Introducing Q Methodology to Program Evaluators

Harris, Kev; Oatley, Chad; Mumford, Steven; Pham, Phung K.; Nunns, Heather

This document version is the:
Peer reviewed version

The final published version is available direct from the publisher website at:
10.1177/1098214020932227

Find this output at Hartpury Pure

Introducing Q Methodology to Program Evaluators

Introducing Q Methodology to Program Evaluators

This method note presents Q methodology as a useful tool for evaluators to add to their practice toolbox. Q methodology is a system of assumptions and procedures to investigate and analyze a phenomenon that may encompass multiple coexisting perspectives among different people (Stephenson, 1993). According to Brown (2019), at the heart of Q methodology is its focus on "the communicability of opinions and perspectives on any topic" (p. 1). This is important for evaluation because subjectivity (i.e., things we may express internally and externally; Brown, 2019) is central to our perceptions of programs, and the decisions and actions that determine their relative success or failure.

Building on the contributions already made by the likes of Brown (2019) and Militello, Janson, and Tonissen (2016), we present Q as a methodology for understanding human subjectivity and describe its place in evaluation for measuring perception. We first provide an explanation of Q methodology, followed by a brief overview of how evaluators are using it. Next, we walk through the steps for implementing Q methodology. Then for practical demonstration, we describe how Q methodology was used in an evaluation within the UK. We conclude by considering the benefits and challenges of employing Q methodology within evaluative contexts.

Overview of Q Methodology

Q methodology was developed 85 years ago by physicist-psychologist William Stephenson (1935a, 1935b). Stephenson's work correlated and factor analyzed *people* (Stephenson, 1935a, 1952), or more precisely, people's perspectives or viewpoints on a topic of interest at a certain moment in time. At the center of Q methodology is its attention and focus on subjectivity (Brown, 2019), which is at heart of human volition and decision-making. As a

student of Stephenson, Brown (1980) has long contributed to the conceptual foundations of Q

methodology. In regards to subjectivity, Brown (2019) states that through internal and external

conversations within and between individuals, opinions and conjectures manifest themselves.

Appreciation of subjectivity forms the fundamental premise of Q methodology, and its

foundational concept of a "concourse." Brown (2019, p. 2) highlights concourse as the body of

talk and thought about any topic or situation, whereby many "tributaries" feed in. In Q

methodology, the concourse is commonly captured by gathering communications that represent

the opinions, attitudes, and beliefs about a particular phenomenon (Hensel, 2017). In evaluation,

there will be a concourse in existence for any program or intervention, such as prevailing views,

recent reports, academic literature, and even photos and music. This is crucial for evaluation

work, as the concourse forms the basis for measuring perception and understanding programs.

Participants in Q methodology reveal their subjectivity by sorting a representative sub-

set of the concourse, usually in the form of written statements. Sorting patterns are then

statistically analyzed across individual sorts to examine the degree of consensus and divergent

viewpoints within a group (Ramlo, 2015a). In this way, Q methodology is thought to allow

researchers to make subjectivity operant, and thus available for scientific measurement (Ramlo

& Newman, 2011).

Q methodology (hereafter referred to as Q for brevity) entails a unique and

philosophically consistent combination of qualitative and quantitative research approaches that

brings benefits, but also leads to controversy, within social science research communities (Ramlo

& Newman, 2011; Brown, 1980). The focus of Q is on interpreting meaning through the

participant's perspective, traditionally the domain of qualitative methods. At the same time, it

involves statistical methods that lend rigor to its resulting factors, or operationalized viewpoints.

This combination permits researchers to 'measure' subjectivity and empirically group people based on unique sorts that represent their own nuanced perspective (Ramlo, 2016).

On the other hand, critics have dismissed Q as insufficiently scientific, especially from a traditional quantitative research perspective, because it involves elements of subjective interpretation on both the participant's and researcher's parts (Ramlo, 2016). Not surprisingly then, the more procedural components of Q, such as the sorting process or factor analytic techniques, have been adopted by quantitative researchers piecemeal without appreciation for the full methodology.

Nevertheless, Q methodology has been proposed for use in a wide variety of applied social science disciplines, as an approach for uncovering divergent or marginalized perspectives, clarifying complex constructs, grouping stakeholders, and facilitating dialogue and planning processes. Specific endeavours for which Q has been recommended include participatory policy design and appraisal (Militello et al., 2016; Ockwell, 2008; Wolf, 2013; Baker, Thompson, & Mannion, 2006; Robbins & Krueger, 2000; Zabala, Sandbrook, & Mukherjee, 2018), tailoring counselling and social work interventions to individual needs (Shemmings, 2006; Stickl, Wester, & Watcher-Harris, 2017; Ellingsen, Storksen, & Stephens, 2010), informing re-accreditation processes in higher education (Pruslow & Owl, 2012; Ramlo, 2015b), and promoting leadership development in the workplace (Woods, 2011).

Uses of Q Methodology in Program Evaluation

Q has been proposed for use in program evaluation as well, most commonly to enhance authentic stakeholder engagement by identifying and classifying stakeholders' needs, implicit program theories, and changing viewpoints on an issue (Ramlo, 2011; Millitello et al., 2016; Harris, Henderson, & Wink, 2019). In line with these suggestions, evaluators have most often

employed Q to gather stakeholder perceptions of a program – including its feasibility, fidelity of

implementation, perceived outcomes, and critical elements – and suggest formative

improvements (Butler, Hare, Walker, Wieck, & Wittkowski, 2014; McPherson, Sanders,

Schroeter, Troy, & Wiseman, 2016; Pike, Wright, Wink, & Fletcher, 2015; Corr, Phillips, and

Capdevila, 2003; Lazard, Capdevila, and Roberts, 2011; Militello & Benham, 2010).

While stakeholders queried through Q methodology are most often program participants,

they at times include program staff, in order to identify and compare perceptions of the program

across different stakeholder groups. In one participatory evaluation (Lazard et al., 2011),

program staff administered the Q sorts to program participants, promoting process use.

Evaluators have also used Q to explore stakeholder engagement in participatory evaluation, in

order to identify participants' initial motivations (Militello et al, 2016; Thompson, 2015),

perceptions of the process (Danielson, Webler, & Tuler, 2009; Cotton & Mahroos-Alsaiari,

2015), and their methodological preferences (Kelly, Moher, & Clifford, 2016).

Indeed, Q has been widely used by evaluators to improve their own evaluation processes

and systems. Q has been employed to gather stakeholders' differing perceptions of performance

measurement systems and resulting data use (Velez, 2006; Frayne, 2014; de Jonge et al., 2017;

Militello, Bass, Jackson, & Wang, 2013), identifying divergent priorities for reconciliation.

Evaluation researchers have used Q to explore stakeholder perceptions of process use resulting

from participation in evaluation (Baptiste 2010, 2011; Akanban, Darko-Osei, & Atengdem,

2013), as well as stakeholder self-assessment on different aspects of evaluative thinking

(Mumford, 2018; Nunns, 2017; Pham, 2018).

Q methodology has been used for summative program evaluation, albeit less commonly.

For instance, reported in a series of articles, Cuppen (2013, 2012) employed pre and post Q sorts

as part of a quasi-experimental evaluation of a dialogue intervention. The initial Q results were used to select stakeholders for participation in dialogue, structure dialogue sessions, and match participants to a comparison group that did not receive the dialogue intervention. Then, Q was used as a repeated measure at the end of the evaluation with both dialogue participants and the matched comparison group, to compare changes in each group's Q sorts over time and to one another. The evaluator found that dialogue participants changed their sorts (and thus their perspectives) more than the comparison group, and showed broader awareness of the issue's complexity after the dialogues.

Q offers practical benefits and opportunities related to these diverse aims. Because it inverts traditional factor analysis (i.e., variables are people, and items comprise the sample that in turn represents the concourse), Q can be employed with samples of 30 or fewer participants. Dziopa and Ahern (2011) suggest that a 1:1 ratio of people to items is appropriate. Q is seen as more engaging and sensitive than a traditional Likert-scale survey for eliciting participants' complex and idiosyncratic viewpoints, and it may help attenuate participant bias and enhance validity of the resulting perspectives (Ho, 2017; Fluckinger, 2014), although it is also potentially more time-consuming to implement than traditional survey methods.

Q is focused on the measuring of individual perceptions, but its contributions to generalizability are theoretical rather than statistical (Watts & Stenner, 2012). That said, Q sort results have frequently been used to strategically sample participants for follow-up interviews, and compare perceptions of program implementation across multiple sites. Q results can also form the basis for developing traditional surveys (Danielson, 2009), such as on the effectiveness of partnerships and collaborations (Jeffares & Dickinson, 2016).

In summary, Q methodology has been employed in evaluation both formatively and summatively, and as an approach to developing new programs. Evaluators have also used Q to gather stakeholder feedback and reflections on the evaluation processes and performance measurement systems they produce, and their construct definitions and measurement tools. There is thus strong precedent for adopting Q methodology within evaluative contexts, and perhaps with additional awareness and guidance, its popularity among evaluators will continue to grow.

Q procedures

An overview of Q procedures is provided in Table 1. In summary, Q explores subjective meaning via statistical procedures. Participants in a Q study typically sort statements (or pictures or objects) about a specific topic, which provide a representative view of all the discourses related to that topic. Participants are asked to sort or rank the statements according to a specific instruction, to reflect on their subjective interpretations of the statements in relation to the instruction. The completed sorts are then compared and grouped via factor analysis using specialized (often free) software (e.g., PQMethod).

[Insert TABLE 1 Here]

The resulting factors represent statistically distinct groups of participants' perspectives about the topic of interest (Brown, 1980; Watts & Stenner, 2005). Using abductive reasoning, the evaluator interprets each factor to understand what, how, and why differences exist in points of view, attitudes, or opinions about the topic. For more complete treatments of the steps to implement Q methodology, including sample conditions of instruction and sorting grids, see, for example, Watts and Stenner (2012), Militello et al. (2016), and a host of other tutorials on Q cited in this article and elsewhere.

Demonstration Case: Use of Q in an Evaluation from the United Kingdom

What now follows is an illustration of a UK-based evaluation that implemented Q. The evaluation was conducted (by two of the authors) to evaluate a UK public and voluntary sector partnership, pseudonymized as Partnership4People (P4P) to protect confidentiality.

Funded by a large organization in the UK, the P4P program was implemented via a 'hub and spoke' approach whereby a key agency triaged referrals and coordinated a partnership of stakeholders across the public and voluntary sector to deliver tailored support to people living in crisis (for example, financial poverty and unemployment). The project focused on achieving outcomes at both an individual level for the person in crisis, and at an inter-organizational level to operate effectively and coherently as a network of providers in partnership to support the person(s) in taking action to address the crisis. There was motivation among the P4P stakeholders to obtain insight and understanding about the program outcomes, the causal processes that lead to them, and how they could use the findings to inform future practice to enhance effectiveness in responding to crises.

The evaluation focused on mechanisms of change, and therefore it was essential to investigate how the network's resources (mechanisms) provided through P4P partners were helping network stakeholders foster the intended outcomes (Pawson and Tilley, 1997). To capture this, Q was mobilized to tease out what worked, and how, for the different stakeholders. The overriding evaluation aims were addressed across three key phases: *Phase 1: Developing the concourse (program theory), Phase 2: Selecting participants and Q sorting exercise,* and *Phase 3: Interpreting and planning for improvement* based on evidence synthesis. These phases are discussed below with reference to Table 1.

Phase 1: Developing the Concourse (Program Theory)

This phase focused on developing the concourse about P4P, which we did collaboratively with practitioners. They were not comfortable with the term 'concourse' because it confused them. To address this, we referred to this phase as theories about how and why the P4P program *could* work to achieve the outcomes envisaged. We made sure that, while we referred to this as the program theory development stage, we were capturing all of the important aspects of concourse central to Q. Developing program theory usually involves engaging with key stakeholders to gain their own perspectives about how they think the program works, alongside the analysis of existing literature relevant to the program at multiple levels. This usually leads to the creation of candidate program theories (Pawson and Tilley, 1997) which are open to investigation.

Harris et al. (2019) have argued that the development of program theory and the concourse in Q are mutually compatible because they both place significant emphasis on context. Pawson (2013) categorises context into four layers: 1. individuals (their characteristics and capacities), 2. interpersonal relations (stakeholder relationships), 3. institutional settings (rules, norms and customs), and 4. infrastructure (social, economic, and cultural setting). This overview of context identifies the key circumstances, boundaries, and backgrounds, which are synonymous with establishing the concourse in Q methodology.

Specifically, we explored the volume of discussion (Stephenson, 1982) about P4P, making use of primary and secondary data to establish the concourse. This involved interviewing key stakeholders involved in delivering the P4P program, as well as examining industry and academic literature associated with P4P and similar programs. This concourse phase helped establish a host of "if/then" and "because" narratives about the P4P program.

This process informed the development of the Q statements (Table 2, consisting of 40 statements from the concourse). We constructed the statements iteratively and carefully to ensure that each one reflected the program theories about P4P. Collectively the statements focused on key contexts and outcomes associated with the P4P program. The statements were validated in two ways: first, among the evaluation team through our reflexivity and engagement in phase 1 of the evaluation; and second, the statements were presented to a small sample of practitioners leading P4P to confirm that they made sense and reflected the program.

[Insert TABLE 2 Here]

Phase 2: Selection of Participants and Q sorting

Our evaluation questions focused on gaining insight into how practitioners from different organizations (including referrers to P4P) understood P4P, how it worked as a network, and what it contributed towards supporting people in crisis. Given the reliance on good inter-organizational partnerships, including for learning and knowledge transfer, it was necessary to select a range of practitioners with varying roles and experiences to complete the Q sorting exercise. Participants in the P-set were made up of 18 practitioners, consisting of stakeholders from various government departments and non-governmental organizations.

Each participant ranked 40 statements (Table 2) into the Q-sort grid, completing the Q sort activity individually at times that were convenient for them. This was important because the Q sorting procedure required a lot of explanation and was very new to most of the participant sample. While the participants completed the Q-sort, the evaluator made notes of questions asked, comments made, and relevant contextual information about the participant's role in relation to P4P, as this data could inform understanding and explain why participants sorted statements in the way they did.

Once all practitioners had completed their Q-sorts, we entered each sort into PQMethod, a free statistical software program designed for Q, and selected by-person principal components analysis. This analysis identified eight possible groups for qualitative interpretation across the 18 completed Q-sorts. Having examined the statistical relevance of each group, we selected four groups for full qualitative interpretation based on how many people loaded on to each factor, study variance, eigenvalues, and 'feel' for the data (Table 1). This was made up of Group 1 with four participants, Group 2 with three participants, Group 3 with six participants, and Group 4 with three participants.

Phase 3: Interpretation and Planning for Improvement in P4P

This final phase consisted of interpreting each resulting Q sort score for each of the four groups to produce a "realist holistic narrative" (Harris et al., 2019) of how and why P4P works. In order to foster use of the evaluation findings, the evidence needed to be presented in a way that was easy to interpret and understand for stakeholders and funders. In considering the different ways to interpret and communicate the findings of the Q sort (Table 1), Watts and Stenner's (2012) approach was adopted, as it allowed the Q sorts within each Q group, and between the different Q groups, to be qualitatively interpreted to explain the statistical relationships in narrative form. Watts and Stenner's (2012) approach was followed by drawing upon the data presented from the comprehensive factor analysis output file for the 4 subgroups. Their approach (crib sheet method) specifically involved identifying where participants in each subgroup had ranked their statements, as well as distinguishing statements (from other subgroups) which we used to build our abductive holistic narratives.

These narratives were produced around the different subjective viewpoints, drawing comparisons with each partner's organizational role and context. This was important given that

we were exploring subjectivity at the organizational level of operation, where a lot of inter-agency networking was taking place. This helped the P4P funders make sense of the different ways P4P worked in relation to the different views of the stakeholders' Q sorts and subsequent groupings. In doing so, we provided insight into how to work with different stakeholders based on their interpretation of P4P to better target, communicate, and integrate P4P, and achieve better outcomes for people in crisis.

What now follows is the presentation of findings from two of the subgroups that emerged from the Q sort activity (we have selected subgroups 2 and 3). As indicated in the previous section, the stakeholders who completed the Q activity loaded across four subgroups. We presented these factors using the Watts and Stenner (2012) format, where the story of the narrative is supported by the statement rankings. For example, "25+3" would mean that those in this particular subgroup would have rated statement 25 as "plus 3" on their respective Q sort. This use of statement rankings helped us to justify the relevance and credibility of the story behind the narratives which was important to the client commissioning the evaluation. For the purpose of this method note, we present subgroups 2 and 3 in the narratives below to illustrate the data informing the establishment of the subgroups (we have not supplied the full factor analysis output data file due to length and size). This is then followed by Table 3, which captures the distinction between each subgroup. At the top of each column, a summary descriptor is provided about the subgroup. Underneath the descriptor are key bullet points that represent the subjective shared viewpoints of the practitioners who loaded into each subgroup.

Subgroup 2 narrative

"Collaboration and smart working facilitates individual and system level outcomes, but is the

approach sustainable?"

This subgroup represented a total of 3 participants, and is formed by a manager of an

organisational partner of P4P, a key partner in supporting people referred to them and a volunteer

at an organisation which supports people to access P4P.

This subgroup strongly agrees that through the adoption of tailored action plans, the P4P

programme helps those who are in genuine hardship (5 +4). These action plans help to motivate

individuals to take control of important aspects of their lives in the longer term (12 +5) including

their own finances (11 +1). Furthermore, this subgroup feels the expert advice from the P4P

programme helps people to make immediate changes to their lives (14 +3) which leads to less

dependency on resources in the longer term (21 0). Interestingly, this motivation does not extend

to empowering people to take positive action against the cause(s) of their crisis (13 -3).

Subsequently, like most subgroups, this subgroup feels that more work needs to be invested in

ensuring the P4P model is a sustainable approach to preventing crisis from reoccurring in

peoples' lives (10 -4). However, this is also dependent on factors outside of the networks control

such as broader problems in line with the socio-political context.

 The ability of P4P to take time to listen and understand the causes of why people are in

crisis is well recognised within this subgroup (31 +4), which places importance on the role of

listening to understand individual circumstances . However, this subgroup does not feel the

programme has the capacity to effectively build peoples' skills and knowledge of the appropriate

services (4 -4) around them to successfully move out of crisis. In addition, in contrast to any

other subgroup, this subgroup did not feel that the financial support provided through the P4P

programme provides the opportunity for people who have fallen into crisis to meet their basic

needs (2 -2). Thus, this subgroup feels the programme does not ensure benefit issues are dealt

with in an efficient and reactive way to avoid individuals falling into further crisis (32 -3). In

addition, the theory of individuals accessing the programme and sharing knowledge and

solutions with others who may be experiencing a similar crisis (8 -3) is not something that

regularly happens.

Overall, this subgroup feels that the support P4P provides for individuals reduces

pressure (22 +2) and strain on system resources (37 +3). This subgroup distinctively recognised

that the ability for the P4P network of partners to work within a collaborative environment (24

+2) led to them effectively working together to facilitate the holistic needs of every person in

crisis being met (23 +3). This subgroup shows evidence of the programmes achieving a range of

individual, organisational and system level outcomes.

Subgroup 3 narrative

"The key stakeholders are informative experts and provide immediate support"

This narrative represented 6 participants (more than any other subgroup), included staff

working within departments across the system, and organisational partners of P4P. This

subgroup is comprised of staff whom are connected to P4P and understand how it works to

support the system at a strategic and operational level.

This subgroup feels that the P4P network consists of all the key organisations (25 +3)

necessary to alleviate crisis. The capacity of these organisations to have the knowledge and

understanding of the wider system results in an effective and efficient model that deals with

people in immediate crisis (30 +4) within 24 hours (19 +4) by signposting people to the

appropriate services (9 +3). This subgroup shares commonality with other subgroups that P4P

helps people in genuine hardship (5 +2) and provides help in an empathetic way to support

people (3 +2). Similarly, this subgroup identifies the contribution of P4P in being able to take

time to listen, understand the causes of individual crisis and work with the individual to make

changes to resolve the issue(s) causing crisis (31 +3). In addition to this, this subgroup shares

viewpoints around how P4P works; through being able to signpost people to the appropriate

services that they need (9 +3) and providing financial support which helps to build individual

capability (1 +1), and more strongly, creating an opportunity for people (2 +5) who have fallen

into crisis to meet their basic needs.

 This subgroup identifies how the current eligibility criteria results in the programme

working with people who are most in need of their support (6 +2) and operates in a responsive

way to people in crisis within 24 hours (19 +4). As such. This supports to ensure benefit issues

are dealt with in an efficient and reactive way to avoid individuals falling into further crisis (32

+1).

 There is scepticism about how the P4P model is a sustainable approach to preventing

crisis from reoccurring in peoples' lives (10- 2), supported by how P4P helps people to stay out

of crisis longer-term (15 -4). This is linked to why this subgroup disagrees that the expert advice

provided through P4P programme will not always helps people to make immediate changes to

their lives (14 -4) which prevent crisis from reoccurring. Further, some causes or crisis are by

larger system changes such as the introduction of Universal Credit which this subgroup feel

strongly that P4P is constrained in helping to reduce complexity around for people in crisis (38 -

3). In addition, this subgroup is apprehensive about how P4P ensures housing issues (33-1) and

domestic living issues (35-1) are dealt with in an efficient and reactive way to avoid individuals

falling into further crisis. The reason for why this subgroup feels less strongly about the capacity

of P4P to have achieve a wider positive impact with system partners could be due to the current

funding model and perceived sustainability of the P4P network which impacts on how it is

valued by other service areas (28 -1).

Building on this, like subgroup 2, this subgroup does not feel that people who have been

supported by the P4P programme currently have to capacity give something back through

volunteering with P4P partners (7 -5) and therefore unsurprisingly the programme does not

provide individuals with opportunities to share knowledge with others who may be experiencing

similar crisis (8 -3). This could be because within the context of attempting to reduce the

complexity around the new universal credit system (38 -3) people are still concentrating on

taking control of their own finances in the shorter term (11 -2). However, this is not to say that

P4P could not achieve this, if it was financially resourced in an improved way.

Overall, this subgroup feels it is the financial support provided through the P4P

programme which helps people who have fallen into crisis to meet their own basic needs (2 +5),

as opposed to the expert advice from the P4P programme which helps people to make immediate

changes to their lives (14 -4). This is an important finding and places emphasis on financial

support as the key mechanism to ensuring benefit issues are dealt with in an efficient and

reactive way (32 +1). Generally, this subgroup captures the programme as a simple and efficient

programme which uses a collaborative model effectively to address short term outcomes.

[Insert TABLE 3 Here]

Reflections on using Q in evaluation

The above demonstration case provided a snapshot of the steps taken and findings in

using Q to evaluate P4P. On reflection, the evaluators mobilizing this approach found the process

of conducting Q iterative and clear. The methodological steps it requires enabled us to capture a

program theory through concourse-building, which drove the creation of the statements, all the

way through to the sorting and analysis, where we were able to build a story with depth as to

how and why each stakeholder saw P4P working and why. This comprehensive process instilled

confidence in the evaluation, given how much time we afforded to the use of Q.

Essentially, carrying out a Q study enabled us to assess the merit and the worth of the

program under study because it was able to showcase the subjective viewpoints of those integral

to the program concerning how they saw it working. Further, the quantitative factor analysis

procedures, followed by the qualitative interpretation, allowed for the groupings of shared

viewpoints, which helped us to tell a story about certain individuals and specific contexts around

the program across multiple levels. The synergy of qualitative and quantitative approaches also

helped us to mitigate the ongoing quandary a researcher or evaluator finds themselves in when

choosing between qualitative or quantitative methodologies.

Moreover, when conducting any type of evaluation, it is always a danger that in the face

of any academically respected and innovative evaluation methodology, the client and

stakeholders become alienated (Harris et al., 2019). For example, in the knowledge translation

discourse, it is crucial that the clients requesting and paying for the evaluation understand what is

being done, and why it is being done in the way it is, to avoid alienation and subjugation (Harris

& Adams, 2016).

Therefore, in accordance with Shula, Whitmore, Cousins, Gilbert, and Hudib (2016),

there should be motivation for collaboration on the part of the stakeholders from the beginning,

and the findings should foster use to inform improvements and refinements to practice in the

future. This was important for us in relation to the use of Q in this evaluation, and we did not see

or experience any confusion on behalf of the client throughout. In fact, there was a lot of interest

and motivation to be part of the process, as we highlighted in the generation of the statements.

In terms of fostering use (Shula et al, 2016; Harris, 2020), Q is useful for clients because

the holistic narratives (the Watts and Stenner approach we took) can provide a succinct story and

reflection about those shared viewpoints that can share practice. In a realist sense, it was crucial

to be able to communicate to the client that P4P works for different individuals in different ways,

taking into account varying subjective viewpoints (Harris, Henderson and Wink, 2019; Pawson

and Tilley, 1997).

Therefore, the findings informed P4P at both a practical level – in how to tailor and

communicate the benefits and address misunderstandings between partners in delivering P4P –

and at a program level, that changes or adaptations to the program should not be generalized in

one way, and should be adapted and refined in accordance with what is illustrated within the

holistic narratives we presented. We argue that this can only be positive for program refinement

and development, informing practice moving forward.

Challenges to Using Q

However, when using a new, innovative methodology within program evaluation such as

Q, there will be limitations and/or implications which arise in practice. We present here some

practical challenges that occurred when conducting Q with participants and in following the

methodology.

It is likely that participants will not have heard of Q, as was the case in this evaluation.

This created our first challenge to support the participants in feeling capable and willing to do

something different to what they are used to. Secondly, when explaining the process of Q to

participants, they perceived Q as a simple, process-oriented task (as shown in Table 1). In

practice, participants became confronted with a series of challenges in deciding between statements (as shown in Table 2). For example, when conducting the Q-sort, some participants did not feel comfortable with ordering statements within the negative (-) scoring of the Q grid as they felt they were being negative about the program.

Within the demonstration case, participants generally agreed with more statements than they disagreed with, and therefore did not have space to fit them all neatly within the "+" spaces available on the Q grid. This presented a challenge to the evaluator in encouraging the participant to choose the statements which were most prominent to them, without negatively impacting on their motivation or influencing sorting of the statements. Further, the time it takes to conduct a Q sort can vary: for some participants it took 30 minutes, whereas for others it took close to 60 minutes. Participants took longer when they re-ordered their statements as their reasoning changed during conducting the Q, or as a result of having to choose between statements when (re)ordering. This can result in participant fatigue.

Beyond practical challenges with participants completing the Q sort, the methodological process of Q – from generating the concourse to conducting the Q sort, factor analysis, and interpretation – can span a duration of months, with each phase requiring considerable time to ensure there is a smooth transition between the phases described in Table 1. Further, in considering the technical requirements of Q methodology, the evaluator should be competent in the use of statistical software packages such as PQMethod in order to conduct the factor analysis (Table 1). This is a technical phase in the Q process and requires evaluators to be careful with their data entry and competent with the software to conduct the factor analysis, in order to then proceed to identify subgroups and build the holistic narrative.

Conclusion

This paper has sought to demonstrate the diverse potential uses of Q in evaluation, either by itself or with other methods. The case-study example has illustrated the value of Q in evaluation, crucially the capacity of Q to quantify participant viewpoints, meaning participant voices are more than anecdotal evidence in an evaluation. It is valuable when working with relatively small sample sizes, and offers practitioners an appropriate methodology for evaluations in which the human factor and/or context may influence an intervention's success. This holds promise in fostering use through better informing program funders and decision-makers on where to target, adapt, and refine the program in relation to the holistic narratives.

In reflecting on using Q methodology within the case study example, the authors highlight practical challenges in particular phases, especially when participants are not familiar with Q and are challenged by the number of statements to rank and having to place statements into negative (-) scoring on the grid. As such, the authors recommend appropriate participatory opportunities to be promoted with program practitioners and other stakeholders throughout the Q process: supporting concourse development, refining the Q statements, interpretation of Q factor analysis, and how findings can be best presented to foster evaluative thinking and use. We accept that the approach we have demonstrated in this method note is not the only way to mobilise Q. Indeed there are a number of different ways to carry out the interpretation of the factors (see Militello et al, 2016), and as reflective practitioners we are keen to explore these different approaches in the future. Nevertheless, we anticipate that the approach we have followed and demonstrated can make a suitable contribution to the evaluation community.

Interested readers can learn more about Q methodology through the International Society for the Scientific Study of Subjectivity (ISSS), a professional community of Q researchers. They offer an annual conference, website (qmethod.org), an active listserv, and a peer-reviewed

journal, *Operant Subjectivity*. In addition, the "Sue-Z Q" YouTube channel, created and

maintained by Prof. Susan Ramlo of the University of Akron, provides detailed tutorial videos on

every step of Q methodology, including use of software for factor analysis. Finally, we hope that

evaluators interested in collaborating about Q methodology will reach out to the authors, so that

we might establish our own community of practice around this innovative methodology and

promote its broader adoption. After all, subjectivity is central to program evaluation, and Q

provides a useful methodology for measuring it.

References

Akanbang, B. A. A., Darko, R. O., & Atengdem, P. B. (2013). Programme implementers' experiences of process use types in three evaluation contexts in Northern Ghana. *Operant Subjectivity: The International Journal of Q Methodology, 36(4)*, 297-319.

Baker, R., Thompson, C., & Mannion, R. (2006). Q methodology in health economics. *Journal of Health Services Research & Policy, 11(1),* 38-45.

Baptiste, L. J. (2010). *Process use across evaluation approaches: An application of Q methodology in program evaluation* (Doctoral dissertation, Kent State University).

Baptiste, L. (2011). What educators learn when they evaluate students. *Operant Subjectivity: The International Journal of Q Methodology*, *34(2),* 104-23.

Brown, S. R. (1980). *Political subjectivity: Applications of Q methodology in political science.* New Haven, CT: Yale University Press.

Brown, S. R. (2019). Subjectivity in the human sciences. *The Psychological Record, 69,* 565-579. doi:10.1007/s40732-019-00354-5

Butler, H., Hare, D., Walker, S., Wieck, A., & Wittkowski, A. (2014). The acceptability and feasibility of the Baby Triple P Positive Parenting Programme on a mother and baby unit: Q-methodology with mothers with severe mental illness. *Archives of Women's Mental Health, 17(5)*, 455-463.

Corr, S., Phillips, C., & Capdevila, R. (2003). Using Q methodology to evaluate a day service for younger adult stroke survivors. *Operant Subjectivity: The International Journal of Q Methodology*, *27(1),* 1-23.

Cotton, M. D., & Mahroos-Alsaiari, A. A. (2015). Key actor perspectives on stakeholder engagement in Omani Environmental Impact Assessment: An application of Q-Methodology. *Journal of Environmental Planning and Management*, *58(1),* 91-112.

Cuppen, E. (2012). A quasi-experimental evaluation of learning in a stakeholder dialogue on bio-energy. *Research Policy*, *41(3),* 624-637.

Cuppen, E. (2013). Q methodology to support the design and evaluation of stakeholder dialogue. *Operant Subjectivity: The International Journal of Q Methodology*, *36(2),* 135-161.

Danielson, S. (2009). Q method and surveys: Three ways to combine Q and R. *Field Methods, 21*, 219-237.

Danielson, S., Webler, T., & Tuler, S. P. (2009). Using Q method for the formative evaluation of public participation processes. *Society & Natural Resources*, *23(1),* 92-96.

De Jonge, L. P., Timmerman, A. A., Govaerts, M. J., Muris, J. W., Muijtjens, A. M., Kramer, A. W., & van der Vleuten, C. P. (2017). Stakeholder perspectives on workplace-based performance assessment: towards a better understanding of assessor behaviour. *Advances in Health Sciences Education, 22(5),* 1213-1243.

Dziopa, F., & Ahern, K. (2011). A systematic literature review of the applications of Q-technique and its methodology. *Methodology, 7(3),* 39-55.

Ellingsen, I. T., Størksen, I., & Stephens, P. (2010). Q methodology in social work research. *International Journal of Social Research Methodology*, *13(5),* 395-409.

Fluckinger, C.D. (2014). Big five measurement via Q-sort: An alternative method for constraining socially desirable responding. *SAGE Open*, 1-8.

Frayne, D. (2014). *Nonprofit leader perceptions of effective organizational performance measurement: A Q methodology study* (Doctoral dissertation, University of Phoenix).

Harris, K and Adams, A (2016) Power and Knowledge in the field of sport for development. *Sport Management Review* 19 (2), 97-106.

Harris, K., Henderson, S., & Wink, B. (2019). Mobilising Q methodology within a realist evaluation: Lessons from an empirical study. *Evaluation*, 1-19.

Harris, K. (2020) Building Capacity in Practitioner Realist evaluation through CAE principles IN J. B. Cousins (Ed.). *Global Test Drive of Principles for Collaborative Approaches to Evaluation (CAE)* Thousand Oaks: Sage

Hensel, D. (2017). Using Q methodology to assess learning outcomes following the implementation of a concept-based curriculum. *Nurse Educator, 42(5),* 250-254. https://www.ncbi.nlm.nih.gov/pubmed/28045739.

Ho, G.W.K. (2017). Examining perceptions and attitudes: A review of Likert-type scales versus Q-methodology. *Western Journal of Nursing Research, 39(5),* 674-689.

Jeffares, S., & Dickinson, H. (2016). Evaluating collaboration: The creation of an online tool employing Q methodology. *Evaluation*, *22(1),* 91-107.

Kelly, S. E., Moher, D., & Clifford, T. J. (2016). Expediting evidence synthesis for healthcare decision-making: exploring attitudes and perceptions towards rapid reviews using Q methodology. *PeerJ, 4*, e2522.

Lazard, L., Capdevila, R., & Roberts, A. (2011). Methodological pluralism in theory and in practice: The case for Q in the community. *Qualitative Research in Psychology, 8(2),* 140-150.

McPherson, K. E., Sanders, M. R., Schroeter, B., Troy, V., & Wiseman, K. (2016). Acceptability and feasibility of peer assisted supervision and support for intervention practitioners: A Q-methodology evaluation. *Journal of Child and Family Studies*, *25(3),* 720-732.

Militello, M., & Benham, M. K. (2010). "Sorting Out" collective leadership: How Q-methodology can be used to evaluate leadership development. *The Leadership Quarterly, 21(4),* 620-632.

Militello, M., Bass, L., Jackson, K., & Wang, Y. (2013). How data are used and misused in schools: Perceptions from teachers and principals. *Education Sciences*, *3(2),* 98-120.

Militello, M., Janson, C., & Tonissen, D. (2016). InQuiry: A participatory approach for understanding stakeholder perceptions. *The Foundation Review, 8(1),* 88-107. doi:10.9707/1944-5660.1286

Mumford, S. W. (2018). *Ways of knowing in participatory program evaluation* (Doctoral dissertation, The George Washington University).

Nunns, H. (2016). *The practice of evaluative reasoning in the Aotearoa New Zealand public sector* (Doctoral dissertation, Massey University, New Zealand).

Ockwell, D. G. (2008). 'Opening up' policy to reflexive appraisal: a role for Q Methodology? A case study of fire management in Cape York, Australia. *Policy Sciences*, *41(4),* 263-292.

Pawson, R., & Tilley, N. (1997). *Realistic evaluation.* London: Sage Publishing.

Pawson, R. (2013). *The science of evaluation: a realist manifesto.* London: Sage Publishing.

Pham, P. K. (2018). *Evaluative thinking in two parallel contexts: Evaluation and Emergency Medicine* (Unpublished master's thesis, Claremont Graduate University).

Pike, K., Wright, P., Wink, B. & Fletcher, S. (2015). The assessment of cultural ecosystem services in the marine environment using Q methodology. *Journal of Coastal Conservation, 19(5),* 667-675. https://search.proquest.com/docview/1719537634.

Pruslow, J. T., & Owl, R. H. R. (2012). Demonstrating the application of Q methodology for fieldwork reporting in experiential education. *Journal of Experiential Education*, *35(2),* 375–392.

Ramlo, S. (2011). Facilitating a faculty learning community: Determining consensus using Q methodology. *Mid-Western Educational Researcher, 24(1),* 30-38.

Ramlo, S. (2015a). Theoretical significance in Q-methodology: A qualitative approach to a mixed method. *Research in the Schools, 22,* 73-87.

Ramlo, S. (2015b). Student views about a flipped physics course: A tool for program evaluation and improvement. *Research in the Schools, 22(1),* 44-59.

Ramlo, S. (2016). Mixed methods lessons learned from 80 years of Q methodology. *Journal of Mixed Method Research, 10(1),* 28-45.

Ramlo, S., & Newman, I. (2011). Q methodology and its position in the mixed-methods continuum. *Operant Subjectivity: The International Journal of Q Methodology, 34(3),* 172-191.

Robbins, P., & Krueger, R. (2000). Beyond bias? The promise and limits of Q method in human geography. *The Professional Geographer*, *52(4),* 636-648.

Schmolck, P. (2015). *PQMethod manual (revised).* Retrieved from http://schmolck.userweb.mwn.de/qmethod/pqmanual.htm

Shemmings, D. (2006). 'Quantifying 'qualitative data: an illustrative example of the use of Q methodology in psychosocial research. *Qualitative Research in Psychology*, *3(2),* 147-165.

Shula, L, Whitmore, E, Cousins, B, Gilbert, N, & Hudib, H. (2016). Introducing evidence-based principles to guide collaborative approaches to evaluation: Results of an empirical process. *American Journal of Evaluation, 37(2),* 193-215.

Stephenson, W. (1935a). Technique of factor analysis [Letter to the editor]. *Nature, 136,* 297.

Stephenson, W. (1935b). Correlating persons instead of tests. *Journal of Personality, 4,* 17-24.

Stephenson, W. (1952). Some observations on Q technique. *Psychological Bulletin, 49,* 483-498.

Stephenson, W. (1982). Q-methodology, interbehavioral psychology, and quantum theory. *The Psychological Record, 32(2),* 213-230. https://search.proquest.com/docview/1301194158.

Stephenson, W. (1993). Introduction to Q-methodology. *Operant Subjectivity: The International Journal of Q Methodology, 17,* 1–13.

Stickl, J. E., Wester, K. L., & Wachter Morris, C. A. (2018). Making sense of subjectivity: Q methodology in counseling research. *Counseling Outcome Research and Evaluation*, 1-13.

Thompson, A. J. L. (2015). *Why get involved in program evaluations?: Toward a model of stakeholder involvement motives* (Doctoral dissertation, Carleton University, Canada).

Velez, R. (2006). *Perceptions of school performance measures: A study of principals in the United States and head teachers in the United Kingdom using Q methodology* (Doctoral dissertation, University of North Florida).

Watts, S., & Stenner, P. (2005). Doing Q methodology: Theory, method and interpretation. *Qualitative Research in Psychology, 2,* 67-91.

Watts, S., & Stenner, P. (2012). *Doing Q methodological research.* Los Angeles, CA: Sage Publishing.

Wolf, A. (2013). Wellbeing for public policy: Roles for Q methodology. *Operant Subjectivity: The International Journal of Q Methodology*, *36(3),* 203-226.

Woods, C. E. (2011). Using Q methodology to explore leadership: The role of the school business manager. *International Journal of Leadership in Education*, *14(3),* 317-335.

Zabala, A., Sandbrook, C., & Mukherjee, N. (2018). When and how to use Q methodology to understand perspectives in conservation research. *Conservation Biology*, *32(5),* 1185-1194.