

Exploring the Potential for a Consolidated Standard for Reporting Guidelines for Qualitative Research: An Argument Delphi Approach

Karin Hannes¹, Mieke Heyvaert², Karin Slegers³, Stef Vandenbrande⁴, and Marc Van Nuland⁵

Abstract

Background: Consolidating a standard for reporting qualitative research remains a challenging endeavor, given the variety of different paradigms that steer qualitative research as well as the broad range of designs, and techniques for data collection and analysis that one could opt for when conducting qualitative research.

Method: A total of 18 experts in qualitative research participated in an argument Delphi approach to explore the arguments that would plead for or against the development and use of reporting guidelines (RGs) for qualitative research and to generate opinions on what may need to be considered in the further development or further refinement of RGs for qualitative research.

Findings: The potential to increase quality and accountability of qualitative research was identified as one of the core benefits of RGs for different target groups, including students. Experts in our pilot study seem to resist a fixed, extensive list of criteria. They emphasize the importance of flexibility in developing and applying such criteria. Clear-cut RGs may restrict the publication of reports on unusual, innovative, or emerging research approaches.

Conclusions: RGs should not be used as a substitute for proper training in qualitative research methods and should not be applied rigidly. Experts feel more comfortable with RGs that allow for an adaptation of criteria, to create a better fit for purpose. The variety in viewpoints between experts for the majority of the topics will most likely complicate future consolidation processes. Design specific RGs should be considered to allow developers to stay true to their own epistemological principles and those of their potential users.

Keywords

reporting guidelines, Delphi study, qualitative research

Background

Reporting guidelines (RGs) have successfully been developed and disseminated for experimental research (Schulz, Altman, & Moher, 2010), longitudinal research (Von Elm et al., 2008), and nonrandomized research (Des Jarlais, Lyles, & Crepaz, 2004). Such RGs include criteria on how to report the background, methods, findings, and discussion of original research projects. The advantages of RGs are twofold. First, they assist the research community in achieving consistency between research reports (Tate & Douglas, 2011). RGs provide authors with clear instructions on what type of information should be included in the report. It follows that RGs can serve as a guide to protocol development for qualitative studies. Second, a

¹ Centre for Sociological Research, Faculty of Social Sciences, KU Leuven, Belgium

² Methodology of Educational Sciences Research Group, Faculty of Psychology and Educational Sciences, KU Leuven, Belgium

³ Institute for Media Studies, Faculty of Social Sciences, KU Leuven, Belgium

⁴ Research Unit of Theological and Comparative Ethics, Faculty of Theology and Religious Studies, KU Leuven, Belgium

⁵ Academic Centre for General Practice, Faculty of Medical Sciences, KU Leuven, Belgium

Corresponding Author:

Karin Hannes, Centre for Sociological Research, Faculty of Social Sciences, KU Leuven, Belgium.

Email: karin.hannes@kuleuven.be



Table 1. Core List of Standards for Reporting Original Research Retrieved From the EQUATOR Network Online Resource.

Reporting guidelines	Study design or method	Developed by
CONSORT	Randomized controlled trials	Schulz, Altman, & Moher, 2010
TREND	Nonrandomized trials	Des Jarlais, Lyles, Crepez, & The TREND Group, 2004
STROBE	Observational research	Von Elm et al., 2008
STARD	Diagnostic research	Bossuyt et al., 2003
COREQ	Qualitative research	Tong, Sainsbury, & Craig, 2007
SRQR	Statistical analysis and methods	O'Brien et al., 2014
SQUIRE	Quality improvement reporting	Ogrinc et al., 2008
CHEERS	Economic evaluations	Husereau et al., 2013
CARE	Clinical case reporting	Gagnier et al., 2013

Note: CONSORT = Consolidated Standards of Reporting Trials, STROBE = Strengthening the Reporting of Observational Studies in Epidemiology, STARD = Standards for Reporting of Diagnostic Accuracy, COREQ = Consolidated criteria for reporting qualitative research, SQUIRE = Standards for QQuality Improvement Reporting Excellence, CHEERS = Consolidated Health Economic Evaluation Reporting Standards, CARE = Clinical Case Reporting Guideline, SRQR = Standards for reporting qualitative research.

detailed audit trail of methodological choices made by authors facilitates critical appraisal of methodological quality and adequacy of content of original studies (Hannes, Lockwood & Pearson, 2010). The latter is particularly beneficial for authors of systematic reviews. Systematic reviews aim to provide an exhaustive summary of literature relevant to a particular research question by identifying, appraising, selecting, and synthesizing results or findings from original studies with particular attention to minimizing bias (Montori, Wilczynski, Morgan, & Haynes, 2003). Transparency in reporting enables them to judge whether a study has been conducted according to the state of the art or not. As a consequence, several RGs have been developed by researchers involved in conducting systematic reviews. Most of them can be retrieved from the website of the EQUATOR network.¹ This website serves as an international forum for the development of RGs. A core list of RGs for reporting original research as identified by the EQUATOR network is presented in Table 1.

Many of these RGs have been developed based on a consensus between international, methodological experts, mainly from health-care disciplines and are supported by a broad range of researchers. Several RGs have been adopted by high-impact scientific journals, including, for example, the *Lancet* and the *BMJ*. Previous studies have shown that RGs increase the transparency of reporting for quantitative research designs (Bastuji-Garin et al., 2013; Möhler, Bartoszek, & Meyer, 2013). We expect that RGs supporting qualitative researchers would have a beneficial effect on the transparency of qualitative research reports. The Consolidated criteria for reporting qualitative research (COREQ) statement developed by Tong, Sainsbury, and Craig (2007) and the SRQR statement developed by O'Brien and colleagues provide guidance to researchers drawing on qualitative methods, more specific interviews, and focus groups. To date, these statements have not been subject to a formal consensus procedure among experts in qualitative research, hence adoption by researchers and major journals has

been limited. Consolidating a standard for reporting qualitative research remains a challenging endeavor, given the variety of different paradigms that steer qualitative research as well as the broad range of designs, data collection, and analysis techniques that one could opt for when conducting qualitative research. It is therefore important to identify potential areas of agreement and conflict among qualitative researchers, in an attempt to produce RGs that have a high degree of acceptance. A growing number of systematic reviews of qualitative evidence has been published in recent years (Hannes & Macaitis, 2012). These syntheses draw on the findings of original qualitative research. The uptake of RGs for qualitative research may contribute to the overall quality of reporting in the field of qualitative research methodology and as such facilitate authors conducting qualitative evidence syntheses. This paper contributes to the current discourse on the value of RGs for qualitative research.

Objectives and Research Questions

We conducted a pilot study to explore the potential for a consensus on RGs for qualitative research, by consulting experts familiar with this type of research. We further aimed to identify aspects that are important to consider when developing guidance for qualitative inquiry, both from a generic point of view and from a design-specific point of view. Commonly used designs include but are not limited to grounded theory, case study approaches, ethnographic study designs, biographical, phenomenological, arts-based, and narrative approaches. The lead question in this research study was what are the arguments that would plead for or against the development and use of RGs for qualitative research? In addition, we were interested in the experts' ideas on what may need to be considered in the further development of RGs for qualitative research. The study does not involve any invasive manipulations on human subjects and was therefore exempted from mandatory ethical review in Belgium. It follows that the proposal has not been subjected to ethical approval prior to the launch of the research project. The protocol has been examined post hoc by the ethical review board of the faculty of psychology and educational sciences

¹ <http://www.equator-network.org/>

Table 2. Definition of Key Terms Used in the First Questionnaire (Based on Pearson, Wiechula, Court, & Lockwood, 2005).

Key term	Definition
Effectiveness	Effectiveness is the extent to which a reporting guideline or standard, when used appropriately, achieves an intended effect. It is about the relationship between the guidance and a potential positive or negative outcome of it.
Feasibility	Feasibility is the extent to which a reporting guidance might be practical and practicable. It also refers to whether or not using such guidance is possible within the context of qualitative research (opportunities as well as threats that relate to its potential success or failure).
Appropriateness	Appropriateness is the extent to which such guidance might fit with or is apt to fit in current (methodological, epistemological and ethical) debates on and trends in qualitative research.
Benefits	A benefit is the extent to which reporting guidance might be helpful, advantageous, or good to something or someone. It refers to persons who might benefit from reporting standards as well as, e.g., the nature and field of qualitative research itself.
Meaningfulness	Meaningfulness is the extent to which the idea of a reporting standard is positively or negatively experienced by you. Meaningfulness further relates to your personal opinions, values, thoughts, beliefs, and interpretations.

per the requirements of the journal. The board is formally unable to provide retroactive approval of a research protocol but did not see any reason why the protocol at hand would not receive favorable evaluation in keeping with present rules and regulations.

Method

Research Team

Our multidisciplinary research team consisted of two educational scientists, a marketing researcher and one researcher from each of the following disciplines: health care, social sciences, theology/religious studies, and criminology. All of them have been trained in qualitative research methods. The team was responsible for the selection of experts, the construction of the questionnaires, the feedback sessions to respondents, the data collection and analysis process, and the critical input in every phase of the research project.

Selection of Experts

We approached approximately 30 experts in qualitative research, based on personal knowledge of expert profiles within our respective fields of science. Eighteen experts agreed to participate in the study. The main reasons for experts to decline were a lack of time, personal issues, or a critical attitude toward the Delphi procedure. The experts represented a broad range of different disciplines, such as education, social sciences, criminology, health care, history, literacy, arts, and architecture. A maximum variety in methodological expertise was achieved on three different levels. Firstly, the experts differed in research paradigm or school of thought: postpositivism, interpretivism, constructivism, critical theory, and pragmatism. Secondly, they covered a rich pallet of methodologies, including grounded theory, systematic review methodology, phenomenology, action research, arts-based methods, ethnography, and mixed methods. Thirdly, the experts mastered different techniques for data collection: focus groups, observation, interviews, and unobtrusive measures such as documents.

Delphi Procedure

An argument Delphi approach was chosen to facilitate the research process. The main advantage of such an approach is that it can stimulate anonymous discussion and bridge the geographical distance between the different experts (Hasson, Keeney, & McKenna, 2000). It is meant to identify ideas, commonalities, and differences in opinions between experts. Experts are also stimulated to evaluate items in terms of pros and cons. Furthermore, counterproductive group processes such as monopolization of the discourse, marginalization of deviant opinions, dominant positions of certain authorities, and group thinking can be avoided. The argument Delphi technique is characterized by three important features: (a) the researchers draw on the knowledge and experience of experts in the field, (b) the method is an iterative process consisting of several survey rounds, and (c) the group interaction process is anonymous and runs via questionnaires. Two Delphi rounds were conducted. Both of them followed a similar procedure: (a) development of the questionnaire, (b) sending it out to the experts, and (c) analyzing the answers. In the first round, the questionnaire was open ended to allow us to generate qualitative data. In the second round, the questionnaire contained a Likert-type scale to allow us to generate percentages of agreement and disagreement among experts. Both questionnaires were sent out to the experts via mail and returned to an administrator. Identification information was removed before it was delivered to the research team for analysis. A small financial token was provided to experts that completed the Delphi exercise.

First Delphi round. For the first Delphi round, the research team developed an open-ended questionnaire that explored how the experts generally felt about RGs for qualitative research, either in general or for design-specific issues. General RGs were defined as a set of general criteria that holds true for different methods, methodologies, designs, or paradigms of qualitative research. Specific RGs were defined as a set of specific criteria for each particular method, methodology, design, or paradigm. Experts were asked to provide us with their viewpoint on potential positive and negative aspects of both types of RGs and the conditions under which they would use such RGs. We

also asked them to indicate for which particular qualitative methodologies, methods, or approaches these guidelines should specifically be considered or, on the other hand, might be counterproductive. Finally, we asked for their opinion on the potential effectiveness, feasibility, appropriateness, benefits, and meaningfulness of RGs, in terms of both advantages and disadvantages. For a definition of these five key concepts, we refer to Table 2. Their answers to the open questions generated qualitative data used for further analysis.

Data analysis first Delphi round. The analysis of the responses from the first Delphi round was descriptive in nature. Statements provided by an expert on a particular topic were extracted by two independent researchers from the open text boxes. These statements were then cross compared by at least three researchers. When statements from different experts were roughly addressing the same content or idea, we simplified them by removing duplicates. The statement that provided the best metaphors or the clearest description of ideas was adopted for the questionnaire. In case of doubt about the similarity in meaning between statements, both statements were withheld for the second round of the Delphi procedure. Apart from this adoption strategy, we applied an adaptation strategy, rephrasing particular statements for clarity due to the complexity of understanding, the use of jargon, multiple layers of meaning in one statement, or style issues. Statements that were irrelevant in terms of our research objective, ambiguous, too vague, or too broadly defined were omitted from the study.

Second Delphi round. The objective of the second round in this “argument Delphi study” was to feed back to the whole expert panel the rich pallet of information obtained after the first round and to have it judged by all experts. It allowed us to validate individual opinions and meanings of individual experts. In the first part of the questionnaire we listed the advantages and disadvantages of RGs. It also included a list with conditions that might facilitate the uptake or use of RGs. The experts were asked to mark all items they agreed on and comment on the items they disagreed on. In the second part, we addressed the form of the RG as well as some process-related issues to be considered for the development and implementation of RGs. These issues were presented as a comprehensive list of items that were subject to the approval or disapproval of the experts.

Data analysis second Delphi round. The analysis of the responses to the questionnaires from the second Delphi round was quantitative. From the answers to our questionnaires, we calculated percentages of agreement and disagreement between experts for each of the topics or statements that were included. Detailed results are discussed in the findings section and are depicted in Figures 1–7. We did not select an a priori cutoff point for agreement or disagreement. Attention was paid to the distribution of all responses. Issues that (a) generated large agreement between experts, (b) were in line with previous research, or (c) were “unexpected” were picked up for further discussion.

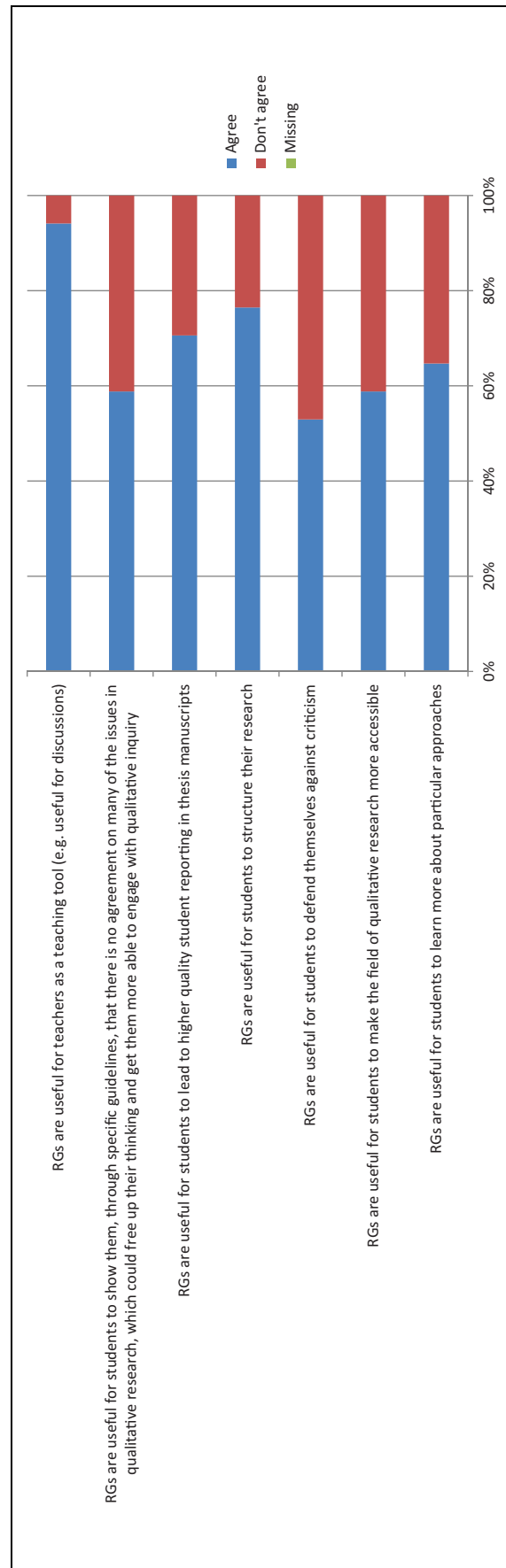


Figure 1. Reporting guidelines for qualitative research useful for students/teachers?

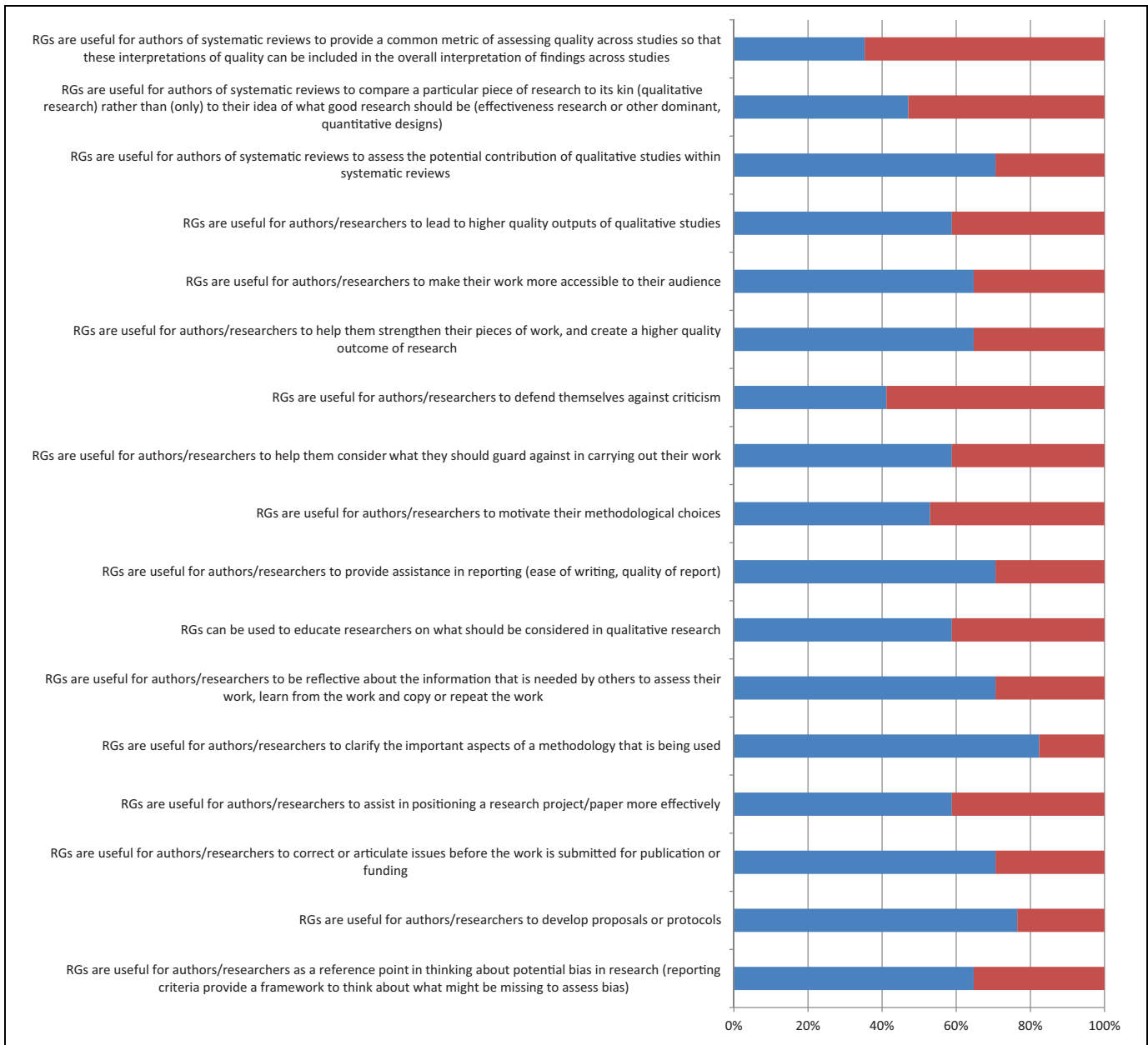


Figure 2. Reporting guidelines for qualitative research useful for authors/researchers?

Findings

Participants

We provided an outline of our study to all experts who were invited to participate. A positive response to our invitation was considered as a consent to participate. Consent to be acknowledged as an expert in the paper was achieved via a tick box (Yes/No) included in the demographic information sheet handed out to each of the participants at the start of the project. Of the 18 experts who agreed to participate, 67% were female. The age of the experts ranged from 24 to 67. The experts in our sample further represented six different countries: Australia, Belgium, Canada, the Netherlands, the United Kingdom, and the United

States. The average proportion of the time the experts spent on qualitative research ranged from 20% to 95%. Their experience with qualitative research methods ranged from 2 to 40 years, suggesting that junior and senior profiles were represented in the sample.

First Delphi Round

The experts identified four different situations in which RGs may be considered: (a) never; (b) for all approaches; (c) only for well-known approaches to qualitative research such as grounded theory, document analysis, ethnographic, or phenomenological research; or (d) only for specific parts in a research

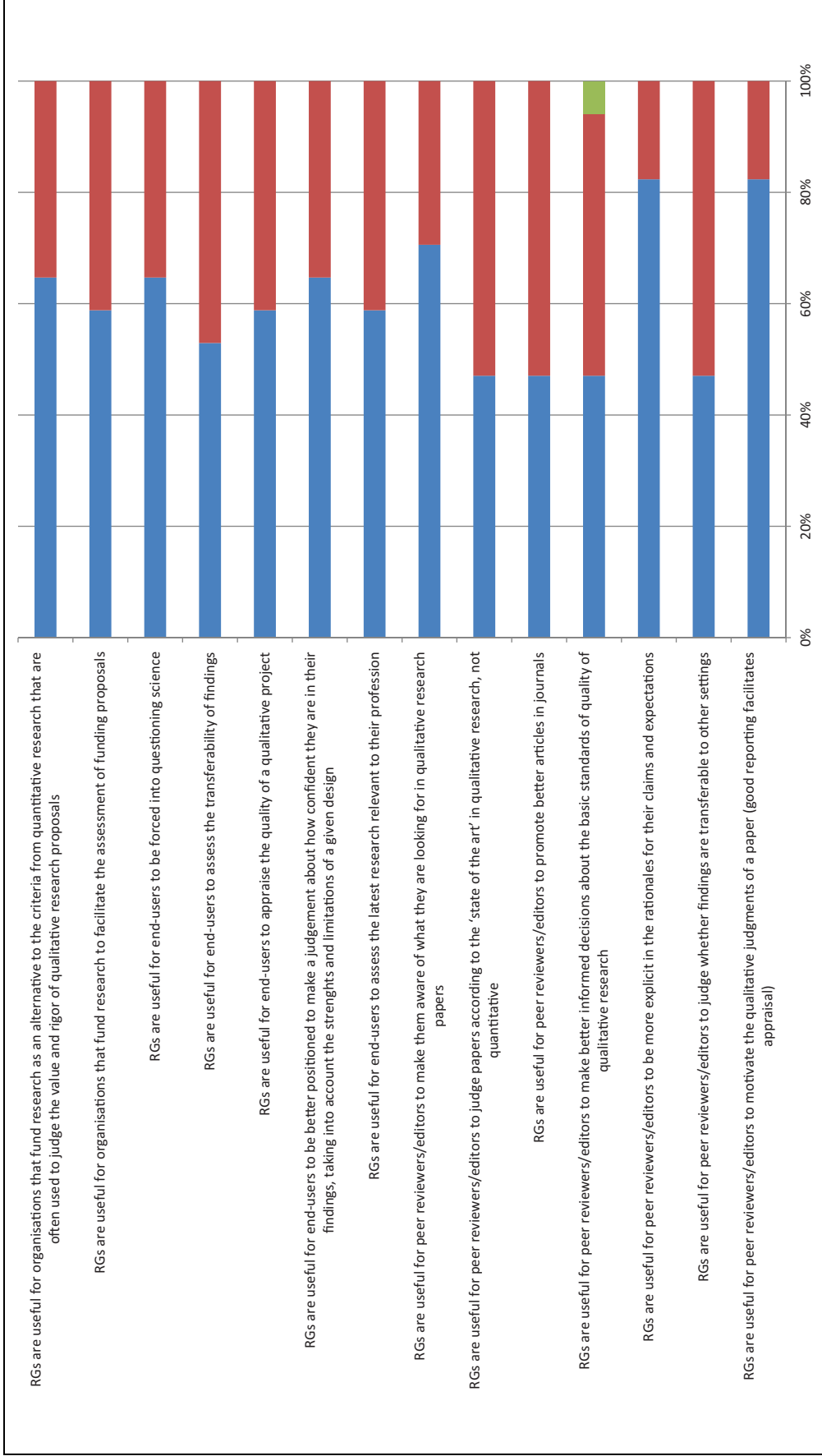


Figure 3. Reporting guidelines for qualitative research useful for reviewers/editors/funders/end users?

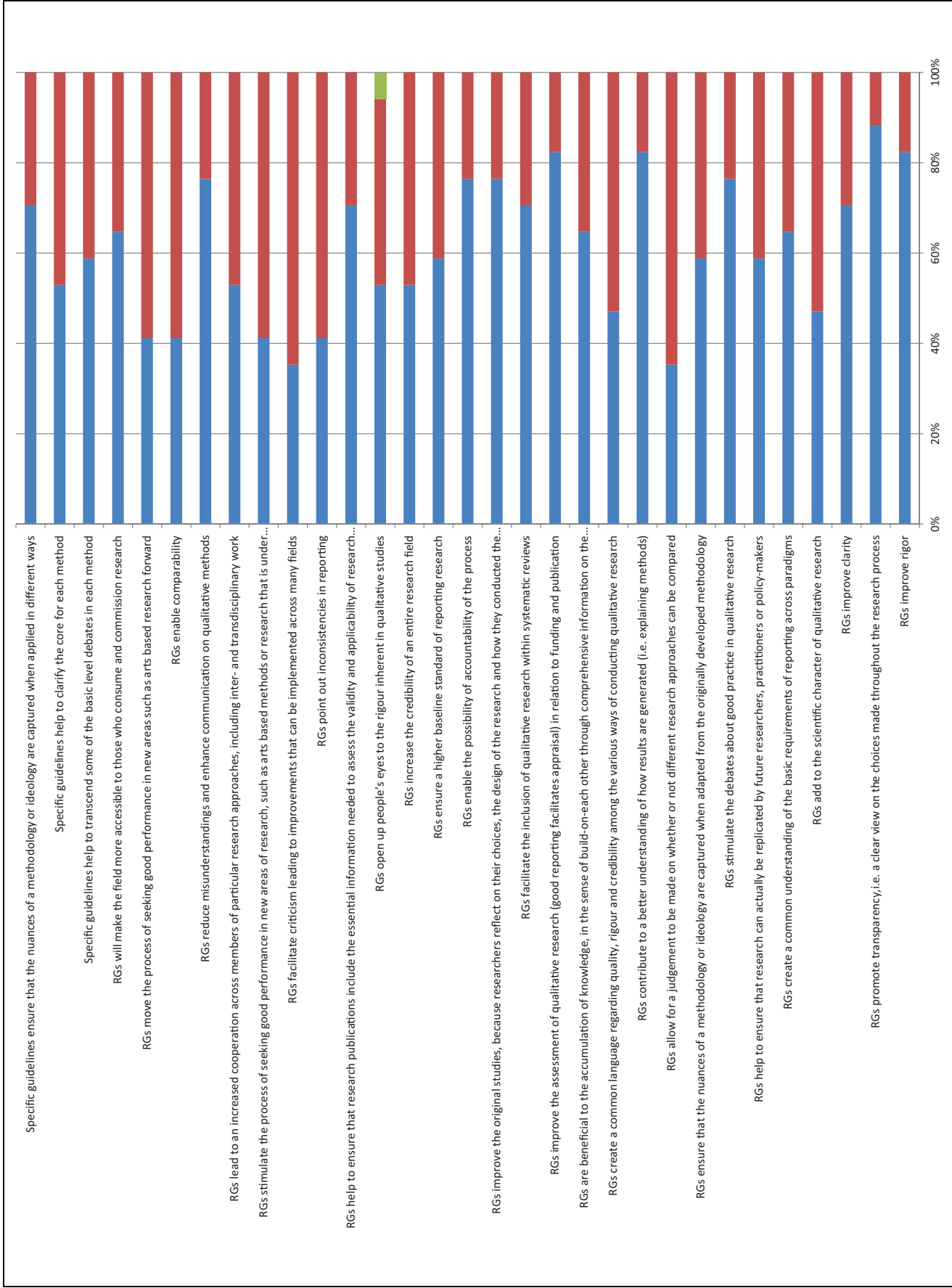


Figure 4. Positive aspects of reporting guidelines for qualitative research?

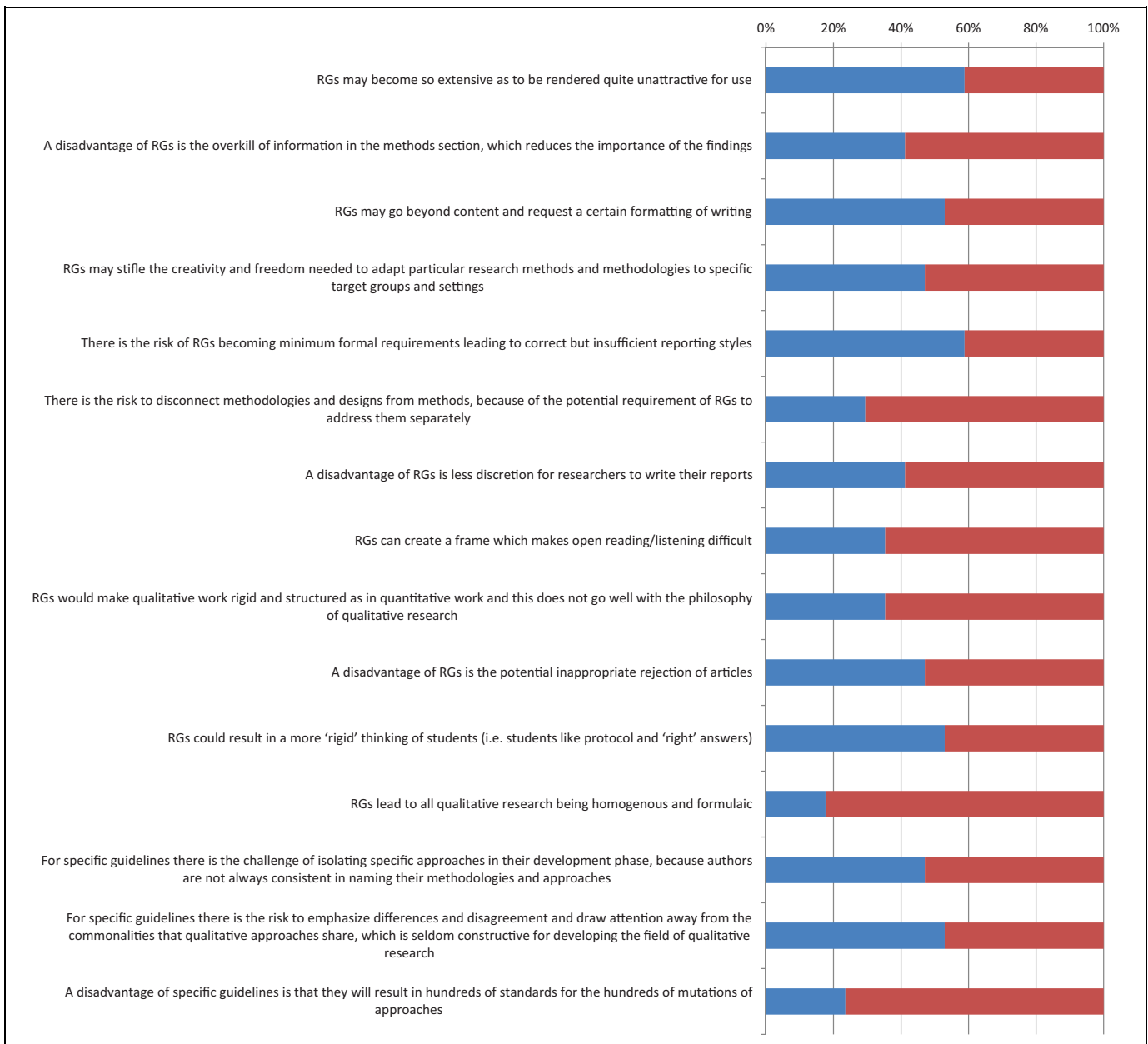


Figure 5. Negative aspects of reporting guidelines for qualitative research?

process such as sampling strategies, description of setting, type of analysis, or data collection techniques. They further identified 16 facilitating factors for the use of RGs. Six different target groups who may benefit from RGs were mentioned: students, teachers, peer reviewers and editors, authors/researchers, end users of research, and funding organizations (see Figures 1–3). The experts identified more advantages ($n = 31$; see Figure 4) than disadvantages ($n = 15$; see Figure 5) of RGs. Six different alternatives for the form of RGs for qualitative research were retrieved from the data: (a) no RG at all, (b) an RG listing a set of general principles outlining the minimum essential features of qualitative research, (c) a comprehensive general guideline or an overall framework that can be suited to encompass different qualitative methodologies and can

be used across different study designs, (d) a general guideline with an annex listing out specific features for specific designs, (e) a fully comprehensive RG that includes all proper elements from all types of qualitative research methodologies, and (d) one particular RG for each particular design or methodology. The group further identified a list of 16 conditions that may facilitate the use of RGs (see Figure 6) and 17 process-related issues or concerns to be considered in the development and implementation of RGs for qualitative research (see Figure 7).

Second Delphi Round

One of the experts dropped out and was no longer considered as a unit for analysis in the second round of the Delphi procedure.

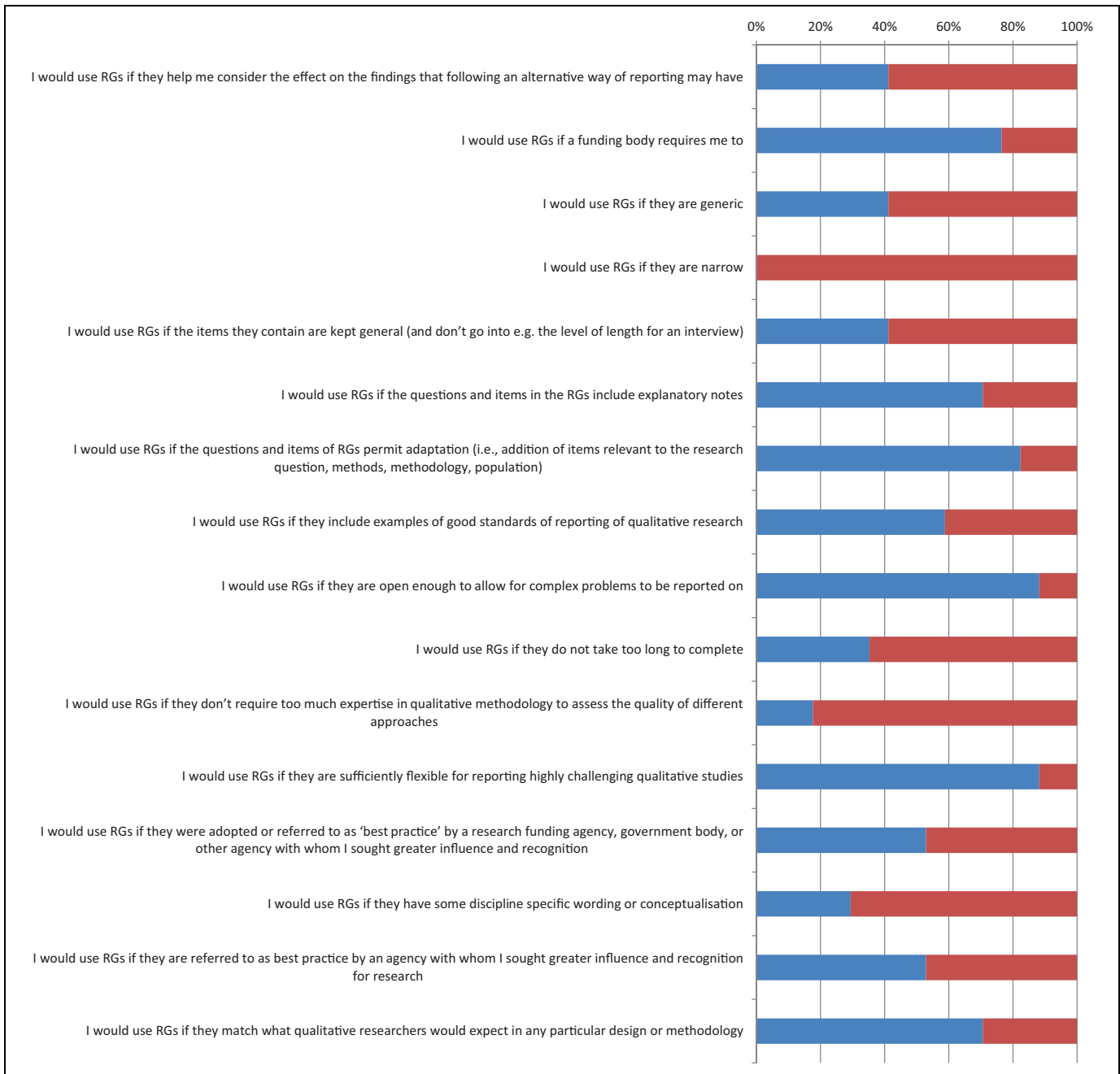


Figure 6. Conditions that may facilitate the use of reporting guidelines for qualitative research?

The second round provided us with information on the extent to which the experts agreed or disagreed with the viewpoints of their colleagues. It further revealed some of the factors that would facilitate the use of RGs and preferences of the experts regarding the form of an RG. In what follows, we report on the percentage of agreement between experts on the topics listed for discussion in the second questionnaire. A full list of figures is available from <http://ppw.kuleuven.be/home/english/research/mesrg/documents/paper-supplements/annex-reporting-guidelines-for-qualitative.docx>.

When do we need to consider RGs for qualitative research? Experts had different opinions on when they would use RGs for qualitative research. Only one expert would never use them. Forty-seven percent would always use them and another 47% would use them for particular research approaches or designs. There was a 100% agreement that RGs would not be used to conduct research within specific research traditions, such as grounded theory, phenomenology, arts-based or practice-based research, nor to conduct research within particular theoretical frameworks, for example, symbolic interactionism. However, some

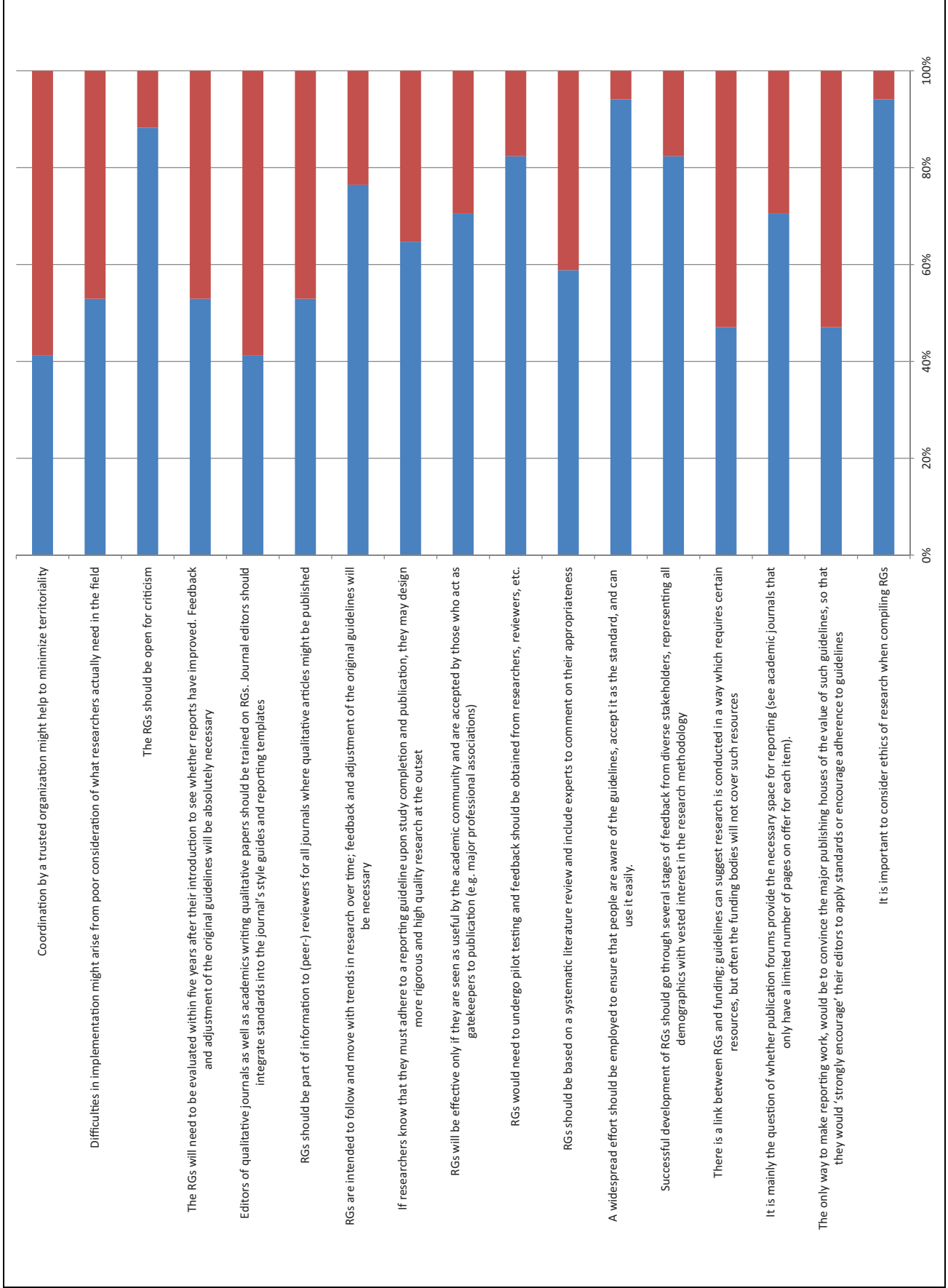


Figure 7. Process related concerns to be taken into account when developing and implementing reporting guidelines for qualitative research?

experts indicated that they would consider using them to assist their reporting on data collection procedures, such as observation, interviews, or focus groups (3 out of 17 experts for each of these techniques). Only one expert would consider using them for research involving document analysis, ethnographic fieldwork, and discourse analysis. Equally, only one would use them to assist in assessing the validity of interpretations or to describe methodological issues such as sampling and recruiting strategies, descriptions of settings, and type of data collection or data analysis.

What are the facilitating factors for the use of RG? Experts generally agreed on the fact that RGs for qualitative research should be flexible and open enough to allow for the reporting of complex problems and challenging studies (88% agreed) and the adaptation of criteria to better match the research approach opted for (82% agreed with this statement). In addition, they should include explanatory notes (71% agreed) and examples of good practice in reporting (59% agreed). Several experts stated that they would use them if a funding body required it (76%), if they were adopted or referred to as best practice by a funding agency or if they were referred to by an agency with whom the experts sought greater influence (53% agreement for both items). The level of conceptualization or expertise needed did not seem to facilitate the use of RGs. Less than 30% of the experts agreed on this issue. Forty-one percent of the experts would consider using RGs if they would help them consider the effect on the findings that following an alternative way of reporting may have.

Who benefits from RGs for qualitative research?

Teachers and students. Over 90% of the experts agreed that RGs were useful for teaching purposes. Six different advantages of RGs for students were identified, which all received support from at least half of the group of experts in our sample. RGs improve the overall quality of reporting in thesis and other manuscripts (71% agreement), assist in structuring research (76% agreement), create a learning potential about particular qualitative approaches or designs (65% agreement), make the qualitative research field more accessible (59% agreement), and help the student to defend himself or herself against criticism (53% agreement). It may also be used to showcase that there is no general agreement on many of the issues in qualitative research, which could free up their thinking (59% agreement).

Funders and the intended audience (further referred to as end users). The statements made by experts on the usefulness of RGs for funders of research and end users link into the general idea that RGs increase the understanding of qualitative research by end users (76% agreement). They force end users into questioning science and to better position themselves in making judgments about how confident they are in the findings. Several statements launch the idea that RGs enable end users and funders to critically assess qualitative research, with criteria that are sensitive to qualitative inquiry (65% agreement on all of these statements), and that they are useful in the context of

evaluating funding proposals (59% agreement). About half of the experts further supports the idea that such RGs assist in evaluating the transferability of findings from research to the end user's context.

Researchers/authors, including authors of systematic reviews. The statement that received the highest percentage of agreement among experts (i.e., 82%) expressed the idea that RGs could be used by researchers to clarify the important aspects of the methodology they used. Over 70% of the experts acknowledged a role for RGs in developing proposals, correcting an article before it is submitted to a journal, and being reflective about what information end users need to judge the quality of the article. The majority (71%) of the experts in our sample supported the claim that RGs for qualitative research are useful for authors of systematic reviews to assess the potential contribution of qualitative studies within systematic reviews. By contrast, only 35% of the experts agreed with the statement that RGs would provide authors with a common metric of assessing quality across studies or to compare a particular piece of research to its kin (in this case, other types of qualitative research approaches). The use of RGs to challenge the idea of what good research should look like (inspired by the dominant positivistic discourse in science) was not fully supported either. However, 65% of the experts agreed that RGs are a good starting point to think about bias in research. RGs further improve the overall quality of reporting of authors and make the reports more accessible to their audiences (65% agreement for both statements). They potentially help authors to consider what is important and to position their papers more effectively (59% agreement on both statements). Only 41% of the experts supported the claim that RGs would help researchers to defend themselves against criticism.

Editors and peer reviewers. There was over 80% agreement among the experts that RGs would stimulate editors and reviewers to be more explicit in the rationale of their claims and expectations and to motivate the qualitative judgment of a paper. Seventy-one percent of the experts supported the claim that RGs raise awareness in editors and peer reviewers of what they are looking for. There was a 47% agreement that RGs allow them to make better informed decisions about the basic standards in qualitative research and judge papers according to the state of the art in qualitative research and help them to promote better articles in journals. The agreement rate for the statement that RGs facilitate a judgment on the transferability to a particular setting by editors was larger than the support for the same statements in the target group researchers and authors (47% vs. 53%).

What are the advantages of RGs for qualitative research? Statements related to the advantages of qualitative research RGs mainly addressed the potential of these RGs to create a common understanding regarding quality, rigor, and credibility among qualitative researchers. The biggest advantage of RGs appears to be the transparency of research, which was agreed on by 88% of the experts. Over 80% supported the idea that

RGs improve rigor of research, contribute to a better understanding of how research results are generated, and improve the assessment of qualitative research. Most experts agreed that RGs reduce misunderstanding and enhance communication. They stimulate debates on what constitutes good practice in qualitative research (76% agreement), help ensure that the essential information is included in a report and as such, improve clarity (71% agreement). Transparency in reporting further facilitates the uptake of qualitative research in systematic reviews (71% agreement). Several experts believe that RGs help ensure that the nuances of a methodology or ideology are captured when applied in different ways (71%). At the same time, most of the experts agree that RGs create a common understanding of the basic, generic requirements in reporting qualitative research (65% agreement). The idea that RGs may also stimulate debate on what a qualitative study should look like was supported by only 35% of the experts. However, RGs seem to facilitate replication of studies by future researchers, practitioners, or policy makers (59% agreement). Authors could further improve their original studies by reflecting on their choices as well as the design of the research and how they conducted the study beforehand and by eliminating inconsistencies in reporting. This may secure a higher baseline standard of reporting research (59% agreement), for example, in pointing out inconsistencies (41% agreement), and as such increase the credibility of an entire research field (53% agreement). There was modest support for the idea that the development of RGs could stimulate the process of seeking good performance in new areas of research, such as arts-based methods or research that is under scrutiny (41% agreement), could lead to an increased cooperation across members of particular research approaches, and could increase the credibility of an entire research field (53% agreement for both items). To a certain extent, RGs may help to transcend some of the basic-level debates in each method and help clarify the core for each method (59% agreement on both items).

What are the disadvantages of RGs for qualitative research? Agreement between experts on the disadvantages of RGs was generally lower than for other topics addressed. None of the disadvantages inventoried were supported by more than 60% of the experts. The highest agreement among experts was reached for the statement addressing the risk that RGs would become minimum formal requirements leading to correct but insufficient reporting styles (59%). About half of the group of experts supported the claim that RGs would lead to more rigid thinking and that they could move attention away from the commonalities between different approaches, by focusing on the differences between designs and approaches (53% agreement on both topics). Several experts shared the concern that RGs could lead to a potential inappropriate rejection of articles, if they were forced upon reviewers (47% agreement). They also stated that they would make qualitative research more rigid and structured like quantitative research and would complicate open reading or learning (35% agreement for both statements). In addition, RGs may stifle creativity and the freedom needed

to adapt research methods to specific target groups and settings (47% agreement). There was some support for the statement that RGs would give qualitative researchers less discretion in writing their reports (41% agreed) and that RGs that require addressing designs and methods separately would lead to disconnecting them (29% agreed). According to some experts in our sample, the focus on methods in most types of RGs further increases the likelihood that the content of qualitative articles would become marginalized (41% agreement). The majority of experts in the sample did not support the statement that trying to develop RGs for qualitative research would result in numerous standards to respond to all sorts of mutations of approaches (24% agreement) or that it would result in qualitative research becoming more homogenous and formulaic (18% agreement).

What form should RGs for qualitative research take? Most experts in our sample preferred an RG with minimum essential features (41%) or a general framework that could be used across different designs (24%). Separate guidelines for particular qualitative research approaches or designs, general guidelines including an annex with specific criteria targeted toward particular designs and the idea of no RG at all received one vote each. A fully comprehensive RG with all criteria for all different types of designs appeared the best option for only 18% of the experts. Most experts thought it was feasible to develop general RGs for qualitative research (59% agreed, with 18% not answering the question). Only 24% thought that design-specific guidelines could be developed (with 35% not answering the question).

What are process and implementation factors that need to be considered in developing RGs for qualitative research? About half of the experts related the failure to implement RGs to the poor consideration of what researchers actually need in the field. The experts in our sample almost unanimously agreed on the fact that developers of RGs should consider ethics of research. Once these RGs have been developed, they should be disseminated to increase awareness among researchers (94% agreement for both issues). Over 80% of the experts stressed the importance of consulting a variety of vested stakeholders in the area of qualitative research and piloting the RGs before making them public, with 53% of the experts supporting the idea that such an exercise should be reconducted after 5 years and supported by a literature review (59% agreement). The overall expectation that RGs would change rapidly over time explains the high agreement on items suggesting that RGs should be open for criticism (88% agreement). Less than half of the experts supported the opinion that major publishing organizations should be mobilized to increase the uptake of RGs or that journal editors should integrate these standards into the journal style or that these RGs should be part of the information package sent to peer reviewers assessing the quality of qualitative research reports. Most experts (71%) believed that researchers would design more rigorous research when they are expected to adhere to RGs, that is, when these RGs are accepted or perceived as useful by the research community and

the gatekeepers to publication: the major professional associations. Seventy-one percent of the experts was concerned about the limited word count of several journals, preventing researchers from successfully following the guidelines.

Discussion

A broad variety of different viewpoints and preferences on RGs for qualitative research has been identified and analyzed in our Delphi study. Our sample of experts was small but adhered to the principle of maximum variety in expertise on the level of research paradigm, design and techniques, and scientific discipline. Consequently, we were able to extract a rich pallet of challenges, concerns, obstacles, and facilitators from these experts. For further research, the questionnaire could be adapted for a large-scale roll out, for example, a survey study in which a representative sample of experts in qualitative research spanning a broader geographical region is consulted on the topics of interest identified. Our study reveals that there is a considerable amount of topics where individual experts oppose each other in whether or not RGs are desirable, in what particular situations they should or should not be considered and in what form they should be presented. This variety in opinions was expected in advance. We opted for a maximum variation type of sampling approach. Maximum variation samples aim at capturing and describing the central themes or principal outcomes that cut across a great deal of participant or program variation. We acknowledge that a great deal of heterogeneity can be a problem when dealing with small samples, mainly because individual cases are very different from each other. The maximum variation sampling strategy, however, turns that apparent weakness into a strength (Patton, 1990). Any common patterns that emerge from the variation between experts are of particular interest and value in capturing the core opinions about and potential of RGs for qualitative research. Our findings reveal the many different ways of thinking about qualitative research. The researchers included in our sample utilize a variety of perspectives when studying their phenomena of interest, with some subscribing to a particular methodology and others taking a more eclectic approach that incorporates multiple techniques of data collection and analysis. One particular approach is not necessarily “better” than another, each simply emphasizes different views on how to methodologically approach things. The expected variety in the viewpoints between the experts is the main reason why the research team opted for an argument Delphi study rather than trying to gain consensus (i.e., the aim of traditional Delphi approaches) and moving toward a saturation point. The logic behind our choice for a maximum variation strategy was valorized in practice through spotting the tendency toward agreement between experts on a proportion of statements concerning RGs. The percentages revealing the support or nonsupport of experts for particular criteria and statements led us into developing a line of argument based on emerging areas of consensus and opposition that would otherwise have remained hidden.

The (Im)Possibility of RGs for Qualitative Research: Critical Notes

Only one expert mentioned that the RGs would probably not be used. Most experts mentioned that an increase in transparency, facilitated by RGs, would increase the overall accountability of qualitative research. This is in line with the argument for RGs outlined by some of the lead conveners of the EQUATOR network mentioned previously (Simera, Altman, Moher, Schulz, & Hoey, 2008; Simera, Moher, Hirst, et al., 2010; Simera, Moher, Hoey, Schulz, & Altman, 2010). The potential to increase quality, rigor, and credibility of qualitative research was identified as one of the core benefits of RGs for a variety of different target groups. Several experts claimed that RGs would lead to an improved quality of reporting and structure of reports for students and would facilitate a clear audit trail of methodological choices or judgments made by authors as well as peer reviewers. It did not automatically lead the experts into concluding that it would allow students or researchers to defend themselves against critics. However, there was considerable agreement among experts that RGs would stimulate them to reflect upon and to learn about their skills, particularly in the context of assessing the bias of qualitative research.

Despite the high support for transparency of reporting, experts in our sample seemed to resist the idea of a fixed, extensive list of criteria that would be developed to assist qualitative researchers. RGs with minimum essential features or a general framework that could be used across different designs were the most popular forms of RGs. This feeds into the argument that qualitative research is, in fact, considered as “an umbrella cross- and interdisciplinary term, unifying very diverse methods with often contradicting assumptions . . .” (Gabrielian, 1999); that there is indeed a variety of methods, approaches, and strategies that can be applied to successfully conduct qualitative research. It also suggests that the experts in our sample considered this variety a strength of qualitative research, particularly in relation to trying to understand the complexity of a particular phenomenon. Most experts emphasized the importance of flexibility in developing criteria for RGs, to respond to newly emerging methods, and the possibility to adapt criteria to better match the research approach opted for.

The 100% agreement between experts on the statement that they would not use RGs to facilitate research reporting within a particular methodological tradition, such as grounded theory or phenomenology, can be interpreted in various ways. Firstly, it may suggest that qualitative researchers adopt methodological design labels in their reports, but intentionally or unintentionally do not adhere to the theoretical principles underpinning these designs and creatively adapt them to create a better fit for purpose. Therefore, RGs with design-specific, restrictive criteria would make them uncomfortable. Secondly, researchers may want to adopt features from different designs, because they set out to achieve different things. When these different perspectives are put together, they can provide us with a multifaceted view of the phenomenon of study that may deepen our

understanding of it. In this case, design-specific criteria would be inappropriate, unless they are presented as an annex to more generally applicable criteria. The preference for a list of general criteria may further link into the idea that RGs that are too narrow would restrict the amount of creative research being conducted as well as the possibilities to publish reports on unusual, innovative, or emerging research approaches. Rigidity of RGs would most likely motivate particular methodological subdisciplines to write their own guidelines, thereby increasing fragmentation and lack of large-scale standardization. Consequently, the idea that RGs could be used to stimulate debates on what constitutes good practice, particularly for innovative research methodologies that are under scrutiny (e.g., arts-based research), received little support. Most experts appeared to agree on a set of criteria that stimulates authors to provide a decent audit trail of whatever approach they use or combine, so that each study can be assessed on its own merit.

Several experts acknowledged that they would respond to the criteria outlined in RGs if a funder or an organization with whom they sought greater influence would require them to adopt these criteria. The external pressure on researchers to subscribe themselves to a business-like scientific endeavor producing certain outcomes and meeting the utility expectations of their funders may compromise some of the values outlined in the previous paragraph, emphasizing flexibility and openness to individual choices made by researchers. This is particularly the case when the list of criteria is applied rigidly by funders, editors, or peer reviewers who conduct their quality control from the point of view of satisfying the end user, rather than from a critical reflection of what it is that a researcher intends to achieve or what the potential and value is of the qualitative research approach opted for. RGs, like critical appraisal tools can only be considered technical tools that increase the transparency of decisions made by users. They may help us to determine the extent to which we may have confidence in the target group's competence in being able to conduct research that follows established norms (Morse, Barrett, Mayan, Olson, & Spiers, 2002). However, to be able to use them appropriately, one needs to have a more than just a basic understanding of qualitative research.

Over 90% of the experts in our sample would welcome RGs as a training tool. When used appropriately, RGs may serve a teaching purpose. In the last decade, the use of qualitative research has become "fashionable" among students as well as researchers whose primary research background is quantitative. In some situations, the interest of students in qualitative research has run ahead of faculty expertise. RGs may unintentionally encourage supervisors with limited expertise to refer their students to the criteria outlined in the guideline, instead of giving them a proper training. The poor research resulting from such exercises undermines the credibility of qualitative research in general. RGs would provide little support to researchers when making methodological decisions, nor would they invite researchers to motivate the rationale behind their decisions. They would add to the transparency of research but would not assist the researcher or student in evaluating

methodological coherence or congruity between paradigms that guide the research project and the methodology and methods chosen. RGs would be limited in their capacity to assist researchers or students in actually defining their analytic stance or theoretical position. Most RGs would only remind users that they are required to report on these issues. The misperception that qualitative research is easy and does not require much training is very persistent among researchers (Dingwall, Murphy, Watson, Greatbatch, & Parker, 1998). It feeds into the inappropriate use of RGs. The use of a list of criteria to assist in reporting on qualitative research, without having been taught the skills of critical thinking, will lead to a false claim to have provided good qualitative research because "everything is in the report." In addition, peer reviewers that are insufficiently trained in qualitative methods may think a particular research project has been conducted according to the state of the art "because all categories have been used." Responding to the criteria in RGs does not automatically guarantee proof of quality. This may explain the rather modest support for the claim that RGs should widely be adopted by major publishers, integrated in the journal styles, and part of the information packages sent out to peer reviewers. They may consider using it as a technical tool only, without further reflection.

Optimizing Existing RGs for Qualitative Research

The COREQ statement and the SRQR (see Table 1) previously introduced match some of the expectations of the experts consulted. The COREQ statement claims to support researchers using focus group and interview techniques for data collection, however it precludes generic criteria that are applicable to all types of research reports. It invites researchers to report on their personal characteristics and relationship with participants and to make their theoretical framework explicit. It further emphasizes the importance of providing information on the participant selection, setting, data collection, and analysis. Overall, it does a good job in guiding researchers toward transparency of reporting. However, it fails to identify criteria that may be more central to the qualitative research tradition and may truly empower researchers to defend themselves against criticism. Examples of such criteria include resonance with the readers of qualitative research reports, the degree of sensitivity to the utility of research in its ability to empower people, theoretical sensitivity, or the researchers' ability to relate findings to the existing knowledge base, and the disclosure of personal values, assumptions, and motivations for choosing a particular design or approach (Elliott, Fischer, & Rennie, 1999), which is different from just revealing your personal impact on the research procedure or from stating potential conflicts of interest as promoted by O'Brian and colleagues in the SRQR. This highlights the importance of consulting relevant stakeholders and experts and creating opportunities for debate, as supported by many experts in our sample. One should bear in mind though that these processes take time, on average 20 months for the development and another 11 months to get the outcome published (Simera et al., 2008), without evaluation component. It is

recommendable though that such development processes evaluate the uptake of the RG by journals and assess the impact of the guideline on the reporting of, in this case, qualitative research.

Conclusions

Based on the findings of this exploratory study, we would recommend the further development and refinement of RGs for qualitative research. They are perceived as valuable by almost the entire group of experts in our sample and serve a variety of different target groups, albeit not as a substitute of proper training in qualitative research methods. The preference to focus on a more generic set of criteria that can be used across different approaches reveals a reluctance of qualitative research experts to comply with criteria based on theoretical principles outlined for design-specific qualitative research approaches. It suggests that they feel more comfortable with RGs that are flexible. Flexibility means that the criteria selected for RGs should enable us to respond to methodological changes as well as the nature of qualitative researchers to adapt methods and techniques in order to create a better fit for purpose for the often complex questions that need to be answered. For example, conventional criteria such as “has a saturation point been reached” may work well for authors that claim to produce a theory that is transferable to similar settings as the ones discussed in their own research paper, but it may be counterproductive for studies that present detailed narratives of one individual. In such cases, a more general criterion evaluating thickness of description might work better. We acknowledge that there is variety in the viewpoints for a considerable amount of topics discussed. This will most likely complicate a consolidation process for RGs, particularly if we follow the logic of the experts in trying to compile a generic set of criteria. In adopting such a strategy we may risk to end up with an RG that will hardly add anything more to the transparency of an article’s outline than a list of “issues to write about.” It may not give us any insight into the rationale and epistemological understanding of the authors. The current trend to develop RGs for different types of meta-synthesis approaches, as opposed to, for example, the more general Enhancing transparency in reporting the synthesis of qualitative research statement (Tong, Flemming, McInnes, Oliver, & Craig, 2012), confirms the trend toward producing design-specific guidelines (Wong, Greenhalgh, Westhorp, Buckingham, & Pawson, 2014) or even guidelines on methodological parts in a report, such as how to report on a literature search or develop an abstract (Booth, 2004; Hopewell et al., 2008). We believe this is a worthwhile endeavor. It allows RG developers to align criteria with particular schools of thought adopted by researchers and encourages authors to make their epistemological point of view more explicit. RG developers may benefit from familiarizing themselves with the arguments provided by the experts involved in this study. It will help them to consider what researchers actually need in the field, how to respond to criticism, how to create face validity, and how to align the RGs to the ideas of people in order for them to overcome their difficulties with RG.

We are confident that pragmatic issues such as limited word count of journals will be solved in the short term, with more and more journals providing an opportunity to post papers or annexes online.

Acknowledgments

The authors thank the experts who participated in this pilot study: removed for peer-review purposes and also thanks to Elke Emmers for the administrative assistance and to Marlies Vervloet for her statistical advice.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Financial support provided by Academische Stichting Leuven.

References

- Hannes, K., Lockwood, C., & Pearson, A. (2010). A Comparative Analysis of Three Online Appraisal Instruments’ Ability to Assess Validity in Qualitative Research. *Qualitative Health Research, 20*(12), 1736–1743.
- Bastuji-Garin, S., Sbidian, E., Gaudy-Marqueste, C., Ferrat, E., Roujeau, J. C., Richard, M. A., & Canoui-Poitrine, F. (2013). Impact of STROBE statement publication on quality of observational study reporting: Interrupted time series versus before-after analysis. *PLoS One, 8*, e64733. doi:10.1371/journal.pone.0064733. eCollection 2013
- Booth, A. (2004). ‘Brimfull of STARLITE’: Toward standards for reporting literature searches. *Journal of the Medical Librarian Association, 94*, 421–429.
- Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, P. P., Glasziou, P. P., Irwig, L. M., . . . De Vet, H. C. (2003). Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Annals of Internal Medicine, 138*, 40–44.
- Des Jarlais, D. C., Lyles, C., Crepaz, N., & The TREND Group. (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health, 94*, 361–366. doi: 10.1093/intqhc/mzm042
- Dingwall, R., Murphy, E., Watson, P., Greatbatch, D., & Parker, S. (1998). Catching goldfish: Quality in qualitative research. *Journal of Health Services Research & Policy, 3*, 167–172. doi:10.1177/135581969800300308
- Elliott, R., Fischer, C. T., & Rennie, D. L. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British Journal of Clinical Psychology, 38*, 215–229. doi:10.1348/014466599162782
- Gabrielian, V. (1999). Qualitative research methods: An overview. In G. J. Miller & M. L. Whicker (Eds.), *Handbook of research methods in public administration* (pp. 167–205). New York, NY: Marcel Dekker.

- Gagnier, J. J., Kienle, G., Altman, D. A., Moher, D., Sox, H., Riley, D., & the CARE Group (2013). The CARE guidelines: Consensus-based clinical case reporting guideline development. *BMJ Case Reports*. doi:10.1136/bcr-2013-201554 PMID: 24155002
- Hannes, K., & Macaitis, K. (2012). A move to more transparent and systematic approaches of qualitative evidence synthesis: update of a review on published papers. *Qualitative Research, 12*(4), 402–442.
- Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing, 32*, 1008–1015. doi:10.1046/j.1365-2648.2000.t01-1-01567.x
- Hopewell, S., Clarke, M., Moher, D., Wager, E., Middleton, P., Altman, D. G., Schulz, K. F., & the CONSORT Group (2008). CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Medicine, 5*, e20. doi:10.1371/journal.pmed.0050020
- Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., . . . Loder, E. (2013). Consolidated Health Economic Evaluation Reporting Standards (CHEERS)—Explanation and elaboration: A report of the ISPOR health economic evaluation publication guidelines good reporting practices task force. *Value Health, 16*, 231–250.
- Möhler, R., Bartoszek, G., & Meyer, G. (2013). Quality of reporting of complex healthcare interventions and applicability of the CREDEC list: A survey of publications indexed in PubMed. *BMC Medical Research Methodology, 13*, 125.
- Montori, V. M., Wilczynski, N. L., Morgan, D. R., & Haynes, B. (2003). Systematic reviews: A cross-sectional study of location and citation counts. *BMC Medicine, 1*, 2. doi:10.1186/1741-7015-1-2
- Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification strategies for establishing reliability and validity in qualitative research. *International Journal of Qualitative Methodology, 1*, 13–22. Retrieved from <https://ejournals.library.ualberta.ca/index.php/IJQM/article/view/4603/3756>
- 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. *Acad Med, 89*(9), 1245–1251.
- Ogrinc, G., Mooney, S. E., Estrada, C., Foster, T., Goldmann, D., Hall, L. W., . . . Watts, B. (2008). The SQUIRE (Standards for Quality Improvement Reporting Excellence) guidelines for quality improvement reporting: Explanation and elaboration. *Quality and Safety in Health Care, 17*, i13–i32.
- Patton, M. (1990). *Qualitative evaluation and research methods*. Beverly Hills, CA: Sage, pp. 169–186.
- Pearson, A., Wiechula, R., Court, A., & Lockwood, C. (2005). The JBI model of evidence-based healthcare. *International Journal of Evidence-Based Healthcare, 2*, 207–215.
- Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMC Medicine, 8*, 18. doi:10.1136/bmj.c332
- Simera, I., Altman, D. G., Moher, D., Schulz, K. F., & Hoey, J. (2008). Guidelines for reporting health research: The EQUATOR network's survey of guideline authors. *PLoS Medicine, 5*, e139. doi:10.1371/journal.pmed.0050139
- Simera, I., Moher, D., Hirst, A., Hoey, J., Schulz, K. F., & Altman, D. G. (2010). Transparent and accurate reporting increases reliability, utility, and impact of your research: Reporting guidelines and the EQUATOR Network. *BMC medicine, 8*, 24. doi:10.1186/1741-7015-8-24
- Simera, I., Moher, D., Hoey, J., Schulz, K. F., & Altman, D. G. (2010). A catalogue of reporting guidelines for health research. *European Journal of Clinical Investigation, 40*, 35–53. doi:10.1111/j.1365-2362.2009.02234.x
- Tate, R. L., & Douglas, J. (2011). Use of reporting guidelines in scientific writing: PRISMA, CONSORT, STROBE, STARD and other resources. *Brain Impairment, 12*, 1–21. doi:10.1375/brim.12.1.1
- Tong, A., Flemming, K., McInnes, E., Oliver, S., & Craig, J. (2012). Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. *BMC Medical Research Methodology, 12*, 181. doi:10.1186/1471-2288-12-181
- 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 of Quality in Health Care, 19*, 349–357. doi:10.1093/intqhc/mzm042
- von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2008). Strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Journal of Clinical Epidemiology, 61*, 344–349. doi:10.1016/j.jclinepi.2007.11.008
- Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J., & Pawson, R. (2014). RAMESES publication standards: Meta-narrative reviews. *BMC Health Services Research, 14*, 20. doi:10.1186/1741-7015-11-20