

Qualia

CONTENTS

[Background on data collection with Qualia](#)

[The potential of AI interviewing and conversational research](#)

[Our paper on an inductive workflow to gather and analyse evidence at scale.](#)

[Case study – our 'seamless stories' workflow in practice](#)

[Case study – Qualia asks about USA problems, again](#)

[Qualia the multi-linguist](#)

[People are often more candid with machines than with other people. Why?](#)

Background on data collection with Qualia

22 Oct 2025

[Qualia](#) is our AI interviewer.

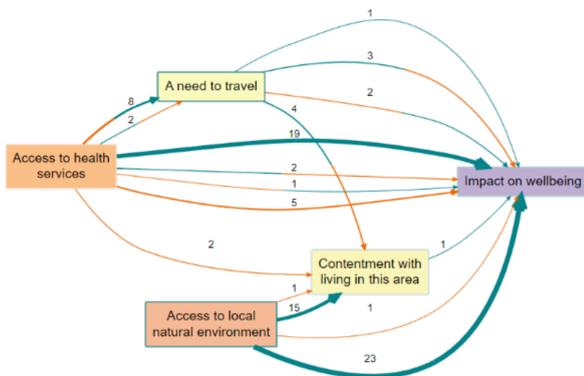
You can use Qualia to collect data for causal mapping: [Task 1 -- Introduction](#). You can use it for other kinds of data collection too.

If you want to know what buttons to press for setting up, sharing and managing your interviews, look at the [Qualia technical documentation](#), including How do you [create a really great interview?](#) and [How to manage languages?](#)

QualiaInterviews

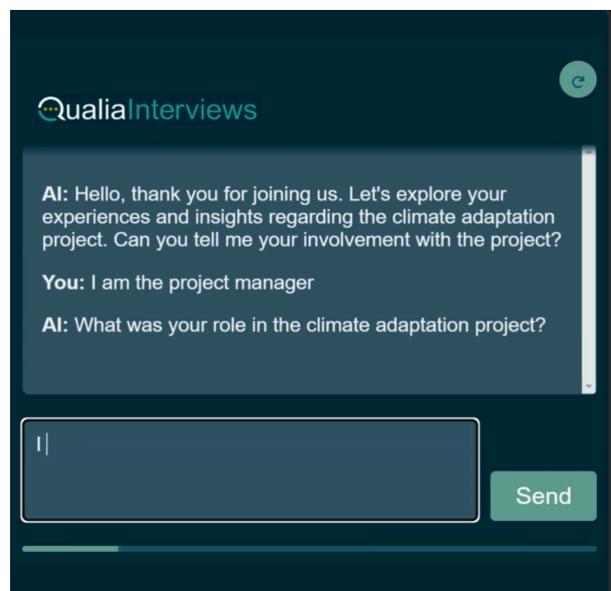
Go beyond questionnaires and “free text” answers. Find out what people are really thinking, in their own words:

- What worked and why?
- What didn't work, and why?
- What new issues are emerging?



How does it work?

1. Design and test your Qualia interview online.
2. Send your respondents a link to the interview.
3. Your respondents click on the link. Our friendly AI-interviewer guides them through the whole interview. They can type or speak their answers.
4. Get an automated report. Or download the transcripts for further analysis.



This chapter (which is a constant work in progress) gives you the background, including research have we done on Qualia and some case studies.

The potential of AI interviewing and conversational research

23 Aug 2025

Qualitative approach: You can use Qualia any way you want, but we prefer a conversational approach which has just a few loosely-defined high-level research questions, and leaves Qualia to manage the discussion. This way, the interviewer approaches the stakeholders and their stories as far as possible **without preconceived templates**, to remain open to emerging and unexpected changes in respondents' causal landscapes.

Scalability and reach: The AI's ability to communicate in many [languages](#) and even to **switch languages** presents an opportunity to reach more places and people, subject to internet access and the AI's fluency in less common languages, and to include representative samples of populations.

Consistency: The interview and coding processes are machine-driven and use zero temperature. An AI interviewer can be made to be more consistent than any individual human could manage, let alone a team of humans. This means the interview behaviour should be **mostly reproducible**. Reproducibility opens the possibility of **comparing results across groups, places and timepoints**.

Reporting: The Qualia platform has its own automated [reporting](#). This helps you analyze your interview transcripts using AI to generate different types of insights, from identifying typical "Tribes" within your respondents to generating social network and Power-Influence diagrams.

In summary, this kind of semi-automated pipeline opens up **possibilities for monitoring, evaluation and social research which were unimaginable just three years ago** and are well suited to today's challenging, complex problems like climate change and political and social polarisation. Previously, only quantitative research claimed to produce generalisable knowledge about social phenomena validly and at scale, by turning meaning into numbers. **Now perhaps qualitative**

Qualitative causality: We like to use Qualia to help researchers and evaluators **answer evaluation questions which are often causal in nature**, like: understanding stakeholders' mental models; judging whether "their" ToC matches "ours"; investigating "how things work" for different subgroups of stakeholders; tracing impact from mentions of "our" intervention to outcomes of interest; triaging the key outcomes in stakeholders' perspectives.

When Qualia uses a causal focus, we can use causal mapping (and our own Causal Map app) to rapidly make sense of stakeholder world views.

research will eclipse quantitative research by bypassing quantification and dealing with meaning directly, in somewhat generalisable ways.

Our paper on an inductive workflow to gather and analyse evidence at scale.

22 Aug 2025

In our recent paper (Powell et al. 2025) we demonstrated that it is possible to gather evidence at scale about program theory and contribution simultaneously. Here is a [preprint](#), and below is a summary.

Abstract

This article presents an artificial intelligence-assisted causal mapping pipeline for gathering and analysing stakeholder perspectives at scale. Evidence relevant to constructing a programme theory, as well as evidence for the causal influences flowing through it, are both collected at the same time, without the evaluator needing to possess a prior theory. The method uses an artificial intelligence interviewer to conduct interviews, automated coding to identify causal claims in the transcripts, and causal mapping to synthesise and visualise results. The authors tested this approach by interviewing participants about problems facing the United States. Results indicate that the method can efficiently collect and process qualitative data, producing useful causal maps that capture respondents' views as they evolve across time points. The article discusses the potential of this approach for evaluation. It also notes limitations and ethical concerns, emphasising the need for human oversight and verification.

Method

Our method comprised the following steps (following Tasks 1-3 according to [Powell, Copestake, et al. \(2023\)](#), p. 108-112):

Step 1: Conducting the chat interviews

This was a proof-of-concept analogue study. We employed online workers as respondents, recruited via Amazon's MTurk platform³ (Shank, 2016). We decided to investigate respondents' ideas about problems facing the United States, as this generic theme was likely to elicit opinions from randomly chosen participants. This unsophisticated way of recruiting respondents means that the results cannot be generalised to a wider population in this case.

We had no specific evaluative questions in mind. We aimed to demonstrate a method which can be easily adapted to a specific research question.

A short semi-structured interview guideline was designed on the theme of 'What are the important current problems facing the USA and what are the (immediate and underlying) reasons for those problems?'. We aimed to construct an overall collective 'ToC' around problems in the United States. As it does not encompass a specific intervention this theory is not an example of a programme theory.

This interview guideline was implemented via an online interview 'AI interviewer' called 'Qualia',⁴ which uses the OpenAI application programming interface (API) to control the AI's behaviour. Qualia is designed to elicit stories from multiple individual respondents, in an AI-driven chat format. Individual respondents are sent a link to an interview on a specific topic and, after consenting, are greeted by the interviewer. Rather than following a set list of questions, the interviewer is instructed to adapt its responses and follow-up questions depending on the respondents' answers, circling back to link responses and asking for more information as appropriate, focusing on the interview's objective mentioned above. These behaviours are based on the instructions written by the authors.

The respondents, who had the level of 'Master'⁵ on Amazon's MTurk service, each completed an interview. The Amazon workers were given up to 19 minutes to complete the interview.

We repeated this interview at three different time points in September, October and November 2023, inviting approximately $N = 506$ respondents each time. The data from the three time points were pooled.

- The Research Question for Step 1 is: can an automated interview bot successfully gather causal information at scale?

Step 2: Coding the interviews

Step 2a: Constructing a guideline

Once the interviews were completed, we wrote instructions to guide the qualitative causal coding of the transcripts, in a radical zero-shot style: without giving a codebook or any examples. The assistant was told not to give a summary or overview but to list *each and every causal link or chain* of causal links and to ignore hypothetical connections (e.g. ‘if we had X we would get Z’). We told the AI to produce codes or labels following this template: ‘general concept; specific concept’. We gave no examples, but expected the AI to produce labels like: ‘economic stress; no money to pay bills’. We call the combination of both parts a (factor) label.

The assistant was told also to provide a corresponding verbatim quote for each causal chain, to ensure that every claim could be verified. Codings without a quote which matched the original text were subsequently rejected, thus reducing the potential for ‘hallucination’.

Step 2b: Coding

The final instructions were human-readable and could have been given to a human assistant. Instead, we gave these instructions to the online app ‘Causal Map’, which used the GPT-4 OpenAI API. As the transcripts were quite long (each around a page of A4 in length), each was submitted separately. The ‘temperature’ (the amount of ‘creativity’) was set to zero to improve reproducibility. The Causal Map app managed the housekeeping of keeping track of combining the instructions with the transcripts, watching out for any failed requests and repeating them, saving the causal links identified by the AI, and so on.

Step 2c: Clustering

The coding procedure resulted in many different labels for the causes and effects, many of which overlap in meaning. Even the general concepts (e.g. ‘economic stress’) were quite varied. The procedure for clustering these labels (including both the general and specific parts of the label) into common groups with their labels was a three-step process based on assigning to each of the original labels an embedding. An embedding is a numerical encoding of the meaning of each label ([Chen et al., 2023](#)) in the form of a vector (often visualised as a point in a high-dimensional space). For any two embedding vectors, cosine similarity can be calculated (measuring the angle between them) to quantify the semantic similarity between the labels they encode:⁷

1. *Inductive clustering*. First, we grouped the labels into clusters of similar labels using the `hclust()` function from the stats package of base ([R Core Team, 2015](#)).

2. *Labelling*. We then asked an AI to find distinct labels for each cluster. We also manually inspected these labels with regard to the original labels within each cluster and adjusted some of them.
3. *Deductive clustering*. We then discarded the original clustering, created embeddings for the new labels, and formed a new set of clusters, one for each of the new labels, assigning each original label to one of the new labels, the one to which it was most similar, providing the similarity was at least higher than a given threshold. This additional deductive step ensures that each member of each new cluster is sufficiently close in meaning to the new cluster label, rather than just to the other members of the cluster.

After each sub-step, we checked the AI's results to ensure that the instructions were being followed correctly and, if they were not, the instructions were tweaked or rewritten and tested again to ensure quality and consistency.

- The Research Question for Step 2 is: can automated causal mapping successfully code causal information at scale?

Step 3: Making useful syntheses of causal mapping data to answer evaluation questions

Standard filters (details on request) can be applied to the resulting data set of causal claims to create overview causal maps as a qualitative summary of the respondents' 'causal landscapes'. The primary aim is to construct a simple map with a not-overwhelming number of links and factors which captures a large percentage of the information given by the respondents. In addition, network metrics like centrality can be used to identify the factors which are most central within the network. To weigh up the evidence for the contributions made to a specific factor, we can list the evidence (the specific quotes from specific respondents) for direct and indirect links leading to it.

- The Research Question for Step 3 is: can automated causal mapping help answer evaluation questions?

Results and discussion

- *Question for Step 1. Can an AI interviewer successfully gather causal information at scale?:* Our AI interviewer was able to conduct multiple interviews with no researcher intervention at a low cost, reproducing the results of (Chopra & Haaland 2023). The

interview transcripts read quite naturally and the process seems to have been acceptable to the interviewees.

- *Question for Step 2. Can automated causal mapping successfully code causal information?:* Automated coding was able to identify causal claims made by respondents. The coding was noisy, with 35 per cent dropping at least one quality point, but with no evidence of *systematic* errors. This level of precision is adequate for sketching out ‘causal landscapes’ but would not be for high-stakes evaluations without additional manual correction. The accuracy can also be substantially improved by getting the AI to revise its work, (see Powell et al., forthcoming). This procedure still involves the researchers making significant high-level decisions in the formulation of the coding instructions as well as, before analysis, in clustering similar factor labels into groups. We believe this coding approach using genAI represents a significant improvement over the more hard-coded approaches for identifying causal relationships expressed in text ([Dunietz, 2018](#); [Dunietz et al., 2017](#); [Jiang et al., 2023](#); [Hooper et al., 2023](#); [Yang et al., 2022](#)), and provides a more detailed, section-by-section coding which relies less on using AI as a black box to identify themes for initial coding ([Jalali and Akhavan, 2024](#)) or to identify a global map ([Graham, 2023](#)).
- *Question for Step 3. Can automated causal mapping help answer evaluation questions?:* An overview map was produced which included over 40 per cent of the causal claims identified within the transcripts, using just 11 relatively broad factor labels.

The most central factor with the highest number of citations was economic stress, which is a plausible result, with plausible connections to other factors.

We can use the map to identify and weigh up the evidence for contributions from and to individual factors. For example, the major contributions to economic stress are government policy and Covid-19, as well as ‘self-loops’ mentioned by 46 sources, that is, where one aspect of economic stress was seen as causing another.

All such results depend on the (not automated) decisions made during the clustering process: how many clusters to use, whether to intervene in labelling, and so on. This situation is closely parallel to decisions facing a statistician who has to identify variables for, say, structural equation modelling ([Goertz, 2020](#): 136 ff).

Comparison of citation frequency across time points was able to show that some links were mentioned significantly more than others, illustrating how this kind of map could be used to explore changes in systems (or in mental models of systems) over time.

Caveats

Ethics, bias and validity

This kind of AI processing is not suitable for dealing with sensitive data because information from the interviews passes to [OpenAI's \(2024\)](#) servers, even though it is no longer used for training models.

[Head et al. \(2023\)](#) and [Reid \(2023\)](#) raise concerns about bias and the importance of equity in AI applications for evaluation, which have led to questions about the validity of AI-generated findings ([Azzam, 2023](#)). The way the AI sees the world, the salient features it identifies, the words it uses to identify them, and its understanding of causation are certainly wrapped up in a hegemonic worldview ([Bender et al., 2021](#)). Those groups most likely to be disadvantaged by this worldview are approximately the same who have least say in how these technologies are developed and employed.

AI is developing quickly: new models and techniques become available every month. However, we believe that any tools which genuinely add to knowledge should use procedures which are broken down into workflows consisting of simple individual steps, so that, humans can understand and check what is happening.

Interviewing

Researchers should carefully consider whether the interview subject matter is compatible with this kind of approach. For example, the AI may miss subtle cues or struggle to provide appropriate support to respondents expressing distress ([Chopra and Haaland, 2023](#); [Ray, 2023](#)). We recommend that interview guidelines are tested and refined by human interviewers before being automated. No automated interview can substitute for the contextual information which a human evaluator can gain by talking directly to a respondent, ideally face-to-face and in a relevant context.

There is likely to be a differential response rate in this kind of interview: some people are less likely to respond to an AI-driven interview than others, and this propensity may not be random.

Causal mapping

Causal mapping is not at all suited for estimating the strength of causal effects: it can reveal the *strength of the evidence* for the influence of X on Y but this is not to be confused with the *strength of the effect* itself. There can be strong evidence for a weak link and vice versa.

Auto-coding

The work of the AI coder and clustering algorithms are not error-free. The coding of individual high-stakes causal links should be checked. In particular, there is a danger of accepting inaccurate results which look plausible.

This approach does not nurture substantive, large-scale theory-building of the kind expected, for example, in grounded theory ([Glaser and Strauss, 1967](#)). However, it can do smaller-scale theory-building in the sense of capturing theories implicit in individuals' responses.

This pipeline relieves researchers of much of the work involved in coding, but it is not fully autonomous. The human evaluator is responsible for applying the techniques in a trustworthy way and for drawing valid conclusions.

Potential

Qualitative approach

These procedures approach the stakeholder stories as far as possible without preconceived templates, to remain open to emerging and unexpected changes in respondents' causal landscapes.

Scalability and reach

The AI's ability to communicate in many languages presents an opportunity to reach more places and people, subject to Internet access and the AI's fluency in less common languages, and to include representative samples of populations.

The interview and coding processes are machine-driven and use zero temperature, so this approach should be mostly reproducible. Reproducibility opens the possibility of comparing results across groups, places and time points.

The low cost of coding large amounts of information means that it is much easier to develop, compare and discard hypotheses and coding approaches, something which qualitative researchers have previously been understandably reluctant to do.

Qualitative causality

These procedures have the potential to help evaluators answer evaluation questions which are often causal in nature, like: understanding stakeholders' mental models; judging whether 'their' ToC matches 'ours'; investigating 'how things work' for different subgroups of stakeholders; tracing impact from mentions of 'our' intervention to outcomes of interest; triaging the key outcomes in stakeholders' perspectives.

In summary, this kind of semi-automated pipeline opens up possibilities for monitoring, evaluation and social research which were unimaginable just 3 years ago and are well suited to today's challenging, complex problems like climate change and political and social polarisation. Previously, only quantitative research claimed to produce generalisable knowledge about social phenomena validly and at scale, by turning meaning into numbers. Now, perhaps, qualitative research will eclipse quantitative research by bypassing quantification and dealing with meaning directly, in somewhat generalisable ways.

Further work

We have tried to demonstrate a semi-automated workflow with which evaluators can capture stakeholders' emergent views of the *structure* of a problem or programme at the same time as capturing their beliefs about the *contributions* made to factors of interest by other factors. We have presented this approach via a proxy application but have since applied it in real-life research. Many challenges remain, from improving the behaviour of the automated interviewer through improving the accuracy of the causal coding process to dealing better with valence (e.g. distinguishing between 'employment', 'employment issues' and 'unemployment'). Perhaps, most urgently needed are ways to better understand and counter how LLMs may reproduce hegemonic worldviews ([Head et al., 2023](#); [Reid, 2023](#)).

References

Chopra, & Haaland (2023). *Conducting Qualitative Interviews with AI*.

<https://doi.org/10.2139/ssrn.4583756>.

Powell, Cabral, & Mishan (2025). *A Workflow for Collecting and Understanding Stories at Scale, Supported by Artificial Intelligence*. SAGE PublicationsSage UK: London, England.

<https://doi.org/10.1177/13563890251328640>.

Case study – our 'seamless stories' workflow in practice

9 Apr 2025

Automating chat interviews with **Qualia**. Then using **Causal Map** to make sense of them. In-depth research was never this easy! A case study from Chile.

At Causal Map we're thrilled because our [seamless AI-supported workflow](#) is finally coming together. Recently we helped colleagues at a University in Chile to complete a qualitative, explorative evaluation of the impact of a programme, using our automated interviewer **Qualia** to conduct the interviews and **Causal Map** to make sense of them.

This workflow means you can do **in-depth research** so much more **quickly** and **cheaply** than before while maintaining depth and quality, opening up new possibilities for understanding complex social issues.

Background

DuocUC, a higher education institution in Chile, hired our consultancy to conduct QuIP-style interviews with Qualia and analyse them using the Causal Map app. The interviews were motivated by concerns about the gender gaps faced by women pursuing STEM careers at the university.

This study has been developed in the quality assurance department, as part of the institutional evaluation strategies, led by Felipe Rivera, Head of Academic Quality Evaluation.

We had a first meeting to understand what they wanted to find out, their research questions and the scope of the study and to determine the domains in which the interviews would be conducted.

After this, we started writing the instructions for Qualia to conduct the interviews, having a few iterations with the client's team to come up with an interview structure that would suit them.

The process

Step 1: Setting up the interview in Qualia

- The instruction for the AI interviewer was similar to the instructions you could give to a human interviewer. And both the interview instructions and the interviews itself were conducted in Spanish.
- The AI asked questions about changes in 3 domains: educational experiences, professional development and relationship dynamics.
- We used GPT-4o which is the best AI model to date.

Qualia: Hola, ¿te parece bien empezar la entrevista? Esto tomará entre 15 y 30 minutos en completarse.

😊 **You:** Si

Qualia: ¡Perfecto! La mayoría de nuestras preguntas se refieren a lo que ha sucedido en el último año. Para comenzar, ¿puedes recordar algo importante que te haya ocurrido hace 12 meses?

Type here! (Press Enter to send.)



Step 2: Collecting stories with Qualia

- We sent the interview link to 50 people and were able to collect 32 interviews.
- We created special individual links to be able to track the interviews:
 - At Qualia, we don't store personally identifying information at all. But we can add a personalised key like `&key=0003` to the end of the URL for each individual invitation.
 - And this allowed the researchers to keep track of who they sent which invitation to, so that they knew that e.g. key 0003 belongs to Claudia.
- We downloaded the interview results from Qualia and uploaded them into Causal Map.

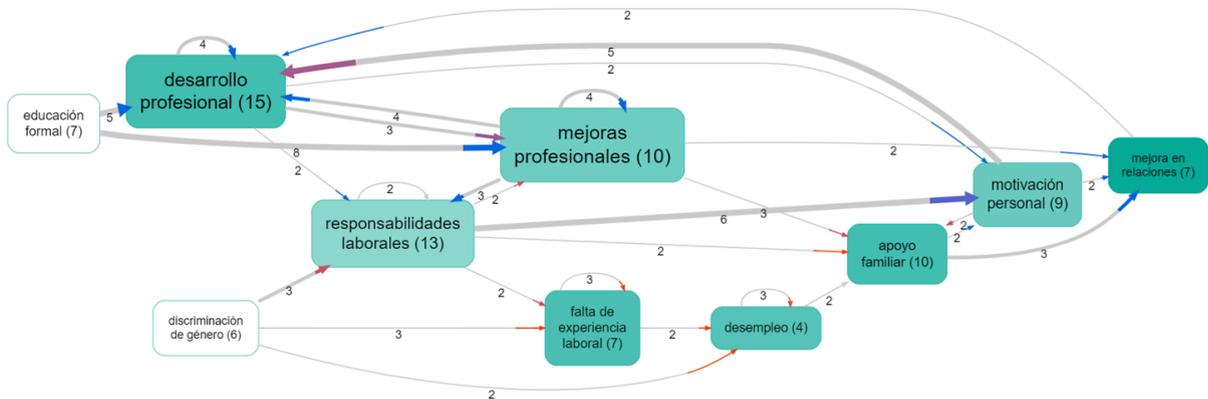
Step 3: Analysing stories with Causal Map

- We used AI (GPT-4o) to identify each and every causal link in the interviews, and for each link, to label the cause and effect.

- We used a “radical zero-shot” approach in which the AI is given no codebook and is simply told to invent its own codes (in Spanish). We gave the AI context about the project.
- We found **251** causal links mentioned by the respondents
- Then we also auto-coded the sentiment of each link in order to show which contributions were “positive” (blue arrowheads) and which were "negative" (red arrowheads).

Step 4: Answering research questions with Causal Map

- Once the coding was done, we used the filters in the app to create different maps that answered their research questions:
 - “What was the immediate impact on the respondents’ lives because of gender discrimination?”
 - “What is the causal network from gender discrimination?”
 - “What are the most mentioned factors by the sources?”



- We also used the ‘AI Answers’ feature to help us understand more about the interviews
 - This functionality allows you to ask questions about all the text in your file.
 - It is completely independent of causal coding. It will work just as well without causal coding.

See what **Javiera Cienfuegos**, Senior Researcher of the evaluation project, has to say:

👤 "The type of questions that were asked "what causes what", were equally linked to methodological innovation. The results were able to portray how gender barriers are intertwined in domains ranging from higher STEM education to the performance of new professionals and technicians once they enter the labour market, reaching deeper explanations and social impact."

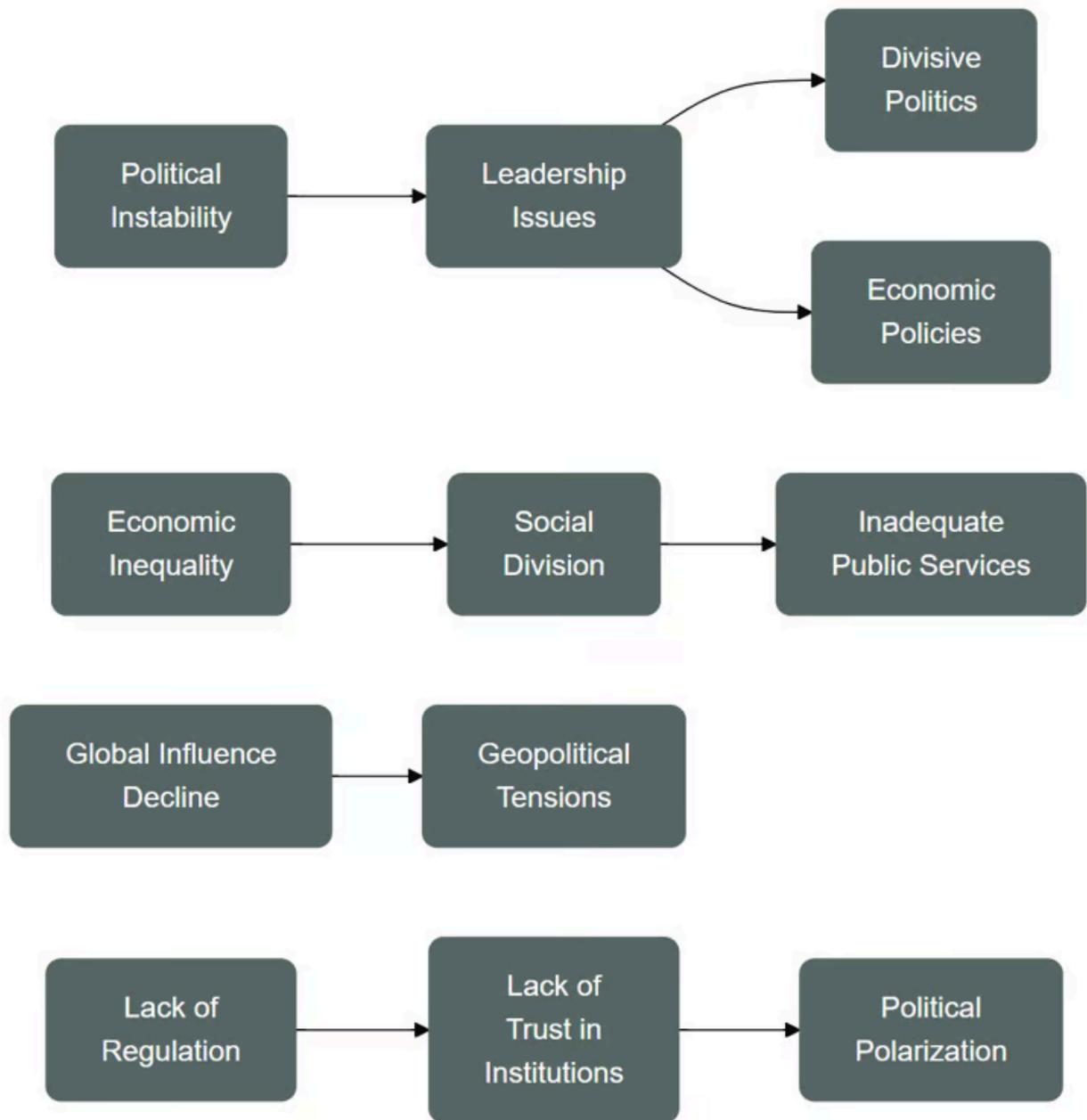
Case study – Qualia asks about USA problems, again

9 Apr 2025

Feb 27, 2025 at EES

How can we capture and visualise people’s mental models of a complex situation like the state of a nation? This week, as part of an EES [webinar](#) demonstrating our automated AI interviewer Qualia, we asked the participants to spend a few minutes being interviewed about problems facing the USA and the reasons for them, and the reasons for the reasons. Over 90 people did, with a mean of 13 messages per conversation. Details below.

The Qualia platform provides an instant overview of the transcripts. For some reason, we didn’t think to show it at the time, but I’ve pasted it in at the bottom of this post. Qualia also provided a simple causal map:

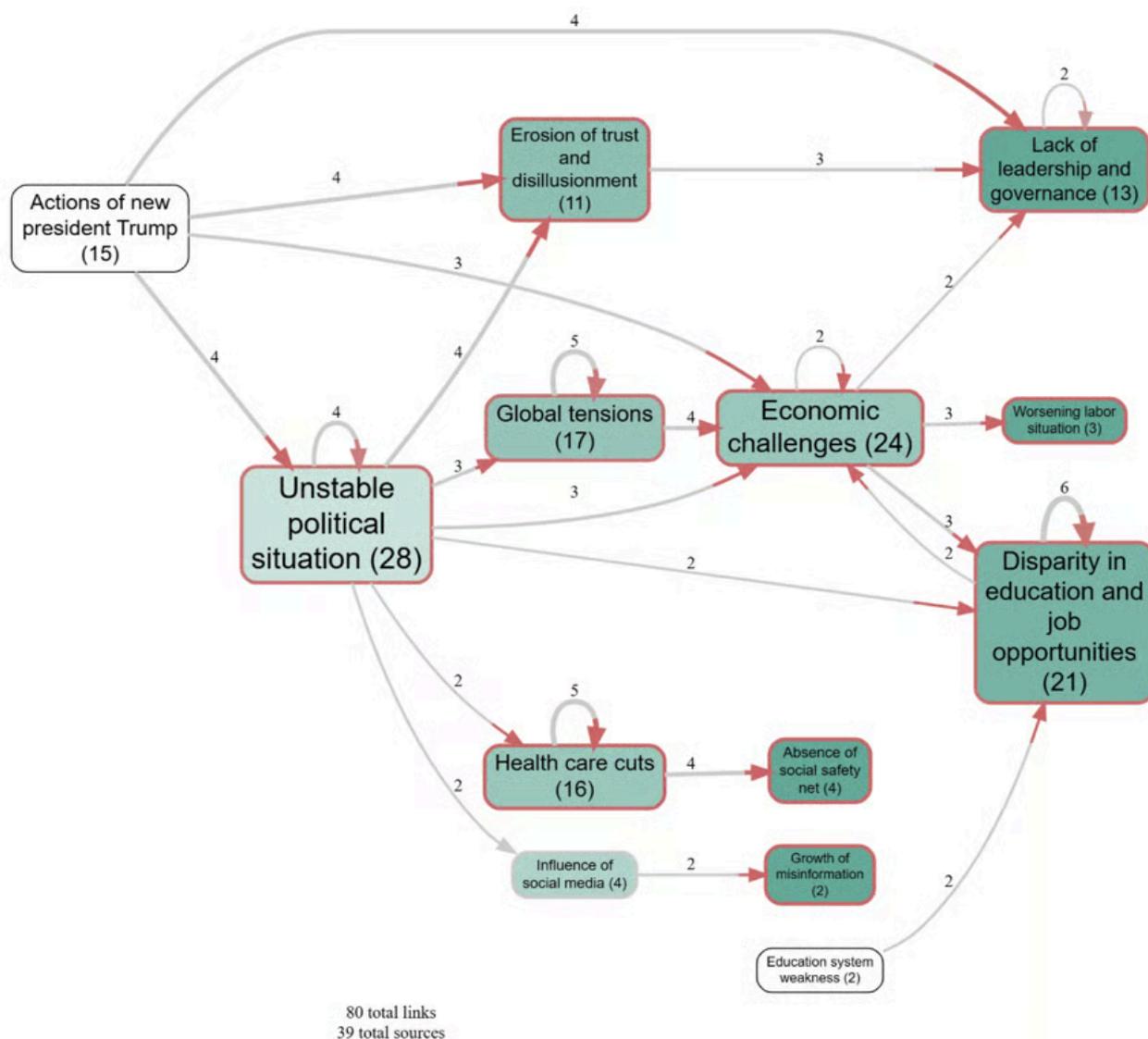


Because this was a demo interview and many respondents only started it and only a few finished the conversation, we are not taking this analysis so seriously, it's just an example of the types of outputs you can get with the Overview Tab in QualiaInterviews — but although we can't make any claims to be doing fundamental social science here, the results are still worth a look.

The Overview in the Qualia Workspace app is just a simple hack which is basically like uploading all the transcripts to ChatGPT and saying “make sense of this please”. [We've already talked at length](#) about the dangers of that: basically you are entrusting a whole load of evaluative

judgements to a black-box AI, which is not only completely non-transparent but is cutting corners everywhere in the attempt to come to a plausible enough result as quickly and cheaply as possible.

A much better way is to break up the vague, high-level task into multiple simple, transparent ones, in this case, identifying all the causal claims in the transcripts, where someone said that one thing leads to or influences another, and aggregating them. The result looks like this:



A “Factor” is any box, including outcomes, drivers and things in between The map is filtered to show most important links and/or factors: many other links and factors are hidden Numbers on factors (boxes): number of mentions Sizes of factors (boxes): number of mentions Numbers on links: number of sources mentioning it Darker backgrounds: higher “Outcomeness”: a bigger proportion of incoming links Deeper red arrowheads: the effect was more negative in significance/sentiment

Some things to note:

Many people mentioned Trump as a driver of changes (white background, positioned at left)

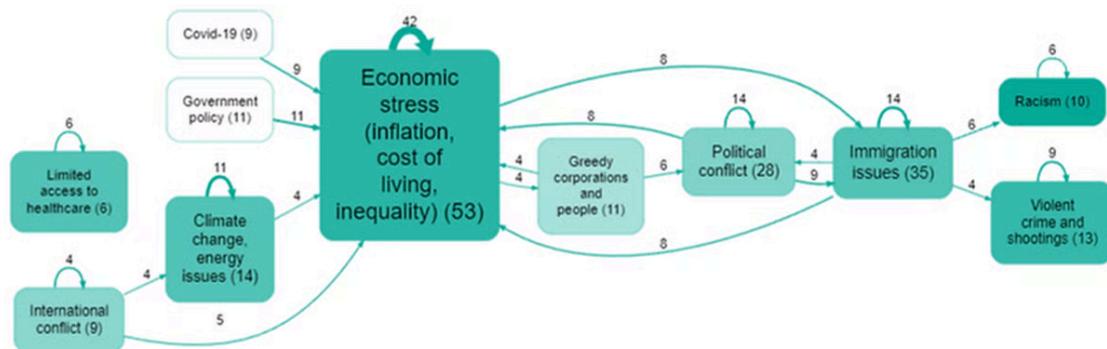
Most frequently mentioned factor was “Unstable political situation”, whose only significant driver was Trump’s actions.

We shouldn’t fall into the "transitivity trap" of thinking that, because Trump is linked to Unstable political situation which is linked to Health care cuts that many or any individual sources told us about all the sections of this chain: the information for each section might have come from different sources (in fact, it mostly did).

We have done this type of interview several times before. Here is a map from 2023.

EES 2023: Sharing our journey on AI’s application in qualitative research

This was a completely different method and sample. The difference between these two maps has substantial face validity, but that is about all we can say at this point.



Filename: usa-problems-merged. Citation coverage 53%. 343 of 653 total citations and 83 of 88 total sources are shown here. Numbers on factors show source count. Factor sizes show citation count. Darker factor colours show greater outcomeness. Numbers on links show source count.

Zooming out to level 1 of the hierarchy. Auto clustering factors with granularity of 56% at level 1. Showing only factors with at least 6 sources. Showing only links with at least 4 sources.

Technical annex

The interview

The instructions for the AI interviewer are to conduct a fairly mechanical interview which simply asks the respondents to list the main problems facing the USA today, and for each problem, to ask for reasons and reasons for the reasons, and to ask how these are interconnected. It also asks the respondent to confirm the causal connections identified before finishing. It isn't a very chatty or explorative interview, but it works.

Producing the maps

We asked our AI to code the transcripts one at a time and to use codes for each cause and effect which were already a little abstracted, using "concepts from a social science textbook" but with no more specific codebook. This resulted in 325 links using 458 factor labels with overlapping meaning, which we then clustered using a procedure detailed in a [forthcoming publication](#).

Instant overview of transcripts provided by Qualia (global, "black box" approach).

Overview of Interview Transcripts

The interviews highlight a wide array of challenges currently facing the USA, with themes revolving around political, social, and economic issues. Common problems include divisive politics, leadership concerns, economic inequality, governance issues, and the influence of certain political figures like Donald Trump.

Common Features:

- **Leadership and Political Instability:** Many respondents highlight issues with leadership and governance as pivotal problems. Concerns include the actions and influence of Donald Trump, new administrations, and an unstable government.
- **Economic Concerns:** Economic inequality, unemployment, and budget deficits are frequently mentioned, pointing towards systemic economic challenges.
- **Social Issues:** Polarization, racism, and loss of public services are recurrent themes, along with critiques on social justice and inequality.
- **Foreign Policy and Geopolitics:** Geopolitical tensions, particularly involving countries like Russia and China, and the USA's diminishing global influence are noted.

- **Environmental and Social Policies:** Climate change and health care are cited as critical problems needing more attention.

Differences:

While some respondents focused on specific figures, like Trump, and their impact on the national and international stage, others highlighted systemic issues such as the neoliberal economic model or the enduring influence of oligarchs.

Perspectives vary on underlying causes, with some pointing to specific policy decisions and others referencing broader societal trends, such as populism or media influence.

Qualia the multi-linguist

21 Aug 2025

Qualia can switch between arbitrary languages on request. So if a user says "koennen wir auf deutsch weitermachen?", Qualia will just switch to German without missing a beat. It's a pretty amazing skill, which can go a long way to improve the reach and inclusivity of conversational research. Lots of respondents talk multiple languages but may be more comfortable in one than another.

That's the theory, and often it works. It is **less likely to work when using EU-only mode at QualiaInterviews for GDPR compliance and at the same time allowing respondents to use audio**.

- The simplest solution is to send users language-specific invitation URLs. These can be pre-prepared for different groups of users or we can use generic invitations which offer a choice, so if they want to speak French, they click the French link.
- Alternatively, users can select a language from a dropdown. This is more general but a bit less intuitive.

Understanding how Qualia manages with different languages

There are two things to think about:

- 🗣️ the transcription service (necessary only if we enable the option for people to speak instead of type)
- 🧑 the AI interviewer service which provides interviewer responses.

Bearing that in mind:

- The top 50 or so languages in terms of how present they are on the internet should all work fine for both transcription and interviewing.
- Major languages and their variants like Brazilian Portuguese should be fine for both.
- Many quite common languages like Kurdish would require us using dedicated services for both.
- For Arabic variants (beyond Modern Standard), the situation is more tricky, but probably similar for both. As I understand the current state of affairs the problem the models have with Arabic variants is more about cultural adaptation rather than the language itself. Voice transcription would probably require us to install a special model. But we can't guarantee this would work.

- Although Qualia does a very good job of detecting / guessing the respondent's preferred language and adapting to that, we get best results if we don't do that but tell it in advance which language will be used -- but this means people who we expect to use, say, Portuguese are not then able to switch to, say, English.
- **Are you a Language Capability Doubter?**
 - Are you sceptical whether Qualia can effectively handle interviews in their target language or across multiple languages.
 - Qualia supports approximately the top 50 languages present on the internet, with particularly strong capabilities in major languages like Brazilian Portuguese.
 - For optimal results, we can configure Qualia to specifically operate in your target language rather than relying on automatic detection.
 - The system combines both transcription services (for spoken responses) and AI interviewer capabilities customized to your language needs.
 - Less common languages may require special considerations, we can evaluate feasibility for your specific language requirements.
 - Qualia's language capabilities allow for consistent interview quality across different markets, ensuring comparable data collection.
- **Are you an AI Reliability Sceptic?**
 - Are you sceptical about the reliability, quality, and authenticity of AI-conducted interviews compared to traditional human methods?
 - Qualia operates on the best available new generative AI technology, producing consistent and friendly interviews that eliminate human interviewer variation.
 - The system can be precisely configured to follow your interview protocol, ensuring methodological rigor.
 - We can provide demonstrations showing how Qualia handles different respondent types and interview scenarios.
 - The AI interviewer can adapt to respondent answers while maintaining your research objectives, combining flexibility with consistency.
 - Using Qualia allows you to conduct more interviews within your budget, significantly increasing sample size and explanatory power.

People are often more candid with machines than with other people. Why?

23 Jun 2025

(Gambino et al. 2020)

People are often more candid with an AI interviewer than with a human interviewer, especially about sensitive or embarrassing topics. This is not a magic property of “AI truth serum”; it’s a predictable consequence of how people manage impressions, respond to perceived judgement, and adapt their social scripts to machines. In practice, increased candour can be a **feature** (less socially desirable responding) and a **risk** (over-disclosure, misplaced trust, and biased measurement).

Mechanisms (why candour often increases)

Reduced evaluation pressure (less impression management)

When the “other” is perceived as non-judgemental (or at least less socially consequential), respondents can feel less need to perform competence, morality, status, or emotional stability. One way to phrase it: the *cost of looking bad* is lower, so people spend less effort on impression management and can answer more directly. Evidence from virtual-human interviewing suggests that simply believing the interviewer is automated can reduce fear of self-disclosure and impression management, and increase willingness to disclose in health-screening contexts (Lucas et al. 2014). Relatedly, in a hiring context, chatbot-based personality assessment was found to be **less susceptible to social desirability bias** than traditional psychometric tests (though with trade-offs in predictive validity) (Dukanovic & Krpan 2025).

Different “social script” for machines (CASA, updated)

The CASA paradigm argues that people often apply social rules to computers (e.g., politeness, reciprocity), even when they know the system is not a person (Gambino et al. 2020). The important twist for candour is that the script can be *social-but-not-socially-risky*: respondents may treat the AI as a conversational partner while simultaneously perceiving reduced judgement, reduced gossip risk, and reduced reputational spillovers.

Greater control over pacing and phrasing

Many AI interview modalities (especially text chat) give respondents more control: time to think, revise, and choose precise wording. That can reduce anxiety and improve articulation, which can

look like “candour” because the answer is clearer and less hedged. (This can also enable strategic responding—see risks below.)

Psychological distance and “as-if anonymity”

Even when the system is not truly anonymous, the *felt* experience can be closer to anonymous disclosure: no facial expressions, fewer micro-judgements, and less immediate social feedback. That psychological distance can be particularly salient when the alternative is a high-status interviewer or a context with strong power asymmetries (e.g., workplace, clinic, immigration, legal).

Differential effects (who becomes more candid, when)

Topic sensitivity matters

The effect is strongest when questions touch stigma, shame, taboo behaviours, or norm violations (e.g., mental health, sexual behaviour, substance use, unethical conduct, politically sensitive views). If the question is neutral and low-stakes, you should not expect large candour gains—people already answer candidly.

Trust, literacy, and perceived surveillance can flip the effect

Candour can *decrease* if respondents suspect monitoring, future consequences, or data misuse (“this is going on my record”). In those cases, an AI interviewer can feel like a **surveillance interface** rather than a safe partner, and respondents may become more guarded than they would with a trusted human.

The interface design changes the psychology

More humanlike embodiments (voice, face, avatar) can increase rapport, but may also increase evaluation pressure. Conversely, a plain text interface can reduce social pressure, but might reduce emotional safety for some people. CASA-style responses can be triggered by subtle cues, so small design changes can shift disclosure dynamics (Gambino et al. 2020).

Population differences are real (and can become bias)

People differ in comfort with technology, prior experiences with institutions, cultural norms around authority, and baseline social anxiety. That means “AI increases candour” is a **distributional claim**, not a universal one: you can get systematic differences in who discloses what, which then affects your dataset.

Risks and things to think about

Over-disclosure and informed consent

If respondents feel unusually safe, they may disclose more than they later feel comfortable with. Make consent and data-handling salient at the moment of disclosure (not just in a long preamble), and be explicit about whether a human will read the transcript.

Strategic responding and “gaming” the system

Lower social desirability pressure does not imply higher truthfulness. Some respondents may optimize for what they think the algorithm rewards (or what they think the organization wants). The hiring study is a useful reminder: reduced social desirability bias can coexist with weaker predictive validity and new failure modes (Dukanovic & Krpan 2025).

AI introduces its own interviewer effects

AI systems can inadvertently lead respondents (tone, follow-up choices, perceived empathy), and can behave inconsistently across demographic groups or writing styles. Even if the interview is consistent, downstream AI-assisted analysis can introduce non-random errors: work on LLMs for qualitative analysis shows serious risks of biased annotation and misleading inferences when errors correlate with respondent characteristics (Ashwin et al. 2023).

Safety boundaries

If the topic includes trauma, self-harm, abuse, or crisis, an “always-on, friendly” conversational agent can create ethical and duty-of-care hazards. Decide up front what the AI should and should not do in those scenarios, and ensure participants are not misled about the system’s capabilities.

We do not recommend using any AI interviewer for employment interviews or in a clinical setting without human oversight and significant safeguards.

Practical takeaway

Treat “AI increases candour” as a **design variable**: you can dial it up or down via framing (“AI” vs “automated”), embodiment, reminders about audience (private vs reviewed), and question style.

The key is to choose the level of candour you want, *and* to document the conditions so you can interpret the data responsibly.

1. my comment [↩](#)

References

Ashwin, Chhabra, & Rao (2023). *Using Large Language Models for Qualitative Analysis Can Introduce Serious Bias*. <https://doi.org/10.48550/arXiv.2309.17147>.

Dukanovic, & Krpan (2025). *Comparing Chatbots to Psychometric Tests in Hiring: Reduced Social Desirability Bias, but Lower Predictive Validity*. *Frontiers*.
<https://doi.org/10.3389/fpsyg.2025.1564979>.

Gambino, Fox, & Ratan (2020). *Building a Stronger CASA: Extending the Computers Are Social Actors Paradigm*. *Communication and Social Robotics Labs*.
<https://doi.org/10.3316/INFORMIT.097034846749023>.

Lucas, Gratch, King, & Morency (2014). *It's Only a Computer: Virtual Humans Increase Willingness to Disclose*. <https://doi.org/10.1016/j.chb.2014.04.043>.