee for review. Steering committee members suggested combining of similar issues (9 items), removal of items that were considered off the subject (e.g., item that regarded ethical testing or reporting) (2 items), and rephrasing of particular items for clarity. The list of 21 remaining proposed items can be found in appendix 1 and

_____

[2] We could not access two guidelines, but have reason to believe we did not miss much as they regarded traffic safety and lower back pain studies.

formed the basis for the first Delphi questionnaire. We used Qualtrics (Qualtrics, Provo, UT, USA) to create and disseminate the survey.

Procedure

Round 1

We sent an invitation email to panelists in February 2020, the email linked to the questionnaire, the study protocol, and the project's website. The questionnaire for round 1 took between 40 and 60 minutes to complete.

Panelists were directed to a personal online survey where they received additional information on the background and content of the questionnaire. They then had to indicate informed consent before starting the actual questionnaire. After the digitally written informed consent, they read the questionnaire instructions.

The questionnaire comprised 21 proposed items (a proposed item consisted of a proposed term and a proposed elaboration, e.g., "Sampling Rationale –Please describe your sampling rationale, typical strategies are maximum variation, purposive, theoretical, convenience, snowball, random or mixed", see appendix 1). These proposed items were split between four parts labelled with the headings: Study information (6 items), Design plan (4 items), Data collection (6 items), and Analysis plan (5 items).

Per proposed item, participants answered four questions: (1) to what extent do you agree with the suggested term, (2) to what extent do you agree with the elaboration of the term, (3) to what extent do you agree that the topic is relevant for qualitative preregistration, and (4) under which heading should the topic be placed?

Questions 1 and 2 regarded comprehensibility and question 3 regarded the relevance of the proposed items. Finally, after having read all the proposed items that belonged to one of the major headings, we enquired whether there was any other topic that they would

9

consider relevant (e.g., "Is there any topic you consider relevant for qualitative preregistration that you would like to add under 'study information'?") to assess comprehensiveness.

Questions 1-3 were answered using a 5-points Likert scale with response options 'Totally agree', 'Somewhat agree', 'Neither agree nor disagree', 'Somewhat disagree', and 'Strongly disagree'. Questions had a text box underneath where participants were encouraged to give arguments for their ratings. For every question, panelists could indicate 'no expertise' as an option. When panelists indicated 'no expertise', their vote did not count for the calculated percentage of agreement (this option was eventually used only twice).

The answer options for question 4 listed the four headings (Study information, Design Plan, Data collection, and Analysis plan) and a fifth option 'Miscellaneous'. Again, participants could indicate 'no expertise' if they felt unable to answer the question.

Finally, we enquired panelists' professional background by asking about their disciplinary field, type of qualitative data they work with, the paradigm that guides their work, and the extent to which they saw the potential of qualitative preregistration, respectively. The full questionnaire for round 1 can be found [here](#).

Round 2

The same 295 participants received the invitational email to participate in the second round in March 2020. Attached to this email they received a feedback report of round 1 with all ratings and arguments from this round provided. In round 2, we had compiled all responses and focused on issues left from round 1 where panelists did not agree on the best phrasings. We revised or removed items in accordance with participants' ratings and arguments. In the questionnaire, we briefly summarized why and how particular arguments influenced revisions. Panelists were then invited to rate and comment on the suggested

revised items. The questionnaire for round 2 took between 15 and 25 minutes to complete and can be found here.

Finally, participants received a feedback report with all ratings and arguments from round 2. The feedback reported listed how many panelists chose a particular answer and included all pseudo-anonymized arguments. If panelists had indicated that they wanted to receive the final preregistration form, this was sent to them after final analyses.

Analyses

We preregistered our protocol on OSF (see here). Results (i.e., ratings and arguments) were analyzed pseudo-anonymously, meaning that we removed all identifying information from the arguments that panelists provided. Rating data are open and can be found here.

We defined 'agreement' as 68% or more of the participating panelists agreeing. This meant that if 68% or more rated "strongly agree" or "somewhat agree" on the five-point Likert scale, we treated that as an item on which participating panelists converged. We chose 68% as an indication that a substantial group of panelists was on board with what is proposed, whilst retaining room for panelists with dissenting views. There is no agreement as to what is the 'best' criterion to use, as these cut-offs are fundamentally arbitrary (Diamond et al., 2014). We reflect on the robustness of our agreement in the discussion.

We revised proposed terms and elaborations that the panel did not agree on based on the arguments provided. Items that had a relevance score below the cut-off were omitted from the next round, as it was our goal to make a concise preregistration form. TH read all arguments and then provided the steering committee with proposed suggestions based on these arguments that committee members commented on.

Pilot-testing of the surveys

11

For round 1, we asked researchers (*n* = 3) that used qualitative research methodology in their work to pilot-test the survey to see if was it was comprehensible. This led to minor modifications in wording to improve clarity of the proposed items or elaborations and gave us an estimate of the expected survey duration. For round 2, we asked one researcher who was familiar with open science and qualitative research to pilot-test the survey. This again led to minor modifications. Pilot testers were not included in the panel and we removed these ratings prior to analysis.

Results

Response rate

From the 295 invitees in total, 48 participated in our Delphi (response rate: 16%). We invited all eligible panelists for both rounds. In round 1, 35 researchers completed the survey (and 4 partially). In round 2, 31 researchers completed the survey (and 4 partially). There were 25 panelists that participated in both rounds.

From the 48 panelists, 19 were researchers from our own network, 3 were authors of reporting guidelines, 23 were actively publishing qualitative researchers, and 3 were suggested by invitees. Demographic characteristics of the respondents appear in Table 1.

Table 1. Demographics of the panelists[1]

Discipline		Type of research		Paradigm	
Anthropology	4	Archival/documentary evidence	22	Critical theory	7
Biomedicine	6	Ethnography	12	Positivism	7
Education	6	Focus groups	26	Post positivism	9
Environmental studies	1	Interviews	40	Social constructivism	23
International relations	1	Observation	17	Other	10
Linguistics	1	Secondary literature	17		
Nursing	9	Self-reports	13		
Political science	10	Other	5		
Psychology	9				
Public health	6				
Sociology	10				
Other	2				

Relevance of the items

In round 1 panelists considered 14 of the 21 proposed items as relevant to include in the preregistration form. The panel indicated 7 items were not relevant enough, of which 2 proposed items (related to the duration of the study and the researcher's background) had relevance scores just below the cut-off value of 68% (65% and 66%, respectively) with mixed arguments. We revised these items in accordance with panelists arguments and they were put forth again in round 2 (see appendix 2), preceded by a summary of panelists' arguments. At times, panelists suggested for different topics to be merged, as they were tapping into similar issues. We combined three items, namely 'sort of sample', 'sampling strategy', and 'sampling rationale', into a single revised item called 'case selection strategy'. Two items, 'origin of data' and 'type of data' were revised into a single item called 'data source and type'. This resulted in 11 items instead of 14 that were presented in round 2.
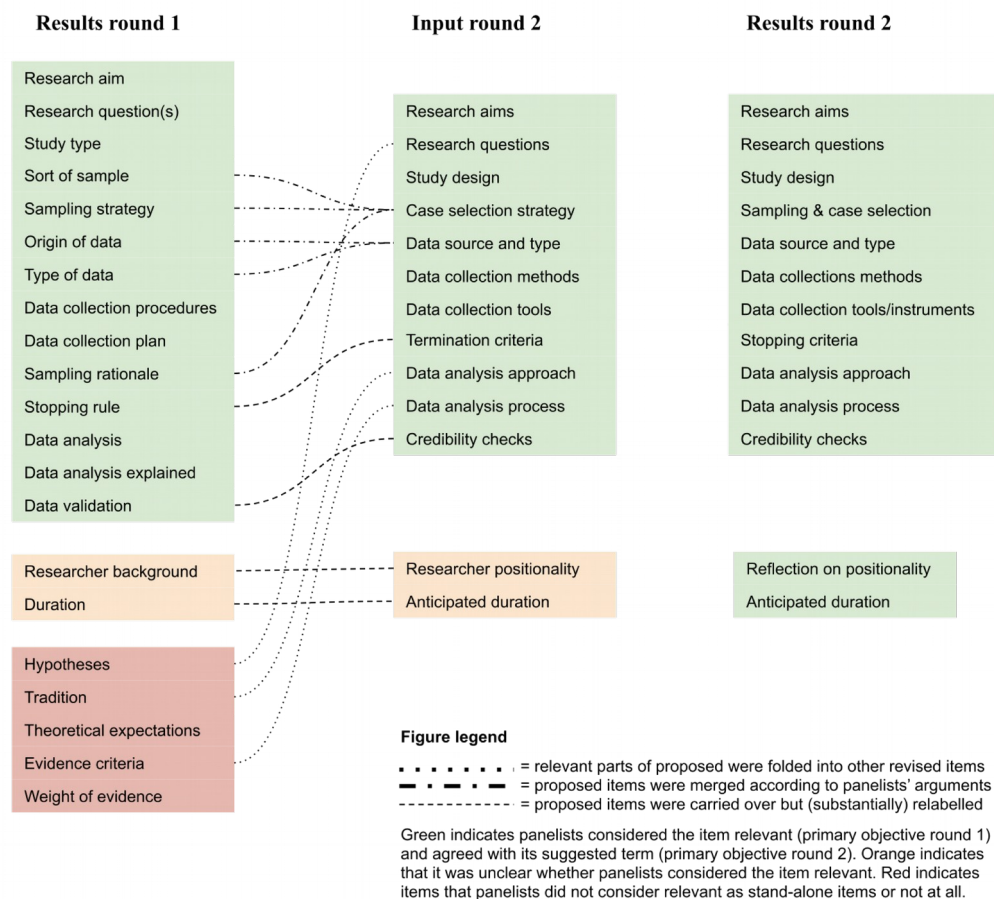
The remaining 5 items had relevance scores between 29% and 60%. Panelists argued that these proposed items were only relevant for a subset of qualitative research that explicitly investigates hypotheses such as process tracing (Collier, 2011). Therefore, we either removed these items or removed them as a stand-alone item and folded them into other elaborations where possible and in accordance with panelists' views (e.g., "hypotheses" got folded into "research questions" to accommodate those who would want to preregister hypotheses).

In round 2, panelists converged on including the two terms that had a 65% or 66% relevance score in round 1. In their revised version, 'researcher(s) positionality and potential biases' and 'anticipated duration' obtained relevance scores of 77 and 87%, respectively. Participants argued that whilst it can be hard to indicate precisely how long a project will

take, knowing the anticipated duration (and time when it was conducted) can be helpful. These were added to the form, so the final form consists of 13 items, see figure 2.

Figure 2. Overview of items through the two rounds.

Results round 1

- Research aim
- Research question(s)
- Study type
- Sort of sample
- Sampling strategy
- Origin of data
- Type of data
- Data collection procedures
- Data collection plan
- Sampling rationale
- Stopping rule
- Data analysis
- Data analysis explained
- Data validation

- Researcher background
- Duration

- Hypotheses
- Tradition
- Theoretical expectations
- Evidence criteria
- Weight of evidence

Input round 2

- Research aims
- Research questions
- Study design
- Case selection strategy
- Data source and type
- Data collection methods
- Data collection tools
- Termination criteria
- Data analysis approach
- Data analysis process
- Credibility checks

- Researcher positionality
- Anticipated duration

Results round 2

- Research aims
- Research questions
- Study design
- Sampling & case selection
- Data source and type
- Data collections methods
- Data collection tools/instruments
- Stopping criteria
- Data analysis approach
- Data analysis process
- Credibility checks

- Reflection on positionality
- Anticipated duration

Figure legend

∙∙∙∙∙∙∙∙∙∙ = relevant parts of proposed were folded into other revised items
▬ ∙ ▬ ∙ ▬ = proposed items were merged according to panelists' arguments
‑ ‑ ‑ ‑ ‑ ‑ = proposed items were carried over but (substantially) relabelled

Green indicates panelists considered the item relevant (primary objective round 1) and agreed with its suggested term (primary objective round 2). Orange indicates that it was unclear whether panelists considered the item relevant. Red indicates items that panelists did not consider relevant as stand-alone items or not at all.

Terms & elaborations

Panelists suggested alternative terms for the different topics addressed. They argued that our initially proposed terminology was either too quantitatively-oriented (e.g., 'data validation') or not in accordance with the qualitative canon (e.g., 'sampling'). Panelists emphasized that our terminology should not clash with existing discipline-specific reporting guidelines for qualitative research and suggested small changes.

Panelists added clarifications or new examples to our suggested elaborations and at times pointed out that the elaborations were somewhat repetitive. In addition, the use of the

14

word "typical" in the elaborations was criticized as it had normative connotations to it. Instead, panelists recommended to use a more refined phrase, such as "examples include, but are not restricted to…".

In round 1, there were 11 proposed items (out of the 13 relevant items) where participants agreed with the suggested terminology. Despite this convergence of opinions, we carefully reviewed all arguments, substantially revising terms ($n = 4$) that the panel just agreed upon and implementing small suggested revisions (e.g., 'data analysis' was revised to 'data analysis approach') for terms ($n = 7$) that the panel firmly agreed upon. For the 2 items where there was agreement that our initially proposed term did not work well, we proposed new terms based on panelists' arguments.

At the end of round 1, panelists also proposed new items. They consistently pointed out that items regarding ethical review, or ethics or ethical approval should be included in a preregistration, as well as funding. The Center for Open Science was working on integrating these items as meta-data fields into OSF registrations, so these were not asked in round 2, but panelists received a note explaining that these items will be added as meta-data fields to the form.

In round 2, panelists converged on 11 of our 13 revised terms, meaning they broadly agreed (agreement ranged from 68% to 100%). Panelists were dissatisfied with the term 'termination criteria' (65%). We received many clear arguments to get rid of the 'termination' part and to use words such as 'stopping' or 'stop' instead.

Panelists also argued that our revised term, 'case selection strategy', in round 2 was unclear (68% agreement) as the term conflated recruitment (where a researcher can use strategies like snowballing) and case selection (where the researcher can use selection criteria, such as typical case, or most deviant case (George & Bennett, 2005)). Furthermore, different panelists suggested to include 'sampling' into the term, because 'sampling' is

15

commonly used to denote the process of selecting participants. In the previous round,

panelists argued that participants are selected in variation to the phenomenon in question and

not selected to representative in the statistical sense. The term now reads 'Sampling & Case

Selection' strategies, see figure 3.

Figure 3. Visualization of the Delphi feedback process.
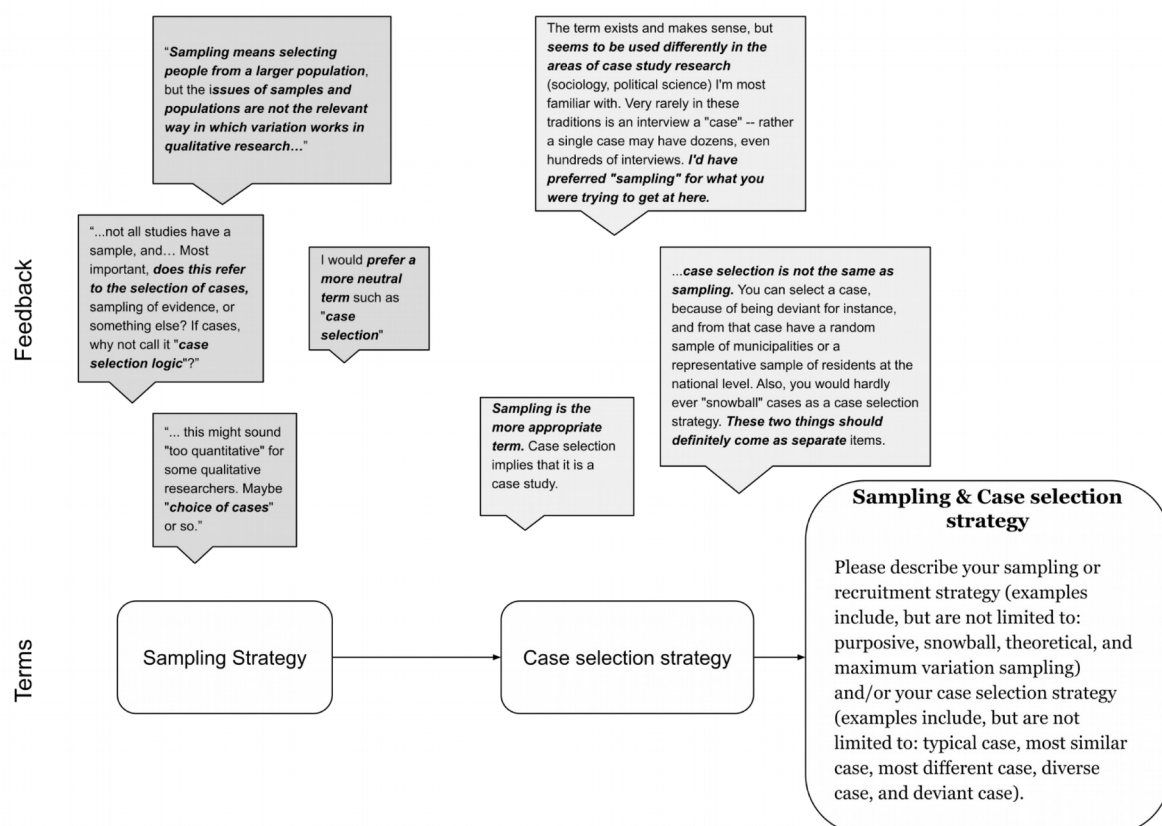
Finally, the revised elaborations yielded high agreement among the panel (most

agreement ranged from 77% to 97%) in round 2. Yet, just 71% of the panelists agreed with

the elaboration for the term 'study design'. Panelists argued that the elaboration's examples

were insufficiently precise to get an indication of the level of detail with which one should

respond to the item. We refined the examples and adjusted the elaboration after round 2,

focusing on examples of designs that are more broadly used (e.g., 'case study'). The final

version of the preregistration form as it will be made available on OSF appears in table 2.

Table 2. Final preregistration form

Heading	Term	Elaboration	*Example*
Study information	Research aim(s)	Please specify the overall purpose(s), objective(s), or aim(s) of the research. If helpful, please select the type(s) of aim. Examples include, but are not limited to: * exploring, * describing, * theory evaluating * comparing *understanding In addition, please reflect on whether your aim is different across different domains (e.g., knowledge generation, policy development, community resourcing). If so, specify your aim for each domain that is relevant for your study.	
	Research question(s)	Please specify your research question(s) as they are guiding your research now. If relevant, you may also specify here any hypotheses to be assessed. The research questions may break down your aim into smaller, distinct inquiries. If relevant, you may distinguish between primary and secondary research questions or hypotheses.	For example, if it is your aim to explore the attitudes of caregivers towards Alzheimer patients in a local ward, your research questions could specify precisely what you plan to study; for instance, how ward staff tries to treat the patients with dignity or how the relationship between the patient and their family members or loved ones evolved since that patient was admitted to the ward.
	Anticipated duration	Please indicate the estimated project start date (mm/yyyy) and estimated project end date (mm/yyyy).	
Design Plan	Study design	Please provide a brief, overarching characterisation of the study design. Your response might consist of a succinct label (e.g., "case study" or "ethnography") and/or a brief elaboration of that label's meaning. A study may involve a combination of different designs, including a mix of quantitative and qualitative methods.	

	Sampling & case selection strategy	Please describe your sampling or recruitment strategy (examples include, but are not limited to: purposive, snowball, theoretical, and maximum variation sampling) and/or your case selection strategy (examples include, but are not limited to: typical case, most similar case, most different case, diverse case, and deviant case). *Please provide a short rationale for why you selected this type of strategy.*	
Data Collection	Data source(s) and data type(s)	Please describe the source(s) and type(s) of data you will be using. In describing the data, distinguish between data that existed prior to your study (e.g., archival documents, newspaper articles, [social] media, secondary literature, or data collected for a different purpose than the current study) and original data (i.e., data that will be collected/generated for the current study).	
	Data collection method(s)	Please describe your method(s) of data collection or data generation. Examples of methods include, but are not restricted to: interviews, focus groups, enabling techniques, self-reports, field notes, diaries, (participative) observation, archival research, or mixed methods. *Please provide a brief rationale for why you plan to use each particular data collection/generation method in your study.*	
	Data collection tool(s), instrument(s), or plan(s)	Please describe or upload the tool(s), instrument(s), or plan(s) you will use in collecting or generating your data. Examples could be, but are not limited to: topic guide, interview questionnaire, focus group guide, observation scheme, creative tools (e.g., photos, videos, musical pieces, paintings, etc.), or a description of your archival search plans.	
	Stopping criteria	Please describe the criteria or rationale behind when you will stop data generation or collection. Possible criteria include, but are not limited to: data saturation*, when inclusion criteria are satisfied,	

18

		resource constraints (e.g., time/funding), or when the analysis has produced an enriching answer to the research question(s). ** We follow Fusch & Ness (2005) and interpret saturation to be reached when there is enough information to replicate the study, the ability to obtain new information has been attained, and further coding is no longer feasible.*	
Analysis Plan	Data analysis approach	Please specify the type and details of your data analysis approach. Examples of approaches include, but are not limited to: narrative analysis, phenomenological analysis, thematic analysis, content analysis, psychoanalytic analysis, grounded theory, process tracing, comparative analysis, or discourse analysis. If multiple interpretations of your approach exist, please specify the version you will be using. *Please provide a rationale for why your selected data analytic approach is appropriate given your study's aim(s).*	For example, if you indicated 'phenomenological analysis', you may want to specify the theorist whose approach you are following, e.g., "We use a phenomenological approach as explained by Colaizzi (1978)"; or if you indicated 'content analysis', a specification could be: "We apply inductive content analysis as described in Elo & Kyngäs (2008)".
	Data analysis process	Please describe what your process of data analysis will look like. Questions to keep in mind could be, but are not limited to: * who will be involved in the data analysis, and in what role? * if relevant, indicate any procedures that will be used to turn "raw" data into analysable form (e.g., a coding scheme) * if relevant, indicate any evidentiary criteria that will be used to assess any hypotheses (e.g., what evidence will count as consistent or inconsistent with a given proposition) * if relevant, what software or analytic tools will you use and how will you use them?	

	Credibility strategies	Please specify the strategies, actions, or measures you will employ to assure methodological integrity. Examples include, but are not limited to: * member checking * triangulation with other data sources * bringing in different perspectives * have different researchers analyse the data * consensus building among team members or 'interrater reliability' * negative case analysis * peer debriefing * cross-checks for rivalling explanations * bring in an 'auditor' * reflexivity * verisimilitude, * emotionality, * personal responsibility, * an ethic of caring, * political praxis, * multivoiced texts, * dialogues with subjects *Please provide a short rationale for why you selected particular strategies and how they are appropriate given your study's aim(s) and approach.*	
Miscellaneous	Reflection on your positionality (optional)	Feel free to reflect on your relation to* or association with the studied phenomenon and your position in the research setting/field, including your academic/personal standpoints, assumptions and values. In addition, if there is a potential conflict of interest that can arise, you may want to report that here. ** whether you have a previous relationship with the studied phenomenon and if you consider that there are previous positions or assumptions that may influence the present study.*	For example, if you study the lives of detained immigrants, you might want to talk about your political viewpoints, experience working with detained immigrants, relevant policy work, or perhaps your own experience as a detained immigrant.

Discussion

We developed a form for the preregistration of qualitative studies in consultation with

a panel of qualitative researchers. The form consists of 13 items spread over four headings

and will be made available as a registration option on OSF to enable it to be broadly accessible to be used and piloted.

Comparison with existing literature

The preregistration form presented here differs from existing suggestions for preregistration forms (see here) for qualitative preregistration in two ways. First, it can be used for a broader range of qualitative approaches, as it has been designed with experts from different disciplines and traditions. Second, it uses the same overall structure of four headings as the existing OSF preregistration forms, although the heading labels are slightly different.

The proposed preregistration form overlaps with parts of reporting guidelines for qualitative research, such as the Consolidated Criteria for Reporting Qualitative Research (COREQ) in biomedical research (Tong et al., 2007) and the Journal Article Reporting Standards (JARS) from the American Psychological Association (Levitt et al., 2018). To illustrate, COREQ enquires whether there was an established relationship between the researcher and participants prior to the study (COREQ, p. 352, Table 1), and we pose a similar question when asking the user to reflect on their relation to or association with the participants or phenomenon they will study. Similar to JARS, we ask the user to provide an overview of the research design and a rationale for applying this particular design (JARS, p. 34, Table 1). These similarities will be useful as ideally study preregistration facilitates reporting of the final findings in accordance with reporting guidelines.

Limitations

Our response rate of 16% might seem too low to base any broad agreement-related conclusions on; however, Boulkedid and colleagues found that the median number of Delphi panelists was 17 (Boulkedid, Abdoul, Loustau, Sibony, & Alberti, 2011), whereas ours was

nearly double. One reason for this low response rate is that we combined invitation and inquiry of interest, whereas some Delphi authors first poll their potential invitees for interest in participating and then only invite those who expressed interest, leading to a higher response rate.

We did not reach our predefined criterion (68%) for one proposed item. Some may argue that staying true to the Delphi approach would involve a third round. However, we only planned two rounds and it is known that response rates tend to decline after more than two rounds (Hasson, Keeney, & McKenna, 2000). We felt confident that a third round was not necessary because panelists provided very clear arguments in round 2 for how the term needed to be revised that we implemented accordingly.

Strengths

We attained high agreement for the items in round 2, despite the fact that our panel was very diverse (i.e., they worked in a broad range of disciplines and employed different approaches in their research, and ground their work in widely differing epistemological perspectives and paradigms (Creswell, 2014)) and slightly skewed toward non-positivists. We regard the latter as an advantage, given that the discussion about preregistration tends to involve empiricists interested in generating production transparency, and thus critical insights from non-positivists may broaden the appeal of the preregistration form. That said, we acknowledge that the panel setup may also have led to the exclusion of some more positivist terms.

This is the first preregistration form that was developed in direct consultation with the potential user-community that can now be piloted. Other preregistration templates were developed either by the registries themselves or by small research teams that (informed by the literature) published their suggested templates, such as van 't Veer & Giner-Sorolla

preregistration template for social psychology (van 't Veer & Giner-Sorolla, 2016) or Brandt and colleagues replication recipe preregistration template (Brandt et al., 2014). These templates were peer reviewed, so to some extent the community was involved in their development, but their views were not explicitly solicited in the template's design.

Future research

Note that a majority agreement-based form is not per se the best form (Jünger, Payne, Brine, Radbruch, & Brearley, 2017). In a similar vein, the qualitative study preregistration form proposed here may not be the best fit for a specific study. Future research could look into 'modules' that could be added to this form and that would be relevant for their specific type of inquiry, but may not be relevant for qualitative inquiry across the board. For example, more deductively oriented approaches such as process tracing may be able to define specific criteria for how to 'weigh' evidence before commencing data collection.

There is currently no empirical evidence to show that preregistering qualitative research will improve the quality of the work. A first step would be collecting good examples of qualitative study preregistration, two early examples of qualitative preregistration can be found [here](#) and [here](#). Crucially, the developed form here will have to be piloted to test individual parts and modules for applicability, and to understand where the form needs to be expanded or reformulated.

Implications

It is pivotal to assure that the strength of qualitative research, which is its flexibility to adapt, adjust and respond, is not lost in preregistration (Tong et al., 2007). Analogous to quantitative preregistration, a preregistration is a plan, not a prison (DeHaven, 2017). The preregistration should provide a systematic starting point that can be updated as the study

23

evolves. Although further investigation is needed, the use of this qualitative preregistration form can facilitate more transparent and credible research for various types of qualitative research.

Conclusion

We showed that despite the large variety of qualitative research methodology, it is possible to develop a template where the participating qualitative researchers converged on an opinion of what to include in a qualitative study preregistration template. This template can now be piloted, taking an important step concerning the place of qualitative research in the broader open science debate.

References

ArnoldFoundation. (2020). Guidelines for investment in research. Retrieved March 3, 2020, from https://www.arnoldventures.org/guidelines-for-investments-in-research

Bonniaud, P., Fabre, A., Frossard, N., Guignabert, C., Inman, M., Kuebler, W. M., … Kolb, M. (2018). Optimising experimental research in respiratory diseases: An ERS statement. *European Respiratory Journal*, *51*(5). https://doi.org/10.1183/13993003.02133-2017

Boulkedid, R., Abdoul, H., Loustau, M., Sibony, O., & Alberti, C. (2011). Using and reporting the Delphi method for selecting healthcare quality indicators: A systematic review. *PLoS ONE*, *6*(6). https://doi.org/10.1371/journal.pone.0020476

Bowers, B. (2019). Pre-registering qualitative research: benefits, unintended consequences, and unanswered questions. *Innovation in Aging*, *3*(S1), 2019.

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., … van 't Veer, A. (2014). The Replication Recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, *50*(1), 217–224. https://doi.org/

10.1016/j.jesp.2013.10.005

Cohen, D., & Crabtree, B. (2006). Qualitative research guidelines project. Retrieved March 3,

    2020, from http://www.qualres.org/HomeAudi-3700.html

Collier, D. (2011). Understanding process tracing. *PS - Political Science and Politics*, *44*(4),

    823–830. https://doi.org/10.1017/S1049096511001429

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed method*

    approaches (4th ed). Los Angeles: Sage.

DeHaven, A. (2017). Preregistration: A plan, not a prison. Retrieved April 14, 2020, from

    https://cos.io/blog/preregistration-plan-not-prison/

Denzin, N., & Lincoln, Y. (1998). *Handbook of qualitative research* (Paperback). Thousand

    Oaks, CA SE.: Sage Publications.

Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M., &

    Wales, P. W. (2014). Defining consensus: A systematic review recommends

    methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology*,

    67(4), 401–409. https://doi.org/10.1016/j.jclinepi.2013.12.002

George, A. L., & Bennett, A. (2005). *Case studies and theory development in the social*

    sciences. Cambrigde: MIT Press.

Glasziou, P., Altman, D. G., Bossuyt, P., Boutron, I., Clarke, M., Julious, S., … Wager, E.

    (2014). Reducing waste from incomplete or unusable reports of biomedical research.

    The Lancet, *383*(9913), 267–276. https://doi.org/10.1016/S0140-6736(13)62228-X

Graham, B., Regehr, G., & Wright, J. G. (2003). Delphi as a method to establish consensus

    for diagnostic criteria. *Journal of Clinical Epidemiology*, *56*(12), 1150–1156.

    https://doi.org/10.1016/S0895-4356(03)00211-7

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey

    technique. *Journal of Advanced Nursing*, *32*(4), 1008–1015.

https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x

Haven, T. L., & van Grootel, L. (2019). Preregistering qualitative research. *Accountability in Research*, *6*(3), 1–16. https://doi.org/10.1080/08989621.2019.1580147

Humphreys, M., Sanchez de la Sierra, R., & Van der windt, P. (2013). Fishing, commitment, and communication: A proposal for comprehensive nonbinding research registration. *Political Analysis*, *21*(1), 1–20. https://doi.org/10.1093/pan/mps021

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), 0696–0701. https://doi.org/10.1371/journal.pmed.0020124

Isaacowitz, D. M., & Lind, M. (2019). Open science is for aging research, too. *Innovation in Aging*, *3*(4), 1–10. https://doi.org/10.1093/geroni/igz028

Jacobs, A. M. (2019). Of bias and blind Selection: Pre-registration and results-free review in observational and qualitative Research. In *The Production of Knowledge: Enhancing Progress in Social Science* (pp. 1–40). Cambridge University Press. https://doi.org/10.1017/CBO9781107415324.004

Jünger, S., Payne, S. A., Brine, J., Radbruch, L., & Brearley, S. G. (2017). Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: Recommendations based on a methodological systematic review. *Palliative Medicine*, *31*(8), 684–706. https://doi.org/10.1177/0269216317690685

Kern, F. G., & Gleditsch, K. S. (2017). *Exploring pre-registration and pre-analysis lans for qualitative inference*. *Pre-print*.

Koch, T. (2006). Establishing rigour in qualitative research: the decision trail. *Journal of Advanced Nursing*, *53*(1), 91–103. https://doi.org/10.1111/j.1365-2648.2006.03681.x

Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., Josselson, R., & Suárez-Orozco, C. (2018). Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The APA publications and

communications board task force report. *American Psychologist*, *73*(1), 26–46. https://doi.org/10.1037/amp0000151

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

Linstone, H. A., & Turoff, M. (2002). *The Delphi Method: Techniques and Applications*. https://doi.org/10.1007/s00256-011-1145-z

Malterud, K. (2001). Qualitative research: Standards, challenges, and guidelines. *Lancet*, *358*(9280), 483–488. https://doi.org/10.1016/S0140-6736(01)05627-6

Miguel, E., Camerer, C., K., C., Cohen, J., Esterling, K. M., Gerber, A., … Van der Laan, M. (2014). Promoting transparency in social science. *Science*, *343*(6166), 30–31. https://doi.org/10.1021/nn300902w

Miles, M. B., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA SE: Sage.

Miller, D. L. (1997). One strategy for assessing the trustworthiness of qualitative research: Operationalizing the external audit. *American Educational Research Association Annual Conference*, 1–50.

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie Du Sert, N., … Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 1–9. https://doi.org/10.1038/s41562-016-0021

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine*, *89*(9), 1245–1251. https://doi.org/10.1097/ACM.0000000000000388

Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: An example, design considerations and applications. *Information and Management*, *42*(1), 15–29. https://doi.org/10.1016/j.im.2003.11.002

OSF. (2020). Top Implementors. Retrieved March 3, 2020, from https://osf.io/2sk9f/?_ga=2.41262050.1198546445.1583248675-1434101306.1582838143

Picciotto, M. (2018). Analytical transparency and reproducibility in human neuroimaging studies. *Journal of Neuroscience*, *38*(14), 3375–3376. https://doi.org/10.1523/JNEUROSCI.0424-18.2018

Piñeiro, R., & Rosenblatt, F. (2016). Pre-Analysis plans for qualitative research. *Revista de Ciencia Política*, *36*(3).

Powell, C. (2003). The Delphi technique: Myths and realities. *Journal of Advanced Nursing*, *41*(4), 376–382. https://doi.org/10.1046/j.1365-2648.2003.02537.x

Rowe, G., Wright, G., & Bolger, F. (1991). Delphi: A reevaluation of research and theory. *Technological Forecasting and Social Change*, *39*(3), 235–251. https://doi.org/10.1016/0040-1625(91)90039-I

Schwandt, T., & Halpern, E. (2011). Linking Auditing and Metaevaluation. *Linking Auditing and Metaevaluation*, 17–41. https://doi.org/10.4135/9781412984768.n1

Tamminen, K. A., & Poucher, Z. A. (2018). Open science in sport and exercise psychology: Review of current approaches and considerations for qualitative inquiry. *Psychology of Sport and Exercise*, *36*(August 2017), 17–28. https://doi.org/10.1016/j.psychsport.2017.12.010

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., De Vet, H. C. W., Westerman, M. J., Patrick,

D. L., … Mokkink, L. B. (2018). COSMIN standards and criteria for evaluating the

content validity of health-related Patient-Reported Outcome Measures: a Delphi study.

Qual Life Res, *27*, 1159–1170.

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative

research (COREQ): a 32-item checklist for interviews and focus groups. *International*

Journal for Quality in Health Care : Journal of the International Society for Quality in

Health Care, *19*(6), 349–357. https://doi.org/10.1093/intqhc/mzm042

van 't Veer, A. E., & Giner-Sorolla, R. (2016). Pre-registration in social psychology—A

discussion and suggested template. *Journal of Experimental Social Psychology*, *67*, 2–

12. https://doi.org/10.1016/j.jesp.2016.03.004

Wagenmakers, E.-J., & Dutilh, G. (2016). Seven selfish reasons to preregister. Retrieved

March 3, 2020, from https://www.psychologicalscience.org/observer/seven-selfish-

reasons-for-preregistration

Wu, S., Wyant, D. C., & Fraser, M. W. (2016). Author guidelines for manuscripts reporting

on qualitative research. *Journal of the Society for Social Work and Research*, *7*(2), 405–

425. https://doi.org/10.1086/685816