Bath Social & Development Research Ltd

QuIP Qualitative Impact Assessment Protocol

QuIP methodology note: Outcome-led interviewing and blindfolding

The Qualitative Impact Protocol is an approach to collecting, analysing and sharing narrative statements about the causal pathways leading to intended and unintended outcomes, from planned activities alongside incidental drivers of change. This briefing focuses on one of its core features of the QuIP - the style of interviewing employed to collect credible and useful information about the causal pathways. Its aim is to reflect on the theory behind the QuIP approach to interviewing in the light of more than a decade of practical experience, drawing on concrete examples. The briefing is also a response to our experience of frequently being asked about deliberate 'blindfolding' – or the practice of not telling interviewers and interviewees more than is necessary about the activity being evaluated in order to encourage broader reflection on drivers of change. The key message of the briefing is that full blindfolding is not necessary (as well as often not feasible) for QuIP interviewing, whereas framing interviews through explicit reference to intended outcomes, and not planned activities or interventions, is central to the QuIP approach to data collection. The brief includes reference to several real evaluation examples, with honesty about the reality of how the approach was used in very varied contexts to try to provide guidance to QuIP users.

This note was put together by Fiona Remnant with substantial contributions from James Copestake, Hannah Mishan and Rebekah Avard. Illustration by Strawberry Grace Designs.

Contents

Figures and tables

1. Introduction

QuIP is a qualitative approach to impact evaluation, best described as a form of contribution analysis. For an overview of QuIP see the entry in Better Evaluation or our briefing paper. It is primarily defined by two key aspects which set it apart from other similar theory-based approaches; an outcome-led approach to eliciting information from respondents in interviews - including using various degrees of blindfolding, and the way that narrative data is coded using causal mapping. This paper will focus on the first of these aspects - aiming to share practical experiences of how outcome-led interviews are carried out. Most of what we say also applies to collecting data using focus groups alongside as is often the practice with QuIP studies. For more on using causal mapping in data analysis please see other papers in our resources library.

Development of the QuIP included a period of collaborative action research over three years led at the University of Bath, funded by UK government, and based on pilot studies in Ethiopia and Malawi. This was followed by the founding of a non-profit enterprise Bath Social & Development Research to mainstream the QuIP, and to continue testing and refining its design in a range of contexts. How to collection narrative data with minimal prompting has been both a recurring, challenging and rewarding part of this journey.

This paper briefly summarises the theoretical reasoning behind our emphasis on outcome-led interviews, and the 'blindfolding' (whether partial or full) to achieve it. It then provides an overview of the key features of outcome-led interview design, practical strategies on how to achieve it, and scope for flexibility. Section 4 uses real examples to illustrate how the approach has been adapted in a range of contexts over the last decade. Finally, we briefly consider some of the ethics involved in conducting this type of work.

Please note that we use the terms 'field team', 'researchers' and 'interviewers' interchangeably to mean the people responsible for conducting interviews with respondents.

2. Blindfolding - a means to an end?

The action research phase of QuIP (2012-2015) started with the premise that people with lived experiences of the impacts of policies and programmes are best placed to give us direct information about what works, for whom and how - and what doesn't. However, we were also cognisant of the challenges that evaluators face when relying on direct narrative testimony in qualitative evaluations, particularly the extent to which donors and commissioners find such testimony credible and informative. The approach to QuIP interviews therefore focused on achieving two main aims: mitigating the risk of confirmation bias and making interviews as open-ended or 'exploratory' as possible. To achieve this, QuIP operationalises the principles of Goal Free Evaluation (GFE), particularly the strategy of withholding programme goals from the data collector to avoid 'tunnel vision' - or bias that arises when the researcher only looks for intended effects and misses unintended consequences or unanticipated causes (Youker, 2024: 102).

While QuIP shares the analytical ambition of Realist Evaluation and Process Tracing - specifically the need to open the 'black box' of causality to understand *how* outcomes are generated through specific mechanisms - it differs in how data is collected. Standard applications of Process Tracing or Realist Evaluation often begin with a known theory to be tested; in contrast, QuIP applies a GFE-style 'blindfold' during the interview stage. This ensures that the causal pathways and Context-Mechanism-

Outcome type configurations articulated by respondents emerge inductively, rather than being prompted or 'confirmed' by an interviewer asking leading questions based on a pre-defined Theory of Change. Blindfolding therefore serves as a tool both to enhance perceptions of rigour in self-reported causal mechanisms, and to increase the chances of reporting on unexpected connections and outcomes - positive or negative.

2.1 Tackling confirmation bias

Concerns about potential biases in relation to using narrative statements as evidence of causal data, include confirmation bias, or a tendency for respondents to say what they think the interviewer expects or would like them to say. These mean that qualitative methods are often discounted in favour of quantitative methods that rely on statistical inference based on variable exposure to observed 'treatments' and reported outcomes. To mitigate this possibility of bias design of QuIP interviewing schedules and the choreography of interviewing avoided referring explicitly to projects or interventions at any point, with information about these projects withheld from those conducting the interviews as well as respondents. We landed on the term 'blindfolding' as a short-hand for this approach (and 'double blindfolding' when interviewer and interviewee were both informed as little as possible about what activities were being evaluated) to help avoid leading questions, intentional or otherwise. Blindfolding rather than blinding reminds us that this is not a permanent state, and the blindfold can be taken off at any point to help facilitate more open discussions. Indeed, QuIP guidelines strongly encourage eventual reversal of blindfolding before the end of the evaluation both for ethical reasons (see below) and to allow for fully transparent sharing and discussion of findings (referred to in QuIP material as sensemaking).

The QuIP briefing paper written shortly after the action research period summarises the rationale,

> *"There are strong ethical grounds for asking people directly about the effect of actions intended to benefit them but doing so involves finding credible ways to address potential response biases. The QuIP does this by arranging for qualitative data collection to take place with as little reference as possible to the specific activity being evaluated, and by giving equal weight to all possible drivers of change in possible domains of impact. This is achieved by working, where possible, with field researchers who are completely independent of the organisation responsible for the actions being evaluated. Indeed, where possible, field researchers are 'blindfolded' from knowing the identity of the organisation being evaluated, the details of project implementation and the theory of change behind its actions. …The purpose of this blindfolding is primarily to reduce potential for pro-project bias on the part of respondents, including their response to cues from the researchers."* QuIP Briefing Paper

This lays out the mitigation of confirmation bias as the primary driver for choosing to take this approach, but arguably the most important driver is in fact to ensure the evaluation is exploratory - open to discover unexpected impacts and causal relationships.

2.2 Taking an exploratory approach

The first article about the QuIP action research published in 2014 focuses particularly on the role that an outcome-led interview can play in accommodating both confirmatory and exploratory approaches to impact evaluation. An exploratory interview is more open-ended *"in the sense of explicitly limiting prior theorisation on the part of the researcher"* (Copestake, 2014: 9), in other words limiting

preconceptions interviewers may have about what the evaluation needs before approaching an interview. Avoiding these preconceptions encourages both parties in the conversation to explore all possible drivers and outcomes related to a particular outcome domain, leaving the respondent free to propose alternative explanations for change, or indeed no change. A more confirmatory approach will be testing assumptions from a theory of change, with the interview either confirming or refuting those assumptions (as is most explicit in Realist Evaluation). Whilst this doesn't rule out discovering alternative explanations, it is less likely to encourage unprompted answers. The concept of goal-free evaluation (GFE) (as opposed to goal-based evaluation) was introduced by Michael Scriven in a paper first published in 1972, 'Pros and cons about goal-free evaluation'. In this approach the evaluator is 'blinded' from the intended outcomes of a programme or intervention in an effort to focus on the real and experienced consequences of a programme or intervention rather than only on what was initially intended. Vedung describes the full knowledge of the aims of an intervention as a *"mental corset impeding [the evaluator] from paying attention to side effects, particularly unanticipated side effects"* (1997: 57).

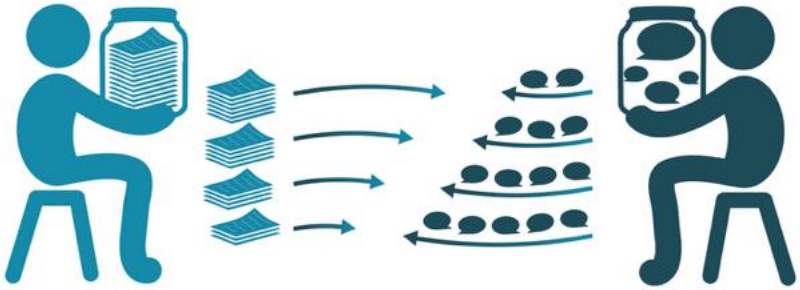
In contrast, Process Tracing tends to be more explicitly theory-led, often operating deductively to determine if the evidence within a specific case matches a hypothesised causal mechanism. In contrast, QuIP aims to collect data that can first facilitate inductive theory-building, while remaining open to theory-testing at the analysis stage - see Copestake, Goertz and Haggard 2020 for extended discussion of this point. While the evaluation design remains *theory-informed*, guided by a Theory of Change known to the lead evaluator and analyst (to help with both questionnaire design and confirmatory coding and analysis), the data collection itself is intentionally not theory-led. In this way QuIP is a form of what Youker describes as Goal Dismissive Evaluation rather than purely GFE (Youker, 2024b:29). Unlike a fully goal-free approach where objectives might be genuinely unknown or irrelevant to the evaluator, the evaluator in a QuIP acknowledges the specific goals in the design of the interview (see section 3.1 for more) but intentionally "dismisses" them from the interview protocol. This strategic blindfolding ensures that the causal mechanisms identified by respondents are emerging via inductive process tracing, rather than the interviewer seeking to confirm a pre-determined single-case theory.

This aligns with Copestake's (2025) distinction between 'quant-led' and 'qual-led' models of impact evaluation. While quant-led models often test specific variance-based hypotheses, the qual-led model relies on generating evidence that is *"not collected to fit a predetermined conceptual framework or coding pattern"* (Copestake, 2025: 2). Use of blindfolding is the practical application of this principle: by dismissing the 'predetermined framework' of the intervention's goals during the interview, we allow the 'qual-led' analysis to remain more open to unexpected causal mechanisms.

The design of the interviews - including decisions about what information needs to be provided and what can be withheld – depends on a mixture of considerations, including what is practically and ethically achievable, and what kind of evidence the commissioner is seeking, for whom, and subject to what threshold level of credibility or rigour.

The illustration below attempts to summarise the rationale for an outcome-led, or goal-dismissive, approach to interviews. The more information the interviewer keeps locked in the 'jar', the more information the respondent shares from their 'jar' of experiences. However, this is not a neat on/off switch - there are options and variations in how this can 'exchange of information' is conducted, including variation in the duration of the interview.

Figure 1: Information exchanges between interviewer and interviewee



More articles and papers covering the theoretical rationale for this exploratory approach are available at our Resources Library: www.bathsdr.org/resources.

3. Application and adaptation of QuIP in context

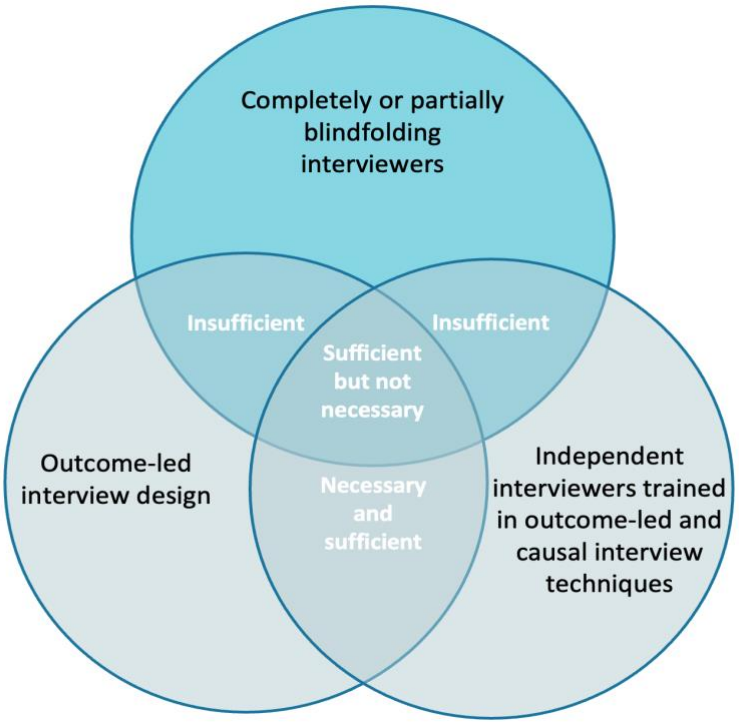
Beyond the theory, the question of practical application is where interested users may struggle. This is where important decisions need to be made about what is feasible and desirable, and this usually means **not** using 'full' blindfolding in its original sense. For example, researchers may know who has commissioned the research and even something about the intervention, and some of this information may also need to be passed on to the respondents - without abandoning the outcome-led style of QuIP interviews. In fact, when going back through the 100 evaluations Bath SDR have worked on over the last ten years, it was notable that full blindfolding was only used in a minority of cases. There are usually mitigating circumstances which require some adaptation - and examples of this follow in the next section.

From the early stages of the action research period, we encountered and addressed the practical challenges of managing this process in Malawi and Ethiopia. In the 2014 article (Copestake, 2014) we discuss how "*clear limits to the extent and sustainability of such blinding emerged"* and how blindfolding should not be relied upon or assumed to replace the importance of researchers receiving good training on how to use an outcomes-led approach.

> "*It was both practically necessary and ethically important for the field team and respondents to have a broad understanding of the reasons for the research. In the case of the Malawi pilot the explicit rationale was to gain a better understanding of recent changes in rural livelihoods and food security in selected localities, and the main causes of these changes. In the light of information thereby generated it would have been easy for the lead researchers to confirm the identity of the specific project being evaluated. Hence while useful, partial blinding is ultimately not a substitute for researchers' skills, integrity and professionalism. In addition, respondents need some label to attach to visiting researchers, and if not associated with a specific project or NGO then they can be expected to ascribe another label which may also prompt strategic bias - e.g. pro-authority bias if respondents are perceived as representatives of government."* Copestake, 2014: 18-19

A later paper titled ('Managing relationships in qualitative impact evaluation of international development: QuIP choreography as a case study' Copestake et al., 2018), specifically aimed to describe the practical challenges of managing this process. In the paper we describe how researchers were recruited, and how they set about accessing communities, approaching people and conducting interviews without the support of the NGOs managing the projects. The process was not straightforward, and the challenges the research team faced are detailed in the paper which concludes that it will likely not always be possible to use blindfolding in a formulaic manner. Indeed, this has proved to be the case, and most QuIP evaluations are what we would now call **'partially blindfolded'** (more on this to follow). However, being aware of confirmation biases and keeping interviews open to unexpected findings is still possible with good training and advice for interviewers, and well-designed interview schedules. To this end, the Venn diagram below attempts to summarise the necessary and sufficient elements involved in achieving a good exploratory QuIP interview - demonstrating that blindfolding may be desirable and helpful in reaching the goal, but it is not necessary (or sufficient) without the other two foundational elements - outcome-led interview design and good training of researchers. This section unpacks each of these elements.

Figure 2: Elements involved in achieving an exploratory QuIP interview



3.1 Outcome-led interview design

A QuIP interview must be designed around the intended outcomes of the subject of the evaluation, guided by a 'working backwards' approach to elicit stories of change from respondents. Whilst it is important not to direct the interview specifically around the interventions, some structure is necessary to ensure that the interview captures the type of outcome stories we are expecting to hear (and to limit the length of the interview). To this end the interview is structured around key domains of change (e.g., income, health, education, food security) identified in the intervention's Theory of Change - typically 3-5 domains.

The overarching guiding principle is to work backwards from the outcomes rather than starting with the activities of the intervention. Questions begin by asking respondents to describe the main changes - positive or negative - they have experienced within a specific domain over a predetermined recall period (sufficient to capture how things were before the intervention). Most questions are open-ended (more exploratory) to encourage the respondent to provide rich stories of change in their own words, with optional supplementary questions provided to the interviewers to encourage more detail. For example,

> Has the variety and quantity of foods your household eats changed in the last two years? Please explain what has changed, and why.
>
> - *If there isn't enough food in the house, are there members of the family who eat less? Or would everyone eat less?*

Most will also include some closed (more confirmatory) questions to round off discussion of a domain before moving onto the next, and to give the respondent the opportunity to summarise what has been discussed (not leaving this up to the analyst to infer from what could be a complex answer). For example,

> Overall, would you say that the quantity of food your household eats has:
>
> Increased/ Stayed the same/ Decreased
>
> Why is that - what is the most important reason?

Probing for more detail about drivers and outcomes (a process sometimes called 'back-chaining') is crucial to elicit the respondent's own perception of why the change occurred and to whom or what they attribute the change - and this forms a key part of QuIP researcher training.

3.2 Training interviewers in QuIP interview techniques

In terms of practical application, if any sort of blindfolding is to be used, it is very important to recruit independent researchers to conduct the interviews. These people should be local to the context of the evaluation and not have any connection with the organisation or programme being evaluated. This may be the hardest part of applying a QuIP, but using internal staff will be unlikely to yield the independent results that a QuIP aims to achieve. Training independent interviewers is not only an opportunity to practice conducting effective causal interviews, but also to understand what is feasible and advisable in terms of the approach to blindfolding. Discussions about how they can introduce themselves to respondents without too much reference to the project being evaluated should be guided by their local knowledge as well as by the commissioner's own requirements.

Our training includes covering the theory behind QuIP evaluation design, how to set up and introduce QuIP interviews to put respondents at ease, practical tips and exercises on encouraging back-chaining in interviews with good probing questions, and practice coding of pilot interviews. Practice coding is an essential part of cementing understanding of what makes a full causal story and can be done very simply with sticky notes on flipchart paper.

Feedback on the pilot interviews conducted during the training not only helps to confirm if the questions and domains are sufficient to elicit the expected stories of change but also helps researchers to understand if they have missed opportunities for probing and getting at the root causes of change.

Example instructions from QuIP training:

You may not know much about the project or issue which is being assessed, however, you need to make sure that you do not accidentally encourage a respondent to elaborate on one particular area because you think that might be of interest to the commissioner or you have any other reason to pursue that particular path. The respondents should guide the conversation. You are simply helping the conversation along using the supplementary questions and to keep probing for the reasons for change. If you're not sure what they mean or they could be referring to more than one driver, then double check. Do not assume that you know. The closed questions will help to get an opinion about whether what they are talking about is positive or negative in their own opinion. However, if it's not clear, keep asking and find out more, the analyst will need all your notes to help them understand better how the respondent was trying to describe that story.

Stick to the questionnaire as it is designed but allow respondents to expand on an area as the conversation flows naturally. If they answer one of the next questions naturally in conversation, you do not need to ask it again. If people talk about positive or negative change without any explanation, then you need to use probing questions and keep asking 'why' until you understand the root cause of change.

The open-ended questions are there to stimulate that free speech and the free narrative so let respondents talk for as long as they want in those areas. Then close that domain down with the closed question before moving on to the next topic. The closed questions help you to use the structure of the questionnaire to move from one topic to another and helps the respondent feel that they are making progress through the interview. You may also find it easier to say at the beginning of the interview how many and which domains are going to be covered, so that the respondent knows what to expect.

Feedback from QuIP researchers:

> *"The feedback after our initial pilot was really helpful - I started asking more "how and why" questions. I probed further to exhaust all the answers from the respondents. I didn't know anything about the intervention so I couldn't lead respondents on in the first place, and all my how and why questions were based on their specific responses just to understand the rationale behind the statements made by respondents."*
>
> *"At first, it was a bit destabilising not to know the project in detail, because we are used to wanting to master our subject. But very quickly, I understood the value of this approach."*

Outcome-led interviewing is a skill which comes more naturally to some researchers. We recommend spending time exploring and developing this skill during researcher training. It is key that researchers, do the following during interviews:

1. Let respondents lead the conversation. Even when using some form of blindfolding we all have some preconceptions or interests, and it is important that respondents are encouraged to discuss what is important to them.
2. Probe to better understand the full story. Capturing details such as who was involved, and all the root drivers is key to a holistic understanding of the changes in respondents' lives.
3. Avoid assumptions, always clarify to make sure that potentially confusing aspects such as attribution and sentiment are understood correctly.
4. Record faithfully what is said, either using verbatim transcripts or detailed summaries from recordings or a note taker.

Researchers are instructed to make sure respondents feel comfortable with the process and are informed of the approach. This includes:

1. Being clear on the rationale behind not being able to share full information on the evaluation (if this is the case) and how they will be unblindfolded.
2. Explaining the interview structure from the start, explaining what domains will be covered so they know what to expect
3. Encouraging respondents to talk freely and spend time discussing what they see as important in their lives.

3.3 The blindfolding design space

When considering context, we often talk about blindfolding as a design space - a spectrum which can range from complete lack of knowledge through to completely open interviews, with 'partial blindfolding' sitting somewhere in between. Partial blindfolding means that some amount of information is shared with either or both the interviewer and interviewee - and this is a negotiable space. Some of the range of possible approaches to using blindfolding are summarised in the table overleaf, with options for how much information is shared with both the interviewer and the respondent, moving from as little as possible at level 1, through to most open at level 5.

Information levels:

1. Broad topic area, e.g. livelihoods, health, education
2. Name of commissioner
3. Name of project, policy, intervention
4. List of interventions
5. Theory of change (mechanisms)

Although the table doesn't contain a comprehensive list of all possible options, it is indicative of the flexibility available to those designing a QuIP evaluation. We will discuss examples using variants along the spectrum in the next section of this paper.

The question of when a QuIP is no longer a QuIP can be difficult to answer. However, this table may go somewhere towards this. The last line of the table is as open as an interview would be; there is not an example of a case where all five levels of information are shared with both interviewer and interviewee. At this point, a much more structured interview would be more akin to a Realist Evaluation, where the interviewer's role is to clearly check mechanisms from the theory of change with the interviewee. At the other end of the spectrum, an interview with little to no structure would also not be a QuIP, and would be closer instead to what Youker describes as intentional Goal Free Evaluation (GFE), *"in which the evaluator deliberately and proactively avoids the stated goals and objectives"* (Youker & Ballard, 2024: 100). As referred to earlier, Youker in fact lists QuIP as a form of the second of his goal-free typologies, goal-dismissive GFE which, *"typically ask program participants and stakeholders about any changes or outcomes they experienced or witnessed, and then the evaluator explores whether these reported changes are attributable to the intervention. The evaluator collects these data without referencing the intervention's goals"* (Youker & Ballard, 2024: 100).

Figure 3: Options for framing QuIP interviews

Options for framing QuIP interviews			
	Information given to interviewers	**Information given to respondents**	**Rationale for this approach**
Blindfolded	1. Topic area	1. Topic area	Interviewers are able to gain access through professional affiliation, and respondents are likely to agree to participate. Double blindfolding is important to the commissioner's level of trust.
↕	1. Topic area 2. Name of commissioner	1. Topic area	Interviewers are likely to need a letter or introduction to gain access to communities, but respondents are likely to agree to talk about a broad topic area without knowing the commissioner's name.
Partial blindfolding	1. Topic area 2. Name of commissioner (3. Project name optional)	1. Topic area 2. Name of commissioner	A letter or introduction likely needed for access and to gain consent from respondents, but no project details unless necessary.
↕	1. Topic area 2. Name of commissioner 3. Project name 4. Interventions	1. Topic area 2. Name of commissioner (3. Project name optional)	As above, but interviewers need to know more about what to probe for if information is not provided on certain interventions of interest. This would work using a 'funnel' approach, with the interview becoming narrower and more specific over time, and only if information is not provided voluntarily.
Unblindfolded	1. Topic area 2. Name of commissioner 3. Project name 4. Interventions 5. Expected mechanisms in ToC	1. Topic area 2. Name of commissioner (3. Project name optional)	As above, but where the interviewer is also the lead evaluator and/or wants to probe for specific mechanisms. However, if this information is also shared too openly with the respondent, then this becomes a Realist interview where mechanisms are openly explored together.

3.4. Application in varied contexts

As we have acknowledged, blindfolding can add additional logistical and ethical challenges. Gaining access to respondents and permission to complete interviews may be difficult, but in some locations and contexts it is possible to approach respondents without the direct involvement of commissioning organisations. Researchers will often need to have open conversations with whichever gatekeeper will allow you access to the respondents (community leaders, team managers etc), explaining the concept of outcome-led interviews as well as the thematic focus of the study, the commissioner (if necessary and if they know), and who they hope to talk to. With their permission, and if necessary, an introduction, individuals are then invited to interview either via phone calls, emails or in-person visits. However, the initial communication is made, researchers should open with an explanation of the format of the interview and that the respondent is free to describe whatever drivers and outcomes they want to without fear of reprisal or identification. Although sometimes met with initial scepticism, once people understand the reasons behind this approach, and why the interviewer may not know much about the details or the 'goals' of the evaluation, they are often willing to participate.

Participants should also know how they will be given access to the findings at the end of the evaluation. In an email, links can be provided to a holding page with more information which will be updated at the end of the evaluation. In rural contexts where advance communication may be difficult, researchers are equipped with written letters introducing the study for their discussions with community leaders and respondents. Where possible these use the headed paper of the organisation managing the research in-country (and in our case of Bath SDR) and avoid reference to the commissioner but will always include a contact telephone number and email address for the lead in-country researcher, a contact at Bath SDR and a whistleblowing email address for any concerns. A QR code to a holding page is also now commonly used on the cards left with respondents as many will have access to a smart phone.

Figure 4: Example logistical planning for a QuIP interview

Below are examples of how we have organised interviews in three broad types of scenarios.

World Bank Family Dialogue programme, Mauritania (Commissioned by Trinity College Dublin)
Location: Remote villages in Maghama District
Sampled respondents: Women in communities where the training intervention had been delivered - to be interviewed alone
Means of introduction: Via village elders
Format of interviews: Face to face
Blindfolding: Interviewers Level 2, Respondents Level 1 (but assume a link made to Level 2)
Logistics: The QuIP was conducted alongside an RCT meaning that once a sampling strategy was agreed on, World Bank staff were able to provide the all-female QuIP field team with lists of names, villages and GPS locations in selected locations. Due to the remoteness of the villages, it was not possible to contact people in advance, so researchers were instructed to quota sample opportunistically in situ - up to a specified number from each list. They required letters of permission from Bath SDR and the World Bank in order to travel, and knew that this was linked to a national cash transfer programme but they were not informed of the name or details of the training intervention which formed the 'plus' element of this programme, or what treatment group each village was in (there were four groups in total). Bath SDR received transcripts with

pseudonymised codes, which were then changed. The key to link the codes was deleted at the end of the analysis process to ensure full anonymisation.
Researchers were instructed to gain verbal and written consent from respondents before starting with a form which explained that they were supporting research into changing lives and livelihoods among people living in the area and which listed the domains contained in the interview. They were told that anonymous information collected would be used in social research and shared with commissioners Bath SDR and Trinity College Dublin and ultimately international organisations including the World Bank. Since the World Bank is the main organisation working in the area, we assumed that respondents would make this link - but not necessarily to the specific programme of interest. Researchers knew that they could use their WB permit to help gain access if needed. The main challenges that the field team encountered were physical as distances were long and roads very poor. In most places women were happy to speak with the researchers and men accepted to not be involved since the researchers were women. If there was any resistance or interruptions the researchers abandoned the interview to avoid putting women at risk. The data received did not indicate that blindfolding was compromised; the drivers were wide ranging and the interventions not always mentioned in relation to expected outcomes.

Community building organisation, Netherlands
Location: Online
Sampled respondents: Members who had participated in an activity within the last 5 years
Means of introduction: Email
Format of interviews: Video calls
Blindfolding: Interviewers Level 4, Respondents Level 2
Logistics: Since limited blindfolding was possible in this evaluation the interviews were conducted by Bath SDR staff - hence a high level of knowledge about the programme. Possible respondents were split into three groups by the organisation based on their known engagement. The organisation made initial email contact with all potential respondents to explain that the organisation was undertaking research and inviting them to take part - no personal details were shared with Bath SDR at this stage. Suggested wording was:
We would like to invite you to participate in a discussion with an independent interviewer to talk about what influences people like you in terms of choices about xxx [domains of interest]. This is not a test of you or of xx [organisation], but it is very important to help us understand more about the people who engage with us. All responses will be processed and analysed in complete anonymity; xx [organisation] will not know who participates.
Three separate Google Form links were created by Bath SDR for the organisation to send to each group to sign up and complete consent forms. This gave permission for Bath SDR to contact them independently to arrange interviews and gave us the information about which group they were in to enable quota sampling. Respondents were not told which group they were in to avoid prompting them about the engagement/intervention we were interested in. Interviews were arranged and conducted independently of the organisation and the same interview schedule was used with all respondents. They were told at the start that the interviewers deliberately had limited information about the content of any engagement to allow them to share what was of interest to them. This also aimed to ensure that they shared sufficient detail and didn't assume that the interviewers would know the content of any engagement.

St Mary's University, Twickenham, UK (Commissioned by TASO)
Location: On university campus
Sampled respondents: Students who had participated in an activity within the last year
Means of introduction: Email
Format of interviews: Face to face
Blindfolding: Interviewers Level 4, Respondents Level 3-4
Logistics: TASO were keen to keep interviews open and exploratory as the intended outcomes outlined in the theory of change were likely to be influenced by a range of factors. However, accessing students who had taken part would preclude blindfolding the researchers or respondents. The implementing team at St Mary's University provided a list of participants who Bath SDR contacted via email to invite to participate. The introduction explained the involvement of TASO and St Mary's, and that we were interested in speaking to students about what they had found enjoyable and helpful to the student experience over the last year, what they might have found difficult – and where they may have looked for support. However, low take-up of interviews meant that the researchers had to attend some intervention sessions to directly invite students to participate. This meant that most participating students would have been aware that the interview was related to the intervention.
The interview guide was kept outcome-led and open-ended, and researchers were encouraged to always prompt for additional drivers of change. TASO requested that we include a 'safety-net' question at the end of the interview to ask directly about the interventions to mitigate the potential risk that students wouldn't mention the project at all. The analysis flagged responses that had been prompted in this way and there wasn't a marked difference between the stories reported, which may have been partly because the interviewees already knew the interview was linked to the specific intervention. However, the range and sentiment in the data seemed to show that the open-ended and outcome-led approach was effective in reducing confirmation bias.

If there is *any* likelihood that these logistical challenges could place the researchers or the respondents at risk, then it should not be used. Consideration of trade-offs is an important part of the design process. While *"blindfolding may increase the credibility of respondents' voices from the perspective of the … primary audience, this must be offset against the potentially disempowering effect of not immediately revealing to respondents everything that could be revealed about the intervention being evaluated."* (Copestake et al, 2019: 34)

In low trust environments which may be politically sensitive, where the interviewers may be presumed to be associated with a political party or be perceived to be any kind of potential threat then full blindfolding is certainly not appropriate. An example of this was an evaluation conducted for the Aga Khan Foundation in communities living in the Pamir Mountains of Tajikistan and Pakistan. These are isolated communities where a rare outsider visiting would pose questions and unease without explanation. In this case the field team explained clearly that the research was commissioned by AKF but that they were working for an independent organisation to conduct the interviews. The team did not know any details about the programmes which had been implemented by AKF in the area.

Stating the name of the organisation or funder behind an evaluation (often the main prerequisite for gaining access) is not the same as providing full details of the specific programme or policy you are evaluating. This type of partial blindfold in the form of restricted information can provide the researcher with a suitably low level of information which helps them to maintain an open-ended interview and avoid probing respondents in a particular direction. There is still value in outcome-led

questioning to help prevent the interview from focusing too heavily on the organisation or project of interest and to capture other potential causal factors alongside any impact the project may be having.

There are other potential issues to navigate. If there are very **few organisations or programmes** operating in the area, and the questions are fairly specifically on a theme, respondents may know or at least suspect what intervention the interview is about. There may also be much stronger cultures of confirmation or social desirability bias in some contexts than others. This will probably be clear when reading transcripts; for example, if all answers revolve around the intervention, other factors are rarely mentioned, and very positive stories of change are not backed up by other sources of information (e.g. monitoring data and observations by the interviewers). In these circumstances it is important to acknowledge these potential biases in reporting and how this may have affected the findings. A skilled and trained interviewer should be able to mitigate the risk of confirmation bias by putting the respondent sufficiently at ease to respond honestly and in detail - but it is important to take the political and cultural context into account before deciding that this type of interview is going to be the right approach.

Finally, who are you interviewing, and how vulnerable might they be? If they are **young or vulnerable** in some other way, or the context of the interview is sensitive, then the interviews are likely to need to be conducted by experts in that field. The need for sensitivity to potentially triggering situations will make blindfolding the researchers almost impossible, and a careful discussion would need to be had about how much information you can safely withhold from the respondents. As an example, Bath SDR has conducted interviews about sexual and reproductive health, and this has required working with researchers who are experts in this field and therefore know how to conduct interviews in a way which would not put respondents in an uncomfortable position. They are likely to need to know something of the context of the intervention to know how far to probe for detail without being intrusive. Similarly, other organisations have conducted QuIP interviews with adolescents, with teams who are trained in working with young people[1].

The table at the end of this paper gives further information on a range of examples of Bath SDR evaluations in which blindfolding was adapted to work with different levels of information in different contexts.

4. Ethical considerations

The main criteria informing much of the discussion of outcome-led interviewing and blindfolding in this briefing has been to maximise the truth, credibility and utility of the data produced. This has implicitly assumed that withholding some information can be justified as the 'means to the end' of obtaining better quality data. However, this is only one way to view the ethics of full or partially blindfolded interviewing, and in this section we reflect more fully on the issue. In focusing specifically on the ethics of blindfolding we do not cover numerous other ethical issues that apply when conducting this type of research, but are not specific to this feature of QuIP interviewing. Here we

[1] For separate guidance on using QuIP with young people see https://bathsdr.org/wp-content/uploads/2025/06/Using-QuIP-with-children-and-young-people.pdf
UNICEF have also created a video on using QuIP with young people and their carers in research on drivers of school attendance in India: https://www.youtube.com/watch?v=JZejxNKa2ZE

focus on ethics as they relate to the person being interviewed, while recognising that with the use of double blindfolding similar issues arise for those conducting interviews too.

4.1 Informed consent

To date we have not encountered issues in getting ethical clearance to use QuIP in a range of countries. With QuIP, informed consent focuses on the nature of the interview rather than on the specific intervention being evaluated: from "consenting to evaluate Project X" to "consenting to share my life experiences regarding Topic Y." Ethical approval then focuses on ensuring that participants are fully informed about how and why interviews are to be conducted, can assess the risks of participation and perceive them to be low, recognise their right to withdraw from the process at any time, and are informed how to provide debriefing or feedback information. Interviewees are informed that:

- The interview will focus on changes in selected aspects of their general life experiences – such as livelihood, health, wellbeing or performance of an organisation to the which they are linked, rather than a specific project or intervention;
- What they say will be treated in confidence and anonymised, and may then be used to provide feedback to particular organisations about specific projects or activities; but that not naming these organisations or projects at this stage helps to ensure that information collected is more holistic and balanced;
- The interviewer is independent of any organisation being evaluated, but (in the case of double blindfolding) has also been informed as little as possible about any specific organisation or project being evaluated so as to protect the integrity of what respondents say, by limiting their ability 'to steer' the interview (knowingly or otherwise) towards issues that may not be so important to the respondent.
- How they can subsequently find out more about the research, who commissioned it and why, should they so wish.

Briefing and training of interviewers emphasises the ethical as well as practical importance of reaching a shared understanding of these issues with respondents. This includes being sensitive to any signs that respondents are not sufficiently reassured and relaxed - both to avoid causing discomfort and to ensure the interview can proceed smoothly. Below is an example suggested response given to researchers to prepare them for fieldwork.

> ***Why will you not tell me the name of the project?***
>
> *We have deliberately been given only limited information about this project so that we hear about everything <u>you</u> want to tell us, rather than you only telling us what you think we might want to hear. We are interested in what you think is important, and this might be about a lot more than one project. We don't know in order that we can't prompt you to go in one direction or another.*
>
> *We would like to understand what individuals, groups or policies, as well as wider changes, have had positive or negative effects on people like you, so that people working in this area can learn more about what works and make improvements in future.*

Explaining blindfolding in this way can reassure respondents, however if respondents or gatekeepers remain concerned, the research team should have recourse to a 'Plan B', e.g. contact a manager who may decide to reveal further information or letters of support to help them progress.

Collecting consent to participate in an interview is still important, and still valid because participants are agreeing to share their *reality* (what actually happened to them), which is the true subject of the study. They are fully informed about the procedures (interview, recording), rights (withdrawal, anonymity), and broad topic. They are not being deceived about what they are asked to do, they are simply not primed to attribute their experiences to a single cause.

4.2 Avoiding overly extractive interview practices

The nature of blindfolding means there is a risk of collecting irrelevant information, which can be considered unethical practice. Since the interviewer and interviewee are unaware of the project intervention they may spend time discussing external factors that have limited value for programme development. The collection of data that does not contribute to the study's objectives also risks undermining the value of respondents' time and the stories they provide. It is therefore important to design the interview guides carefully and to ensure that while they are sufficient to keep conversation open, they are also focused enough on the topics of interest. In certain cases, as seen in the previous section, it may be helpful to make some questions more specific. For example, in a Bath SDR evaluation focusing on a poultry-related intervention, interview guides explicitly distinguished between chicken farming and crop farming to limit the collection of unnecessary information.

It is also important to test the amount of time that an interview takes and to remove any unnecessary questions taking into account the amount of time respondents realistically will be willing and able to offer. When farmers, for example, weather and seasonal working patterns are important considerations. Similarly studies which required speaking to shopkeepers or stallholders have required interviewers to spend more time waiting patiently to speak to respondents in between customers, and to be flexible and accommodating about that.

4.3 Unblindfolding procedures

Respondents who have shared their time and stories with researchers are often curious what they have contributed to and what impact this will have. There is an ethical imperative to share study findings with participants, and this can be done in a number of ways. Facilitating feedback workshops allows respondents to validate findings and give feedback on how their stories have been framed. This is beneficial to both participants and the commissioners of the evaluation. If this is not possible, programme staff may be able to informally share the findings if they regularly visit the study locations or individual respondents. If in-person workshops or visits cannot be arranged, we commit to a minimum of sharing results online with participants as an easy and efficient way of unblindfolding respondents. This is facilitated by leaving respondents with a URL or a QR code to a page where basic information about the evaluation is posted and later will be updated with findings. At the end of an interview respondents may also be given contact details for a local officer from the commissioning organisation, so they can ask follow-up questions or give any comments about the study.

Figure 5: Examples of application of blindfolding in different contexts

The examples below have been selected to give a realistic picture of a range of contexts in which different levels of blindfolding were used, particularly where there were challenges and adaptations made along the way.

Interviewers	Respondents	Evaluation examples	Rationale, application and results
1. Topic area	1. Topic area	**Self Help Africa, CropNuts, Kenya** Cropnuts works to support smallholder farmers in the pyrethrum value chain, a flower which can be crushed and used as an insecticide, including training in climate smart agricultural practices for pyrethrum farming.	While SHA wanted to keep the interviews as open-ended as possible, it was important to ensure we captured specific information on the crop of interest. Therefore, researchers were told that pyrethrum was the crop of interest for the evaluation to ensure that enough detail was captured in the interviews, but they had no knowledge of the hypotheses being tested or the organisations involved in the project. Respondents were not told about the focus on pyrethrum as SHA were also interested in any changes to other crops, so questions about agriculture were left open. However, as insurance, the interview included a supplementary question to be used if respondents had not mentioned pyrethrum by the end of the interview. The question was largely unused as the programme had influenced significant change for most farmers and pyrethrum was therefore discussed unprompted in the open questions.
1. Topic area - - 4. Interventions	1. Topic area	**AgDevCo, Uzima, Uganda** Exploring the impact of the Uzima's Chicken Limited programme which aims to improve the lives and wellbeing of poultry farmers by providing training on good poultry practices and improving access to finance to support business expansion.	AgDevCo were interested in what aspects of the programme respondents mentioned without prompting, however there was concern that without some direction, interviews with farmers might focus on crop farming rather than poultry. For this reason, *researchers* were given a brief overview of the interventions, but without being briefed on the commissioner or the expected outcomes. They arranged access to farmers independently. Whilst this approach worked in terms of eliciting information, there were some issues distinguishing between two different levels of

			training delivered by the organisation. It was not always clear which training the respondents were referring to which made it harder to compare outcomes between them. With more than one similar intervention which needs to be disaggregated, it may be necessary to tell the researchers to probe for sufficient detail to distinguish.
1. Topic area 2. Commissioner	1. Topic area 2. Commissioner	**Better Work Programme, ILO, Sri Lanka** This was a hybrid approach which used aspects of Realist Evaluation in the design and incorporated a process evaluation as well as QuIP to understand the impacts of an ILO programme to improve working conditions in factories.	The respondents were from three main categories; people working for government bodies, companies, and trade unions who were trained to be Master Trainers by the ILO; factory executives involved in Occupational Safety & Health Committees; workers within the factories. The first group were interviewed by the senior researcher in a much more open way, using Realist techniques to provide a process evaluation by asking for direct feedback on the training programme they had participated in, as well as asking about outcomes. The second two groups were interviewed by different researchers who knew that this was an ILO evaluation, but they knew nothing about the theory of change for the programme. They gained access to the factories via a letter from the local ILO office, which also did not refer to the specific programme to help avoid priming them. The interviews with factory managers and workers used the QuIP approach of only outcomes-based questions, based on the expected causal pathways. Using this approach at the factory-level helped to mitigate a very real concern about confirmation bias in the responses managers and workers might give in these situations. Interviews often started with an affirmation that conditions were good in factories and that nothing needed to change, but in unpacking what had happened over the two-year recall period evidence of recent change emerged that could be linked back to aspects of the programme as well as other contextual factors.

1. Topic area 2. Commissioner	1. Topic area Unplanned but presumed in some cases: 2. Commissioner 3. Project	**Save the Children & Give Directly, Maziko, Malawi** Evaluating the impact of targeted cash transfers to families with young children alongside a package of social and behaviour change interventions.	The organisation in Malawi used to conduct the interviews were also involved in conducting the RCT for the same programme, but with different staff. They worked hard to ensure that project details were not shared, but it felt unrealistic to expect that the qualitative researchers would not know that it was for Save the Children given the locations. The team are very experienced in conducting QuIP interviews which gave everyone involved confidence that they knew how to probe without pushing respondents in any particular direction. The researchers did not know the details of interventions or the assumptions in the theory of change. However, once in the field, in some cases members of the research team were referred to by community leaders either as coming from Save the Children or Give Directly as it was clear that the people selected to respond were part of the treatment arms. The team did their best to distance themselves from two organisations during introductions and emphasised that they were interested in broader change. The findings did not indicate strong levels of bias; many different drivers of change and limited or negative outcomes as well as positives were reported - so it was not felt that this affected the outputs.
Started at: 1. Topic area Finished with: 2. Commissioner 3. Project 4. Interventions	Started at: 1. Topic area Finished with: 2. Commissioner 3. Project 4. Interventions	**Washing Machine Project, India** Evaluating the impact of distributing Divya manual washing machines in communities which rely on women handwashing clothes.	This was designed as a double blindfolded study but because there were no other organisations working in the area some respondents immediately presumed that TWMP's delivery partner was involved. By implication, researchers also presumed their involvement after a few interviews when this was the primary focus of the discussions. Blindfolding was therefore not used as intended, but the emphasis on outcomes helped the researchers ask questions about many aspects of respondents' lives and helped them get the full picture. The researchers reported it was an interesting way to interview respondents and helped them reflect on their own biases –

			see this blog called Field Notes from Tamil Nadu: The quiet power of blindfolding and an outcomes based approach about their experiences. Respondents were very open about both the benefits and challenges of the washing machines they had received, implying that they felt comfortable talking to the research team; the fact they were not linked to the delivery organisation may have helped this.
1. Topic area 2. Commissioner 3. Project 4. Overall intervention	1. Topic area 2. Commissioner 3. Project	**Safer Streets Fund, Home Office, UK** As part of a wider evaluation with Verian, the QuIP study added to the understanding of the impact of the Safer Streets Fund Round 4. This funding supported organisations working to reduce neighbourhood crime, anti-social behaviour and violence against women and girls. This evaluation focused on programmes in Swindon and Gwent.	This study interviewed youth workers, teachers and parents with children involved in the programmes and members of the public. Youth workers and teachers had demanding schedules and concerns around confidentiality due to the vulnerable young people they worked with, so informing these participants of the commissioner, project and domains covered in the interview was important for individuals to feel confident participating. In many cases, respondents were contacted through their employers, community groups or local police; therefore, sharing information about the study facilitated accessing potential participants. Reaching respondents was challenging and double blindfolding would not have been possible or appropriate under the circumstances. Interview questions were open-ended to allow the respondents to speak freely about what they believed to be significant changes in their lives and their community. Researchers used probing questions to further establish what the perceived influence factors of these changes were. We found participants discussed a mix of negative and positive changes, linking them to the interventions and wider contextual factors.

More information, example reports and case studies are available at our Resources Library: www.bathsdr.org/resources

5. References

Copestake, J. (2014). Credible impact evaluation in complex contexts: Confirmatory and exploratory approaches. *Evaluation*, *20*(4), 412-427. https://doi.org/10.1177/1356389014550559 (Original work published 2014)

Copestake, J., Allan, C., Bekkum, W. van, Belay, M., Goshu, T., Mvula, P., Remnant, F., Thomas, E., & Zerihun, Z. (2018). Managing relationships in qualitative impact evaluation of international development: QuIP choreography as a case study. *Evaluation*, *24*(2), 169-184. https://doi.org/10.1177/1356389018763243 (Original work published 2018)

Copestake, J., Goertz, G., Haggard, S (2021). The Veil of Ignorance Process Tracing (VoiPT) Methodology. *Qualitative & Multi-Method Research* 18(2): 19-26. https://doi.org/10.5281/zenodo.4046580.

Copestake, J., Morsink, M., & Remnant, F. (2019). Attributing development impact: The Qualitative Impact Protocol case book. Practical Action Publishing

Copestake, J. (2025). Mixed-methods impact evaluation in international development practice: distinguishing between quant-led and qual-led models. Journal of Development Effectiveness, 17(1), 1–20. https://doi.org/10.1080/19439342.2024.2351892

Scriven, M. (1972). Pros and cons about goal-free evaluation. Journal of Educational Evaluation, 3(4), 1-7.

Vedung, E. (1997). Public policy and program evaluation. Transaction.

Youker, B. W., & Ballard, A. (2024a). Goal-free evaluation: An orientation for foundations' evaluations – Updated 2024. The Foundation Review, 16(1). (Original work published 2014.) https://doi.org/10.9707/ 1944-5660.1697

Youker, B. W. (2024b). Scriven's Goal-Free Evaluation. Journal of Multidisciplinary Evaluation, 20(47), 28–31. https://doi.org/10.56645/jmde.v20i47.1005