PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Understanding how and why Quality Circles improve standards of practice, enhance professional development and increase psychological well-being of General Practitioners - a realist synthesis
AUTHORS	Rohrbasser, Adrian; Wong, Geoff; Mickan, S; Harris, Janet

VERSION 1 – REVIEW

REVIEWER	De Groot, Esther Utrecht University, Julius Center for health sciences and primary care
REVIEW RETURNED	19-Nov-2021

GENERAL COMMENTS	Dear authors, Even though this was a long manuscript with a lot of supplementary materials, it was a joy to read and review it. You did an amazing amount of work, in a very thorough manner. It will be really valuable for the field that this manuscript will be published. I have some recommendations for improving it, where some points are minor and some more substantial. Apart from these recommendations, what was complicated for me as a reviewer was that so many things were put in supplementary materials and even as the review is very long there is a risk for losing the reader because the results are very short in comparison with other sections. I am not sure whether the balance is good now, perhaps a bit more beef from the supplementary materials could be included in the results? Perhaps in a final paper, it is less distracting but I found myself skipping and searching through the material instead of just reading from the start to the end. Strengths and limitations of this study In BMJ Open, the recommendation is that strength and limitations are about methodological aspects mainly. What is mentioned now is a mix of that (about methodological aspects) and 'public relations' for the paper. I recommend to stick to the first types of strengths and limitations and as this is a rich manuscript with many methods used, that should not be too difficult. To mitigate the risk of selection bias if researchers choose underlying theories and synthesise them ad hoc, we used stakeholders' mental model and programme documentation as our framework for analysis. □ This is a bit more understandable after reading the paper (but see later on in my remarks), but to lure readers into reading the paper I recommend to adjust it and make it a bit easier to read. Methods

	(Page 8, line 20). It will help the reader (perhaps this is less necessary when the paper is in formal formatting and not in the ScholarOne system) if you point forward after this sentence “We answered our research question in four stages” that details will follow later. In the (clear) introduction, the RQ states that it is about. … GP behaviour and improve psychological well-being and why. … In the methods this is worded slightly different ”improves standard practice and raises professional self-esteem, and increases well-being”. Although one might argue that it is the same, the different wording make the already highly complex manuscript even more complex. (page 8) “information from QC programme documentation and training materials, extracting QC aims, detailed objectives, and roles” □ Unclear where this documentation came from. Perhaps in the Appendix but as a reader I do not want to switch between both parts of the paper. (page 9, r, 35) Since the overlapping theories were complex, we deviated from the original protocol and used the preliminary programme theory (stage one) as an organizing framework. □ More explanation is necessary. A superficial reader might think that you did not use the information from the review later on. In general, for readers who are not familiar with the iterative way of searching that is common in a realist study, a bit more clarification is necessary. After the section about searching for theories, a part follows about selecting articles and the reader is lost in the argumentation then. (page 11, r. 13) Since papers were often closely related, we grouped them based on their kinship, which helped us look for and confirm CMO configurations between papers within the same (family) study. □ This aligns well with recent approaches in doing realist reviews (cluster searching) but the reader wants (at least, I want this as a reader) to have a bit more detail about this ‘based on their kinship’ idea. Perhaps in the appendix but managing large chunks of data is a challenge in this type of work so letting the reader know how you did that is relevant. (Page 12, line 22) To consolidate the … □ Explain why, after the three previous, time consuming, phases, there was a need for a stage 4. I get the idea that it had something to do with policy makers but the reader has to guess. Perhaps the PT in earlier stages had inconsistencies or aspects that were problematic? (page 12, line 32) We then compared and contrasted… □ I had to read it twice to find out that a more overarching step is described here, to bring all the information from the different phases together. At first, I read it as what was done with participants in phase 4, which does not appear to be the case. Perhaps an additional (sub)header would help. Results (page 13, line 12) and also provided programme documentation and training materials … □ See my earlier comment. This sentence does not belong here, but in the methods. And what you have done with those materials is not very clear either. (page 13, line 18) the following programm … □ Apart from the fact that the word ‘theory’ is missing here, the sentences after that are a bit confusing as they sound more like a definition of Quality Circles, and not an initial PT. I

	understand that you had to struggle with the word count but now the reader has to go to Supplemental material 6 to get any idea of the novelty in the results of this phase. (page 13, line 58) including an underlying trial, common themes, common contexts like geographical area, and common methods of organising QCs (e.g., papers that tested similar didactic methods or similar QI tools in QCs). □ See my earlier comment. This are the details that I was looking for and that I expected in the methods. But keeping it here, in the results, is good. At the bottom of page 15 Figure 2 is mentioned, immediately followed by Table 2. This is really minor but confusing as a result of Scholar One lay-out. Perhaps moving the sentence about Figure 2 (Figure 2. Consolidated programme theory on quality circles) a bit higher in the paper, but, as said, this will be solved in the final paper. Page 16.. Table …. their relationships to existing theories □ Most of them are clear but a very global 'Theories about groups' 'Automaticity' or, on the other end of the spectrum, a very specific one 'The PARiHS framework' could be adjusted to make the Table more balanced. (page 18, line 11) The most important contextual requirements for successful QCs are governmental trust in the ability of GPs to deliver QI and appropriate professional and administrative support for QC work. □ The governmental trust as a contextual requirement is introduced here but not mentioned in the results at all. I get the feeling that the results are mostly about mechanisms. The context is missing there (or hidden, this is a long manuscript to review….). (page 19, line 38) We had several insights that had not been reported in current QC literature. □ Rephrase, it sounds as if you just are making things up, which does not do justice to the extensive process of your work. But at the same time, I was surprised that you mention Cognitive dissonance here as a key new finding when this theory was not included in table 2. (page 21, line 2) Not all recommendations will apply to every QC. □ This, to me, sounds as a no-brainer in a realist manuscript. This could be stated with more confidence as this manuscript gives some insights in the context where it can be applied and where not. (page 22, line 34) To mitigate the risk of social desirability bias, AR carefully posed neutral interview questions and tried to avoid embedding assumptions in his questions. □ A bit strange sentence. Why should respondents be inclined to answer with social desirability when asked about program theories? This nearly made me jump to the supplementary material and see whether I misunderstood what you actually had been doing. (page 22, line 40) the risk of selection bias if researchers choose underlying theories and synthesise them ad hoc. □ This is a strange part of the manuscript. I think that selection bias is never entirely possible to be left out and what would be the problem with that? The manuscript is already very detailed and you had to make choices. In making those choices, it is possible that you included some theories and neglected others, but as you have been involving so many people, well, I think that many other papers have a lot more selection bias. If mentioning selection bias is needed, explain what that would mean for your results.

	(page 23) can improve practice, foster professional development, and increase psychological well-being among participants □ See my earlier comment, here it is written differently again. (page 23) The requirements for successful QCs are □ Earlier (at the start of the results section), you use this wording “The most important contextual requirements for successful QCs …”. It looks as if this is a minor detail but in an overall line of reasoning that stresses the importance of context, I find it important to repeat the contextual aspect, otherwise the conclusions sound a bit mundane ‘you need sufficient support’, well, that does not sound too inspiring. □ Overall the conclusion could be adjusted because it reduces the value of your work by making it so concrete that the added value of all the realist work is getting out of view.

REVIEWER	Renmans, Dimitri University of Antwerp
REVIEW RETURNED	08-Dec-2021

GENERAL COMMENTS	Dear authors, Thank you for this very interesting study and this very detailed account of the reasoning process. The supplementary materials are impressive and clearly show how you reach the program theory. It can serve as an example for other realist evaluators on how to write down the process of retroduction. However, as always, there are some comments and concerns from my side. Yet, these are certainly not insurmountable. General remarks: - I am not sure whether a reader not acquainted with the realist approach will be able to follow. It might be useful to expand a little bit more on the approach and also to give your conceptualizatoin of a mechanism. - A bit more thought should go into the limitations section. The first is not really a limitation, yet I think that other limitations might be left out (for example, were you able to really fully grasp the 20 different theories that you used?) - I was thinking whether it might be useful to incorporate some of the information in the results section in the methodology section (e.g. the number of interviews performed, etc.) so that you do not have to repeat certain things and you can have more space for writing about the content. Indeed, it sometimes felt I was only reading about the methodology and everything was in the supplementary material. I understand of course if the word limit does not allow you to do this. Specific comments: I also have some specific comments related to the use of some words and some unclarity in the text. I also went through the different CMOCs and gave some suggestions on how to change them (these are just sugestions!). However, overall, I feel they are very useful and insightful.

	P7, line 51: “Searching for evidence” = Evidence for what? Is it not more logical to put the search for theories first and then the search for evidence? P9, line 5: What is meant with “the larger program theory”? Is this the by theory updated initial PT? P9 line 19: “preliminary programme theory” Is this the 'larger programme theory' mentioned earlier? P10 line 11: “Interaction between context and mechanism to facilitate or constrain QCs” What do you mean by facilitating or constraining QCs? To do what? Reach outcomes? Better to rephrase. P10, line 13: In what way were they closely related? Author, study, context, intervention, theoretical framework? P10 line 20: ‘Predictable’ is not a word I would use in the framework of a realist study. Even if it is only ‘semi’. P10, line 58: Unclear what is meant with ‘underlying reasoning for QC interactions’. Do you mean the underlying mechanisms relevant in explaining QC outcomes? P11, line 34: You say that “formal theories capture a PTs underlying mechanisms”. This is not clear to me, PTs are constructed around underlying mechanisms, they entail and explain the relevant underlying mechanisms, formal theories may give insights into the mechanisms that are relevant and how they work and interact with each other. P12 line 18: What do you mean by “preliminary programme”? Was a programme implemented? Or is it rather a preliminary conceptualization of the programme? P15 table 2: I think the title should be the other way around: Existing theories and their relationships to the CMOCs .... P 18, line 23 : Typo?: “of the individual” P21, line 14: What does “these realist approaches” refer to? P21 line 14: “the resulting theory relies on the detail and depth of the reports we identified in our literature review.”: This is a limitation to every synthesis approach not only the realist. Moreover, isn't the final theory also based on the interviews? At least this is what you write later in this paragraph and which makes this first sentence contradictory to what follows. Figure 2: CMO 1a: Make a difference in what? Is the mechanism here not intrinsic motivation? Especially since you refer to the self-determination theory? CMO 1b: Can this mechanism be called increased self-efficacy? Also related to the theory of planned behaviour. CMO 1c: Is it sufficiently different from CMO 1a? Again, isn’t the mechanism intrinsic motivation? CMO 1d: Again, is the mechanism not self-efficacy? CMO 2a: Is “prefer learning in QCs” really a mechanism that gets triggered by the mentioned context? They prefer it above which other method? If they really prefer it, than why do they need a financial and other incentives? That’s the opposite of “prefer” (which I think refers to intrinsic motivation). CMO 2b: I am doubting about the link between the mechanism and the outcome. “Members get to know each other” is a direct consequence of the introduction and the socially enjoyable contact not of the described mechanism. I also do not see how ‘deciding on rules’ follows from the mechanism “members want to interact with equals”. The contextual condition is already about contact and

	discussions. Maybe it is explained more clearly later on? But her eit is not very clear to me. CMO 2c: I feel the outcome is the same as the mechanism. I would say that the outcome is increased intrinsic motivation to participate. CMO 2d: Is this the mechanism. Or is it more about reluctance to speak in large groups, or the earlier mentioned issue of knowing each other well becomes more difficult in larger groups, or a kind of bystander effect when it comes to intervening (as we also see in our classrooms or at conferences)? CMO 2e: I very much like this one. CMO 2f: Same here, nice insight. CMO 3a: No remarks CMO 3b: I feel the mechanisms are reasons (not reasonings) rather than mechanisms. Or does your conceptualization of a mechanism entail reasons? I would say that these mechanisms are actually contextual conditions (characteristics of the actors) and the mechanism is motivation (one of the non-autonomous motivation forms of the SDTheory) CMO 3c: No remarks CMO 3d: I feel the outcome is not really an outcome but a rephrasing of the mechanism. The because part refers to an argument rather than a causal explanation. CMO 3e: No remarks CMO 3f: I feel the outcome is at another level than the mechanism. The group will create a learning environment but the mechanism 'social learning' happens at the level of the individual. CMO 4a: No real remarks. Although I feel the CMO is incomplete. Isn't more needed to make QCs design creative solutions? This can off course be a consequence of insufficient data on the context. CMO 4b: "Feel unsafe" is somewhat strange here. Why not stick to the feeling of lack of autonomy. CMO 4c: I think the second part of the mechanism is an action rather than a mechanism and the first part seems to be a contextual conditions or at least a bit weak to be seen as the mechanism that generates the outcome here. Not sure how the "act and negotiate cooperatively to achieve a common goal" links with "participants will adapt and generate new knowledge for local use". CMO 4d: Is this the social learning mechanism? CMO 4e: This clearly relates to the self-determination theory, yet according to this theory this leads to more intrinsic motivation. CMO 4f: Not clear how "announcing an intention to change" triggers the mechanism "they and other in the group think it is a good idea". Is this not about peer pressure? CMO 4g: No remarks, but maybe refer to self-efficacy as a mechanism? CMO 5a: Again this is self-efficacy. CMO 5b: No remarks Supplemental material 4: The description of the participants is really detailed which makes that they are no longer anonymous. This is an ethical issue.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1: …..for losing the reader because the results are very short in comparison with other sections. I am not sure whether the balance is good now, perhaps a bit more beef from the supplementary materials could be included in the results?
Reviewer 2: - I was thinking whether it might be useful to incorporate some of the information in the results section in the methodology section (e.g. the number of interviews performed, etc.) so that you do not have to repeat certain things and you can have more space for writing about the content. Indeed, it sometimes felt I was only reading about the methodology and everything was in the supplementary material. I understand of course if the word limit does not allow you to do this
this also concerns Reviewer 1: (page 13, line 12) and also provided programme documentation and training materials … See my earlier comment. This sentence does not belong here, but in the methods. And what you have done with those materials is not very clear either.
we moved suggested contents from the results to the method section (p6/7) and added summarised findings to the results of stages 1, 2 and 3 (pages 13-16), trying to keep them as short as possible.
Reviewer 2 asked us to expand on the realist approach and mention our conceptualisation of mechanism. We added the following text to page 7:
The idea of mechanisms as being the generative power of how and why change occurs is central to realism. In the case of QCs, we looked for mechanisms at the level of human reasoning, because it is individuals who take an action or not, as a result of participating in QCs. When these mechanisms are activated in their context, it can be an immediate or delayed response.
Reviewer 2 P7, line 51: “Searching for evidence” = Evidence for what? Is it not more logical to put the search for theories first and then the search for evidence?
This is about search for evidence that QC had worked in certain settings and contexts to build CMO configurations; iterative searching for theories and evidence was part of the process during the realist review. To provide greater clarity, we swapped the order and put an additional explanation in the beginning of the heading ‘realist review’ (pages 8 and 9)
We performed iterative searches: to become familiar with existing literature; to find possible candidate theories to be tested; to find empirical evidence to refine, refute or confirm CMOs of the emerging programme theory; to look for further empirical evidence or theories to consolidate the programme theory.
Additional titles
Searching for theories
Searching for evidence for QC outcomes

Addressing reviewers’ comments point by point
Reviewer 1
Introduction
Reviewer 1: the same wording in the introduction and in the methods. Text changed as suggested (see page 5): QCs can improve standard practice like prescription patterns and diagnostic habits, enhance professional development and psychological well-being, Methods:
Reviewer 1: p8 line 20: we should hint those details will follow later.
We changed the wording to (see page 6): We answered our research question in four stages with details to follow
Reviewer 1: In general, for readers who are not familiar with the iterative way of searching that is common in a realist study, a bit more clarification is necessary.
We added a short paragraph (see page 8):

We performed iterative searches: to become familiar with existing literature; to find possible candidate theories to be tested; to find empirical evidence to refine, refute or confirm CMOs of the emerging programme theory; and to look for further empirical evidence or theories to consolidate the programme theory.
Reviewer 1: (page 8) Unclear where this documentation came from. Perhaps in the Appendix but as a reader I do not want to switch between both parts of the paper.
We added a short paragraph (see page 8):
Stakeholders provided access to detailed and local information about QC aims, objectives and roles from professional websites, local publications and confidential training material and manuals across different European regions. We co-designed the preliminary conceptualisation of the programme theory, in short, preliminary programme theory, in discussion with the stakeholders, supported by local programme documentation and training material.
Reviewer 1: p9 line 35 More explanation is necessary. A superficial reader might think that you did not use the information from the review later on:
We added a short paragraph (see page 7):
Since the theories overlapped considerably in a complex way, they did not allow empirical testing. Therefore, we deviated from the original protocol and used the preliminary programme theory (stage one) as a starting point for the emerging programme theory. However, we benefitted from these findings in stage 4.

Reviewer 1 (page 11, r. 13) Since papers were often closely related, we grouped them based on their kinship, which helped us look for and confirm CMO configurations between papers within the same (family) study.
This aligns well with recent approaches in doing realist reviews (cluster searching) but the reader wants (at least, I want this as a reader) to have a bit more detail about this 'based on their kinship' idea. Perhaps in the appendix but managing large chunks of data is a challenge in this type of work so letting the reader know how you did that is relevant.
As reviewer 2 also pointed out the need for further explanation at a different location of the text.
Reviewer 2: P10, line 13: In what way were they closely related? Author, study, context, intervention, theoretical framework?]

We decided to put the requested explanation into the methods section (see page 9):
These kinship papers had common contextual features or theoretical backgrounds to the referring studies. We categorised these papers into kinship networks based on common themes, common contexts like geographical area, and common methods of organising QCs (e.g., papers that tested similar didactic methods or similar QI tools in QCs). We broadened the search by examining citations in reference lists and Web of Science and searched manually for closely related papers (kinship papers).

Reviewer 1 (Page 12, line 22) To consolidate the …
Explain why, after the three previous, time consuming, phases, there was a need for a stage 4. I get the idea that it had something to do with policy makers but the reader has to guess. Perhaps the PT in earlier stages had inconsistencies or aspects that were problematic?

We provide an explanation as requested as follows (see page 12/13):
The literature, the interviews and focus groups contained little data about how the national contextual level or how national organisations or reimbursement of PHC affect QCs. Therefore, to consolidate the programme theory at a national and policy level, AR invited five representatives with expertise in QCs from five countries with different PHC provision systems to a one-hour online interview to discuss the ways that different professional associations, institutional settings, and other contexts affect QC outcomes.

Reviewer 1 (page 12, line 32) We then compared and contrasted…
I had to read it twice to find out that a more overarching step is described here, to bring all the information from the different phases together. At first, I read it as what was done with participants in phase 4, which does not appear to be the case. Perhaps an additional (sub)header would help.

Consolidation took place in two steps. To make this clearer we have provided two subheadings (see page 12 and 13)
Interviews with stakeholders across health care systems
Existing theories and their relationships to CMO configurations in the programme theory

Results
Reviewer 1 (page 13, line 12) and also provided programme documentation and training materials …
 See my earlier comment. This sentence does not belong here, but in the methods. And what you have done with those materials is not very clear either.

We moved the part of the sentence to the methods and described in detail where they came from and what we did with them (see page 8).

Reviewer 1 (page 13, line 18) the following programm …
Apart from the fact that the word 'theory' is missing here, the sentences after that are a bit confusing as they sound more like a definition of Quality Circles, and not an initial PT. I understand that you had to struggle with the word count but now the reader has to go to Supplemental material 6 to get any idea of the novelty in the results of this phase.

In response we have revised the text as follows (see pages 13/14):
This co-inquiry along with programme documentation resulted in the following preliminary programme theory: GPs want to meet with their peers, share their problems and exchange ideas. CME credits or requirements from health insurance companies seem to be additional drivers to participate in QCs. Skilled facilitators are key to establish a safe environment where GPs share local data, and exchange experiences and knowledge. Reflection on personal experiences, successes and failures, helps in identifying learning needs. A goal-oriented facilitator helps members to choose the method they want to use to approach an issue and helps them build a learning environment where they adapt or create new knowledge which they then put into practice in a repetitive process.

Reviewer 1 (page 13, line 58) including an underlying trial, common themes, common contexts like geographical area, and common methods of organising QCs (e.g., papers that tested similar didactic methods or similar QI tools in QCs).
See my earlier comment. This are the details that I was looking for and that I expected in the methods. But keeping it here, in the results, is good.
We moved the explanation to the methods section as it confused both reviewer 1 and 2 (p 9) as we revised the text about iterative searches.

Reviewer 1 At the bottom of page 15 Figure 2 is mentioned, immediately followed by Table 2. This is really minor but confusing as a result of Scholar One lay-out. Perhaps moving the sentence about Reviewer 1 Figure 2 (Figure 2. Consolidated programme theory on quality circles) a bit higher in the paper, but, as said, this will be solved in the final paper. 3
We have put the sentence, "Table 2 summarises the theories and their corresponding CMO configurations". after Figure 2, under the subheading Existing theories and their relationships to CMO configurations in the programme theory (see page 16).

Reviewer 1 Page 16.. Table …. their relationships to existing theories
 Most of them are clear but a very global ‘Theories about groups’ ‘Automaticity’ or, on the other end of the spectrum, a very specific one ‘The PARiHS framework’ could be adjusted to make the Table more balanced.

The PARiHS framework could be subsumed under receptive context (p 17/18)

Reviewer 1 (page 18, line 11) The most important contextual requirements for successful QCs are governmental trust in the ability of GPs to deliver QI and appropriate professional and administrative support for QC work.
The governmental trust as a contextual requirement is introduced here but not mentioned in the results at all. I get the feeling that the results are mostly about mechanisms. The context is missing there (or hidden, this is a long manuscript to review….).

In response to the general remarks and specific comments, we have added results in stage 3 and 4, and revised the text as follows (see page 15/16):
Stage three: the refined programme theory
Data from interviews and focus group helped us refine the wording of six CMO configurations and added three new configurations that linked the chains of outcomes. Participants emphasized that the national bodies should entrust QC with QI, but national organisations or professional association should be sufficiently flexible to allow local QCs to implement their plans, giving them a feeling that they had a say and a job to do. At the level of the group, they pointed out that individual character traits and different professional experiences along with differing opinions provide a necessary tension to stimulate lively discussions as long as mutual respect exists. However, there are (a few) individuals who experience critical feedback as threat to self-image and, as a consequence, withdraw or disturb the group process.

Interviews with stakeholders across health care systems
The interview data suggested that QCs can only succeed if they are embedded in a wider system that helps participants to negotiate and sign contracts with governmental bodies or health insurance companies, organises training and supervises facilitators, offers courses on QI in PHC, and facilitates access to educational material and timely data on practice performance (CMO configuration 1b ‘being embedded in a QI system’).

Discussion
Reviewer 1 (page 19, line 38) We had several insights that had not been reported in current QC literature.
Rephrase, it sounds as if you just are making things up, which does not do justice to the extensive process of your work. But at the same time, I was surprised that you mention Cognitive dissonance here as a key new finding when this theory was not included in table 2.

We have rephrased this sentence as suggested and revised for clarity (see page 20):
Literature and interview data provided us with mechanisms that had not been reported in current QC literature. Cognitive dissonance, like conflicting attitudes, beliefs or behaviours that create unease, is a mechanism that compels GPs to reflect on, accept, and adopt new reasoning to resolve inner conflict. This is the starting point of transformative learning.

Implications for policy and practice

Reviewer 1 (page 21, line 2) Not all recommendations will apply to every QC.

 This, to me, sounds as a no-brainer in a realist manuscript. This could be stated with more confidence as this manuscript gives some insights in the context where it can be applied and where not.
Agree – deleted and not necessary (see page 20)

Reviewer 1 (page 22, line 34) To mitigate the risk of social desirability bias, AR carefully posed neutral interview questions and tried to avoid embedding assumptions in his questions.
 A bit strange sentence. Why should respondents be inclined to answer with social desirability when asked about program theories? This nearly made me jump to the supplementary material and see whether I misunderstood what you actually had been doing.
We agree that this sentence seemed to cause confusion and does not add anything to the article – it has been deleted (see page 21).

Reviewer 1 (page 22, line 40) the risk of selection bias if researchers choose underlying theories and synthesise them ad hoc.
This is a strange part of the manuscript. I think that selection bias is never entirely possible to be left out and what would be the problem with that? The manuscript is already very detailed and you had to make choices. In making those choices, it is possible that you included some theories and neglected others, but as you have been involving so many people, well, I think that many other papers have a lot more selection bias. If mentioning selection bias is needed, explain what that would mean for your results.

We agree revisions are needed and have added the following text in response (see pages 22): Our study has some limitations The resulting theory relies on the detail and depth of the reports we identified in our literature review and on the veracity and adequacy of the information stakeholders revealed during 2015-2020 in Europe CMO configurations reflect and explain the complex process in QCs in the current context of European primary health care, and may need to be adapted in response to future changes.
Quality appraisal of relevance and rigour of data that contributed to the emerging programme theory may depend on research team judgements. Another team might have taken differing decisions. We could not include all theories found during iterative searches but had to make choices of the ones that fitted best. Finally, we could relate all aspects of the theories in Table 2 to the CMO configurations to explain how the programme theory's mechanisms interrelate.

Conclusion

Reviewer 1 (page 23) can improve practice, foster professional development, and increase psychological well-being among participants

- See my earlier comment, here it is written differently again.

We have edited the text for consistency (see page 23):
Our consolidated programme theory explains how participation in QCs can improve standard practice, enhance professional development and increase well-being

Reviewer 1 (page 23) The requirements for successful QCs are

- Earlier (at the start of the results section), you use this wording "The most important contextual requirements for successful QCs …". It looks as if this is a minor detail but in an overall line of reasoning that stresses the importance of context, I find it important to repeat the contextual aspect, otherwise the conclusions sound a bit mundane 'you need sufficient support', well, that does not sound too inspiring.

Reviewer 1 Overall the conclusion could be adjusted because it reduces the value of your work by making it so concrete that the added value of all the realist work is getting out of view.

We have edited the text as suggested for clarity (see page 22):
The most important contextual requirements for successful QCs are 1) governmental trust in GPs' abilities to deliver QI and appropriate support like professional facilitation, 2) training in QI techniques, 3) access to educational material and personal performance data; 4) granting protected time, appropriate venues, and financial resources for QC group members.

Reviewer 2
Specific comments:

P9, line 5: What is meant with "the larger program theory"? Is this the by theory updated initial PT? To avoid confusion we have changed it to (see page 10):
emerging programme theory

P9 line 19: "preliminary programme theory" Is this the 'larger programme theory' mentioned earlier?

For clarification, we defined 'preliminary theory' in method / stage 1 / co-inquiry- see page 8) We co-designed the preliminary conceptualisation of the programme theory, in short, preliminary programme theory, in discussion with the stakeholders, supported by local programme documentation and training material.

For sake of clarity we have revised the text and changed this to (see page 10):
We created a data extraction framework based on the preliminary programme theory

P10 line 11: "Interaction between context and mechanism to facilitate or constrain QCs" What do you mean by facilitating or constraining QCs? To do what? Reach outcomes? Better to rephrase.

As suggested, we have rephrased this sentence to (see page 11):
We summarised these configurations into descriptions of interaction between context and mechanisms that either facilitate or hinder QCs to reach their outcomes.

P10, line 13: In what way were they closely related? Author, study, context, intervention, theoretical framework?
As reviewer 1 pointed out the same lack of precision [Reviewer 1 (page 11, r. 13) …a bit more detail], we moved the explanation from the original location in the manuscript to methods section and expanded the explanation about the search process (see page 9):
These kinship papers had common contextual features or theoretical backgrounds to the referring studies. We categorised these papers into kinship networks based on common themes, common contexts like geographical area, and common methods of organising QCs (e.g., papers that tested similar didactic methods or similar QI tools in QCs). We broadened the search by examining citations in reference lists and Web of Science and searched manually for closely related papers (kinship papers).

P10 line 20: 'Predictable' is not a word I would use in the framework of a realist study. Even if it is only 'semi'.

As suggested, we have revised the text to (see page 11):
We iteratively arranged and rearranged the CMO configurations, moving between the papers, their data, and families, and built patterns of outcomes (demi-regularities) to develop the programme theory (see supplemental material 3).

P10, line 58: Unclear what is meant with 'underlying reasoning for QC interactions'. Do you mean the underlying mechanisms relevant in explaining QC outcomes?

For clarity we have revised the text to (see page 12):
After explaining the literature-based programme theory in plain words, AR offered contrasting options for participants to discuss. Then, he asked them to share their understanding of the underlying mechanisms and explain QC outcomes.

P11, line 34: You say that "formal theories capture a PTs underlying mechanisms". This is not clear to me, PTs are constructed around underlying mechanisms, they entail and explain the relevant underlying mechanisms, formal theories may give insights into the mechanisms that are relevant and how they work and interact with each other.
Thank you for the insight, we have revised the text for sake of clarity to (see page 13): Formal theories explain how mechanisms interrelate and how they may work across different disciplines.

P12 line 18: What do you mean by "preliminary programme"? Was a programme implemented? Or is it rather a preliminary conceptualization of the programme?
We defined preliminary theory earlier in the revised text (method / stage 1 / co-inquiry p8) We co-designed the preliminary conceptualisation of the programme theory, in short, preliminary programme theory, in discussion with the stakeholders, supported by local programme documentation and training material.

P15 table 2: I think the title should be the other way around: Existing theories and their relationships to the CMOCs ....

Revised as suggested to (see page 16):
Existing theories and their relationships to CMO configurations in the programme theory

P 18, line 23 : Typo?: "of the individual" Corrected to (see page 19):
The group's capacity for problem-solving surpasses the ability of the individual when members share and pool their experiences and views

P21, line 14: What does "these realist approaches" refer to? Please see our response below.

P21 line 14: "the resulting theory relies on the detail and depth of the reports we identified in our literature review.": This is a limitation to every synthesis approach not only the realist. Moreover, isn't the final theory also based on the interviews? At least this is what you write later in this paragraph and which makes this first sentence contradictory to what follows.

We have done significant revision of 'Limitation' section to address the issues raised by the editors' and reviewers' comments as follows (see pages 21/22):

Our study has some limitations The resulting theory relies on the detail and depth of the reports we identified in our literature review and on the veracity and adequacy of the information stakeholders revealed during 2015-2020 in Europe. CMO configurations reflect and explain the complex process in QCs in the current context of European primary health care, and may need to be adapted in response to future changes.
Quality appraisal of relevance and rigour of data that contributed to the emerging programme theory may depend on research team judgements. Another team might have taken differing decisions. We could not include all theories found during iterative searches but had to make choices of the ones that fitted best. Finally, we could not relate all aspects of the theories in Table 2 to the CMO configurations to explain how the programme theory's mechanisms interrelate.

Figure 2: The programme theory explains how and why QCs reach their outcomes in specific contexts; it is a chain of reasoning. Based on the data, I tried to express the chains of CMOcs in words that refer to the process in QCs based on the data we gathered

CMO 1a: Make a difference in what? Is the mechanism here not intrinsic motivation? Especially since you refer to the self-determination theory?
We can agree with the peer reviewer that motivation is one possible mechanism, but in our interpretation of the data, we wanted to produce a more fine-grained understanding of this motivation; to make a difference in the sense of having a significant effect on a situation.

CMO 1b: Can this mechanism be called increased self-efficacy? Also related to the theory of planned behaviour.
We can agree with the peer reviewer that self-efficacy is the possible mechanism, but in our interpretation of the data, we wanted to produce a more fine-grained understanding of the quite broad concept of self-efficacy.

CMO 1c: Is it sufficiently different from CMO 1a? Again, isn't the mechanism intrinsic motivation? The difference from 1 a is that a decentralised organisation is necessary to allow QC participants to actually make changes according to local needs. We do agree with the peer-reviewer that it is about motivation but we wanted to produce a more fine-grained understanding

CMO 1d: Again, is the mechanism not self-efficacy?
Same response as for 1b

CMO 2a: Is "prefer learning in QCs" really a mechanism that gets triggered by the mentioned context? They prefer it above which other method? If they really prefer it, than why do they need a financial and other incentives? That's the opposite of "prefer" (which I think refers to intrinsic motivation).
This is about the question why GPs meet in QC groups. There were many papers on this topic with overlapping answers GPs gave, even in interviews: there was hardly ever just ONE reason but a combination thereof and depending on the individual with varying emphasis (mechanism as causal tendency): M1 M3 want to learn in groups, but some of them need incentives more than others; some need more administrative support than others ; for some, all the support actually motivates them M2.

CMO 2b: I am doubting about the link between the mechanism and the outcome. "Members get to know each other" is a direct consequence of the introduction and the socially enjoyable contact not of the described mechanism. I also do not see how 'deciding on rules' follows from the mechanism "members want to interact with equals". The contextual condition is already about contact and discussions. Maybe it is explained more clearly later on? But her eit is not very clear to me. I am not sure that people who get introduced to each other automatically get to know each other. Getting to know each other is a process of increasing trust for each other; it is also a process of forming and norming the group. People seem to have the basic needs of becoming related, a need to interact with other human beings. People need to experience a sense of belonging and attachment to other people (Self Determination Theory).

CMO 2c: I feel the outcome is the same as the mechanism. I would say that the outcome is increased intrinsic motivation to participate.

The mechanism is need for autonomy – a feeling of being in control of their own behaviour (Self Determination Theory). if someone just decides or assigns facilitators or topics, then participants will not consider it their QC.

CMO 2d: Is this the mechanism. Or is it more about reluctance to speak in large groups, or the earlier mentioned issue of knowing each other well becomes more difficult in larger groups, or a kind of bystander effect when it comes to intervening (as we also see in our classrooms or at conferences)? Cannot keep up with each other: as I understood the data, people get the feeling of not being able to follow all interactions between members and turn passive. This may also include the lengthy process of collecting opinions on an issue when the you have a large group. People tend to turn off their attention and are not willing to listen to 21 comments.
I didn't have a paper mentioning lack of trust and no interview data either but I can imagine that it may be lack of trust that people feel in large groups that may be a mechanism.

CMO 2e: I very much like this one.

CMO 2f: Same here, nice insight.

CMO 3a: No remarks

CMO 3b: I feel the mechanisms are reasons (not reasonings) rather than mechanisms. Or does your conceptualization of a mechanism entail reasons? I would say that these mechanisms are actually contextual conditions (characteristics of the actors) and the mechanism is motivation (one of the non-autonomous motivation forms of the SDTheory)
According to our interpretation of the data in the literature and in the interviews, participants become involved in the group and share their positive experiences and failures because they want to improve their professional competence (M), gain professional confidence (M), and fulfil their professional potential (M). However, this happens only if you have an experienced facilitator supporting the group and a safe environment (where no one gets emotionally hurt). Mechanism in complex social interventions: reasoning and resources that QCs provide.

CMO 3c: No remarks

CMO 3d: I feel the outcome is not really an outcome but a rephrasing of the mechanism. The because part refers to an argument rather than a causal explanation.
Maybe putting it into words makes it clearer:
if someone teaches GPs about cases they hardly ever see, or at a level of detail that is only important to the specialist (for example, cardiologist), then that knowledge is not relevant to GPs, because it doesn't relate to their everyday work. Therefore, experiential learning, using GPs' own cases seems to make it relevant to GPs. In the literature and in the interviews, they argue that it is relevant as they immediately can use it in their practice.

CMO 3e: No remarks

CMO 3f: I feel the outcome is at another level than the mechanism. The group will create a learning environment but the mechanism 'social learning' happens at the level of the individual. I think they create a learning environment because they want to learn through observation, imitation, and modelling. At a contextual level, QC groups need a skilled facilitator for this.

CMO 4a: No real remarks. Although I feel the CMO is incomplete. Isn't more needed to make QCs design creative solutions? This can off course be a consequence of insufficient data on the context.

There are several papers on social interdependence theory (business and education) explaining why groups may work together towards a common goal.

CMO 4b: “Feel unsafe” is somewhat strange here. Why not stick to the feeling of lack of autonomy. Based on our interpretation of the data feeling threatened in their very professional role was more prominent an issue than lack of autonomy.

CMO 4c: I think the second part of the mechanism is an action rather than a mechanism and the first part seems to be a contextual conditions or at least a bit weak to be seen as the mechanism that generates the outcome here. Not sure how the “act and negotiate cooperatively to achieve a common goal” links with “participants will adapt and generate new knowledge for local use”. I tried to improve this CMOC
Thank you for your insights on this CMO configuration, with further analysis and thought we have revised it to:
If participants maintain a learning environment based on trust that promotes the exchange of knowledge, assisted by facilitators who use professional techniques (e.g., contentious discussion, reaching consensus and role play) (C), then participants will adapt and generate new knowledge for local use (O) because they have a sense of collective responsibility (M).

CMO 4d: Is this the social learning mechanism?
Participants learn from each other, no doubt, more specifically, literature and interview data suggest that if the group addresses barriers then they will more likely implement the intervention. the reasoning seems to be that this is possible because participants support each other and develop strategies to identify and overcome these barriers

CMO 4e: This clearly relates to the self-determination theory, yet according to this theory this leads to more intrinsic motivation.
I understand your comment. According to our interpretation of the data from literature and interviews, it was described as feeling of satisfaction, responsibility, and stewardship (which may reflect high intrinsic motivation to continue)

CMO 4f: Not clear how “announcing an intention to change” triggers the mechanism “they and other in the group think it is a good idea”. Is this not about peer pressure? we have reflected on the comment and agree that a form of peer-pressure is the likely mechanism here.
So, we have rephrased to:
If participants in a QC announce their intention to change (C), then they are more likely to implement the change (O) because they have openly committed to each other to make changes (M)

CMO 4g: No remarks, but maybe refer to self-efficacy as a mechanism?

CMO 5a: Again this is self-efficacy.
We can agree with the peer reviewer that self-efficacy is the possible mechanism, but in our interpretation of the data, we wanted to produce a more fine-grained understanding of the quite broad concept of self-efficacy.

CMO 5b: No remarks

Supplemental material 4: The description of the participants is really detailed which makes that they are no longer anonymous. This is an ethical issue.
We adjusted the tables to address this issue.

We hope you and the reviewers are satisfied with our response. If you have any other questions or suggestions, we will be happy to address them.

Please address all correspondence concerning this manuscript to me at
adrian.rohrbasser@bluewin.ch

Thank you for your consideration of this manuscript.

VERSION 2 – REVIEW

REVIEWER	De Groot, Esther Utrecht University, Julius Center for health sciences and primary care
REVIEW RETURNED	17-Feb-2022

GENERAL COMMENTS	Dear authors, You have answers my earlier questions perfectly. This is a valuable manuscript which will help those interested in this type of realist review, as well as GP (educators) who may be more interested in the content on quality circles. I will certainly share the final publication with my peers.

REVIEWER	Renmans, Dimitri University of Antwerp
REVIEW RETURNED	01-Mar-2022

GENERAL COMMENTS	Dear authors, I would like to thank you for the adaptations you have made and the extra information you have given. It is a very nice and impressive piece of work. All my comments were addressed in a satisfactorily way. Best