Causal Mapping: 97 ideas

CONTENTS

index

💐 Causal Mapping: a Garden of Ideas#

> 🌸 Starting to use the Causal Map app and want to know more about the theory and practice of causal mapping? What is it, what is it good for, how can you do it?

🌻 You've skimmed through a couple of our publications but want to see how it all fits together?

🪻 Here you'll find dozens of one-pagers setting out the key ideas in causal mapping as we see it, curated and assembled from existing publications and blog posts.

The titles of most of these "ideas" pages are expressed in the form of a single thought or claim or idea, like "Causal coding is easy to automate".

This Garden will also serve as a background companion to the Causal Map app. The app has its help which is also integrated into the app itself. Those help pages are also available as the last chapter of this Garden. That chapter is all about how to use the app whereas the rest of this Garden is a more discursive look at the theory and everything that surrounds causal mapping.

We have also included some of our most popular LinkedIn posts, especially on AI.

> ⚒️ This site is very much a work in progress! A lot more ideas are still to come!

Glossary

Some essential terminology for causal mapping.

Term	**Definition**
Causal claim	A section of text where someone or something says or claims that one thing (a cause) influences or influenced another (an effect). There need be no suggestion that the cause is the only cause of the effect, or totally determines it.
Causal link	A causal link is the representation of a causal claim within a causal mapping system or database. Each link consists of a cause, an effect, a quote or other evidence to back up the claim, and a source, for example an interview transcript.
Coding	The process of capturing a causal claim in a text as a causal link in a database or map.
Bundle	Often there will be many links with the same causes and effects. We call this set of links a *bundle* of links. Usually these links are mentioned by multiple sources. In a causal map, we usually display all the links in a bundle as one arrow. A thicker line or a label can be used to show the number of links in the bundle.
Source	Sources are where your causal claims come from: respondents you have interviewed or documents you have collected. Sources are usually given an ID like, say, `Source_001`.
Cause	A causal factor at the beginning of a link, that affects another factor.
Influence factor	See *cause*. Sometimes we prefer "Influence factor" over "Cause" to make it clear that we may be talking only about a partial contribution to an effect, not complete determination.
Effect	A causal factor at the end of a link, that has been influenced by another factor.
Consequence factor	See Effect. Sometimes we prefer "Consequence factor" over "Effect" to make it clear that we may be talking only about a partial contribution to an effect, not complete determination.
Factor	A cause or an effect; something that can influence or be influenced by other factors. We use the word both for the text label within the causal mapping representation as well as for the thing in the world to which it refers.
Codebook	A list of factor labels that can be used when coding causal links and/or other instructions for coding.
Causal map	A set of causal links between a set of factors, usually interconnected, such that a link from X to Y means that someone says or claims that X influences or influenced Y. We can think of a causal map as a visual representation or more abstractly just as a table or database of links which can be visualised.
Self-loop	A causal link from one factor to itself.
Column	Data associated with a link or a source: some columns are essential, e.g. a link must have a cause and an effect, and some are optional.
Field	See *column*

Names of tables and fields ▸

Intro

What is causal mapping? What are its strengths and weaknesses? How does a causal map differ from a systems diagram? This chapter has some answers.

🗺️ What is Causal Mapping?#

Causal mapping is a technique to **visualise what people believe causes what** within a complex system. It creates a "mental map" of the cause-and-effect relationships perceived by an individual or a group.

The process starts with **narratives**—such as interview transcripts, reports, or open-ended survey responses. Causal claims within these texts are systematically identified and structured into a network diagram:

- **Nodes (Boxes)** represent the *factors* or *concepts* (e.g., "Better Training").
- **Links (Arrows)** show the *direction of influence* between them.

🛠️ The Causal Map App#

The specialised **Causal Map app** provides a convenient way to do causal mapping. Users can import interviews or reports and "code" them: highlighting causal claims and adding them to the database. Much of this process can optionally be automated using AI, enabling rigorous analysis of larger datasets.

- **Transparency:** Every link in the map is transparently tied back to the **original source quote**. This ensures that outputs are verifiable and avoids acting as a "black box," maintaining the rigour essential for qualitative work.
- **Querying the Map:** The final map is a dynamic model of **causal evidence** that can be actively explored to answer sophisticated questions, such as tracing all direct and indirect links from a single input to a defined outcome.
- **AI as an Assistant:** Generative AI is optionally used as a **tireless, low-level coding assistant** to quickly extract explicit causal claims from text.

A causal map consists of multiple links where a link from X to Y means someone believes X influences Y

A **causal map** can be defined as a network consisting of links or arcs between nodes or factors, such that a link from factor C to factor E means that someone (P) believes that C in some sense causally influences E. Every link represents one causal claim.

Alternatively we can say that such a link means that there is some *evidence* (P) that C in some sense causally influences E (seeWe can reason about causal maps using a logic of evidence).

- Causal maps encode a belief about a *usually partial* causal *influence* of C on E, and only in special cases encode *total or exclusive* causation such that C entirely determines E or is *the* cause of E.
- Encoding a claim (like 'the heavy rains were one reason the harvest was worse than usual') in causal mapping does not require us to make any judgement about the quality of the evidence or the ability of the source to correctly judge that this link was causal (although we can add this information if we want).

Causal mapping helps make sense of many causal claims from many sources

> From Better Evaluation

Causal mapping helps make sense of the causal claims (about "what causes what") that people make in interviews, conversations, and documents. This data is coded, combined, and displayed in the form of maps. These maps show individuals' and groups' mental models and can support further investigation of causal connections.

Top: two separate single-source causal maps: Beliefs held by sources P and Q about causal influences between various factors. Bottom: A combined, multi-source causal map about beliefs held by sources P and Q about causal influences between various factors

Causal mapping is designed for the analysis and visualisation of qualitative data about causal links. It can be used to test an existing theory of change or create collective empirical theories of change about how a program works based on stakeholders' experiences.

People's narratives and reflections about their experiences provide qualitative data that can be coded and displayed as maps to present the cognitive structures (mental models) of individuals and groups and to support further exploration to understand actual causal connections.

These causal maps can help to answer questions about what people think happened and what they think caused this by building links between different factors, such as different kinds of outcomes and inputs. Mapping the chains of results and their linkages builds pictures of causal pathways showing the intermediate steps and connections between them.

Causal mapping starts from what people actually say and what they do not say

Causal mapping aims to directly understand and collate the causal claims which people make in narrative (and other) data rather than trying deduce causal connections using statistics or other methods. It starts with what people actually say in real-world contexts and does not rely on heavily pre-structured question formats. Urgent, unexpected, and unwelcome information is treated at face value.

The analyst does not need to have any preconceived conceptual framework; types of causal claims are identified in the data inductively and iteratively. This is a partly creative process, however the decisions made by the analyst are transparent as the underlying text is always available.

At least some of the boundaries of causal mapping research are set by the respondents, not the researchers; what are we going to talk about? What are we not going to talk about?

Causal mapping has been used for over 50 years in many disciplines

> From (Powell et al., 2024)

Causal mapping – diagramming beliefs about what causes what – has been used since the 1970s across a range of disciplines from management science to ecology.

The idea of wanting to understand the behaviour of actors in terms of their internal maps of the world can be traced back further to field theory (Tolman, 1948) which influenced Kelly's 'personal construct theory' (Kelly, 1955). A seminal contribution was made by Robert Axelrod in political science, with the book The Structure of Decision (Axelrod, 1976). Causal mapping is largely based on 'concept mapping' and 'cognitive mapping', and sometimes the three terms are used interchangeably, although 'causal mapping' strictly involves maps that only include explicit causal links, rather than, for example, relationships like 'membership'.3 Axelrod's book presents a comprehensive idiographic approach to how individuals make decisions which he himself mostly refers to as 'cognitive mapping' (although his definition makes it clear that all links are causal). An appendix to the book (Wrightson, 1976) gives details about how to code causal links. Bougon et al. (1977) applied a similar approach to a study of the Utrecht Jazz Orchestra as an organisational unit, eliciting 'cause maps' from several individual members and amalgamating them. One strand of literature about causal mapping can be located within the wider literature on sensemaking in organisations pioneered by Weick (1995), and applications within organisations were present almost from the start.

By 1990, there were many different applications of similar ideas, including an edited book (Huff, 1990) that offered a unitary approach to 'concept mapping' in the United States. Most authors (Ackermann and Alexander, 2016: 892; Clarkson and Hodgkinson, 2005: 319; Fiol and Huff, 1992: 268; Laukkanen, 2012: 2; Narayanan, 2005: 2) use a broadly similar definition of a causal map: A causal map is a diagram, or graphical structure, in which nodes (which we call factors) are joined by directed edges or arrows (which we call links), so that a link from factor C to factor E means that someone (P) believes that C in some sense causally influences E. There is a constructive ambiguity (Eden, 1992) about what a collective map is a map of: While maps constructed as a consensus within a group can plausibly be claimed to map 'what the group thinks', this is more problematic for maps constructed post hoc by synthesising individual maps.

We found no significant deviations from this basic definition of a causal map across all the variants of causal mapping reviewed in the following sections, with the caveat that there is variation in how explicit different authors are in describing causal links as representing bare causation as opposed to beliefs about causation.

In the following decades, Eden et al. (1992) applied the approach to understanding and supporting decision-making in organisations, increasingly using the phrase 'causal mapping' rather than 'cognitive mapping', and they subsequently extended the application of causal maps to fields as varied as risk elicitation and information systems development (Ackermann and Eden, 2011; Ackermann et al., 2014), also developing a series of software packages beginning with Decision Explorer (Ackermann et al., 1996). There is now a wealth of literature on using causal mapping for decision support in organisations (including sophisticated approaches to formalise decision support (Montibeller et al., 2008) and even to rank options (Rodrigues et al., 2017)).

Laukkanen (1994, 2012; Laukkanen and Eriksson, 2013) also wrote extensively on causal mapping and developed a software programme called CMAP3 for processing both idiographic and comparative causal maps by importing, combining and analysing factors and links attributed to one or more sources. A broadly similar approach was taken by Clarkson and Hodgkinson (2005) with their Cognizer approach and software.

References

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Do use causal mapping when you have large numbers of claims from multiple sources, and more open research questions

> From [Better Evaluation](#).

When to use causal mapping#

Causal mapping is useful when seeking to understand the causal pathways influencing the outcomes of programs operating in complex settings. It helps make sense of a program and its context in stakeholders' own words. This includes providing ways to make sense of and organise the different, but sometimes overlapping, labels that different groups use to describe the causal factors that are important to them.

Causal mapping can be used to help make sense of large amounts of qualitative data.

Using this method requires expertise in coding and analysis of qualitative data.

✅ So, use causal mapping if you…

- have a relatively large amount of narrative data
- need help to synthesise a large number of links
- have information from more than one source (for example respondents, documents)
- are interested in differences between the sources and groups of sources
- you don't know the contents or boundaries of the map
- want to capture what your sources actually say, systematically and transparently

Do not use causal mapping if you have limited data or want precise models or specific causal links

> From Better Evaluation.

Causal mapping is less frequently used to analyse quantitative data or to do precise mathematical modelling, e.g. of future states of a system under certain conditions.

❌ Do not use causal mapping if you …

- don't place high value on the views of the sources
- only have a relatively small map which you can manage with traditional tools for drawing network diagrams (e.g. PowerPoint, kumu.io etc.)
- need to analyse quantitative data and/or need to do precise mathematical modelling, e.g. of future states of a system under certain conditions
- would like to just sketch out a plan (e.g. Theory of Change or similar) without much reference to the different sources underpinning each link

Causal mapping approaches differ in application, construction, analysis and how they deal with multiple sources

Reference	Main application of causal mapping	Mode of construction	Dealing with multiple sources	Analysis procedures
(Axelrod, 1976)	Understand and critique decision making	Coding documents	Mainly idiographic	Compute polarity of indirect effects in some cases.
(Bougon et al., 1977)	Understand how organisations are constructed and can be influenced.	Semi-structured interview to identify a fixed list of factors aka 'variables'; respondents then say which are linked and give the polarity.	Compare individual maps and combine into global 'average' map.	Identify variables X with high outdegree and Y with high indegree and construct an 'etiograph' to show all the multiple paths from one point to another; discuss how respondents might have influence over some variables.
(Ackermann and Eden, 2004, 2011; Eden, 1992; Eden et al., 1979, 1992)	Decision support and problem solving in organisations. Maps are seen primarily as useful tools rather than research about reality.	Open interviewing of several respondents based on Kelly's Personal Construct Theory. Also map construction directly with groups (1988).	Comparing maps between individuals and analysing group maps directly.	Various structural measures, presence of isolated clusters, hierarchical trees, loops. Simplify individual maps by collapsing X->Y->Z into X->Z.
(Laukkanen, 1994, 2012; Laukkanen and Eriksson, 2013; Laukkanen and Wang, 2016)	Explicitly cognitive, to improve knowledge and understanding in management	Systematic comparative method with semi-structured interviewing: respondents are given anchor topic(s) then asked for causes, effects, causes of causes, effects of effects. Compress the data by standardising factor names. Comprehensive coverage of different map construction possibilities.	Comparative study of different individual maps, combining data into a database.	Display combined maps for subgroups, e.g. all local managers.

Causal mappers believe that humans are good at thinking in terms of causal nuggets

See also Causal mapping has been used for over 50 years in many disciplines

> From Powell, Copestake, et al. (2023

Renewed interest in causal mapping may also be reinforced by the ‘causal revolution’ in quantitative data science initiated by Judea Pearl (Pearl, 2000; Pearl and Mackenzie, 2018), which has fundamentally challenged the almost total taboo placed on making or assessing explicit causal claims, which was dominant in statistics for much of the twentieth century (Powell, 2018), and this has in turn helped rekindle interest in explicitly addressing causation using qualitative methods.

Causal mapping and most related approaches share the basic idea that causal knowledge – whether generalised or about a specific case or context – can be at least partially captured in small, relatively portable ‘nuggets’ of information (Powell, 2018: 52). These can be assembled into larger models of how things worked, or might work, in some cases. More ambitiously, they may contribute to constructing ‘middle-level theory’ theory, useful for understanding causal processes in other contexts, without necessarily reaching the level of overarching scientific laws (Cartwright, 2020). Causal nuggets are also related to the mechanisms that help to explain how people behave in different contexts (Pawson and Tilley, 1997; Schmitt, 2020). These can be thought of as causal schema and linked to the hypothesis that human knowledge is stored in chunks that are activated and combined with others in relevant circumstances. This would suggest that we humans do not have a comprehensive set of causal maps in our heads at any one time, but we do have a set of more basic components and the ability to assemble them when the situation calls for it, including when prompted by a researcher.

Causal mappers believe that humans are the best detectors of causation

See also Causal mapping has been used for over 50 years in many disciplines

Causal mappers believe that humans are good at thinking in terms of causal nuggets

> From Powell, Copestake, et al. (2023

We claim: our everyday causal understanding is as primary as our perception of, say, colour and arises from more than empirical observations of associations between objects or events; our ability to infer causation goes beyond and is not primarily based on noting correlations. And for all its complexity and intuitive brilliance, it is also just as fallible as our perception of colour or size.

This reaffirms our practice as evaluators of taking the causal claims and opinions of humans (experts and non-experts) seriously (Maxwell, 2004a, 2004b); indeed, this kind of information is the bread and butter of most evaluations.

We can thank Judea Pearl for promoting the insight that if you want to thrive in this world, you have to understand causality natively. We humans make causal connections from an early age. We wouldn't survive long if we didn't.

GPT-3.5 just about understood causation. GPT-4 and more recent models understand causal connections within text very well.

Our understanding of the world is drenched with causal understanding: information and hypotheses about how things work (mostly accurate enough, sometimes not). It's really hard for us to *not* think causally: the concept of correlation is much harder to understand than the concept of causation.

Causal Inference?#

Causal inference is the process of determining whether and how one event or variable brings about another.

Some writers mistakenly assume that only a controlled experiment can "really" provide a route to causal inference.

We would go so far as to say that we don't usually in any conscious cognitive sense *infer* causation -- we just see it, all the time, everywhere. We don't have to teach children to infer causation: we have to teach them to question their perceptions of causation and to distinguish causation from correlation.

Why are humans the best detectors of causation?#

1. **Evolutionary Adaptation:**
   Human brains have evolved specifically to detect and act on causal relationships. Survival depends on recognizing which actions lead to which outcomes—e.g., which plants are safe to eat, which animals are dangerous, and how to use tools. This evolutionary pressure has made causal reasoning a core part of human cognition.

2. **Intuitive Causal Models:**
   From infancy, humans build mental models of the world that are fundamentally causal. Children naturally ask "why" questions and seek explanations, not just associations.

3. **Generalization and Flexibility:**
   Humans can generalize causal knowledge across domains. For example, understanding that pushing causes movement can be applied to objects, social situations, and abstract concepts. This flexibility allows humans to detect causation even in novel or ambiguous situations.

4. **Counterfactual Reasoning:**
   Humans often engage in counterfactual thinking—imagining what would happen if things were different. This is a hallmark of causal reasoning and is essential for planning, learning from mistakes, and scientific discovery.

5. **Distinguishing Correlation from Causation:**
   While humans sometimes make errors (e.g., seeing causation where there is only correlation), we are still very good at using context, background knowledge, and intervention to distinguish true causal relationships from coincidence.

6. **Social and Cultural Transmission:**
   Human societies accumulate and transmit causal knowledge across generations through language, stories, and education. This collective causal understanding is a foundation of science, technology, and culture.

7. **Observation and Pattern Recognition:**
   Humans are adept at noticing regularities and anomalies in their environment. We naturally look for patterns—such as temporal precedence (A happens before B), co-occurrence, and changes following interventions—that suggest causal relationships. Even without formal training, people intuitively apply principles like "no effect without a cause" and "causes precede effects."

8. **Intervention and Experimentation:**
   Humans frequently test their causal hypotheses by intervening in the world—changing variables and observing outcomes. This hands-on experimentation, whether in childhood play or scientific research, is a powerful tool for distinguishing causation from mere correlation.

9. **Use of Multiple Sources of Evidence:**
   Humans combine different types of evidence—temporal order, statistical regularities, mechanistic explanations, and observed interventions—to make robust causal inferences. We can weigh conflicting evidence, consider alternative explanations, and update our beliefs as new information arises.

In sum, humans are not just passive recipients of causal information; we are active causal detectives, constantly inferring, testing, and refining our understanding of how the world works. This multifaceted approach to causal inference is what makes us the best detectors of causation.

Conclusion#

Causal reasoning is not just a feature of human cognition—it is its backbone. Our ability to detect, infer, and act on causal relationships is what allows us to navigate, survive, and thrive in a complex world. While formalised, controlled experiments are an incredible tool for causal inference in very particular contexts such as some areas of education and economics where multiple very similar causes are regularly followed by multiple very similar effects, if we had to stick to that kind of knowledge and that kind of context we would never be able to get out of bed in the morning, let alone get the kids to school.

Humans remain the best detectors of causation, both individually and collectively.

See also: 400 realist causation

Causal mapping is part of the qualitative branch of the new causal revolution

See also Causal mapping has been used for over 50 years in many disciplines

> From Powell, Copestake, et al. (2023
>
> Renewed interest in causal mapping may also be reinforced by the ‘causal revolution’ in quantitative data science initiated by Judea Pearl (Pearl, 2000; Pearl and Mackenzie, 2018), which has fundamentally challenged the almost total taboo placed on making or assessing explicit causal claims, which was dominant in statistics for much of the twentieth century (Powell, 2018), and this has in turn helped rekindle interest in explicitly addressing causation using qualitative methods.

Causal mapping differs from related approaches - epistemic, less predictive, unsophisticated, many links, many sources, unclear boundaries

> From (Powell et al., 2024)

Most evaluators are probably more familiar with related approaches under the term 'systems mapping', recently covered by Barbrook-Johnson and Penn (2022). They provide an overview table of relevant methods on pp. 169 ff. – fuzzy cognitive maps (FCM), participatory systems mapping (PSM), Bayesian belief networks (BBN), causal loop diagramming (CLD), systems dynamics (SD) and theory of change (ToC) – which will be briefly mentioned here.

SD, CLDs, FCMs and BBNs are all ways to encode information about networks of interconnected causal links and follow formal inference rules to make deductions based on them, for example, to calculate the strength of indirect effects or to predict behaviour over time. The oldest of the three methods, SD (Forrester, 1971), models flows of a substance (for example, of energy or money) within a network over time, whereas the other three methods model 'bare' causal connections between network elements. SD uses general mathematical functions to model the connections and explicitly models non-linear relationships. CLDs are related but mathematically simpler, modelling causal effects in a semi-quantitative way. FCMs might seem to be of more interest for causal mapping; Kosko's original article on FCM (Kosko, 1986) takes Axelrod's work as its starting point. This tradition (Chaib-Draa and Desharnais, 1998; Khan and Quaddus, 2004; Taber, 1991) was originally introduced to model causal reasoning (Kosko, 1986: 65): If person or group P believes the set of causal propositions making up a map M, the model attempts to predict the strength with which they could or should also believe some other propositions, for example, about indirect effects and how they might change over time. In practice, however, FCM is less interested in cognition than in making predictions about the world. The difference between FCM and the other three methods is more about the fuzzy logic used to make the predictions rather than about the cognitive nature of the data.

BBNs are also designed to make causal inferences by doing calculations with data about causal connections. While FCMs make essentially qualitative predictions such as 'increasing' and 'decreasing', BBNs use directed acyclic graphs (networks without loops) to make quantitative predictions about the probability of events, particularly about the probability that one event was the cause of another.

All four approaches are primarily ways to make predictions about causal effects within a network of factors, and (despite the words 'cognitive' and 'belief' in the names of two of the four) the relative lack of interest in who is doing the reasoning sets FCM, BBNs and SD apart from causal mapping as outlined earlier.

In the last few years, PSM has featured in several publications in evaluation journals and guides (Barbrook-Johnson and Penn, 2021; Hayward et al., 2020; Sedlacko et al., 2014; Wilkinson et al., 2021), alongside mapping of 'systems effects' (Craven, 2020). Indeed, Craven's work (see also Craven, 2017) can be considered causal mapping with a particular emphasis on systems aspects. Barbrook-Johnson and Penn (2022) explicitly exclude causal maps from their overview of systems mapping because they are arguably included via FCM and because they 'sometimes emphasise developing representations of individual mental models rather than representations of systems' (p. 11). Nevertheless, PSM is closer to the tradition of causal mapping (and of more direct interest to evaluators) than the previous four approaches because it is a more concrete and pragmatic intervention to construct a map with specific group of stakeholders to support decisions. A devotee of causal mapping could claim that approaches like PSM are just variants of what they have been doing for the last 50 years, just as a devotee of systems mapping might consider causal mapping as a form of PSM.

Finally, logic models and ToC can be considered causal maps in which they make assertions about past or future causal links that one or more stakeholders believe to be important. They are also political artefacts that aim to justify and inform action by establishing an agreed synthesis of multiple perceptions of change and may also gain legitimacy by being the product of an agreed process of participatory planning and co-design. They do not, however, normally retain information about which stakeholder(s) believe which claim. Reflecting on logic models and theories of change provides one entry point for thinking more carefully both about who actually makes these claims and about the symbols and rules employed to construct them (Davies, 2018).

We think it is useful to distinguish this tradition of causal mapping from related activities in six ways, as set out in the following section. None of these distinctions are definitive, and many are shared with other approaches. To systems people who want to say that causal mapping is just systems mapping and to causal mappers who want to say that systems mapping is just causal mapping (and we have heard both arguments many times), we can only say, perhaps we should all just get to know each other first.

First, the raw material for causal maps comprises claims about, perceptions of or evidence for causal links. Causal maps are primarily epistemic, meaning that their constituent parts are about beliefs or evidence, not facts; yet their logic tends to be parallel to, and based upon, the logic of non-epistemic systems maps and similar diagrams that are broadly used across a range of sciences. Some systems mapping techniques are also sometimes concerned with stakeholder beliefs; causal mapping does this more systematically.

Second, causal maps tend to be unsophisticated about the types of causal connection they encode. To explain this, we should note that causal claims in ordinary language are expressed in an endless variety of ways: 'C made E happen', 'C influenced E', 'C may have been necessary for E', 'C was one factor blocking E', 'C had a detrimental effect on E', 'C had a surprisingly small effect on E' and so on. With a few exceptions, causal mapping analysts do not even try to formally encode this rich and unsystematic range of causal nuance, relying instead simply on the lowest common denominator: A link from X to Y means simply that someone claims that X somehow causally influences or influenced Y.

There is one exception: Many causal mapping approaches do accommodate information about the polarity of links, marking each link as either positive or negative, for example, the claim 'the recession led to unemployment' could be coded as a negative link from 'the recession' to 'employment'.

In general, causal maps usually encode a belief about partial causal influences of C on E and only in special cases do they encode total or exclusive causation such that C entirely determines E. This also means that encoding a claim does not require us to make any judgement about the quality of the evidence or the ability of the source to judge that this link was causal (although it may be very useful to do so).

Third, causal mapping often handles **large numbers** of causal claims, sometimes many thousands. Handling large numbers of claims en masse in this way is made much easier because of the relatively unsophisticated nature of the way claims are coded (as discussed earlier).

Related approaches in evaluation tend to bring more sophisticated tools to bear on a much smaller number of causal links. In process tracing, for example, researchers may produce diagrams depicting claims about causal links but tend to focus on testing the strength of a relatively small number of specific 'high-stakes' causal links, whether through verbal reasoning, application of Boolean logic or Bayesian updating (Befani and Stedman-Bryce, 2017).

Fourth, causal maps may originate from one or many sources, each reporting on one or many cases. In a causal map, the links all originate from one person or document a 'single-source' or 'individual' or 'idiographic' causal map, as in Axelrod's original work (Axelrod, 1976). But we can also draw causal maps that incorporate information from a variety of different sources, as illustrated in Figure 1.

The simplest causal maps refer to only one context and contain information from only one source (which may be the consensus view of several people, treated as speaking with a single voice). Various forms of systems mapping such as PSM could be understood as a special case of causal mapping in this sense.

There are many other variants. One source might give differentiated information about different cases or contexts, or many sources might give information about just one context, as when different water systems experts each give their (possibly differing) opinion about the same water catchment area, for example.

Another frequent type of causal map is drawn from many sources, each reporting on their own situation or context, such as their perception of drivers of change in their own lives. In coding and analysis of this sort of data, one source equals one case and one context; these can subsequently be aggregated across many sources who, for example, all share a similar context.

Fifth, causal maps do not necessarily specify a clear system boundary. The boundaries of a causal map are usually defined more loosely, partly by data collection but also by the sources themselves. Indeed, some systems proponents would say that the term 'systems diagram' simply signals a readiness to use systems approaches (Williams, 2022).

Finally, causal mapping, especially in management sciences and operations research, has nearly always been at least as interested in process as in the result. There is often a focus on the process of reaching consensus as part of the task of solving a business problem, rather than on the universal accuracy or validity of the final map.

References

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Causal mapping has three tasks -- gathering, coding and analysing data

Different approaches to these three tasks are discussed in turn in the following sections.

- Task 1 -- Gathering narrative data
- Task 2 -- Coding causal claims as causal qualitative data analysis
- Task 3 -- Analysing data, Answering questions

Task 1 -- Gathering narrative data

> From (Powell et al., 2024)

How to collect causal claims from which to draw causal maps?

(There is also a whole chapter about this task: Task 1 -- Introduction)

There are a wide variety of options, including in-depth individual interviews (Ackermann and Eden, 2004), reuse of open-ended questions in structured surveys (Jackson and Trochim, 2002), literature reviews (in which 'sources' can be documents rather than individuals) and archival or secondary material within which pre-existing causal claims are already made (Copestake, 2020). Other approaches aim to build consensus by using structured collaborative processes, including Delphi studies and PSM (Penn and Barbrook-Johnson, 2019). Guidelines for causal mapping may include procedures for collecting primary data, with forms of elicitation including back-chaining ('what influenced what?') and forward-chaining (what resulted, or could result, from this?)

With primary data collection, we can distinguish between relatively closed and open approaches and whether respondents are forced to choose between pre-selected optional answers or can formulate their own (see Table 2). Interviewers may also be guided by a chaining algorithm; for example, they may be instructed to iteratively ask questions like 'You mentioned X, please could you tell me what were the main factors that influenced X or led to it happening.'

Table 2. Different approaches within primary data collection for causal mapping, with example questions.

Admissible answers / Scope of questions	**Explicit: factors are explicitly identified**	**Implicit: factors are not explicitly named**
Closed: questions with a predetermined focus	Which factors in this list influenced this particular event?	What influenced this particular event?
Open: a freer discussion	Identify the biggest change you experienced in relation to X, and list three factors that influenced it	Tell me what has changed for you in the last x years

References

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Task 2 -- Coding causal claims as causal qualitative data analysis

> From (Powell et al., 2024)

Some approaches such as that suggested by Markiczy and Goldberg (1995) directly elicit causal links from their sources, perhaps by asking respondents to suggest causal links between a predetermined list of causal factors, and thus, after finishing Task 1, are already in a position to create causal maps.

(See also the dedicated chapters on coding...)

- Task 2 -- Introduction
- Tasks 2 & 3 -- Introduction

More explicitly, qualitative approaches are faced with Task 2: encoding causal claims in the form of explicit causal links and factors. This task is similar to ordinary qualitative data analysis (QDA), whether done manually or using tools like NVivo, Dedoose and AtlasTI. However, these tools are designed to capture general concepts, rather than claimed causal links between concepts, which is what we need for causal mapping. QDA for causal mapping also starts with a corpus of narrative data, but it does not create causal links between independent concepts that might already have been coded using ordinary non-causal thematic analyses. Rather, in causal QDA, the primary act of coding is to highlight a specific quote from within a statement and identify the causal claim made by simultaneously identifying a pair of causal factors: an 'influence factor' and a 'consequence factor'.

The causal factors only exist as one or other end of a causal link and have no meaning on their own. Each claim forms a link in the visual representation of the causal map. The Axelrod school had its own coding manual describing how to highlight areas of text expressing causal connections and code them as links between causal factors, originally inspired by evaluative assertion analysis (Osgood et al., 1956).

Manual causal coding of text data, like ordinary thematic coding, requires a considerable investment of time and expertise to do well. We now use natural language processing to at least partially automate this; however, the process is essentially the same, and discussion of this is beyond the scope of the present article.

Where do the labels for the causal factors come from? As with ordinary QDA and thematic analysis (Braun and Clarke, 2006), approaches vary in the extent to which they are purely exploratory or seek to confirm prior theory (Copestake, 2014). Exploratory coding entails trying to identify different causal claims embedded in what people say, creating factor labels inductively and iteratively from the narrative data. Different respondents will not, of course, always use precisely the same phrases, and it is a creative challenge to create and curate this list of causal factors. For example, if Alice says 'Feeling good about the future is one thing that increases your wellbeing', is this element 'Feeling good about the future' the same as 'Being confident about tomorrow' which Bob mentioned earlier? Should we encode them both as the same thing, and if so, what shall we call it? We might choose 'Positive view of future', but how well does this cover both cases? Laukkanen (1994) discusses strategies for finding common vocabularies. As in ordinary QDA, analysts will usually find themselves generating an ever-growing list of factors and will need to continually consider how to consolidate it – sometimes using strategies such as hierarchical coding or 'nesting' factors (as discussed in the following section).

The alternative to exploratory coding is confirmatory coding, which employs an agreed code book, derived from a ToC and/or from prior studies. QuIP studies mostly use exploratory coding but sometimes supplement labels with additional codes derived from a project's ToC, for example, 'attribution coding' helps to signify which factors explicitly refer to a specific intervention being evaluated (Copestake et al., 2019b: 257). However, careful sequencing matters here because pre-set codes may frame or bias how the coder sees the data (Copestake et al., 2019a). Again, the positionality of the coder matters just as much when doing causal coding as it does for any other form of qualitative data coding.

Combining Tasks 1 and 2

Tasks 1 and 2 result in a coded data set of causal claims, each of which consists of (at the very least) the labels for a pair of causal factors. Those using a more explicit elicitation approach have been able to skip Task 1.

References

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Task 3 -- Analysing data, Answering questions

> From (Powell et al., 2024)

The extensive causal mapping literature provides many examples of its use to answer evaluation questions (see Powell, Copestake, et al., 2023, p. 110), for example:

- Getting an overview of respondents' "causal landscape". This can be useful for orientation or for particular tasks like triaging masses of information to identify key outcomes and possible causal pathways when planning an Outcome Harvesting (Wilson-Grau & Britt, 2012) or Process Tracing (Befani & Stedman-Bryce, 2017) project.
- Weighing up evidence about contribution: in particular, tracing back and comparing the possibly multiple contributory causes of an important outcome or consequence (Goertz & Mahoney, 2006), or examining effects of causes.
- Reporting key metrics of the causal network, for example, to reveal which factors are most central in the whole network or to identify feedback loops.
- Asking whether the empirical ToC matches the plan (Powell, Larquemin, et al., 2023, p. 7).
- Making comparisons between groups or across timepoints.

One way to simplify is to derive from the global map several smaller maps that focus on different features of the data. For example, maps may selectively forward-chain the multiple consequences of a single cause – including those activities being evaluated: effects of causes (Goertz and Mahoney, 2006) – or trace back to the multiple contributory causes of an anticipated or highly valued outcome or consequence: causes of effects. A series of simpler causal maps, each selected transparently to address a specific question, generally adds more value to an evaluation than a complicated, if comprehensive, single map that is hard to interpret. The downside of this is that selectivity in what is mapped and is not mapped from a single database opens up the possibility of deliberate bias in selection, including omitting to show negative stories.

!

Sets of individual links with the same influence and consequence factor (co-terminal links) are usually represented bundled together as a single line, often with thickness of the line indicating the number of citations, and/or with a label showing the number of links in the bundle. The map has not fundamentally changed, but the visualisation is much simpler.

Simplification - factor and link frequency ▸

Another way to simplify a global causal map is to produce an overview map showing only the most frequently mentioned factors and/or links. Care should be taken if this leads to omitting potentially important but infrequently mentioned evidence about, for example, an unintended consequence of an intervention.

Simplification - hierarchical zooming ▸

Another common way to simplify is to combine sets of very similar factors into one. For example, if hierarchical coding has been used, it is possible (with caveats) to 'roll up' lower-level factors (such as health behaviour; hand washing and health behaviour; boiling water) into their higher-level parents (health behaviour), rerouting links to and from the lower-level factors to the parent (Bana e Costa et al., 1999).

Reporting global and local network statistics ▸

Large causal maps can also be analysed quantitatively, including by tabulating which factors are mentioned most often, identifying which are most centrally connected or calculating indicators of overall map density, such as the ratio of links to factors (Klintwall et al., 2023; Nadkarni and Narayanan, 2005). We are wary of the value of summarising maps in this way, not least because results are highly sensitive to the granularity of coding. For example, although a specific factor such as 'improved health' might have been mentioned most often, if two subsidiary factors had been used instead (such as 'improved child health' and 'improved adult health'), these two separate factors would not have scored so highly.

References

Befani, & Stedman-Bryce (2017). *Process Tracing and Bayesian Updating for Impact Evaluation*. http://dx.doi.org/10.1177/1356389016654584.

Goertz, & Mahoney (2006). *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*. Princeton University Press. 12345.

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Wilson-Grau, & Britt (2012). *Outcome Harvesting*.

Strong evidence for a link is not evidence of a strong link

Never confuse the two.

TODO

[Our approach clearly distinguishes evidence from facts and does not automatically warrant causal inferences]

Causal mapping is easier if we are realist about causation

Causal mapping is easier if we are realist about causation. We can say that narrative accounts are full of claims about causal powers, that X had the power to affect Y, and X did exercise that power and Y was affected (perhaps in this particular case in spite of or with the assistance of other things).

Causal realism invites us to say that things **have the causal power** to affect other things.

The weird thing is that most physical and natural scientists think about causation in a realist way, but in the social sciences we tell ourselves not to because it isn't scientific (!).

Causal mapping is good at coping with messiness and complexity

> From (Powell et al., 2024)

... recognising head-on the ambiguity of much narrative causal data, particularly when confronted with large bodies of data collected in disparate ways. Evaluators must contend with messiness: imprecise system boundaries, differing specification of claimed causal influences and lack of clear or consistent information about what case or group of cases claims refer to. Causal mapping can contend with all this ambiguity rather than shying away from it. It can make use of messy operational data, treating urgent, unexpected and unstructured information at face value. This is made possible by distinguishing clearly between two analytical steps in evaluation: The first is to gather, understand and assemble causal evidence from different sources (those in a position to have useful evidence about relevant causal links and chains) to construct, compare and contrast the evidence for and against different possible causal pathways. By focusing on this task, causal mapping lays a more reliable foundation for the second, often critical, task of using the assembled data to make judgements about what is in fact really happening. This avoids the confusion and ambiguity that often arises when evaluators seek to address both steps simultaneously by constraining what data are collected to fit a prior view of reality which other stakeholders may or may not share.

References

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Granularity, generalisability and chunking are coding problems for causal mapping too

> From (Powell et al., 2024)

An illustrative example

!

A positive feature of causal maps, illustrated by the Figure, is that they capture a lot of information in a way that is quick and easy to understand. This example reveals that Source S provided a narrative that connects the intervention to improved feeling of wellbeing as a direct consequence of taking more exercise and via the effect of this on their health. This source also suggests a positive feedback loop, with more exercise making them more physically fit and encouraging even more exercise. The information from Source T is more fragmented; there are two causal statements claiming that improved feeling of wellbeing can result from more exercise and improved health, although T does not link the two causally, nor make any causal link back to the intervention. In addition, T suggests that an additional factor, 'more confidence in the future', also contributes to improved feeling of wellbeing. The two sources of evidence do agree on certain points; there is scope for generalisation beyond either individual source (and can be scaled up from here), both in assessing the multiple outcomes of the intervention and in understanding what explains improved feeling of wellbeing. Generalisability is strengthened when a link is reported by different sources in different contexts. We believe that within causal mapping, we should never make the mistake of thinking that stronger evidence for a causal link is evidence that the causal link is strong; only that there is more evidence for it.

Strong evidence for a link is not evidence of a strong link ▸

. The example also reveals some potential weaknesses of causal maps. First, there is ambiguity about the precise meaning of the labels and the extent to which their use is conceptually equivalent between the two sources. There is also ambiguity about whether they are referring to their own personal experience (and if so, over what period) or speaking in more general terms. Furthermore, the diagram sacrifices details, including how the statements shown relate to the wider context within which each source is situated. To mitigate this, an important feature of any causal mapping procedure is how easily it permits the user to trace back from the diagram to the underlying transcripts and key information about the source (e.g. gender, age, location etc.). Where this is possible, the diagram can be regarded in part as an index or contents page – an efficient route to searching the full database to pull out all the data relating to a specific factor or causal link, in order to validate any conclusions we draw. In particular, we recommend as a technique to mitigate this danger.

The transitivity trap ▸

References

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Consensus versus multi-source data collection#

At Causal Map we are relatively agnostic about data collection. We are most interested in causal evidence and beliefs derived from different sources.

See also:

- Task 1 -- Gathering narrative data
- Intro to data collection with Qualia

You can also create consensus causal maps directly

Consensus versus multi-source data collection#

Some methods related to causal mapping and even some forms of causal mapping itself are not interested in individuals' different causal perspectives: they are primarily aimed at reaching a consensus, expert map straight off without first recording individual viewpoints and attempting to combine them.

TODO

Direct links from intervention to outcome are real links too

You have a causal map with lots of links from an intervention to a final outcome. It's a really impressive chain.

Then a stakeholder comes along and draws a direct link from intervention to outcome, saying "these trainings are great, they caused the law to get changed!"

You can also create consensus causal maps directly, and in these cases these cheeky direct links are often ignored or curated out of the way. When coding individual sources, and source X says `A --> B --> C` and then and source Y says `A --> C` it is very tempting to recode source X as if they were really saying `A --> B --> C`.

For example when source X says

> Thanks to the training, our skills increased and so we were able to manage the outreach work better

.... and source Y only says

> Thanks to the training we were able to manage the outreach work better

... it is very tempting to add the intermediate factor to source Y's story too *even if they did not actually mention it*. If you're going to do that, you should document when and how this is allowed and why.

Causal backchaining#

We have often processed data gathered specifically for evaluation purposes using "causal back-chaining". But we often process secondary data which was not specifically intended for causal coding.

Changes and states#

When gathering primary data, the way in which questions are asked influences the meaning of the maps and their links. For example, in the QuIP, (Copestake et al., 2019b) respondents are asked to identify causes of changes, then causes of the causes and so on. This means that most of the factors are already as *changes in something*, such as 'an improved harvest' or 'reduced hunger'). This has implications for how positive and negative statements are combined, as discussed later.

Task 2 -- Introduction

> Standing on the shoulders of giants

In this chapter we present some of key general principles about how to do causal mapping which we at Causal Map Ltd (and, most of the time, at BathSDR) have adopted.

This is a very restricted yet powerful **minimalist** approach which we have also called "barefoot" or "naïve" coding.

In the next chapter Tasks 2 & 3 -- Introduction we look at specific conventions to make causal coding simple and powerful.

Minimalist coding does not help with blockers and enablers

Thanks for inviting me to the discussion, James. **I'll start by describing the "minimalist" approach** to coding causal statements used for QuIP and developed originally by James, Fiona and colleagues at BathSDR and developed and further formalised at Causal Map Ltd in collaboration with BathSDR. This formalisation lives inside the [Causal Map app](). Then I will try to **answer the question of whether it can help us deal with more complicated constructions like enabling and blocking** and whether this could help us with mid-range theory. As an appendix I'll add a more detailed overview of minimalist causal coding.

The minimalist approach is notable because it is based in our **joint experience of coding thousands and thousands of stakeholder interviews and other data such as project reports**, mostly from international development and related sectors, as well as coding hundreds of thousands of pages with AI-assisted coding. These have nearly always involved **multiple sources talking about at least partially overlapping subject matter**. So this coding produces individual causal maps for each source, which can then be combined in various ways -- rather than constructing single-source maps of expert thinking (Axelrod, 1976) or the collective construction of a consensus map (Barbrook-Johnson & Penn, 2022).

Our experience has been that the vast majority of causal claims in these kinds of texts are easily and satisfactorily coded in the simplest possible form "X causally influenced Y". Explicit invocation of concepts like enabling/blocking, or necessary and/or sufficient conditions, or linear or even non-linear functions, or packages of causes, or even the strength of a link, are relatively rare. The causes and effects are not conceived of as variables, the causal link is undifferentiated, without even polarity, and if any counterfactual is implied it remains very unclear.

This approach is what we call **"Minimalist" or "Barefoot" Coding**.

So what? Can we use minimalist coding to code and make deductions about, say, enablers and blockers?

Using more sophisticated, non-minimalist coding such as DAGs or fuzzy cognitive maps or whatever allows one to code linear or even non-linear causal influences of single or even multiple causes on their effects. One can do the "coding" by simply writing down (using appropriate special syntax) the connections, because one is an expert, and/or one can verify such statements statistically on the basis of observational data. Thus armed, one can make predictions or have sophisticated arguments about counterfactuals. But using minimalist coding we cannot do that, because our claims are formally weaker and therefore our inference rules are weaker. What we *can* do is still really interesting. We can ask and answer useful questions like:

- what are the main influences on (or effects of) a particular factor, according to the sources?
- what are the upstream, indirect influences on (or effects of) a particular factor, bearing in mind The transitivity trap?
- how well is a given programme theory validated by the respondents' narratives? (We can do this basically by using embeddings to get measures of semantic similarity between labels and

aggregate these as a goodness of fit of theory to data.)

That is all exciting and useful. It's a surprisingly simple way to make a lot of sense out of a lot of texts which is, with caveats, almost completely automatable, **but** James suggests that maybe we could ascend from formally weaker but numerically overwhelming minimalist-coded data to make other rich conclusions, in particular about enablers and blockers like the headphones and the rain. However, I don't think this is really possible. In minimalist coding, at the level of individual claims, you can code "The headphones enabled James to answer the question in the Zoom call" as

> The headphones --> James was able to answer the question in the Zoom call

... but we cannot easily get inside the *contents* of the effect. We might like to be able to code this as the effect of the headphones not on a simple causal factor but on *another causal connection*, namely between the question on the Zoom call and James' answer, but we do not have any way at the moment to do this. It might be possible to extend minimalist coding to cope with this, perhaps ending up with three factors (headphones, question, answer) and some new syntactic rules to code their relationship, and some corresponding new semantic rules to be able to deduce more things about these three factors, but I think this would be **missing the point**. I'm not sure what we could do with these kinds of subtle relationships at any scale. Let's guess that within a given corpus, five percent of causal claims are of this form: what are the chances of such claims then overlapping enough in content that we could then apply our new more specialised deduction rules in more than a handful of cases?

It might be the case that certain specific more sophisticated causal constructions become **part of ordinary language**. For example: "Her post mocking Farage went viral, so Farage was forced to respond". Here, the concept of *going viral* is perhaps a kind of shorthand for a quite sophisticated causal claim, yet it might be common enough for us to be able to usefully code it (and reason with it) using only unadulterated minimalist coding, without causally unpacking "her post went viral". So that's useful, and maybe it is even useful in building some kinds of mid-range theory, but without actually understanding or unpacking what "going viral" means.

So that's it, in a nutshell. Sorry to disappoint, James.

Appendix: Minimalist coding#

The 90% rule#

We have found that it is pretty easy to agree how to apply minimalist coding to say 90% of explicit causal claims in texts, without missing out essential causal information, whereas it is very difficult to find appropriate frameworks to cope with the remaining 10%.

Fewest assumptions#

Minimalist coding is perhaps the most primitive possible form of causal coding which makes no assumptions about the ontological form of the causal factors involved (the "causes" and "effects") or about *how* causes influence effects. In particular we do not have to decide if cause and/or effect is perhaps Boolean or ordinal, or if perhaps multiple causes belong in some kind of package or if there is some kind of specific functional relationship between causes and effects.

An act of causal coding is simply adding a link to a database or list of links: a link consists of **one new or reused cause label and one new or reused effect label**, together with the highlighted quote and the ID of the source.

A statement (S) from source Steve:

> I drank a lot and so I got super happy
>
> can be trivially coded minimalist-style as
>
> I drank a lot --> I got super happy (Source ID: Steve; Quote: I drank a lot and so I got super happy)

That's it.

Causal maps#

Crucially, we can then display the coded claims for individuals as a graphical causal map, and we can also display the entire map for all individuals and/or maps filtered in different ways to answer different questions. There is a handful of other applications (Ackermann et al., 1996) (Laukkanen, 2012) for causal mapping which also do this; but as far as we know, only Causal Map also allows direct QDA-style causal coding of texts.

Data structure#

Although we have the option of creating additional tags associated with each link (where many approaches would for example code the polarity of a link) this is not central to our approach.

We don't use a separate native table for factor labels: they are simply derived on the fly from whatever labels happen to be used in the current table of links. This makes data processing simpler and also suggests an ontological stance: causal factors only exist in virtue of being part of causal claims.

We do however have an additional table for source metadata including the IDs of sources, which can be joined to the links table in order, for example, to be able to say "show me all the claims made by women".

Causal powers#

We adopt an explicitly realist understanding of causation, because we think that's what people mean. The outcome occurred in virtue of the causal powers: drinking a lot causally influenced the super happiness in virtue of its causal powers to do so; that's what makes it a causal claim rather than just a remark on a co-occurrence or a sequence of events.

Causal influence, not determination#

We believe that it's rare for people to make claims about causal determination: someone can say that the heavy drinking made them super happy and then also agree that the music had a lot to do with it too, without this feeling like a contradiction.

Not even polarity#

We differ even from most other approaches which are explicitly called "causal mapping" in that we do not even "out of the box" record polarity of links (to do so would involve making assumptions about the nature of the "variables" at each end of the link as well the function from one to the other).

The Focus on Cognition#

In the minimalist approach, **we are quite clear that what we are trying to code is the speaker's surface cognitions and causal thinking**, while the actual reality of the things themselves is simply bracketed off at this stage, either to be revisited later (because we are indeed interested in the facts beyond the claims) or not (because we are anyway interested in the cognitions).

Staying on the surface#

At Causal Map, we rarely make any effort to get beneath the surface, to try to infer hidden or implicit meanings. This is particularly well-suited to coding at scale and/or with AI. Our colleagues at BathSDR do this a bit differently, spending more effort to read across an entire source to work out what the source *really* meant to say.

Closer to the cognitive truth#

It's really easy to code statements like (S) using minimalist coding. The trouble with trying to use more sophisticated frameworks is that they are nearly always *ontologically under-determined*. For example, even a simple approach like Causal Loop Diagramming is strongly functional and requires at least a monotonic relationship between the variables: something like, *the more* I drink, *the happier* I get (in addition to which we have to code the actual, factual claims: I did drink a lot, and: I did get super happy). But is that what the speaker meant? How do we know if the speaker has say a continuous or Boolean model of "drinking"? If Boolean, what is the opposite of drinking a lot? Drinking only a little? If continuous, how do we know what kind of function they use in their own internal model?

We'd say: nonsense. To code most causal claims as meaning some functional relationship between variables is mostly over-specified and psychologically wrong. Trying to apply such non-minimalist models means that even the trivially easy 90% of causal coding becomes suddenly hard. Of course, you can just declare that we are going to use a particular kind of non-minimalist coding for everything, but which? If we code "I got really tired because I have Long Covid", we could perhaps code both cause and effect as Boolean variables, but what about "I got really tired because it was really hot", and "I got really tired because it was really cold" how are we going to code "it was really hot" and "it was really cold"? Is there a moderate temperature which does not have this effect? How moderate? Does this variable pass through zero and come out the other side into minus temperatures? ((Ragin, 2008)) If what want to do is model a system, we can pick any solution we want. But if we want to model *cognition*, any of these answers is usually over-specified.

Unclear counterfactuals#

More formal, non-minimalist coding has clear counterfactuals. These may often be Boolean or continuous (the volume depends on the position of the volume dial; it's a 10, so the volume is maximum, if it had been at 5, the volume would have been about half as loud, and so on). Minimalist coding arguably implies some kind of naked counterfactual, but it is not always clear exactly *what*.

General versus specific#

Minimalist coding focuses primarily on **factual causal claims** which also warrant the inference that both X and Y actually happened / were the case.

Most causal claims in the kinds of texts we have dealt with (interviews and published or internal reports in international development and some other sectors) are factual, about the present or past. Sometimes we see general claims, and we often just code these willy-nilly. In any case, the distinction between general claims and claims about specific events that actually happen is often fractal and difficult to maintain completely when modelling ordinary language.

Minimalist coding as "qualitative causal" coding#

Minimalist coding may be reasonably also called **Qualitative Causal Coding**. It shares characteristics with some forms of coding within Qualitative Data Analysis (QDA), in particular demonstrating an asymmetry between presence and absence.

We don't code absences#

We do not code absences unless they are specified within the text. While codes may be counted, the concept of a *proportion* of codes is challenging because the denominator is often unclear. So if families are talking about reasons for family disputes, and family F mentions social media use, and family G mentions homework, we do not usually assume that family F does *not* think that homework can also be a cause of family disputes.

The labels do all the work#

At Causal Map Ltd, our canonical methodology initially involves in vivo coding, using the actual words in the text as factor labels. This initial process generates hundreds of overlapping factor labels. This part is really easy (and is easy to automate with AI). Obviously, hundreds (or hundreds of thousands) of overlapping factor labels are not very useful, so we need to somehow consolidate them. Arguably, minimalist coding makes the initial coding easy but it just defers some of the challenges to the recoding phase. We can:

- Use human- or AI-powered clustering techniques to consolidate the codes according to some theory

- Use AI-powered clustering techniques to consolidate the codes according to automated, numerical encoding of their meanings
- "Hard-recode" the entire dataset using a newly agreed codebook (see above)
- "Soft-recode" the dataset on the fly using embeddings to recode raw labels into those codebook labels to which they are most similar

None of this really answers all the questions raised above about problematic cases such as what to do with "I got really tired because it was really hot", and "I got really tired because it was really cold" or any other case where we have different factor codes which have shared information. At first blush, this isn't a problem, we can simply code "it was really hot" and "it was really cold" separately, but how to parse the contents to reflect the fact that these two are related? Or, how to parse the contents of "Improved health behaviour (hand washing)" and "Improved health behaviour (using a mosquito net)" to reflect the fact that they are somehow neighbours? We do have some tricks for this, but that would take us beyond the present discussion.

See also:

(Powell et al., 2024)

(Powell & Cabral, 2025)

(Britt et al., 2025)

(Powell et al., 2025)

(Remnant et al., 2025)

References

Ackermann, Jones, Sweeney, & Eden (1996). *Decision Explorer: User Guide*. https://banxia.com/pdf/de/DEGuide.pdf.

Axelrod (1976). *Structure of Decision: The Cognitive Maps of Political Elites*. Princeton university press.

Barbrook-Johnson, & Penn (2022). *Participatory Systems Mapping*. In *Systems Mapping: How to Build and Use Causal Models of Systems*. https://doi.org/10.1007/978-3-031-01919-7_5.

Britt, Powell, & Cabral (2025). *Strengthening Outcome Harvesting with AI-assisted Causal Mapping*. https://5a867cea-2d96-4383-acf1-7bc3d406cdeb.usrfiles.com/ugd/5a867c_ad000813c80747baa85c7bd5ffaf0442.pdf.

Laukkanen (2012). *Comparative Causal Mapping and CMAP3 Software in Qualitative Studies*. https://doi.org/10.17169/fqs-13.2.1846.

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Powell, & Cabral (2025). *AI-assisted Causal Mapping: A Validation Study*. Routledge. https://www.tandfonline.com/doi/abs/10.1080/13645579.2025.2591157.

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.

Ragin (2008). *Measurement Versus Calibration: A Set-Theoretic Approach*. https://doi.org/10.1093/oxfordhb/9780199286546.003.0008.

Remnant, Copestake, Powell, & Channon (2025). *Qualitative Causal Mapping in Evaluations*. In *Handbook of Health Services Evaluation: Theories, Methods and Innovative Practices*. https://doi.org/10.1007/978-3-031-87869-5_12.

Our approach is minimalist -- we code only bare causation

Why we stick to bare causation in causal mapping.

Our rule of thumb: record only that "C causes D." No coding of necessity, non-linearity, moderators, or strength. Just who said what causes what.

The short case#

- **It avoids false precision.** Labelling links as "necessary," "moderator," "non-linear," or assigning strengths suggests evidence we rarely have. We prefer to show what was claimed and how often, then let readers judge. Maps are primarily **epistemic**—repositories of evidence about people's beliefs—not truth machines.
- **It scales and compares.** Bare links plus rich factor labels let us aggregate, filter, and compare across sources, groups, and contexts without fighting about semantics of special symbols. Our tools then summarise with counts (citations, sources) and simple derived measures (like "outcomeness"), instead of speculative link attributes.

What we record#

- **Factors (boxes):** short propositions that do the heavy lifting (e.g., "Not enough money," "Won't take a holiday this year").
- **Links (arrows):** undifferentiated causal influence claims between factors. A link means "P said C influences D." That's it.

What we deliberately don't code on links#

- Necessity/sufficiency
- Non-linear forms or feedback classifications
- Moderator/mediator/inhibitor role
- Polarity or strength

Why? Because (a) respondents seldom state these explicitly; (b) analysts rarely agree on them from text alone; and (c) they reduce inter-coder reliability and slow projects down without very much which we can dependably aggregate.

Our analyses are still useful#

Coding bare links doesn't make maps "impoverished": [Causal mapping produces models you can query to answer questions]

Bottom line#

Most of the time, we code only: "C causes D (as claimed by P)." That minimal, transparent unit is reliable, scalable, and faithful to the data people actually provide. Everything richer belongs in **analysis and interpretation**, not in speculative link types baked into the coding.

Our approach clearly distinguishes evidence from facts and does not automatically warrant causal inferences

> From Better Evaluation.

Causal mapping distinguishes carefully between evidence for a causal link and the causal link itself. It does not provide any specific way to make causal inferences from one to the other. Causal mapping can help the evaluator to identify, code, simplify and synthesise the evidence for causal connections, but the evaluative step to make a judgement about whether one thing in fact causally influences another is left to the evaluator.

But, Causal Mappers are like Janus#

> From (Powell et al., 2024)

..., like Janus, the causal mapper looks in two directions at once: sometimes interpreting maps as perceptions of causation but also often wanting to make the leap to inferences about actual causation. As Laukkanen and Wang (2016: 3) point out, while conceptually poles apart, in practice, the two functions can be hard to distinguish, particularly without sufficient explanation about source information and how this has been analysed. Historically, many causal mappers have been happy with this dual focus and moving from one to the other.

As evaluators, we try to be more rigorous about this distinction. We see the job of the causal mapper as being primarily to collect and accurately visualise evidence from different sources, often leaving it to others (or to themselves wearing a different hat) to draw conclusions about what doing so reveals about the real world. This second interpretative step goes beyond causal mapping per se (Copestake, 2021; Copestake et al., 2019a; Powell et al., 2023).

The elephant in the room -- causal inference ▸

References

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Our approach is minimalist -- factors are not variables

Many or most causal mapping approaches, including Causal Loop Diagrams, also code the perceived strength of a causal link. This means that the factors become variables which can take values between, say, low and high or positive and negative, and we can make a much broader range of inferences using some form of numerical modelling. This can be seen as the extreme reproducible end of our spectrum and borders on quantitative approaches.

However we do not go so far: our causal factors are closer to being propositions rather than variables and we do not jump to code, say, poverty as negative wealth, or unemployment as obviously just the opposite of employment.

The Conventional Assumption#

A foundational assumption, particularly for those approaching causal mapping from a systems dynamics perspective, is that every concept on a map should be treated as a **variable**. This implies that each element is something quantifiable, capable of taking on different values across a defined spectrum, such as from low to high, negative to positive, or from zero upwards. Such a map is backed up by a dataset, a large-ish set of measurements of the state of each variable.

The Discrepancy with Human Narrative#

However, this assumption contrasts sharply with how people actually communicate and describe their experiences. When individuals explain what causes what in their world, they rarely speak in terms of discrete variables. Forcing real-world narratives into a rigid, variable-based structure requires significant and often unnatural contortions.

Constructing variables out of experience is just that: a construction. Quantitative social scientists are really good at it. But people's thinking and language are not inherently structured in this way.

For instance, in an evaluation of a program's effects, the sudden onset of the COVID-19 pandemic presents a significant modelling problem. While the pandemic certainly had a causal impact on countless factors, it doesn't fit neatly into the definition of a variable.

How would one define it? As a binary "COVID vs. no COVID" variable? The concept of a counterfactual -- a world where the pandemic never happened -- is abstract and difficult to operationalize. This example highlights that the way people experience and discuss the world is often event-based, not variable-based, exposing a limitation in traditional modelling assumptions.

1a A minimalist approach to coding helps capture what people actually say

Encoding people's narratives about what causes what always involves a certain amount of modelling or theory-construction. Most approaches -- Causal mapping approaches differ in application, construction, analysis and how they deal with multiple sources. This might involve constructing a codebook of common factors, but it might also involve applying some kind of special logic of causation, for example:

- distinguishing between necessary and sufficient conditions
- identifying special packages of causes which somehow fit together
- coding the strength and/or polarity of causal links

Our experience, together with BathSDR, of coding thousands and thousands of documents every year -- manually and with AI support -- is that most people don't use these special features in their language most of the time, not even scientists. You can go through a whole interview trying to work out if each cause is supposed to be a necessary or a sufficient condition of its effect, but you quite likely won't find a single case where the source actually uses the idea explicitly.

So mostly, we say: don't bother.

At Causal Map Ltd, in our consulting practice, we've taken this minimalist approach even further and mostly code initially without any kind of codebook at all.

1b A minimalist approach to coding makes aggregation easier

We just argued that 1a A minimalist approach to coding helps capture what people actually say. But even if you did succeed in imposing some special logical features on your data -- for example, coding necessity and sufficiency -- you'd probably find that most of your data didn't fit well with these special features. When it comes to aggregating medium or large amounts of coding, you wouldn't find it very useful.

With our minimalist approach, we mostly have just one task: what to do about all those different factor labels.

1c A minimalist approach to coding does not code absences

One thing which makes causal mapping a fundamentally qualitative approach is that we do not code absences.

We do not think that the world, nor the piece of the world we are studying, is essentially a grid of variables and cases (nor a cube of variables and cases and timepoints), in which each case always has a value for every variable (at every timepoint).

If some respondents say that their headaches make them nauseous, and others do not mention headaches, even if they mention nausea, we do not interpret that as meaning that they *did or did not* have headaches. We do not think that having headaches, or not, is a variable which *must* be relevant to everyone's explanations, all the time.

Our approach is minimalist -- we do not code the strength of a link

At Causal Map, we do not endorse coding the strength of causal links. You can't really do it in the Causal Map app (see [Coding with and using link metadata]()).

[Qualitative impact evaluation is less interested in the strength of effects]()

Three types of objections to coding causal strength.[#]()

Objection 1: Variable Construction[#]()

Coding strength requires a massive amount of construction work: it involves thinking about the area of interest in terms of variables. This requires modelling specific entities that go up or down, or show differences in number. Constructing variables like this does allow for capturing and calculating correlations. But this construction process is often difficult and does not fit well with how people actually speak in most situations.

If different people are talking about poverty and wealth, employment and unemployment, to be sure you can try to squeeze this all into a shared model with just a couple of variables like say `household income` and `household employment status`. But that is a massive abstraction.

• So we favour bare propositions over variables.

Objection 2: Translating to Numbers[#]()

• The second objection concerns the difficulty of translating all relationships into actual numbers.

Challenges with Standardization and Polarity[#]()

• If general rules are used, people usually standardize the variables (e.g., ranging from zero to one). • Standardization is difficult for factors like a country's population, especially when numbers may increase exponentially over time. • Such changes in magnitude occur even in quantitative sciences, often requiring arbitrary decisions about log transformations. • A more significant problem involves absences, negatives, and polarities. • Example of a strong positive link: If greater anger leads to greater shouting, this connection can be viewed as a "powerful transmission cable" with a high causal coefficient. • Example of a weak link: If anger ranges from zero to one but shouting remains low (e.g., 0.2), the connection has a very low coefficient of transmission. • This type of modelling becomes difficult when negative numbers are introduced. • Example: If high temperatures cause crop failures, a drop in temperature might see harvest go up. • However, extreme cold temperatures also cause crop failures. • It is difficult to model this complexity using a single, bipolar variable.

Objection 3: Aggregation Difficulties[#]()

• The third objection involves the difficulty of aggregating information from multiple sources, assuming such numbers existed. • This third objection is irrelevant in approaches like participatory systems mapping, where a final number for each link is already agreed upon. In this case you could say there is only one source.

In a causal mapping dataset there is no need for a special table of factors

If you are interested in how to formalise causal mapping or in building software, we'd like to share this insight. If you are not, ignore this.

> Factors are implied by links

We don't need to have a separate table for the factors because the factors can be derived from the links table. If you cannot find a factor X as cause or effect in the links table, it does not exist.

This means that our data model (since version 3 of Causal Map) does not need to have a table for factors. Essentially we just have a table for links, plus a table for sources both to supply the texts and to more conveniently store metadata like gender and district.

It's of course possible to formalise causal mapping in other ways, but we have found dropping a special table for factors to solve a lot of the problems associated with having to keep factors and links tables in sync.

This does bring its own problems: Factor label tags -- coding factor metadata within its label

Factor labels -- a creative challenge

Where do the labels for the causal factors come from? As with ordinary QDA and thematic analysis (Braun and Clarke, 2006), approaches vary in the extent to which they are purely exploratory or seek to confirm prior theory (Copestake, 2014). Exploratory coding entails trying to identify different causal claims embedded in what people say, creating factor labels inductively and iteratively from the narrative data. Different respondents will not, of course, always use precisely the same phrases, and it is a creative challenge to create and curate this list of causal factors. For example, if Alice says 'Feeling good about the future is one thing that increases your wellbeing', is this element 'Feeling good about the future' the same as 'Being confident about tomorrow' which Bob mentioned earlier? Should we encode them both as the same thing, and if so, what shall we call it? We might choose 'Positive view of future', but how well does this cover both cases? Laukkanen (1994) discusses strategies for finding common vocabularies. As in ordinary QDA, analysts will usually find themselves generating an ever-growing list of factors and will need to continually consider how to consolidate it – sometimes using strategies such as hierarchical coding or 'nesting' factors (as discussed in the following section).

The alternative to exploratory coding is confirmatory coding, which employs an agreed code book, derived from a ToC and/or from prior studies. QuIP studies mostly use exploratory coding but sometimes supplement labels with additional codes derived from a project's ToC, for example, 'attribution coding' helps to signify which factors explicitly refer to a specific intervention being evaluated (Copestake et al., 2019b: 257). However, careful sequencing matters here because pre-set codes may frame or bias how the coder sees the data (Copestake et al., 2019a). Again, the positionality of the coder matters just as much when doing causal coding as it does for any other form of qualitative data coding.

Factor label tags -- coding factor metadata within its label

For example you might want to code the respondent's happiness at work as different from yet similar to their happiness at home. With a factor table, you could have a field called `label` = "Happiness" and another, say `context`, which is = either "Home" or "Work". This is what we do with the links table in Causal Map, where we do have some hard-coded (but optional) fields and some user-definable fields.

Hierarchical coding is one way to bring some order to a whole crowd of factors. However, sometimes you don't want to think in terms of a strict hierarchy, or maybe you have an additional set of themes which cut across that hierarchy.

https://vimeo.com/671894620

Tags are useful in either of these cases.

Tags are just sequences of characters within a factor label to which you have given a special meaning, and which are unique and easy to search for. These can include letters, emojis or phrases. You can do coding without any such tags if you want, but it can help when searching and filtering.

Factor tags are just like #️⃣ Link hashtags. Confusingly, a link hashtag doesn't have to actually start with a `#`, and a factor tag can indeed start with a `#`, but we find it easier to keep the names separate like this.

So a tag is nothing more than any sequence of characters which is repeated in several factor labels. Any sequence of characters will do. For example you could consider the letter "a" to be a tag and display the map showing all the factors which contain the letter "a". But this wouldn't be interesting. The trick when using tags is to decide on short, meaningful codes which will not be repeated anywhere else. For example you wouldn't want to use a pair of tags like "women" and "men" to distinguish factors which are only relevant for one or the other gender because the "wo**men**" factors would also turn up when you search for "**men**". That is why we have to be careful when creating tags, for example by preceding a sequence of characters with a tag `"#"`.

A quote like "family situation is better now because of improved food availability" can be coded like this:

> More food –> Improved wellbeing

Now, maybe you are asked also to keep track of any aspects of the project which have to do with nutrition. Nutrition is not really part of your system of factors, but you would like to be able to construct some maps just to look at this aspect. So you can write this:

> More food #nutrition –> Improved wellbeing

Similarly, if Improved Wellbeing is one of the desired outcomes of the project, we might want to reflect that by adding a tag "(Outcome)" like this.

> More food –> Improved wellbeing (Outcome)

Then we can easily search for this and other desired outcomes.

A tag like “men” is not suitable because it is likely to appear elsewhere (e.g. as part of “women” or “management”). To get round this, add additional characters like a hash: “#men”; this makes the tag unique.

If you use curved or square brackets around your tags, you can use one of the app filters to hide the tags for specific maps if desired.

Factor labels -- semi-quantitative formulations can help

It might be tempting to try to formulate all factor labels in a strictly similar way, using for example language like increased probability of … or positive change in … in every case. But it is difficult to identify and agree on a satisfactory template for doing this which will capture enough of the way people really make causal explanations (in the way that quantitative social scientists hope to measure everything just with continuous variables). This is always a balancing act, but we encourage you when in doubt to stick fairly close to the actual language your sources use (so-called "in-vivo" coding), and don't be *too* worried if your factor labels are different from one another grammatically (e.g. some express a difference like improvement in X and some do not).

The formulation of **factor labels** should fit the intended interpretation of the **causal links**. For example, most commonly B ➜ E is supposed to mean that B exerts in some sense an "increasing" or "decreasing" influence on E, then both B and E need to be formulated in a corresponding way. In order to ease interpretation, with a few exceptions, factors should be labelled and understood in such a way that it makes sense to say "more of this" or "this happened as opposed to not happening": we call these semi-quantitative factors.

Consequently you should avoid a factor label like Training courses, which might be understood as a mixed bag of various causal factors to do with training courses. We would usually prefer a label such as Training courses delivered or Quality of training courses which are easier to understand as things which can increase or decrease, or happen or not happen. You may even prefer to use labels like Quality of training courses improved or Improved quality of training courses, in which the *difference made* is already included in the title.

Examples of semi-quantitative factors#

These are examples of factor labels where you can judge whether it happened more or less, whether it is higher or lower, or whether it happened versus not happened:

- Sold cow
- Earthquake happened
- (Had) good harvest
- (Level of) bank account
- (Level of) ethnic tolerance
- Quality of seeds

In some contexts, we can also talk about the *likelihood* of events, so "if people get a good harvest they are less likely to sell their cow."

Non-quantitative factors#

It is also perfectly acceptable and sometimes necessary to use purely qualitative labels, e.g. coping style, see below. However, this may limit some of the analysis and reporting tools available:

- Teaching style
- Coping strategy
- The content of the report

We can even make a link between two such factors, claiming for example that the style of 60’s music influenced the style of 70’s music, without any concept of quantity. That’s ok.

Factor labels -- a creative challenge

In a causal mapping dataset there is no need for a special table of factors

Factor labels -- do not over-generalise

When you are creating factor labels for re-use across different causal claims, you should usually take care to keep them specific: make them no more general than they need to be.

So if you are coding cases where a household has increased income, use a label like Increased household income, not Increased income or even Economic improvement.

This is especially important when using hierarchical factors, when it's easy to fall into the temptation of creating very general top-level labels like Economic improvement even if all your material is actually only about increased income in households and farms.

Coding with and using link metadata

In our implementation of causal mapping in the Causal Map app, Our approach is minimalist -- we do not code the strength of a link.

Providing metadata as a column makes sense when the values of this column make sense across the whole dataset, across all multiple links, like let's say before covid and after covid.

Such a column can function a bit like a *context* variable, for different time periods or applying to different stakeholders. Context in this sense might be seen as functioning a *bit* like a causal factor but not exactly.

But we can also provide metadata as free-form tags. We provide a hard-coded "tags" column for which users can provide comma-separated lists of tags which are made up and adapted on the fly. They don't necessarily make sense across the whole dataset.

In Causal Map 4, as well as a hard-coded Tags column, we do provide a hard-coded sentiment column which can take the values -1, 0 and 1, and which can be averaged to any number between -1 and 1.

Link metadata -- Sentiment ▸

We also provide arbitrary additional free-form, free-text columns for any purpose. We often like to add a column like this:

Link metadata -- Time reference ▸

Link metadata -- quality of evidence ▸

... or simply to code a tag like "#doubtful".

Link metadata -- Sentiment

What is it for?

a hard-coded sentiment column which can take the values -1, 0 and 1, and which can be averaged to any number between -1 and 1.

Link metadata -- Time reference

It is often useful to code a time reference. We often conflate time with hypothetical status, e.g.

- hypothetical past/present
- factual-past/present
- future-planned
- future-hypothetical

For example, if we are to code a whole corpus of reports which also include planning documentation, there might be a lot of causal claims about what is supposed to happen in the future, perhaps interspersed with claims about what actually happened in the past. It will often be important to distinguish these two.

Link metadata -- quality of evidence

Tasks 2 & 3 -- Introduction

In the previous chapter Task 2 -- Introduction we looked at the main ideas of minimalist coding:

- we code links between simple propositions which we do not necessarily conceive of as variables which can take different states at different times.
- we let factor labels do most of the work, so we start off by distinguishing between, say `wealth` and `poverty`.
- we don't use separate metadata columns for factors. In fact, we don't even have a table for factors at all: our coding results only in a table of links.

That approach has general applicability. In this chapter we get down to some specific suggested conventions about how to make this approach more useful for answering concrete questions. These suggestions are all implemented in the Causal Map app, but of course other approaches are possible.

Our approach is algorithmic

All the pioneers of causal mapping used different algorithms to simplify, query and synthesise their data.

What we add: TODO

Simplification - co-terminal link bundles

In most use cases, the data contains many "co-terminal links": links with the same cause and effect. We call these "bundles" of links. It is often bewildering to show all of those links in a map so by default we condense them all into a single visual arrow, usually printing the number of sources and/or citations on the arrow and making the width depend on the number of citations.

✅ Simplifying causal maps with hierarchical coding#

Summary#

You can use the special separator `;` to create nested factor labels, like this:

New intervention; midwife training ➜ Healthy behaviour; hand washing

We can read this separator as “in particular” or “for example”:

New intervention, in particular the midwife training,

Or we can read it in reverse like this:

The midwife training, which is an example of / part of the new intervention

Factor labels can be nested to any number of levels, e.g.

New intervention; midwife training; hand washing instructions

The higher level factors can, within the same coding scheme, themselves be used for coding.

So as well as creating links to and from New intervention; midwife training; hand washing instructions, you can always also use New intervention; midwife training and New intervention as factors too.

e.g. we could code “this whole new intervention has also led to happier health providers” like this:

New intervention ➜ Happier health providers

We can “zoom out” of a causal map containing nested factors to show a simpler, higher-level structure as a summary. This is done by applying an algorithm which re-routes links to and from the lower-level factors into their higher-level parents.

So then, loosely yet informatively and with certain caveats, accepting a loss of detail but affirming that the overall meaning is not distorted, this algorithm can deduce for us, from the first example above, the following causal map:

New intervention ➜ Healthy behaviour

Usually each higher-level factor will each be a summary of *many different* lower-level factors.

Introduction#

An analyst coding text to create a causal map is confronted with the same challenge as any qualitative researcher: identifying recurring themes in such a way as to help a larger picture emerge, while retaining important detail. Expressing factor labels in a hierarchical fashion can help solve this problem. But hierarchical labelling also enables a particular strength of causal mapping: it lets us “zoom out” to view a whole causal map from a higher-level perspective, merging causally similar concepts to give a simpler map with far fewer components. Formally, the process of zooming out produces a map which logically *follows from*, is *implied by*, the original map. This chapter also introduces a smarter way to “zoom out” from a causal map, and explains how these features are implemented in the Causal Map app.

When conducting qualitative coding of any text, there is always a tension between wanting to keep the detail (e.g. hand washing) but also to code in such a way that general themes emerge (e.g. health behaviour). One way to do this is to organise the codes into a hierarchical structure, so that “Hand washing” is nested as part of “Health behaviour”. This can be done (see e.g. Dedoose, saturateapp.com) by using labels in which the hierarchy is directly expressed, for example Hand washing; health behaviour – using semi-colons or some other convenient character to separate the levels of the hierarchy.

This approach is convenient for several reasons:

- A search for “Health behaviour” will find Health behaviour; Hand washing as well as Health behaviour; vaccinating children and other combinations.
- It can be arbitrarily extended to any number of levels, e.g. Health behaviour; Hand washing; Before meals

- Related items appear next to each other when they are listed alphabetically
- The hierarchical structure does not require that the analyst (whether using paper-and-pencil or software) maintains a separate set of “parent” codes; the higher-level codes are simply whatever is visible before the semi-colons. Higher-level codes can be created and changed on the fly without having to open a separate codebook or software interface.
- It is possible to code directly at higher levels, for example using the code Health behaviour where no more details are given.

When reading a nested factor label aloud, the semi-colons could be substituted with “… and in particular ….”, e.g. “Health behaviour, and in particular Hand washing, and in particular Before meals”.

The way factor (labels) emerge during causal mapping is just a special case of the way codes emerge in any qualitative coding process, and nested coding is useful in ordinary qualitative data analysis as well as in causal mapping. However, hierarchical coding in causal mapping is particularly exciting because it allows us to do things like simplify a causal map to give a higher-level version of it with far fewer components.

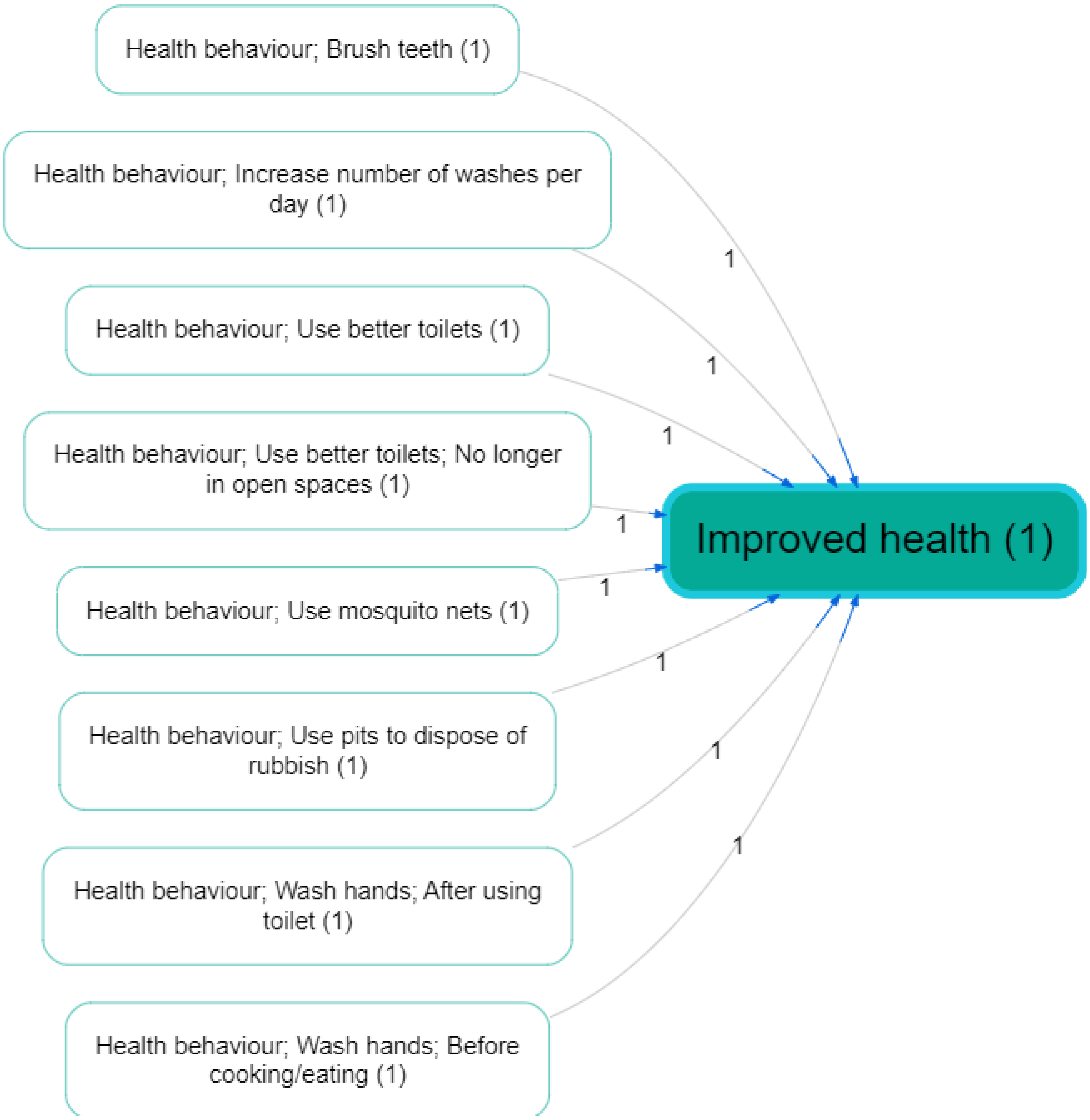
A factor can’t belong to two different hierarchies#

One limitation to this way of expressing the hierarchy as part of the factor label is that you can’t make one factor belong to two different higher-level concepts. For example, you could understand a particular midwife training both as causally part of a new intervention but also perhaps as causally part of an institution’s in-service training programme or an individual’s workload, but you can’t code it as both “New intervention; midwife training” and “In-service training; midwife training” at the same time.

This limitation is because of the meaning of the semicolon: the ; in Y; X means that this label can be replaced as needed with just Y , accepting a loss of detail but affirming that the overall causal story is not essentially distorted. If a hierarchical label had more than one parent, we wouldn’t know which parent to “roll up” the factor into.

If you find yourself wanting to put a factor into more than one hierarchy, try using Factor label tags -- coding factor metadata within its label instead.

Zooming out#



Assuming we have a causal map which has used hierarchical coding, as in the small map shown above, how do we take advantage of this coding to “zoom out”?

If we define the “level” of a factor as the number of semicolons in its label plus 1, here is the same map, zoomed out to level 2 (i.e. a maximum of one semi-colon per factor).

image.png

Here is the same map, zoomed out to level 1 (i.e. there are no semi-colons at all).

image.png

A warning: causal mapping as described here is a *qualitative* process. While zooming in and out can be very useful, it should never be used mechanically.

Zooming out is like making deductions with the ; separator#

A causal map coded using a hierarchical separator can be “zoomed out” given a specific interpretation of the **;** separator, as follows.

If we know

New intervention; midwife training ➜ Healthy behaviour; hand washing

then, loosely yet informatively and with certain caveats, accepting a loss of detail but affirming that the overall meaning is not distorted, we can deduce:

New intervention ➜ Healthy behaviour; hand washing

and

New intervention; midwife training ➜ Healthy behaviour

and even

New intervention ➜ Healthy behaviour

This actually reflects the dilemma of the analyst often referred to as *granularity*: with how much detail should I code the beginning (or the end) of this causal story? Expressing a factor as Health behaviour; Hand washing; Before meals shows that this is indeed to be understood as a kind of health behaviour, although of course not the whole of it. By using this approach, the analyst says: if you are just looking for the global picture, I am happy for this factor to be understood as Health behaviour.

When factors are nested like this within one another as part of a hierarchy, it is possible to give an overview and ‘zoom out’ of the detailed data. This helps to simplify some of the analysis, enabling the user to focus on the links between the top-level groups rather than all the details. It follows that two factors like Y; X and Y; Z are *causally similar enough to one another to merge into Y* at a more general level.

Expressing a factor in a form like Y; X **means** it can be replaced as needed with just Y, accepting a loss of detail but affirming that the overall meaning is not essentially distorted. If you wouldn’t be happy to accept this replacement, don’t use the “;” to code this factor.

Semi-quantitative formulations work best#

We already saw that causal mapping often works best when the factors are semi-quantitative. The hierarchical approach also works best when the higher-level factors are themselves labelled such that also they are *semi-quantitative*, *causal* factors which could be used on their own – in a way which themes or categories see here could not. Good examples would be:

- Social problems
- Social problems; Unemployment
- Social problems; Addiction
- Psychosocial stressors

- Psychosocial stressors; Fear of job losses
- Psychosocial stressors; Pre-existing mental health issues

Here, Social problems and Psychosocial stressors are higher-level causal factors in their own right; they are not just themes or boxes to put factors into.

So we might have:

“The problem of unemployment is a psychosocial stress for many”

Social problems; Unemployment ➜ Psychosocial stressors

“When people get stressed they often turn to drugs“

Psychosocial stressors ➜ Social problems; Addiction

These could be combined into this story:

Social problems; Unemployment ➜ Psychosocial stressors ➜ Social problems; Addiction

If we zoom out of the above story, we could focus in on the higher-level factors and in this case we would get a vicious cycle:

Social problems ➜ Psychosocial stressors ➜ Social problems

Higher-level factors are *generalisations*#

Usually, we don’t use higher levels merely to organise factors into themes which are not causally relevant.

Health; vaccinations law is passed

Health; mortality rate

These two items can be grouped into a *theme* “health” but can hardly be understood as special cases of a more general causal factor, so it would be a mistake to use the semi-colon. Instead, it would be more suitable to include the word “Health” just as a **hashtag**, without the semi-colon:

Vaccinations law is passed #health

Mortality rate #health

In other words, **causal factors in hierarchies should usually be semi-quantitative**.

Don’t mix desirability!#

All the factors within one hierarchy should be desirable, or undesirable, but not both.

Generally speaking, make sure that the **sentiment of a more detailed factor is interpretable in the same way as the factor higher up in the hierarchy**. Ideally *all* the detailed factors within a hierarchy should be broadly *desirable,* or all *undesirable*, but not both. For example, you don’t want to nest something undesirable into something desirable. E.g. you don’t want to formulate a factor like this:

Stakeholder capacity; Lack of skills.

Because capacity would normally be understood as something desirable, and lack of skills would not. If you zoom out to level 1, this factor will be counted as an instance of Stakeholder capacity which is surely not what you want.

Try to use opposites coding and the ~ symbol to reformulate as:

~Stakeholder capacity; ~Presence of skills.

If you zoom out to level 1, this factor will *not* be counted as an instance of Stakeholder capacity.

Using higher level factors for "Mixed bags"#

In spite of what we just said, sometimes you find you *have* use higher-level factors just to group a mixed bag, like this:

Politics; increase in populism

Politics; shift to the left

Politics; shift to the right

Politics; more engagement from younger voters.

The higher-level factor Politics is not in any sense a generalisation of these very disparate factors. However, we can at least think of it as a 'mixed bag'. If we roll the map containing these stories up to level 1, we'll see this 'mixed bag' factor Politics as a cause and effect of other factors. It will be a bit hard to interpret, but we can live with it as long as we remember that it is a mixed bag rather than a semi-quantitative summary.

Hierarchical coding as a way of coping with a large number of factors#

Usually analysts are faced with the quandary of either having too many factors which they lose track of, or merging them into a smaller number of factors and losing information. With hierarchies, you can have your cake and eat it; it is similar to the process of recoding an unwieldy number of factors into a smaller number of less granular items, but with the advantage that the process is reversible; the information can be viewed from the new higher level but also viewed from the original, more granular level. For example, we can count that the higher-level factor component "Health behaviour" was mentioned ten times, while retaining the information about the individual mentions of its more granular components.

Don't use a hierarchy when a hashtag will do#

When the analyst wants to group certain factors into a theme (like "health") which is not itself a causal factor, hierarchical coding should not be used. Instead, text hashtags like "#Environment" or "[Environment]" or just "Environment" can be used to create themes simply by adding the text hashtag to the factor label, e.g.

Soil loss (Environment)

Eco training courses for NGOs (Environment)

Re-usable factor components as hashtags#

Sometimes your factors relate to each other in ways which are not just hierarchical. For example:

- Activities completed; Training; Health
- Activities completed; Distribution; Health; First-aid kits
- Outcomes achieved; Health; First-aid skills

These are three (hierarchical) factors in which "health" appears in different places, and at different levels of the hierarchy.

This is not ideal, but sometimes it's just the best way to organise a tricky set of factors.

In this example, "Health" appears only as a "component" of other factors. Although on its own it might look like a mere theme rather than a causal factor, it plays a role in differentiating the causal factors in which it participates, e.g. "Activities completed, in particular training, in particular on health"; and because "Health" is used across hierarchies, it can *also* be treated as a hashtag and can be used as part of searches, lists and counts of factors, etc.

Isn't that a contradiction? Didn't we just say that "Health" is not to be used on its own as a factor because it is just a theme and is not expressed in a semi-quantitative way? No, because here the word "Health" does not function as only a theme but as a way of differentiating causal factors: all the actual factor labels in which it participates are correctly expressed in a semi-quantitative way. So Activities completed; Training; Health is intended as a causal factor about the extent of completion of activities, in particular training activities, and in particular health training activities.

Causal mapping looks for linearity first

Causal mapping most often looks for linearity first, while of course being on the lookout for feedback loops and circular shapes. Whereas most systems approaches do the opposite.

Can you spot a complex system when you see one?#

Version 1#

The network pictured above, even though it is quite small, looks pretty tangled. We're not going to fully understand it, so we'd better get out our tools for dealing with complexity? But wait, look at the boring, old-fashioned hierarchy below.

Version 2#

Did you spot that they have exactly the same structure? Now it is easier to see that it is just a

hierarchy. D, E and F have one contributor each, whereas G and H share I, J and K as contributors, and feed only into B, whereas D, E and F all feed into both B and C, which feed into A. Easy. Nothing which should be too hard to predict, no balancing feedback loops.

"Complex" and "System" are very buzzy buzz-words at the moment. We should check we don't throw them around too much without thinking. I'm just reading Moore, Parsons and Jessop in the American Journal of Evaluation. They quote Magee and de Weck (2004) who define complex systems as systems "with numerous components and interconnections, interactions or interdependence that are difficult to describe, understand, predict, manage, design, and/or change." Well yes, kinda. But what if you find a system difficult to describe, etc, just because you didn't look hard enough?

Yes, causal maps are just concept maps with only one type of connector, and that connector means "... causes....". Whereas concept maps can have any type of connector you like. Historically, causal maps come from concept maps.
Laying out causal maps is a challenge! Most folks from the systems tradition like swirly circular layouts which make them look like everything is one big feedback loop. If there is a more linear structure, we recommend showing that linear structure.

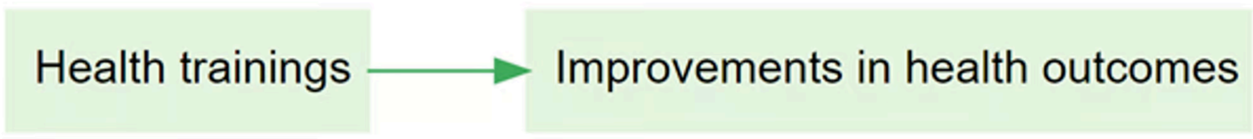
➕➖ Opposites

Introduction#

In the first part of this Guide we have dealt only with undifferentiated links which simply say “C causally influences/influenced E” or more precisely “Source S claims/believes that C causally influences/influenced E.” We call this “barebones style” causal mapping. There is nothing more to this kind of causal map than links between factors. No other features are used.

Barebones-style mapping can be interpreted in two different ways:

- As in QuIP, to show causal influences between past events. It is an open question to what extent these causal claims can be generalised. It is possible to parse a link from C to E as saying not only that C is something which has the causal power to influence E but also that in some sense C happened and did in fact influence E, i.e. made a difference to it.
- To show only causal influences between factors, without recording what did or does happen.



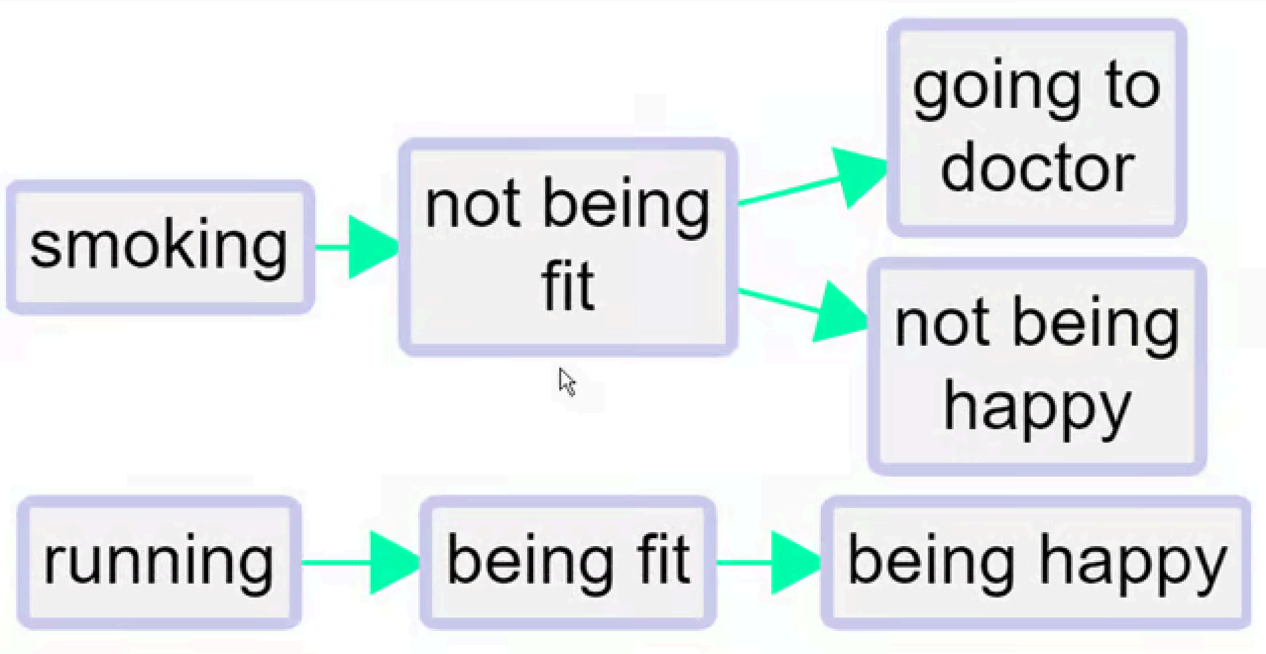
In Part 1 of this Guide we did in fact also introduce an additional convention, hierarchical coding, in which the “;” separator is used to encode the idea that C; D can be read “D, an example of C” and that at a suitable level of abstraction we can approximate C; D as C.

Combining opposites#

This section presents a simple and powerful way of dealing with “negative” factors such as `poor health` which are in some sense the opposite of existing “positive” factors such as `good health`.

In some kinds of causal mapping and systems diagramming, more sophisticated approaches are used in which the factors are considered to be variables and the links between them can have positive or negative strength. The approach we present here is a simpler alternative.

Here is an example of quite minimalist QuIP-style coding. There are the beginnings of some ideas about (and issues with) polarity: for example, we have `fit`and `not fit`.



We've all done this kind of coding, with classic examples being coding for both employment and unemployment or for both health and illness. This could for example be two different stories about two different people; or it could be different aspects of or periods within one person's life.

This minimalist style on its own is unsatisfactory. We haven't told the app that being fit is represented with a somehow positive and somehow negative factor. So can't join them up. We can't compare the way that being fit leads to happiness and on the other hand not being fit leads to unhappiness (and to visiting the doctor). **We can't for example deduce that running might make visits to the doctor less likely.** Also, if we produce a table or do other analyses focused on healthy habits, we might miss data on the closely related unhealthy habits.

The first step forwards is to follow this convention:

> To signal that two factors are opposites we formalise the idea we already instinctively used in the above example, where we used the word "not" for one of each pair. Formally, we will code them in the form "~Y" and "Y." The ~ may appear at the start of a factor label. This already ensures that when we search for "Y" we will also find "~ Y."In the Edit Multiple Factors panel, these two factors will be listed next to each other - the alphabetical listing will ignore the ~.

We talk about *opposites* rather than positive/negative or plus/minus because that frees us from any implications about valence or sentiment: smoking is the opposite of not smoking, health is the opposite of not health / ill health / illness.

Where there is some kind of valence or sentiment involved, we do suggest using the `~` sign for the negative member of the pair. But it wouldn't make any difference to the app.

So the same map would look like this, using

```
~
```

instead of

```
not
```

.

image.png

Non-hierarchical coding with opposites is easy:

- Eating vegetables
- ~Eating vegetables
- Smoking
- ~Smoking

When you use the “combine opposites” filter (switched on),

Untitled

the app tries to combine any pairs of factors which are opposites. It looks at all the factors which begin with a ~ and takes off the ~ where there was one. But it only does this if there is in fact such an opposite already coded in the file as currently filtered, otherwise there wouldn’t be any point.

image.png

So now there are for example two factors combined into the “fit” factor and two into the “happy” factor. The “not” factors have their incoming and outgoing links preserved, but when a factor is flipped to match up with its opposite, the part of the link next to that factor are now coloured pink. So the lower link from fit to happy is pink because the factor at each end of the link has been flipped from “~Y” form to “Y” form; the influence factor was originally *not fit* and the consequence factor was originally *not happy*. So there is no danger of thinking that this is really just another case of the other link, i.e. of fitness leading to happiness.

So, a link has two polarities: a *from* polarity and a *to* polarity. If the signs of the two polarities are opposite, then the effect of the influence factor on the consequence factor is reversed.

Both links from fit to happy have the same overall polarity (normal, not reversed) but they do not represent the same information.

No information is lost when you press the “combine opposites” button; you can still always read off the original map from it.

Opposites coding within a hierarchy#

When using hierarchical coding, the sign “~” may appear at the start of a factor label *and/or at the start of any component in a factor label*.

Here is a similar story, now coded hierarchically. In this example, we only see ~ at the beginning of the factors, not yet within them.

image.png

When you press "combine opposites," the app tries to combine any pairs of factors which are opposites. It looks at all the factors which begin with a ~ and *flips each component*, taking off the ~ where there was one, and inserting one where there was not. But again it only does this if there is in fact such an opposite already coded in the file as currently filtered, otherwise there wouldn't be any point, because there is nothing to combine.

image.png

Here is the same example, but also "zoomed out" to the top level.

image.png

A quantitative social scientist might solve this problem by flipping the polarity of the negative examples, coding them as positive but using minus strengths for the connections. So smoking influences good health but with a minus strength. However this always seems somehow unsatisfactory and is complicated to do. It is particularly unsatisfactory when *both* ends of the arrow are flipped in this way so that we code the influence of being unfit on being unhappy as a green arrow from fitness to health!

By default in print view, links in the same bundle, i.e. with the same from and two factors, are no longer always displayed as one, with the frequency noted as a label. If we were using the quantitative approach, some of the links in the bundle may have minus rather than plus strength, etc, and we would have to somehow form some kind of average to arrive at an overall strength score, which is not at all satisfactory. Now, the links are only counted together if they have the same *from* and *to* polarities. So there can be up to four different links from one factor to another in Print view.

We are deliberately **not** falling Into the trap of somehow trying to aggregate the different *strengths* to say for example "there are 6 plus links from advocacy to compliance and 1 minus link so this is like 5 plus links because 6 - 1 = 5." We don't have evidence for an aggregated strength; we have aggregated evidence for a strength. Aggregating different pieces of *evidence* for links with different strengths is not the same as aggregating links with different strengths. So our more conservative approach preserves information.

It's also possible that someone says "I know this intervention works not only because the intervention made me happier but also because I saw the people who didn't get it and they are definitely not happier as a result." In this case, we might code both arrows, intervention ➜ happy and not intervention ➜ not happy.

Opposites coding within components of a hierarchy#

Sometimes we need to use the ~ sign within the components of a hierarchy.

- ~Healthy habits; ~eating vegetables

is the opposite of

- Healthy habits; eating vegetables

Not eating vegetables, which is an example of unhealthy behaviour, is the opposite of eating vegetables, an example of healthy behaviour.

- ~Healthy habits; smoking

is the opposite of

- Healthy habits; ~smoking

Smoking, which is an example of unhealthy behaviour, is the opposite of not smoking, an example of healthy behaviour.

So here we add one more causal claim to our above example, at the bottom:

image.png

The healthy habit of not smoking leads to being fit.

So the app correctly detects that not smoking is the opposite of smoking:

image.png

The two arrows at bottom left (one all green, one all pink) show that there is one example of this particular healthy habit leading to fitness, and the complementary example in which the opposite of this habit leads to the opposite of fitness.

Zoomed out to the top level:

image.png

Untitled

Bivalent variables?#

Note that this is the mutually exclusive condition from classical logic, but we don't have any mention of an exhaustive condition: for us, it is not the case that everything has to be either wealthy or poor.

We could imagine that in a causal map, even before we even think about which specific factors might be opposites of other specific factors, each and every factor is really a kind of two-value factor like wealthy (as opposed to non-wealthy) or receiving tuition (as opposed to not receiving tuition). This is the quantitative way of thinking, in which every factor is really a variable which takes at least two values. But this is an unnecessary complication for us. The important point is that usually people don't think of the absence of something as having causal powers (though there are exceptions).

Suppose we are trying to code someone's understanding of income. They tell us that if when people are wealthy, they tend do certain things, and when they are poor they tend to do other things. To some extent, these two sets of things are themselves the reverse or contrary of one another. For example wealthy people will probably have the best education and poor people the least. Sometimes the sets of things we associate with the two opposite poles are not obviously a simple reflection of one another. For example, poor people are often hungry whereas non-poor people are not. But in an affluent country, not being hungry is probably not an important feature we would think of to describe being wealthy as opposed to not-wealthy, if we can assume that non-wealthy people are *overall* rarely hungry. We are more likely to think of things like eating often in posh restaurants.

So being wealthy and being poor are a good example of factors which we *should* consider coding as opposites. We can call the resulting factor, when we press the combine opposites button, *bivalent*.

, we could imagine there was a middle point too:

- Wealthy
- Not-wealthy / Not-poor - a kind of midpoint
- Poor

This is the usual case.

Which pairs of factors should we consider for opposites coding?#

Short answer: use opposites coding for a pair of factors X and Y (i.e. recode Y as ~X or recode X as ~Y) if both X and Y naturally occur, separately, in the coding, but they can be considered, broadly speaking, as opposites of one another:

- in particular, it wouldn't normally make sense to apply both of them at the same time in the same situation (you can't be both wealthy and poor in the same sense at the same time)

If in doubt about which of the pair to recode, we usually pick X as the primary member of the pair if it is:

- usually considered as positive / beneficial / valuable
- and/or usually associated with "more" of something rather than "less" of something.

⚒️ Colour of factor borders (not currently implemented)#

But there is more: the border colour could reflect the overall polarity of the factor.

One possibility is: The more *plus* incoming and outgoing arrows there are, the greener it is, and the more *minus* arrows there are, the redder it is. If the balance is equal, the border is light grey. If a factor has a red border, that means that at least mostly, its *opposite* was mentioned. So in this example, happy has a grey border because it was mentioned once and its opposite was mentioned once. Fit has a slightly pink border because it has one incoming and one outgoing plusses, and one incoming and two outgoing minuses.

However this reflects information which is actually already visible in the diagram. A second possibility is to use the same green-red scale to reflect the proportion of factors contributing to a factor which have been flipped: this information can not otherwise be deduced from the map. It becomes particularly relevant when coding opposites within a hierarchy, see below. However this colour scheme might be misinterpreted, as it is only affected by the number and polarity of factors which have been collapsed and not by the number of links.

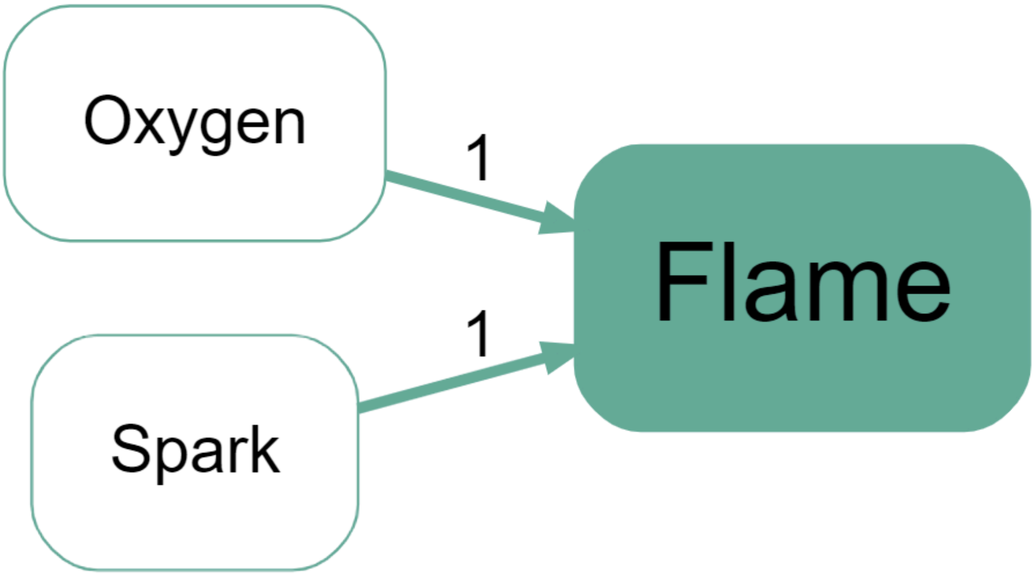
The total of the numbers on the arrows (here all the numbers are 1 so they are not shown) into and out of a factor is its citation count. But now we have additional information and its citation count is also equal to the number of times it was mentioned in plus form and the number of times it was mentioned in the opposite form.

Context

In this section, we'll look at "Context" as it appears to us at Causal Map when we do actual (qualitative) causal mapping: taking causal claims which real-life stakeholders actually make and trying to encode them in as systematic a way as possible. We believe that doing qualitative causal mapping is a really good testbed for theoretical ideas in evaluation and social science: do they fit with what people actually say?

So, how do we code this in causal terms: "When enough oxygen is present, a spark will always cause a flame"? (How) can we distinguish between a context like Oxygen and just some causal factor like Flame? How do we encode a context in a causal map?

We can draw a line from Spark to Flame, but what do we do about Oxygen? Drawing a line also from Oxygen to Flame doesn't seem to capture the context-ness of Oxygen.



Version 1: context is just a causal factor

In general terms this diagram is ok.

But we know the relationship between Spark and Oxygen to Flame is a *partial* function:

Oxygen	**Spark**	**Flame**
Yes	Yes	Yes
Yes	No	No
No	Yes	?
No	No	?

We know how Spark controls Flame only *given* Oxygen, but not when there is no Oxygen. This suggests that we cannot deal with a contextual factor as an ordinary causal factor. It is a special one which causally *enables* causal link(s) between other factors.

This *lack* of information about what happens in the *absence* of the contextual factor means that it acts like a *sufficient* condition for the causal relationship. (The defining characteristic of a sufficient condition is that no claim is made about what happens in its absence, just as the defining characteristic of a necessary condition is that no claim is made about what happens when it is present (only when it is absent).

Inside the context Oxygen, Spark is a necessary and sufficient cause of Flame (or so our respondent tells us). Outside the context C, Spark never makes any difference to Flame (or, we have no information about the effect, which isn't the same thing, but it doesn't matter here.)

Perhaps it is this very absence of information about what happens with no Oxygen which makes Oxygen feel more like a context rather than an ordinary causal factor. You can't see that in this first diagram.

Of course there are other things which you could mention as part of the context: there is fuel, it's dry, etc.

There are many ways to encode contextual information in a Causal Map; here we suggest storing the information inside each link.

Using context in the Causal Map app#

At the moment, you can encode context in the Causal Map app simply by creating a hashtag for it. You can use a family of hashtags by using some characters in common, like "Context:", like this:

Untitled

You can also search by and filter for contexts ✨ Transforms Filters: Include or exclude hashtags, so you can show a map only in the context of oxygen or without that context.

Plain coding

Summary#

Causal mapping doesn't usually deal with the kind of non-causal themes which are the focus of ordinary QDA (like in NVivo!). However sometimes it can be really useful to be able to simply note the presence of something without any causal connection.

We call this "plain coding". You can use it for:

1. Noting the presence of something which is not mentioned as part of a causal link in the statement you are coding but does appear *elsewhere* as a causal factor as part of a causal link.
2. …Or noting the presence of something which is "nothing but" a theme and never appears in causal coding.

Causal Map makes this possible in a simple way: we define a "plain coding" as simply a link from a factor to itself with has the hashtag `#plain_coding`.

Factors involved in plain codings (whether some of the time or all of the time) can of course still be involved in hierarchies and opposites coding just like any other factor.

So by default, plain codings will still appear in maps, as self-loops from the factor to itself, although it is possible to exclude such links by using the appropriate transforms filter.

Doing plain coding like this has the big advantage that factors coded with plain codings will still appear in the tables in the ordinary way, so for example the plain coding to (and from) the factor increased income will still count towards its source and citation counts.

It also means that you can easily delete a plain coding from a map by clicking on the link.

The hashtag `#plain_coding` distinguishes plain codings from actual causal links where a factor indeed influences itself, for example if someone says that increased income led to even more increased income – in this case the hashtag is not used. This hashtag can also be used together with other hashtags for the same link in the usual way.

How the Causal Map app implements this:#

See [[#causal-overlay]]

Table features -- Statistical tests of group differences

Summary#

How can we compare different groups like districts, gender or questionnaire sections, within maps and tables?

Reporting global and local network statistics

TODO

Simplification - factor and link frequency

Simplification - hierarchical zooming

TODO

Causal mapping produces models you can query to answer questions

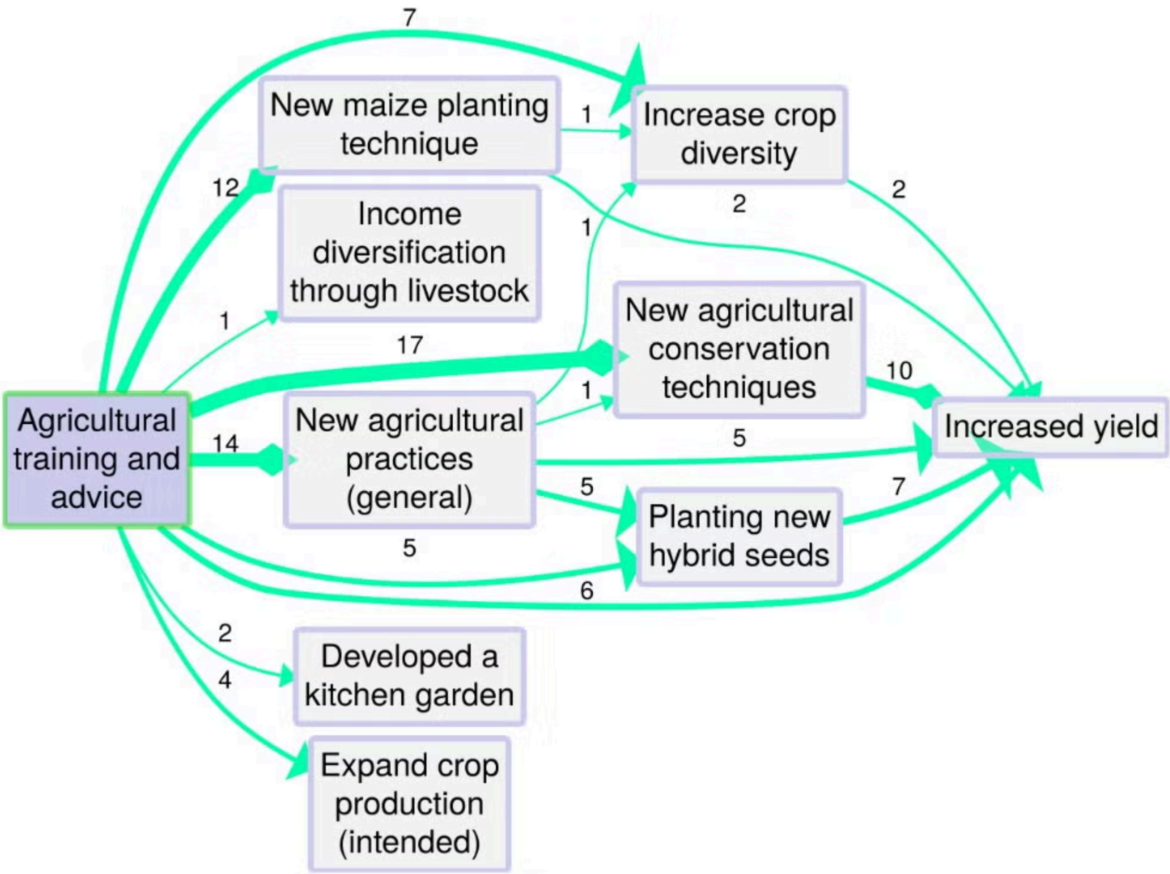
The fundamental output of causal mapping is a database of causal links. If there are not too many links, this database can be visualised "as-is" in the form of a causal map or network. But usually there are too many links for this to be very useful, so we apply filters.

By applying filters and other algorithms, a causal map can be queried in different ways to answer different questions, for example to simplify it, to trace specific causal paths, to identify significantly different sub-maps for different groups of sources, etc.

As explained on the [Causal Mapping website](#):

"*A global causal map resulting from a research project can contain a large number of links and causal factors. By applying filters and other algorithms, a causal map can be queried in different ways to answer different questions, for example to simplify it, to trace specific causal paths, to identify significantly different sub-maps for different groups of sources, etc.*"

The figure below shows a map from the application Causal Map, showing coded causal statements for a project that provided farmers with agricultural training and advice in order to increase crop yields. The map has been filtered to show only outcomes downstream of the influence factor 'Agricultural training and advice'. Numbers shown indicate how many times the links were mentioned across all interviews.



Source: BDSR, 2021, p 4

Outputs of QDA

The logic of QDA: you've done your analysis, now what?

The result of qualitative data analysis can be understood as, at least, some kind of qualitative theory or model at least of the sources' beliefs, with at least some possibility of generalising beyond them. But it can be hard to know what to do with the results of an emergent qualitative text analysis. There is no clear decision procedure: we can ask the author, and the answer is: some explanation, i.e. more text. In more reproducible approaches we do get some more structured outputs such as tables of frequencies. Some authors such as Mayring see these kinds of outputs as an important analysis result. QDA software is often used to capture and structure and even make inferences with these kinds of outputs.

In the logic of (non-causal) reproducible QDA, we can do things like this:

- - count occurrences of concepts, and use ordinary arithmetic to report eg which of two concepts was more common
- - count co-occurrences of two concepts, and construct measures like association between concepts, and more generally combine and query occurrences with boolean logic
- - create case/code matrices
- - report relationships between sources and concepts, for example to compare codings of one concept for different genders
- - reason about concepts, for example to deduce that an occurrence of "lion" is also an occurrence of "mammal", either relying on our implicit understanding of the concepts or through the explicit declaration of a parent-child relationship.

Of course frequency statistics are notoriously unstable, because they depend on our decisions about granularity and chunking. If I have a codebook which has 100 different codes for cats and only 1 code for dogs, we may conclude that dogs were mentioned in the text more often than any other animal-concept even if cat-concepts were mentioned more often in combination. This is one reason why reasoning with these kinds of outputs can never be merely automated. There always has to be a "human in the loop".

Nevertheless the point is that we can understand the output of QDA coding as some proportion of "more text", which itself needs to be interpreted by humans, and a complementary proportion of machine-readable, structured output which can be used to ask and answer questions (Which are the overarching themes? How much does climate anxiety come up as a theme? Who mentions it most?) at least somewhat independently of human guidance.

QDA logic can also be extended beyond the simple logic of frequencies and occurrences to apply (special kinds of) codes which have additional explicit rules associated with them, such as code weighting (as for example in MaxQDA). This means we can for example apply codes like "somewhat happy" or 'very unhappy' which enable us to say that the expression of happiness in one case is stronger than the other, or (if we also allow coding for time) that happiness increases or decreases over time. These extra deductions we can make come free with the (implicit or explicit) underlying ordinal logic of comparison of intensity.

QDA without coding#

Coding does not have to be central to qualitative data analysis (Morgan, 2025); (Nguyen-Trung & Nguyen, 2025). ...

References

Morgan (2025). *Query-Based Analysis: A Strategy for Analyzing Qualitative Data Using ChatGPT*.

Nguyen-Trung, & Nguyen (2025). *Narrative-Integrated Thematic Analysis (NITA): AI-Supported Theme Generation Without Coding*. https://doi.org/10.31219/osf.io/7zs9c_v1.

The fundamental property of causal maps is transitivity

How does causal inference work in a causal *network*?

When is a pathway not just a link?

The logic around how links might combine into pathways and what that means for evaluation, that's the most exciting part. e.g. how might this intervention influence an outcome which might be multiple steps downstream of it?

From

> a -->b

and

> b --> c

what can we conclude about

> a --> c. ?

For example, if the relation --> means "causes", when and under what circumstances can we conclude that a causes c?

Once we know the inference rules for a network, in particular the transitivity rule, we can infer all kinds of useful things about it.

There is a whole library of thinking about causal reasoning within a statistical or probablistic network.

There is less written about qualitative causality within a qualitative causal network.

But our problem is harder again: to reason with what we call a causal map, where the links are about **beliefs about** or **evidence for** a causal connection.

[We can reason about causal maps using a logic of evidence]

Epistemic logic does not help us with reasoning about causal maps

(An example of kind-of qualitative causal logic, with a focus on groups: @castellaniCasebasedSystemsMapping])

> From (Powell et al., 2024)

Seen as models of the world, causal maps, like systems maps, are fallible but useful: We can use inference rules (which are explicitly set out in FCMs, SDs, BBNs and CLDs and are implicit in other related approaches), and in particular, transitivity rules, to make deductions about the world.

There are at least three problems of transitivity which we need to think about

1. Given that A influences B and B influences C, does A influence C?
2. Given that P believes that A influences B and P believes that B influences C, does P believe that A influence C?
3. Given that someone believes that A influences B and someone else P believes that B influences C, does someone (who? we? the people?) believe that A influence C?

So if A causes B and B causes C, causal logic might tell us the answer to 1) under what circumstances A causes C.

Seen as models of individuals' causal beliefs, we can arguably use analogous rules to make deductions about what individuals believe, or ought to believe, given what else they believe.

There is a thing called epistemic logic which is a strange shadow of causal logic. Can it help us answer 2 and 3?

But epistemic logic is a strange thing.

If a person P believes that A causes B and B causes C, epistemic logic tells us what P believes about A causing C *if they were a rational person*. Whereas, facts about what people actually do believe is a branch of psychology.

In the last decades, thinkers like Daniel Kahneman have shown that in this sense, humans are so far from rational that it does not make sense even to start off with a rationality assumption and then add some corrections.

It would be great to use causal maps to infer, given a bunch of information about different people's causal beliefs, what they believe about *other* causal connections. That would be really useful. But it is hard.

There is a much easier way to reason with causal maps which is also vital for evaluators: to reason about **evidence**.

We can reason about causal maps using a logic of evidence

References

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

We can reason about causal maps using a logic of evidence

> From (Powell et al., 2024)

Evaluators can break the Janus dilemma and make the best use of causal maps in evaluation by considering causal maps not primarily as models of either beliefs or facts but as repositories of causal evidence. We can use more-or-less explicit rules of deduction, not to make inferences about beliefs, nor directly about the world, but to organise evidence: to ask and answer questions such as:

• Is there any evidence that X influences Z?

• . . . directly, or indirectly?

• . . . if so, how much?

• Is there more or less evidence for any path from X to Z compared to any path from W to Z?

• How many sources mentioned a path from X to Z?

• . . . of these, how many sources were reliable?

We also argue that this is a good way of understanding what evaluators are already doing: gathering and assembling data from different sources about causal connections in order to weigh up the evidence for pathways of particular interest, like the pathways from an intervention to an outcome.

References

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Context is critical to the logic of evidence

These questions may depend on a somewhat hidden assumption: that all the causal claims (the links) come from a single context. Such as, in most cases, when all the claims are all agreed on by a group as in participatory systems mapping (PSM). For example, when we wrote "which factors are reported as being causally central?", can we really answer that by simply checking the network? From:

> Factor X is central within these claims

Can we deduce:

> Factor X is claimed to be central

Or, can we go from:

> There are two relatively separate groups of causal claims?

Can we deduce:

> It is claimed that there are two relatively separate groups of causal claims?

In general, no. It is easy to think of counter-examples. If we ask a parents and children about the causal network surrounding family disputes, we might get two relatively separate causal networks with only a little overlap. From this we cannot conclude that these respondents taken together claim that there are two relatively separate systems. It might be that the parents and children indeed are giving information about relatively separate systems about which they each have the best information, or it might be that the two groups are telling conflicting and perhaps incompatible stories.

We could express this as, say, the first axiom of causal mapping:

> If a network of causal evidence from context C has property P, we can conclude that there is evidence that the corresponding causal network has property P, but again only in context C.

P(E(N)) --> E(P(N))

We often assume that contexts are sources and sources are contexts. But this is not always the case. For example one respondent might give two sets of information, one from before losing their job and one set from afterwards, without trying to encode the job loss as a causal factor within the network of claims.

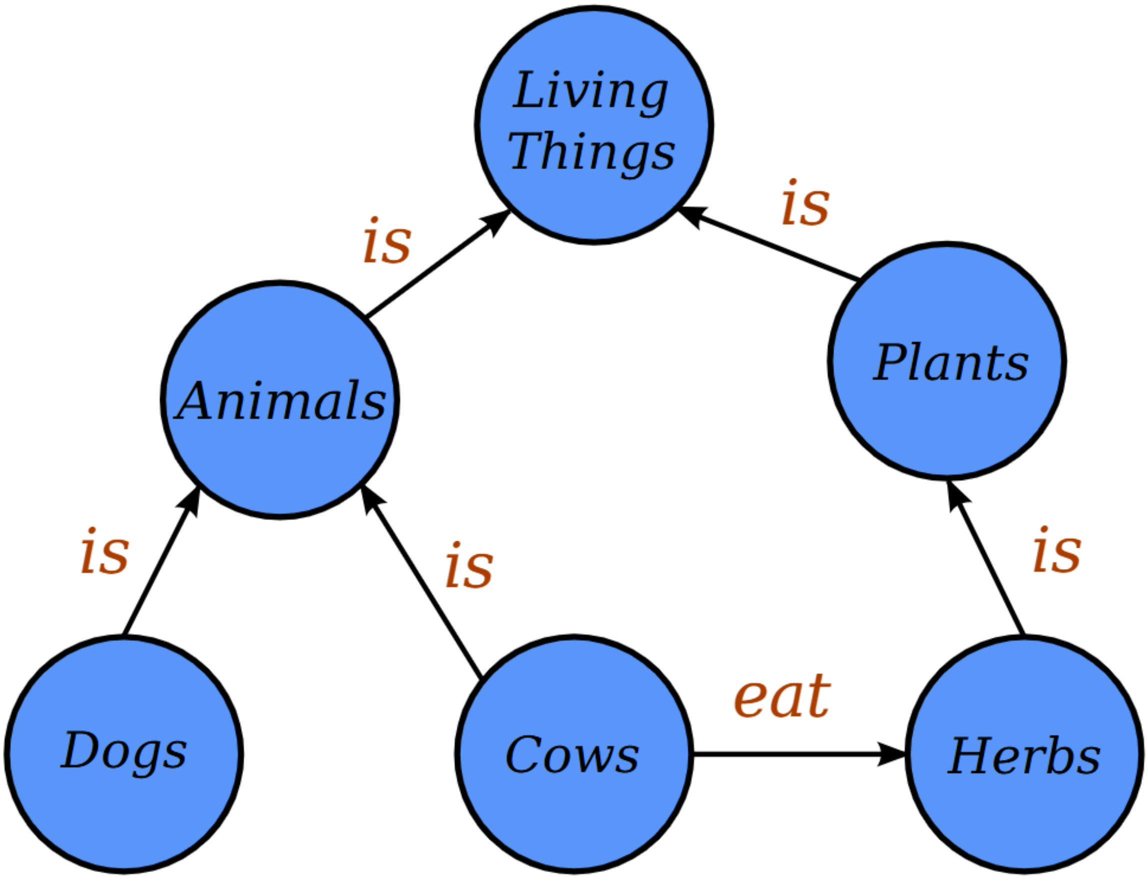
In a PSM workshop, there may be multiple respondents but, as long as they construct a consensus map, these are all treated as one source. Part of the job of the moderator is also to ensure that the claims (evidence) all come from one context, which is the same as saying: we can validly make inferences like those above. I don't know whether PSM moderators actually do this.

You might say "this is all pointless because it depends what you mean by context", and that is exactly true. All we have done is

Causal maps are knowledge graphs, but with wings

What is a Knowledge Graph? 🧠

- A knowledge graph is like a **giant mind map for a computer**. It stores information not as text in a document, but as a network of interconnected facts.
- It's built from two main things: **entities** (the "nodes," representing real-world objects, people, or concepts like "Paris" or "Photosynthesis") and **relationships** (the "edges," describing how these entities are connected, like "is the capital of" or "is a process in").
- A single fact has three parts: **(Subject) --- [Relationship] ---> (Object)**. For example: `(Marie Curie) --- [discovered] ---> (Radium)`.
- Why are knowledge graphs specially useful in the age of AI? 💡
- **They create structure from chaos.** AI can read through millions of pages of unstructured text (like news articles or scientific papers) and pull out these factual triplets. This turns a messy sea of words into an organized, queryable database of knowledge.
- **They enable smarter searching and reasoning.** Instead of just searching for keywords, you can ask complex questions that require understanding the relationships between things. For example, "Which scientists who won a Nobel Prize also discovered an element?" A computer can navigate the graph's connections to find the answer.
- **They provide essential context.** A knowledge graph helps an AI understand that "Apple" in a tech article is a company linked to "Steve Jobs," not the fruit. By looking at its connections, the AI gets the right context, which is crucial for accurate understanding and analysis.



[Image](#) from Wikipedia by Jayarathina - Own work, CC BY-SA 4.0.

Why are Knowledge Graphs (KGs) so useful?[#](#)

A major benefit of KGs is we can then apply network logic like transitivity rules to answer meaningful questions. For example, if the relation is "works in the same company as", then if we know

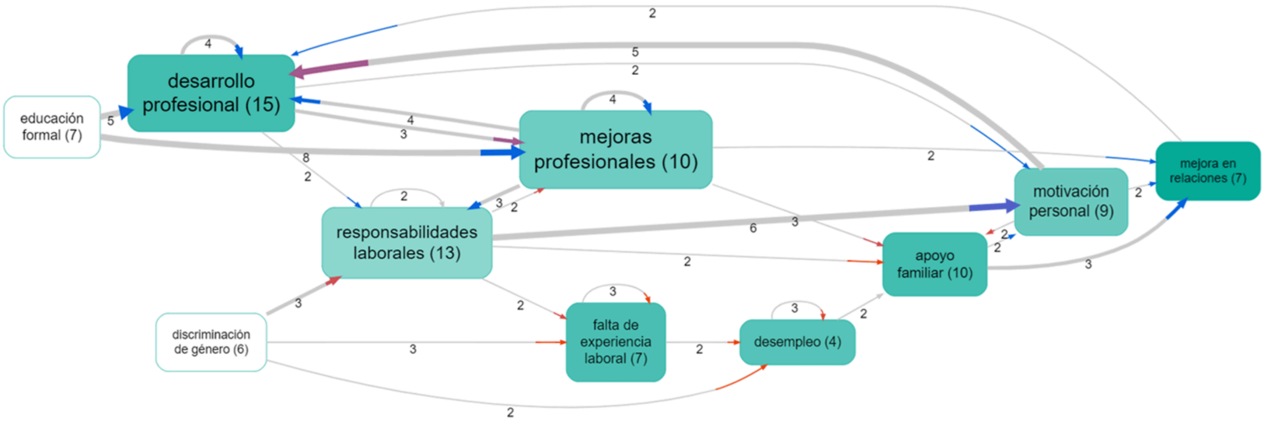
- **A is related to B**
- and
- **B is related to C,**
- then we can conclude
- **A is related to C (A is in the same company as C).**

Challenges with general-purpose knowledge graphs[#](#)

The trick in **constructing** knowledge graphs is to know what relationship(s) to look for. "belongs to?" "is capital of?" "challenges/undermines?" This can be very difficult to decide. on the fly.

Using network logic to answer queries can be difficult where each different type of relationship may have its own logic. It can be very tricky (though potentially rewarding and useful) to design custom queries to answer specific questions.

Causal mapping gives knowledge graphs wings#



A causal map is just a knowledge graph in which there is only one kind of relation: "causes" or "influences". This means:

- It is **much easier to scan and process text data** as we already know what we are looking for.
- We focus on primarily on **exactly the kind of information which is useful for monitoring and evaluation**: what influences what?
- We can **make use of pre-existing logic and queries to help answer common evaluation questions** almost "out of the box". Here we have a whole presentation on *Questions you can answer with causal mapping* which gives plenty of suggestions.

Can only doing causal mapping answer all the questions you might want to ask about a text? Of course not. But it can help answer a lot of the most interesting and important ones.

What about social network analysis?#

Yes, social networks can also be constructed as knowledge graphs with just one (or a small number of) relationships, such as "works with".

So can we use causal mapping tools to construct general network graphs?#

You might ask if the reverse is also true: can you use causal mapping software like Causal Map to also do your AI-supported knowledge graphing for you? The answer is yes! Concretely, in the new version of Causal Map, version 4, which is arriving very soon, you can manually code any type of link, not just causal, and you can also guide the AI to do this too.

The product of (causal) qualitative coding can be a model you can query

The result of qualitative coding of texts: is it a model?#

Every type of Qualitative Data Analysis (QDA) is a bit different....

1) The most important result of coding might be a **deeper understanding of the text**. Something the researchers have in their heads. Perhaps something shared within a group. They can answer new questions about the data and write summaries of it from a new angle. In a sense you could say they have more or less formally developed a **model** of the data or even a theory about it, or, in the sense of Grounded Theory, a theory which goes beyond that data but is partly inspired by it. If you yourself read some of these different outputs and engage with them, maybe you yourself can start to build such a model in your head.

2) Apart from the intangible products in researchers' heads there are also **tangible products such as research reports**. Different schools of QDA give different importance to tangible as opposed to intangible outputs.

3) Thirdly, there may be **tabular outputs** such as tables of coding frequencies, cross-tabulations, etc. The definitive set of tables can be queried to answer additional questions like "how many women who mentioned working from home also see television as a an outdated medium". These kinds of tabular outputs can be quite useful to answer different questions.

So when you do an evaluation, what's the product? What do you get?

Obviously, you can think of the evaluation report, which might answer predetermined questions, but it may also include material that goes beyond the specific questions we were tasked with answering—for example, to address unanticipated issues or simply to describe or contextualize. Another important output is relational: hopefully, people have come together in a way that helps to expand learning and perhaps develop projects or relationships.

But today I want to talk about something different.

A statistical analysis not only answers questions but gives you the whole model -- can a qual analysis do that too?#

When you do quantitative research, you might have specific research questions, but often one of the major outputs is a statistical model of the phenomenon. (There might be an effort to go beyond the data and hope that the model generalizes more widely, but that's not my focus here.) In the simplest case, the model might represent a suspected causal relationship—say, between the amount of screen time in the evening and difficulty falling asleep.

At the very least, the model allows us to look at a case in the dataset and say: on this day, this person looked at a screen for three hours and rated, let's say, a difficulty of four out of five falling asleep on some self-rating scale. Because we have the model, **we can explain that**: yes, this is quite a high level of difficulty, and it's explained at least partly by a high level of screen time, at least in this individual case. The model might also enable us to **make predictions**, like: people who, at least in this context, spend more than three hours on screens in the evening are, on average, going to experience a higher level of difficulty falling asleep.

A more sophisticated model will probably capture more variables, and many models—like directed acyclic graphs—link up these kinds of connections into a causal network, so you can explain or even predict how tweaking one variable will affect another variable downstream of it. Of course, there are other kinds of statistical models apart from causal models, but if you're reading this, you love causal models, don't you?

There's a relatively small but extremely well-funded section of evaluation activity based around this kind of statistical causal model, with randomized control trials (RCTs) as one facet. Ideally, an RCT is tasked not primarily with producing a model, but with answering a specific question, like: which is the better of these two interventions? Or: does this treatment work better than a placebo? But these are calculations conducted on the underlying model, which, from the point of view of workflow, is the major output of the work.

To generalise or not to generalise#

There are two ways you might want to use that kind of model.

- One is to answer further questions about the same dataset—for example, to ask whether a particular subset (say, people over 70) differ in how screen time influences difficulty falling asleep, compared to other subgroups.
- If it's a sophisticated model, it might allow us to *generalize beyond* this specific context, perhaps by including more general variables like attention style or eye movement speed, which might help explain and predict behavior in other contexts.

Most quantitative researchers make a big deal about generalizing the model beyond the specific use case or context. In fact, the whole point of the study is normally to do that, and there's a whole armoury of tools, concepts, and arguments about how and under what conditions you can generalize a model to other people, other years, or even other countries with different kinds of screens, etc.

Can qualitative research do that?#

The majority of evaluations aren't like that, although they might include a specific quantitative question somewhere in their terms of reference. In most cases, we think of the research output as a report in which the original (possibly modified) list of questions is answered, with additional narrative to summarize and link these sections.

Now, going beyond strictly quantitative paradigms, some evaluation projects will also include what we might call **qualitative modelling**. For example, if we're using QCA (Qualitative Comparative Analysis), apart from answering specific questions, we've likely also produced QCA-style tables, which could help us answer other questions beyond those we were actually tasked with. You might see those tables as annexes to the report. The same goes for causal loop diagrams and other techniques, which are essentially quantitative models but with a more restricted set of numbers. For example, in causal loop diagrams, we might model a variable like inflation with a number from -1 through 0 to +1, and do the same for variables like unemployment or military threats, building models of the relationships between these things using simplified numbers.

What I want to argue here is:

> any halfway decent evaluation, which at least implicitly gathers qualitative information about how things within the evaluation influence one another, can be considered as constructing a qualitative causal model. This is irrespective of the specific methods used—even if it doesn't include something explicitly called causal pathways analysis or causal mapping.

Qualitative models as products: theory, model, or something else?#

In quantitative research, the idea of a "model" as a product is well established: you build a model, and then you can query it to answer new questions—even ones you didn't anticipate at the start. But what about qualitative research? Can the result of a qualitative analysis be a model in this sense, rather than just a set of answers to specific questions or a summary?

Of course (some) qualitative research produces models. Just don't call them that.#

Some qualitative researchers do indeed conceptualize their results as models. For example, grounded theory often produces a theoretical model that explains the underlying processes or relationships within the data. These models can be revisited and "queried" to generate new insights beyond the initial research questions.

However, many qualitative researchers are more comfortable with the term "theory" rather than "model." "Theory" aligns more closely with the interpretive and conceptual nature of qualitative work, emphasizing explanation and understanding rather than prediction or parameterization. Still, the distinction is often more about language than substance: both models and theories can serve as frameworks for making sense of data and for generating new questions.

How are qualitative models used?#

In qualitative research, especially in fields like grounded theory and narrative inquiry, the focus is less on "prediction" and more on generating understanding or insight. Researchers talk about "theorizing" from the data—developing concepts and frameworks that explain the phenomena under study. Once a theory or model is developed, it can be revisited, interrogated, and applied to new data or different contexts. This iterative process allows for continual refinement and deeper insight.

Importantly, a qualitative model or theory can also be used to answer new questions about the same dataset. For example, after developing a grounded theory, researchers (or others) can return to the theory and use it as a lens to interpret further cases, refine concepts, or generate new insights—

without having to re-examine all the original data. This practice is sometimes referred to as "secondary analysis" or "theoretical application," where the theory or model functions as a standalone analytical tool.

In causal mapping, for instance, the model might consist of a network of causal links derived from qualitative data. Even if there are no quantitative parameters on the links, the model can still be queried: "Is there evidence for a causal pathway from A to B?" or "What are the main factors influencing outcome X?" This allows the model to be used flexibly, supporting both anticipated and unanticipated lines of inquiry.

Why does this matter?#

Thinking of qualitative research outputs as models (or theories) that can be queried and reused has several advantages:

- **Transparency:** It makes explicit the structure of the findings and how they relate to the data.
- **Reusability:** Others can use the model/theory to answer new questions or apply it in new contexts.
- **Iterative learning:** The model can be refined and expanded as new data or perspectives emerge.
- **Bridging paradigms:** It helps bridge the gap between qualitative and quantitative traditions, showing that both can produce structured, interrogable outputs.

In summary, while qualitative researchers may prefer the language of "theory" over "model," the idea is the same: a well-constructed qualitative analysis can produce a framework that is more than just a set of answers—it is a model of the phenomenon, one that can be queried, shared, and built upon.

The transitivity trap

> From (Powell et al., 2024)

Granularity, generalisability and chunking are coding problems for causal mapping too

Transitivity is perhaps the single most important challenge for causal mapping. Consider the following example. If source P [pig farmer] states 'I received cash grant compensation for pig diseases [G], so I had more cash [C]', and source W [wheat farmer] states 'I had more cash [C], so I bought more seeds [S]', can we then deduce that pig diseases lead to more cash which leads to more seed (G à C àS), and therefore G à S (there is evidence for an indirect effect of G on S, i.e. that cash grants for pig diseases lead to people buying more seeds)?

The answer is of course that we cannot because the first part only makes sense for pig farmers, and the second part only makes sense for wheat farmers. In general, from G à C (in context P) and C à S (in context W), we can only conclude that G à S in the intersection of the contexts P and W. Correctly making inferences about indirect effects is the key benefit but also the key challenge for any approach which uses causal diagrams or maps, including quantitative approaches (Bollen, 1987).

> For want of a nail the shoe was lost,
>
> For want of a shoe the horse was lost,
>
> For want of a horse the rider was lost,
>
> For want of a rider the battle war lost,
>
> For want of a battle the kingdom was lost,
>
> And all for the want of horseshoe nail.

(Thanks to Gary Goertz for remembering this one!)

Frog thinks: eating salad leads to health (less scurvy), and health (general fitness) leads to better sprinting ability, therefore if I eat this yummy lettuce – AARGH!

One of the key features of causal maps is that you can draw inferences, make deductions, from them. One of the most exciting is to be able to trace causal influences down a chain of causal links. BUT, when you are drawing conclusions from causal maps, beware of the transitivity trap:

from

> B → C

and

> C → E

we can only conclude

> B → E in the intersection of the contexts of 1 and 2

… and in general with any causal mapping, you'll never be sure that these two contexts do intersect. You actually have to look at each chain and think about it, and hope you've been told all the relevant facts.

For example:

If

> Source P [pig farmer]: I received cash grant compensation for pig diseases (G), so I had more cash (C)

and

> Source W [wheat farmer]: I had more cash (C), so I bought more seeds (S)

can we deduce

> G → C → H

and therefore

> G → S

(cash grants for pig diseases lead to people buying more seeds)?

No, we can't, because the first part only makes sense for pig farmers and the second part only makes sense for wheat farmers.

There are thousands of different kinds of transitivity trap. It isn't just a problem across subgroups of people. It can apply for example in different time frames.

If

> Child does well in year 13 (A) → Child has improved academic self-image (C)

and

> Child has improved academic self-image (C) → Child does better in year 9 (D)

can we deduce

> A → C → D

and therefore

> A → D

(child doing well in year 13 leads to child doing well in year 9)?

Of course not - even though these claims might be true of the same child. The problem arises as soon as we generalise one causal factor to apply to different contexts. We have to do this, to make useful knowledge. But there are always pitfalls too.

Not just a problem for causal mapping#

This is also true, isn't it, of any synthetic research / literature review?

And in statistics, knowing the effects from B → C and C → E means you can calculate the indirect effect of B on E but not the direct effect. You have to have additional data just for that. This is one source of various so-called paradoxes in statistics.

Can we mitigate the trap with careful elicitation protocols?#

Sometimes, we might know that all the information in one particular chain came from the same source, and all this information was explicitly given as a series of explanations of the factor which was initially in focus. But even here, we have to be careful. We might have to ask again, having reached the end of the chain, "did B really influence C which influenced D which influenced E? Was this all part of the same mechanism?" Are we sure we know exactly what we mean by this, and are we sure that our respondents do too?

In any case, part of the point of causal mapping is the synthetic surprises which we can discover by piecing together fragments of causal information which were *not* necessarily provided in this way. This is the situation every evaluator is in when piecing together information from, say, experts for Phase 1 and experts for Phase 2. We just always have to be aware of the transitivity trap.

Transitivity trap, or identity trap?#

We can talk about the *identity trap* as more fundamental than the transitivity trap. It comes down to saying, how can you be sure that the way in which this factor is exemplified in one particular context is the same as the way that this similar seeming factor is exemplified in a different context: whether to use "the same" factor to code two different things.

References

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Individual questions -- introduction

Because the output is a structured network, we can apply a range of queries to explore the data. This gives us a library of **pre-existing approaches** to ask **practical questions** about the causal landscape described by the participants...

Individual views -- How does the system work according to individual sources?

This question asks about **individual perspectives** on mechanisms and outcomes. By viewing the results for each source in turn you can identify sources reporting specific and unusual pathways of change and analyse individual viewpoints in depth.

This can help tailor a program approach to address specific concerns or leverage unique knowledge.

To get this view, filter your map to show only causal claims made by this particular respondent.

Summarising -- How do the sources claim that the system works, in summary?

Creating **overall maps** provides a **top-level view** of what your sources report about the system.

This can for example help in understanding the mental models of your stakeholders, which may differ from your own or your project's theory of change. This allows you to identify what your stakeholders collectively believe are the project's most important mechanisms and outcomes.

While simply listing the main factors and links is interesting, it can be much more informative to view the map which remains when we filter the links table to retain only the most important rows. There are eight ways to do this, resulting as a combination of these three choices:

- Filter factors *versus* links
- Filter with respect to sources *versus* citations
- Filter by retaining only the top n items *versus* retaining only the items with a count of at least n
- So for example we might filter to retain only the *top* 10 *factors* by *citation count*. Often we find it useful to use two frequency filters: an initial filter for only say the top 10 factors may still result in so many links between these factors that the map is still hard to understand, so we add another filter, this time filtering on *links* frequency.

Viewing your map like this helps you identify common themes, key pathways, and shared assumptions about a project or system. You can use these insights e.g. for aligning stakeholder expectations and project goals.

Filtering for the top factors and/or links#

- Link frequency filter
- Factor frequency filter

Filtering for the top, say, 5 or 10 or 15 links is a good way to produce relatively consistent maps and tables.

You can also format the links and factors in your map to reflect the data, e.g. you can make factors bigger if they were mentioned more often.

Note that if you select the top 10 links or factors, you may not get exactly 10 due to ties.

Zooming out#

Using Hierarchical coding is a great way to bring structure into a complicated list of factors. If you have done that, you can use to simplify your map.

Vignettes -- What is a typical source and what is their story?

Writing the story for a typical source helps to create an engaging 'human' story, bringing your data to life, providing a **tangible example** that represents the broader trends in your dataset.

There are different ways to do this. One way is to apply an algorithm to the links database to identify the most typical source, and then to construct a **short story** to illustrate the map, complete with quotes from that source.

This is available in Causal Map using non-AI algorithms to identify the most typical source and then applying a simple genAI prompt to create a story based on that source's data.

What are the narratives behind a specific link?

It is important to frequently return to the original quotes associated with each factor or link to

understand how different stakeholders interpret and talk about key concepts. This can be done in Causal Map by clicking on a link in the interactive map, or by printing out quotes for a particular filter (e.g. just for a single bundle of links) with additional context and metadata.

Relevant page from Causal Map help:

Print View ▸

Which factors and links were most frequently mentioned?

Similar to overall frequency counts in non-causal qualitative coding ("which themes are mentioned most often?"), we can count the frequency with which particular factors (causes and effects) are mentioned: which causal factors are mentioned most often?

And we can, with caveats interpret the number of mentions of a particular *link* in a similar way: which causal connections are mentioned most often -- and we can, with similar caveats, interpret the frequency with which a link is mentioned as the strength of the evidence for that link.

We can simply make a table with the number of times each factor was mentioned or cited overall, and sort it in descending order to see the most important ones. In these tables the outgoing (the factor was mentioned as the influence of something else) and incoming (the factor was mentioned as the consequence of something else).

Which factors and links are mentioned by the most sources?

The number of citations can be a useful measure of importance, but it can be distorted if one source mentions some factor a lot of times. A useful alternative is to count the number of *sources* that mentioned a factor or link at least once.

Main outcomes. Which factors are mentioned most often as outcomes?

What do we even mean by “outcome”? A factor might have a lot of incoming links, so it is often mentioned as the consequence of something, but it might also have a lot of outgoing links, so perhaps plays more of an intermediate role. One way to answer this question is to calculate the “outcomeness” for each factor: the number of citations of its incoming links divided by the total number of times it was cited.

High outcomeness means that a factor was often at or near the end of causal stories. A table of factors listed in descending order of outcomeness can help for example, to better understand the perceived impacts of your project or implementation.

Causal mapping looks for linearity first ▸

Outcomeness is a useful measure of whether a factor is more of an outcome or a driver#

Causal Map primarily uses the Graphviz DOT layout engine which does an amazing job of teasing out such a story if there is one. Generally speaking, the "drivers" will be on the left and the outcomes will be on the right, but at the same time trying to maintain readability and avoid the links crossing over factors or over each other. Which is always a compromise. For this reason we also usually use "outcomeness" colouring for the factor backgrounds, which represents the proportion of all the factor's links which are incoming links. So normally factors on the right are darker, except where Graphviz has had to reposition some of the factors for readability. So, does it look like a ToC? Of course that depends on what you expect a ToC to look like and famously there are no standards for that.

If the causal map is more or less neatly hierarchical, then our map will reflect that nicely and therefore "look like a ToC" but of course that's rarely the case.

We often see that for reports, folks often take the original causal map and get a designer to redraw them anyway to match the report styling etc.

In the upcoming version 4 of Causal Map, there is a more interactive style of map where it's possible to drag the factors around to put them just where you want them.

Main drivers. Which factors are mentioned most often as drivers?

Just as we can identify the factors which are in an absolute or relative sense most frequently mentioned as outcomes, we can do almost the same to identify key drivers:

- Identify those with the highest number of outgoing links
- Or identify those with the lowest outcomeness score.

Splitting by groups. Are different groups involved in different ways?

The simplest way to compare groups is to make separate causal maps for each group (e.g.

men/women, project A/B, or by age group), and visually compare them. This allows you to identify **common patterns and context-specific factors.**

This is useful, for example, for understanding how a project impacts different groups.

Comparing groups -- What factors or links were mentioned more by some groups than others, in the same map?

We can directly **compare groups** to find factors or links mentioned more by one group than another

using a statistical test to find the most surprising differences between groups, taking into account the underlying frequencies.

Differences - sample#

Shows which links were preferentially mentioned according to different groups e.g. women more than men. We ask:

> does the proportion of women vs men who mention this link differ from what you would expect (given the total number of mentions of links by both women vs men)?

	Women	**Men**
... other links ...		
Number of mentions of the link from X to Y	10	9
... other links ...		
Total number of mentions of any link	**60**	**10**

In this case we can see that although women mentioned the link slightly more often than men, women altogether mentioned links twice as often as men. So we can compare the number of mentions of the link with the number of "non-mentions" of the link. So we can work out this table (not shown).

	Women	Men
Number of mentions of the link from X to Y	10	9
Number of mentions of any *other* link	50	1

We can do a simple chi-squared test on this table to see if the ratio 10:9 is significantly different from 50:1 (which of course it is) -- this is the same question as to whether 10:50 is significantly different from 9:1 (which of course it is). If this test is significant, the row "Number of mentions of the link from X to Y" is shown in the table, and the intensity of the colouring of each cell reflects its chi-squared residual, i.e. how different is the number it contains from the number you would expect, given the other numbers?

This comparison is agnostic as to whether there are, say, many men or a few men who talk a lot.

The tests for this are chi-squared tests. If the grouping factor is numerical we add an additional correction for ordinal scale so that the chi-squared test is not weaker than it should be.

Table features -- Statistical tests of group differences

Identifying groups -- Are there different subgroups within the data?

Where there are many sources, it may also be useful to identify which sub-groups of sources are substantially different from one another in terms of the stories they tell.

> From (Powell et al., 2024)

For example, (Markiczy & Goldberg, 1995) use dimension-reduction techniques on the table of all links reported by all sources to identify clusters of sources so that the members of each cluster are maximally similar to one another in terms of the links they report, and so that the clusters taken as a whole are maximally different from one another. These groups can also be cross-tabulated with existing metadata, to interpret them as, for example, young city-dwellers versus older rural residents.

References

Markiczy, & Goldberg (1995). *A Method for Eliciting and Comparing Causal Maps*. Sage Publications Sage CA: Thousand Oaks, CA.

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

What are the emerging or unexpected factors?

One way to do identify emerging or unexpected factors is to use the elements from your theory of change as your codebook while coding and only adding other elements when necessary, making a note of these additional elements.

Does the evidence support your theory of change?

One way to do this is to use the elements from your theory of change as your codebook and only add

other elements when necessary. Validated pathways of change (by showing which mechanisms are observed on the ground) and find gaps where expected pathways might be missing or where stakeholders list elements were not anticipated in the theory of change?

We discuss this approach at more length in (Powell et al., 2023).

References

Powell, Larquemin, Copestake, Remnant, & Avard (2023). *Does Our Theory Match Your Theory? Theories of Change and Causal Maps in Ghana*. In *Strategic Thinking, Design and the Theory of Change. A Framework for Designing Impactful and Transformational Social Interventions*.

Assessing systems change

One of the most exciting applications of causal mapping is to assess change over time within a system. If we apply a systematic approach to coding (using blindfolded manual coding or AI-supported coding) we can compare the frequencies with which links or factors are mentioned over time. This becomes particularly interesting when applying inductive coding, so that new and emerging phenomena can be included into the codebook. Re-applying new codes to previously coded data would be very tedious with manual coding but is easy to do with AI-supported coding: Transforms Filters -- Soft Recode with Magnetic Labels

More details are given in this paper: (Powell et al., 2025)

Unfortunately, there is not much consensus about what assessing systems change means. Sometimes we read about *measuring* systems change, which would imply assigning numbers to change, but often just means "assessing".

(Rizzardi, 2025)

References

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.

Rizzardi (2025). *Systems Change | Modern Slavery*. https://www.freedomfund.org/news/systems-change-pathways-measurement/.

Sentiment -- Which changes are perceived as most positive or negative?

- Analysing the sentiment expressed by your respondents and displaying it on your maps using positive (blue), negative (red) and ambivalent (grey) enables readers to **easily identify positive and negative connections.**
- This visual representation can quickly highlight areas of success and potential improvement.

Building on the previous example, you can analyse the sentiment in particular of the immediate impacts of your project or intervention.

Focusing on specific factors. What influences and outcomes are connected to a specific factor?

Focus on a particular **element of your project** to understand the **direct and indirect**

influences leading to a specific factor and all **outcomes it contributes to**. This helps in exploring the role of the factor as both an outcome and an influence within the causal system.

To get this view, filter all the links in your map to show only links connected to this factor.

Asking about all the immediate causes and/or effects of a particular factor or set of factors produces the local neighbourhood or "ego network". What are the immediate causes and effects of traffic accidents? Of melting glaciers?

By default we usually filter for the *immediate* causes and effects of one factor. But we can extend this:

- Show causes and/or effects more than one step removed.
- Suppress causes and show only effects, or vice-versa. This idea then overlaps with .
- Use a more strict algorithm to avoid the The transitivity trap as described here: .

Looking downstream. What are the direct and indirect consequences of one or more factors?

Map out all **direct and indirect paths** flowing from selected factors to:

- Understand both intended and unintended outcomes
- Identify feedback loops and circular relationships
- Understand the causal pathways from a specific factor and how it relates to other causal links and factors

To do this, show only factors which can be linked back to the starting factor within a given number of steps.

Looking upstream. What are the direct and indirect influences on one or more factors?

Looking upstream is just the opposite of looking downstream.

The map above helps you understand the upstream causes and contributing factors that lead to specific outcomes. Looking at the example above, we can see how multiple factors may influence health improvements through various pathways. Here's what this type of analysis allows you to do:

- Trace back all direct and indirect influences that contribute to specific outcomes
- Identify which factors are most influential in driving changes
- Map intervention points that could strengthen positive outcomes
- Recognise interconnections between different influence pathways

To focus on stories told in their entirety by individuals, we can again use source tracing.

Names of tables and fields

Columns and tables#

We can think of a causal map as a database consisting of two tables, the links table and the sources table. We don't need to have a separate table for the factors because the factors can be derived from the links table.

Columns in both the links and factors tables#

Field name	Explanation
Citation count aka Link count	Number of citations of a given factor or link.
Source count	Number of sources mentioning a given factor or link. Source count cannot be higher than citation count and may be a lot lower if some sources mentioned the same factor or link many times.

Columns in the factors table#

Field name	Explanation
incoming_links, indegree	Number of citations of all the incoming links to a particular factor.
outgoing_links, outdegree	Number of citations of all the outgoing links from a particular factor.
outcome-ness (%)	A factor with a high outcomeness percentage is mostly an outcome; it has mostly incoming links. If it has low outcomeness it has mostly outgoing links so it is mostly a driver. Outcomeness is the proportion of citations of incoming links out of all the citations of a particular factor: a normalised version of the Copeland Score (Copeland, 1951). So factors with high *outcomeness* can be thought of as "outcomes". And factors with low outcomeness can be thought of as inputs or drivers.

Columns in the links table#

Field name	Explanation

Columns in the sources table#

Field name	Explanation
source_id	
title	
filename	

Glossary ▸

References

Copeland (1951). *A Reasonable Social Welfare Function*.

Path tracing -- How do one or more causes affect one or more effects, including indirect pathways?

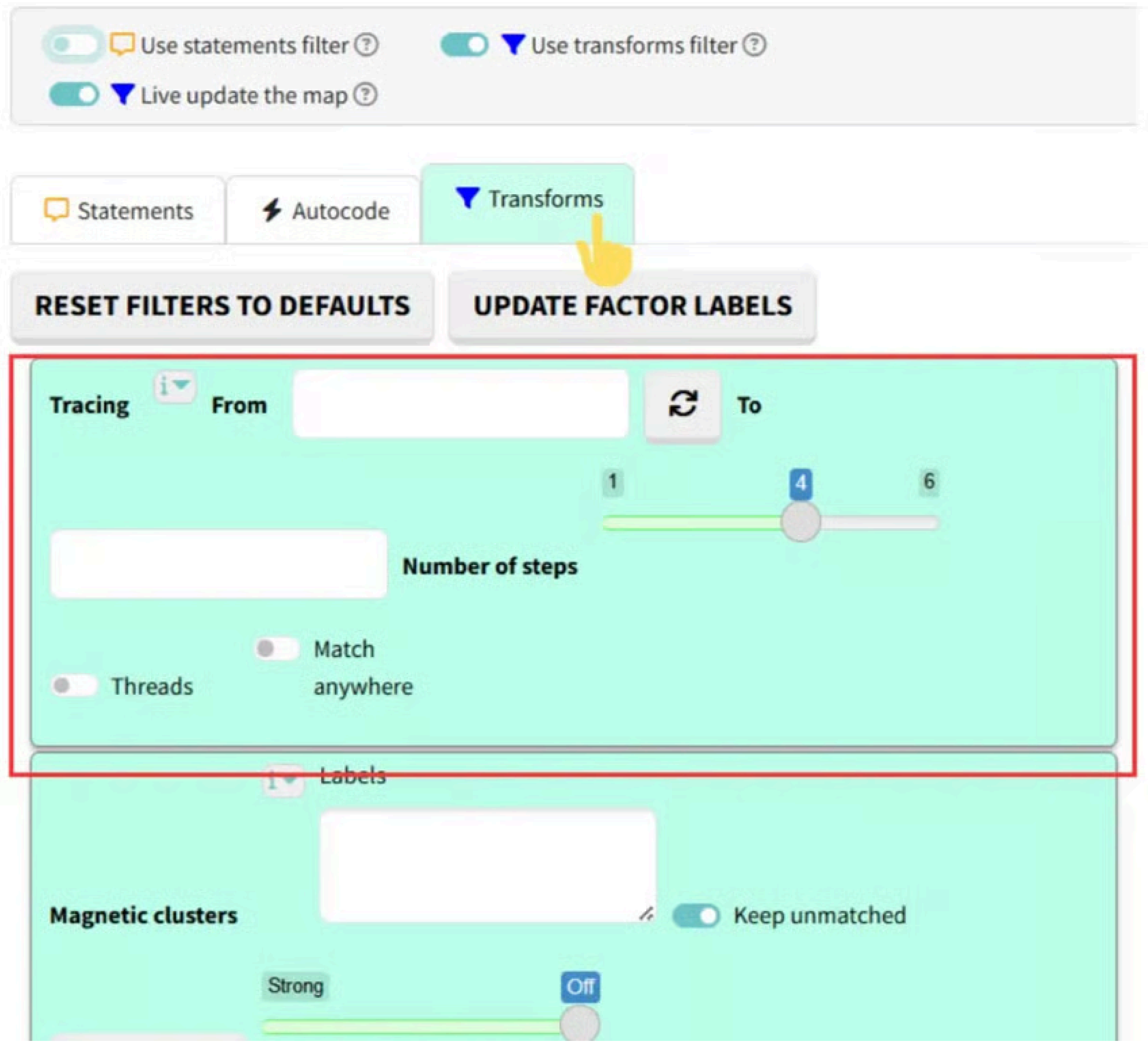
What are the main causal pathways from an intervention to an outcome? We can trace **chains of influence** from a starting point like an intervention to a key outcome, revealing the step-by-step or branching logic described by the sources. We can even compare the strength of evidence for different pathways.

Path tracing is similar to Looking upstream and Looking downstream,: you can specify the number of steps, and you can apply the more conservative source tracing approach. The difference is just that you can specify both the source and the target factors.

Source tracing -- What are the consequences of one or more factors, looking only at stories told in their entirety by individual sources?

How to use

🔍 **Start path tracing:** Click on the Transforms Filters panel and select the 'trace paths' filter.



➡️ **Define your path:** Specify the starting and ending factors for the path you want to trace. You can leave the `end` field blank to show all paths leading *from* a factor, or leave the `start` field blank to show all paths leading *to* a factor.

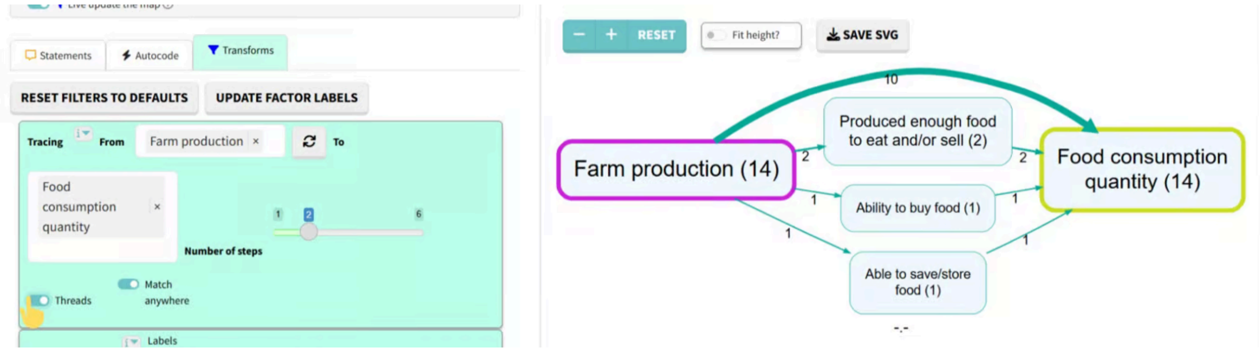
🔢 **Set the steps:** Choose the number of steps (or links) to include in your traced path. Increasing this number will *broaden* the results by including *longer paths*. But keep in mind that, when you're analysing interviews, **people usually don't report causal chains longer than 4 steps.**

The factors which match the filter are shown with **thicker borders.**

Tips for Success 💡 :

🗺️ **Simplify your map:** Consider applying other filters beforehand to format and simplify your map before tracing paths.

⛓️ **Avoid the [transitivity trap](#):** Be careful when drawing conclusions. The presence of links from A to B and B to C does not automatically mean that all respondents indirectly connect A to C: some may have mentioned only A to B and others only mentioned B to C. To avoid this trap, you can trace individual respondents' *threads* within the paths, which filters to show only continuous chains of links from the same source to avoid the transitivity trap issue.



In most cases, we should always trace threads **before** any filter which changes labels: zooming, removing brackets, combining opposites and autoclustering. See this page for more information.

Remember that **order matters**: the order in which the filters are applied makes a difference.

Source tracing -- What are the consequences of one or more factors, looking only at stories told in their entirety by individual sources?

Robustness -- How robust is the evidence for that X influences Y?

How robust is the network of evidence for the influence of one or more "driver" factors on another set of "outcome" factors?

> From (Powell et al., 2024)

We prefer to talk about possible arguments here because this sidesteps the (also interesting) question of whether any individual source made any such argument in its entirety, with all its constituent links. In many circumstances, evidence for a causal path derived from different sources and contexts can be considered to strengthen the argument, whereas heaps of evidence from the same source will not. To address this kind of issue directly, we can use a complementary measure 'source thread count' as a measure of the strength of the argument from C to E: the number of sources, each of which mentions any complete path from C to E.

> From (Powell et al., 2024)

Figure 3. An illustrative example of a very simple causal map.

Even with a simple example like this, we can answer many questions by visually examining the paths. But analysis of larger data sets might be simplified by selecting only all the paths between a selected cause and consequence to produce what Bougon et al (Bougon et al., 1977) called an 'etiograph'. Eden et al. (1992) go so far as to collapse some causal paths into individual links, simply removing intervening factors.

Our causal map software uses the 'maximum flow/minimum cut' algorithm (Erickson, 2019), which quantifies the robustness of longer paths from C to E by calculating the minimum number of causal claims, which would have to be invalidated or lost to remove any possible causal pathway between them. This is simply the idea that the strength of an argument is dependent on the strength of its weakest link, extended to apply to an interconnected network rather than a single chain. In other words, we can express a path through a causal map as a possible argument: An argument can be constructed that C causally influences D, and then E. This also provides ways to formally address questions such as 'how robust is the evidence for the influence of C on E, compared to evidence for the influence of B on E?'. As our causal evidence will often be of varying quality and reliability, we are advised to also construct and compare paths that consist only of evidence from the most reliable sources.

References

Bougon, Weick, & Binkhorst (1977). *Cognition in Organizations: An Analysis of the Utrecht Jazz Orchestra*. JSTOR.

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Counting and comparing influences

How much evidence is there for the influence of our intervention on a valued outcome? Is that a lot? Can we compare these numbers across pathways?

In this map, six sources told stories which start with our intervention and end with Wellbeing (tracing threads). Is this a lot?

Well, we can compare that with the total number of sources to make a source-mention-proportion, to say that 6 out of 24 sources mentioned this pathway.

Or we can compare it with the number of sources mentioning Wellbeing at all, to say that 6 out of the 16 sources mentioning Wellbeing told stories which began with our intervention.

Or we can compare it with the number of sources telling stories ending with Wellbeing and beginning with a different intervention.

Or we can compare any of these figures with the same figures from a previous time point.

Using numbers and proportions like this in a fundamentally qualitative approach like causal mapping can be very useful but we have to be careful. These are quite fragile indicators which can be easily influenced by other factors (for example, how visible was our intervention?) and can be hard to generalise.

We should always also consider the evidence itself behind each link by looking at the quotes.

Properties of the causal map – Which factors are reported as being causally central or causally peripheral?

Because we can understand the map as a network, we can ask network-style questions. For example we can look at measures like "betweenness" to identify factors which are central in the network.

Properties of the causal map -- What is the overall structure of the network?

There is a host of other analytical lenses available, for example, are there relatively isolated islands within the network? Are there sub-systems which operate more on their own? How sparse or dense is it?

Properties of the causal map -- Are there leverage points?

Systems thinking gives us ways to identify possible leverage points within a network.

Where are the potential leverage points? Network analysis metrics can help identify "central" factors that may act as critical nodes or leverage points within the system as seen by the sources.

Properties of the causal map -- Are there feedback loops?

Can we identify positive feedback loops within the network which might serve to maintain or exacerbate phenomena?

(Note that our "minimalist" approach to causal coding means that out of the box we can only identify positive feedback loops; identifying negative feedback loops depends on additional steps: identifying sentiment and/or combining opposites, which we will not cover here.)

TODO

Combining questions

Causal mapping gets really useful when you start to combine the different questions you might want to ask in order to answer more sophisticated questions. We can think of many of the techniques as filters which filter the view in a particular way. Using multiple filters allows you to build up an answer to a question. Usually, order matters.

Intro

Causal mapping has been used in many different fields. In this chapter we look at how it can be applied in evaluation; its strengths and weaknesses.

Causal mapping is particularly useful for evaluations that focus on learning to inform program improvement: visual representation of causal links between context, activities and outcomes can help to facilitate the sharing and collaborative use of findings.

Causal mapping can be used during a program lifespan to inform adaptive management and as part of a final evaluation.

> From (Powell et al., 2024)

Causal mapping offers ways to organise, combine, present and make deductions from a large number of relatively unstructured causal claims – the sort of data that are often collected in evaluations.

References

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Causal mapping can complement contribution analysis

Contribution analysis is about testing and refining a theory of change to build a credible case for contribution. It's focused on a specific theory of change and on the contributions to outcomes. These contributions might happen along a causal chain, but CA tends to be less explicit about how they are to be traced.

Causal mapping is strongly related. It is not an evaluation method in its own right, but more of a tool which can assist with CA. it's about visualizing and interrogating the whole web of causes and effects. It explicitly addresses the challenge of overlap and influence between causes and effects. It doesn't rely on a fixed theory of change but on the factors and links between them which are actually mentioned in documents and interviews.

How causal mapping can help with Contribution Analysis:

- by helping to assemble all the relevant evidence along some causal pathway or pathways from intervention to outcome. Causal mapping has a strong understanding of chains and transitivity. Contribution Analysis can then focus on what it is best at, namely weighing up different explanations for an outcome and how much our intervention really contributed to it.

Causal mapping can complement Outcome Harvesting

Outcome harvesting is about collecting and explaining a (hopefully long and substantial) list of intended and unintended outcomes after the fact, and identifying how the programme contributed to the outcomes. But it's just like a list of cause-effect relationships. It can be a challenge to understand how those causes and effects overlap with one another or influence one another.

Causal mapping is quite similar. It is not an evaluation method in its own right, but more of a tool which can assist with either. it's about visualizing and interrogating the whole web of causes and effects. It explicitly addresses the challenge of overlap and influence between causes and effects. It doesn't rely on a fixed theory of change but on the factors and links between them which are actually mentioned in documents and interviews.

How causal mapping can help with Outcome Harvesting:

- causal mapping can input a pile of Outcome Harvesting data and link them all together into the form of a larger web. (Britt et al., 2025).

Limitations on causal mapping in evaluation - data quality

> From (Powell et al., 2024)

Causal mapping has some limitations. First, the credibility of the causal arguments which can be derived from a map is limited by the credibility of the original data sources. We see the job of causal mapping as collecting, organising and synthesising a large number of claims about what causes what; drawing conclusions about what this actually reveals about the world is a final step that goes beyond causal mapping per se. In specific cases, establishing explicit and context-specific rules of inference may help to make this final step.

References

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

The elephant in the room -- causal inference

[Responding] to our [one-page description of causal mapping], Julian King says the elephant in the room with causal mapping is: can causal mapping really help you get from causal opinions to causal inference?

The short answer is: sure it can help you, the evaluator, make that leap. But it, causal mapping, does not give out free passes.

But in more detail, here are four more responses to the elephant.

1: **Causal mapping ignores the elephant. On its own, causal mapping doesn't even try to warrant that kind of step**: it humbly assembles and organises heaps of assorted evidence in order for the actual evaluator to make the final evaluative judgement. Unlike evidence from an interview, or a conclusion from process tracing or from a randomised experiment, causal mapping evidence isn't a *kind* of evidence, it's an *assemblage* of those other kinds of evidence. It certainly isn't a shortcut to get cheap answers to high-stakes answers by conducting a few interviews with bystanders.

If you have to answer high stakes causal questions like "did X cause Y" and "how much did X contribute to Y" and have just a handful of pieces of evidence, there isn't much point using causal mapping. Causal mapping is most useful for larger *heaps* of evidence, especially from mixed sources and of mixed quality; it gives you a whole range of ways of sorting and summarising that information, on which you can base your evaluative judgements. What it doesn't give you is a free pass to any evaluation conclusions, and especially not the high stakes ones which occupy so much of our attention when we think and write about evaluation.

2: In most actual causal mapping studies, **the elephant usually doesn't even enter the room**. Usually, we aren't dealing with monolithic, high-stakes questions. Most causal mappers are looking for (and finding) answers to questions like these:

- In which districts was our intervention mentioned most often?
- Do children see things differently?
- How much evidence is there linking our intervention to this outcome?
- Does our project plan see the world in the same way as our stakeholders?

All of these are relevant questions for evaluations. Some of them might feed into judgements about relevance, or about effectiveness or impact, and so on. We might notice for example that there is some evidence for a direct link from an intervention to an outcome, and much more indirect evidence, and some of those paths remain even when we remove less reliable sources. We can even compare the quantity of evidence for one causal pathway with the quantity of evidence for a different pathway. We can ask how many sources mention the entirety of a particular pathway, or we can ask which pathways have to be constructed out of evidence from *different* sources. (On the other hand we don't, for example, make the mistake of inferring from the fact that there is a *lot* of evidence for a particular causal link that the link is a *strong* one.)

All of this is bread and butter for evaluators, even though it doesn't answer those elephant questions.

3: Causal mapping pushes back against the elephant.

In *every* evaluation, the evaluator assembles some evidence and makes an evaluative judgement on the basis of it. **All evaluation involves causal mapping** in this sense. Occasionally there is, or seems to be, only a single piece of evidence in the heap – perhaps, evidence from a controlled experiment. But the final judgement is the evaluator's responsibility, and (perhaps implicitly) must take into account other factors: "this is a controlled experiment, it was carried out by a reputable team, … but wait, their most recent study was criticised for allowing too much contamination …. but wait, the effect sizes were calculated with the latest method and controlled experiments seem to be a good way of reaching causal conclusions …", and so on. An essential part of the evaluative process is also careful consideration of how exactly to formulate a conclusion, bearing in mind the context and the audience and how it will be generalised and applied. So, in practice, there is always a heap of factors to consider, often involving different parts of more than one causal pathway, even when the heap seems to be dominated by one or two elephants.

4: Causal mapping embraces the elephant. In most causal mapping studies, we do not in fact simply assemble the evidence we already have but actively gather it systematically. A good example is QuIP, the Qualitative Impact Assessment Protocol. The evidence is "only" the considered opinions of carefully selected individual stakeholders, but it is gathered using blindfolding techniques to minimise bias so that, once assembled and organised with causal mapping, the evaluative leap from opinions about causality to conclusions about causality can be made with more confidence, transparently and with appropriate caveats. Still, it's not the causal mapping itself which makes or warrants the leap, it's the evaluator, using evaluative judgement.

Qualitative impact evaluation is less interested in the strength of effects

Weird image of people counting beans generated by Canva's AI

Soft versus hard impact evaluation approaches? Quant versus Qual? Is there an essential difference?

Summary: quant and qual impact evaluation approaches are different ballparks because quant approaches attempt to estimate the *strength* of causal effects. Whereas qual approaches either don't use numbers at all or only do calculations about the *evidence for the effects*, not about the effects themselves and in particular we don't estimate strength of effects.

Here's the question: how can we distinguish "soft" approaches to impact evaluation like Outcome Harvesting, QCA, causal mapping, Process Tracing, Most Significant Change, Realist Evaluation and so on from statistics-based causal inference (SEM, DAGs, RCTs etc)?

Here are two bad answers:

- We can't distinguish our "soft" approach(es) by saying that we attempt to assess causal contribution and answer questions about for whom and in what contexts etc, because quantitative approaches attempt all of that too.
- We can say that we are focused on complex contexts, but there's nothing to stop someone using say OH in a non-complex context either is there? In any case whether a context is complex or not is also a matter of how you frame it, no? And there's in fact no shortage of examples where quant approaches have been used in complex contexts.

Here's a better answer: these "soft" methods are qualitative, in the sense that where we involve numbers at all, our arithmetic is essentially an arithmetic of *evidence for* causal effects: is there any evidence for one pathway, how much, how much compared with another? For example, Process Tracing sometimes does calculations about the relative certainty of different causal hypotheses. QCA counts up configurations.

Whereas quant causal analysis involves estimating the *strength* of causal effects (as well as having clever ways to reduce the bias of those estimates).

As far as I know, qualitative approaches never attempt this (calculating the strength of a causal effect). We might conclude that the evidence suggests a particular effect is *strong*, for example because we have collected and verified *evidence for a strong connection*. But we don't, say, combine this with another set of evidence for a very weak connection and conclude that the strength of the effect was only moderate (we don't do maths on the strengths).

It's true that qual approaches also do causal inference in the sense of making the jump from evidence for a causal effect to judging that the effect is real. Quant approaches (and, to be fair, some qual approaches) suggest that using their special method gives you a **free ticket** to make this leap. And indeed different methods include different ways to reduce different kinds of bias which mean you can be more confident in making the leap. But I'd say there are no free tickets. No way of an evaluator getting out of the responsibility of making the final evaluative judgement, however clever and appropriate your method.

(You could argue that FCM and Systems Dynamics do arithmetic on the strengths of connections. Perhaps that makes them quant methods.)

Seen this way, in essence qual and quant impact evaluation are not alternatives or competitors. They are different ways to do different things.

A second limitation of causal mapping is the difficulty it has in systematically capturing the strength or type of causal influence. It is relatively rare in open conversation for people to indicate in a consistent way the magnitude of the effect of C on E, or whether C was a necessary or sufficient condition for E or precisely how certain they are about the connection. There is of course scope for framing questions to encourage people to ascribe weights to their answers, which can then be incorporated into the way maps are constructed. But imposed precision risks turning into spurious precision, and stronger framing of questions may distract from other issues and nuances that more open-ended questioning might otherwise have elicited.

Task 3: Answering evaluation questions#

Causal maps help us to assemble evidence for the causal processes at work in specified domains, including the influence of activities being evaluated. They can also help expose differences between the evidence given by different sources and differences between the analysed data and theories of change derived from other sources, including those officially espoused by the commissioner of the evaluation (Powell et al., 2023). The identification of differences in understanding can then feed into further enquiry, analysis and action concerning why people have different views, what the implications of this are and how these might be addressed.

Focusing on causal claims is of course only one way of answering evaluation questions from a corpus of text data. But it is productive because many evaluation questions are at least partly about causation and causal contribution, and we have found that causal mapping points to possible answers to these questions relatively rapidly compared to more generic QDA approaches. Answering questions about efficiency, effectiveness, impact and sustainability, for example, all depend on identifying the causal effects of a specific intervention, be they perceived as positive or negative, intended or unintended (OECD, 2010). Even ‘relevance’ can have a causal interpretation in the sense that an intervention is relevant if it is doing the right thing: Whether it is likely to help to address the needs of stakeholders is at least partly a judgement about its causal powers.

For a data set comprising hundreds or thousands of links, an unfiltered global map of all the links is a bewildering and useless ‘hairball’ that includes everything but highlights nothing.

Causal mapping can help reconstruct a program theory empirically

To evaluate a program, the evaluator can use Contribution Analysis (CA) (Mayne, 2012). We start with a program logic or Theory of Change (ToC), consisting of possible pathways from interventions to outcomes, and collect existing or new evidence for each link. However evaluators can often not assume that the ToC underpinning a program aligns with the realities on the ground, or they may uncover outcomes not anticipated in the original program design - see Koleros & Mayne (2019). We have argued (Powell, Copestake, et al., 2023, p. 114) for a generalisation of CA in which evidence relevant to constructing a program theory, as well as evidence for the causal influences flowing through it, are both collected at the same time, without the evaluator (necessarily) having a prior theory. In this sense, following Mayne, "program theory" need not be something that any person necessarily possessed or articulated at the time, but is something which can be approximated and improved during the evaluation process.

(Re-)constructing program theory empirically in this way is an essentially open-ended, qualitative problem. Closed data collection methods are not suitable because we cannot measure what we do not yet know. Open-ended, qualitative methods to (re-)construct a theory are notoriously time-consuming and are usually heavily influenced by researcher positionality (Copestake et al., 2019).

Powell, Copestake, et al (2023, p. 108) present this task as gathering and synthesising evidence about "what influenced what", evidence which is simultaneously about theory or structure and contribution. Each piece of evidence may be of differing quality and reliability and about different sections of a longer pathway, or multiple interlocking pathways, and may come from different sources who see and value different things.

References

Mayne (2012). *Making Causal Claims*.

The result of an evaluation is a qualitative causal model

Evaluation outputs as a model#

What is the output of an evaluation? We have a report, hopefully answering the evaluation questions. (In the sense of developmental evaluation perhaps some learning has taken place as well, or as a main output, but this is not what I want to address here.)

The final product can be more than a report: it can be a kind of model or knowledge graph which in principle can be queried to answer even new and unexpected questions.

Causal maps are knowledge graphs, but with wings

Intro

Causal mapping is also a kind of Qualitative Data Analysis (QDQ). How does that even work? This chapter explains.

Causal mapping is a simple yet powerful form of qualitative coding

Is this for you? If you're at least a bit familiar with qualitative coding and Qualitative Data Analysis (QDA) as a way of making sense of texts, but you're not yet convinced that causal mapping is a really interesting twist on that, or you've never even heard of causal mapping, this short series on causal QDA is for you! Plus, we are now making available a new version of our software for exploring causal mapping which is free for core functionality.

Quick background on causal mapping: Causal mapping has been used since the 1970s across a range of disciplines from management science to ecology. It helps make sense of the causal claims (about "what influences what") that people make in interviews, conversations, and documents. This data is coded, combined, and displayed in the form of a causal map. Each claim adds a link to the map. These maps represent individuals' and groups' mental models: individual links and (often interlocking) longer chains of causal explanations. For every link, we can store the original sources and quotes. Causal mapping is an approach to coding texts but also to synthesising the coding, so we can think of it as a form of qualitative coding in particular and of QDA more generally. A wider overview of causal mapping as applied in disciplines from ecology to organisation science is presented in (Powell et al., 2024) and in this bibliography.

Our aim is to present causal QDA as a particularly useful way to make sense of text in a way which is straightforward to apply with a highly, though not completely, standardised approach.

Our steps in doing this are:

- We present causal QDA as an important and useful form of QDA
- We explain why causal coding can be easier to apply than non-causal coding.
- We explain how the result of causal coding is a causal network which can be visualised in the form of causal maps and used to answer useful questions.
- We claim that the kinds of questions which causal maps can help answer tend to be particularly useful and practical questions because they are about what causes what, through the eyes of stakeholders.
- Although the focus of this series is not on AI assistance, we also argue that these properties make causal mapping particularly suitable for automation with AI in a way which is verifiable and avoids using the AI as a “black box”.

In two later posts, we will argue:

-
- Causal mapping is easy to automate transparently, so is a great fit for scaling with AI

Why now?[#](#)

If you've been following our work at Causal Map Ltd (and the work of our colleagues at BathSDR), you'll now that we've been talking about and promoting causal mapping for some years now. We've been providing software like the Causal Map app as well as consultancy services for evaluators and researchers. But we were the first to admit that our Causal Map app was sometimes slow and frustrating to use. We're happy to say that the latest version of the app, the **completely new Causal**

Map 4, is not only a vast improvement (and a lot faster!) but also, public projects are **free** for core functionality. We've done this to lower the bar for evaluators, researchers and academics to explore and use causal mapping.

What is causal coding?#

Causal mapping has been used across multiple disciplines for over 50 years (Ackermann et al., 2004; Axelrod, 1976; Eden, 1992; Hodgkinson et al., 2004; Laukkanen, 1994; Narayanan, 2005; Powell et al., 2024), but though some causal mapping practitioners have published guides to causal coding, starting from (Wrightson, 1976) it has not often been presented as (also) a stand-alone form of QDA. We tried to rectify that in (Powell et al., 2024). This post follows on from that explication.

Coding for causal mapping has traditionally been done manually, just like working with Nvivo, Dedoose et al., but instead of coding individual *themes* we code *links between themes* (aka causal factors). For example in the text “the floods destroyed our crops”, we would code “the floods” as the cause, “destruction of our crops” as the effect.

Some forms of causal mapping may involve constructing maps directly, for example in participatory fashion with a group of people. In this post we don’t cover that kind of method, focusing only on cases where an analyst codes causal claims within text -- similar to other forms of qualitative coding.

What makes causal coding special#

Qualitative coding is usually designed to capture **concepts or themes in general**, whereas in causal QDA we code (claimed) causal **links between causal factors**: things, events, phenomena, changes: anything which can affect or influence, or be affected or influenced by, something else.

Article content

Using non-causal coding we can create post-hoc causal links between independent concepts as in the figure above, version 2. This could make it of epistemic network analysis, , see Lazarus (zotero) TODO. Or we can create the causal link as a monolithic concept as in version 1. (This option is mentioned in (Saldaña, 2015).) But then the word “cause” does not really do anything: we can’t automatically deduce anything from it1. We can do ordinary common-sense reasoning on the basis of the coding because we understand the words "cause" and "influence", just as we can do reasoning based on our understanding of floods or crops. But causal coding goes further because it is explicitly designed as an input to causal reasoning algorithms.

In causal QDA, the primary act of coding is to highlight a specific quote from within a statement and identify the causal claim made by simultaneously identifying a pair of causal factors: a "cause" and an "effect". In causal QDA, we *only* code causal claims. The causal factors only exist as one or other end of a causal link and have no meaning except in their role as either end of a causal link. The coding *Floods happened* → *Crops destroyed* IS a causal link, and a set of such links can be immediately visualised as a network or map using suitable software – for example with Excel and Gephi.

In causal coding, the result of each act of coding is a link, a pair of factors (an ordered pair, because B → C is not the same as C → B). It is the factors which are the atomic units, which form the codebook, even though the factors only make sense as part of causal links.

Causal coding can, of course, take account of **context** simply by ensuring it that context is noted separately and included in the interpretation of the resulting causal maps. A more interesting challenge is to include context as part of the coding, for example by using cross-codebook tags (such as, say "post-COVID only") which can appear as part of many different factor labels, or by including additional tags as part of each link, or by introducing contextual factors (such as "COVID pandemic") as additional causal factors in the map itself.

What are the advantages of causal QDA?#

A more restricted, causal, ontology simplifies the coding task#

> What are the causal factors, the cause and the effect, at each end of this causal claim?

This is a very challenging question, but it is a whole ballgame less challenging, less open to interpretation, than this one:

> What themes/concepts are mentioned here?

This fact makes it much easier to get started on and complete our coding, whether we are going to code deductively (from an established codebook) or inductively (developing the codebook iteratively). It is particularly easy when working inductively and using in-vivo labels close to the original text: we simply have to identify each and every causal claim in the text, and for each one, identify the cause and the effect – the "causal factors". (Using in-vivo labels means we will later need another way to consolidate the resulting large number of factor labels.)

In some cases it might be necessary to define in more detail (as part of the codebook) exactly what we mean be "X causes or influences Y", for example to distinguish hypothetical from factual claims. Inter-rater agreement in causal coding can be good (Buzogany et al., 2024; McCardle-Keurentjes et al., 2018). This means that in the simplest case we can reduce most of causal coding to a series of low-level tasks: *code each and every causal claim in the text*. Of course, it is possible to work in a more deductive way, with a more explicit causal codebook.

There are still lots of theoretical, high-level decisions to make when doing causal coding, many of which cannot be easily anticipated, for example, How shall we deal with time, or with tentative or hypothetical statements? What shall we do with claims which one person makes about another's beliefs? And so on. These kinds of problems are familiar across all kinds of qualitative coding. But the focus on causal claims is still a massive simplification of the coding task.

The product of causal mapping is explicitly a theory or qualitative model which can be easily queried to ask and answer important questions#

Just like with any other form of QDA, when we finish the coding we can if we like report a bunch of statistics. In the ontology of causal QDA, we have factors, which are something like concepts, and also links, which are ordered pairs of factors. We can immediately make use of all of the usual frequency/occurrence logic for both factors and links, e.g. to report which was the most commonly mentioned causal factor and/or which was the most commonly mentioned causal link. The usual caveats about granularity and frequency apply (here and also in the following section).

But of course these frequencies are just the tip of the iceberg. The main product of causal mapping is not a set of tables or a static report but a qualitative causal network (a database of links) -- a live object which we can query to bunch of much more interesting and perhaps unanticipated questions, for example: which causal pathways do people identify as leading to their own wellbeing?

Article content

The causal beliefs of a larger number of people will of course be a mass of somewhat overlapping and somewhat contradictory information; maybe some whole sections of the causal map are only mentioned by particular sub-groups, or maybe there are some factors for which different groups disagree about the causes and effects; or maybe different groups mention approximately the same factor but use different language to describe it. A causal map, as a database of links, can contain all of these tensions and contradictions. The challenge is to apply the right queries to it in order to present these differences clearly.

This is possible above all because causal links have the special property of *transitivity*: if we know something about the paths from B to C and from C to D we can start to reason about the path from B to D. Just as a road map can help answer unanticipated questions like ...

> How can I get from Glasgow to Manchester, avoiding motorways?

.... a causal map can help us answer unanticipated questions like

> What is the collected evidence for a causal path from this intervention to that outcome?. By which route? For which pathway is there most evidence? Does everyone agree or are there relevant sub-groups of respondents?

Focusing on causation helps cut to the chase in asking and answering questions which are really useful#

The result of causal QDA is a network, a map, already structured to answer questions about causal beliefs (effects of causes and causes of effects). **Many practical, real-world questions, especially those of interest to program evaluators are at least partly causal in nature:**

- what leads to what?
- why does Z happen?
- what contributed to Q?
- what might happen if X?
- what are the dominant explanations for Y?

Causal mapping makes use of a causal vocabulary which provides some semi-standard ways not only to ask but also to answer causal questions.

In the next post we will look at some of these semi-standard procedures in more detail.

Limitations and caveats#

You want caveats? We've got loads of caveats, as outlined in Powell et al (2023). Here are a few.

- **Causal claims are not causal facts.** *Somebody* claiming that X causes Y is certainly not enough evidence for *us* to believe the same, without further information. We need to remain constantly vigilant to avoid confusing beliefs with facts.
- **Contribution not determination.** In the world of social science we rarely need to even consider anything like total causation: we deal with causal contribution, any form of possible influence, and hardly ever with determination. This goes for beliefs about causation just as much as facts about causation.
- **Beware the transitivity trap.** We have to be especially aware of the transitivity trap: even if we know that someone claims B influences C and someone else claims that C influences D, this does not mean that either of them does or should believe that B (indirectly) influences D.
- **Chunking and granularity.** As with any other form of coding, when creating our causal codes we have to pay attention to problems of chunking (where do phenomena begin and end) and granularity (how big are the chunks). This is important to bear in mind when dealing with coding frequencies, which can be very misleading if not interpreted and presented with care, just as with frequencies in any other form of coding, For example we might report that the factor "children's achievements" was very frequently mentioned, whereas if our codebook had instead included two separate factors for girls' achievements and boys' achievements, the frequencies would have been lower.
- **Always go back to the evidence.** While we can be guided by relative frequencies of links, we always consider the specific (combined) evidence for each link, for example when some evidence might be weak or valid only for a specific context. Frequencies start to become most interesting when a larger number of sources are included; but causal mapping has been very effectively used also for purely idiographic or single-source studies (e.g. in the earliest works by Ackerman and his team), in order to surface, for example, a politician's internal model of how a regional conflict

is fuelled by coding their speeches and writings. (Here too, it can still be important to track the multiple mentions of the same factor or link.)

- **Causal coding cannot answer all questions about a text!** Causal coding will fail to code claims or statements which contain material which never influences anything else or is influenced by anything else. Causal coding, obviously, captures things which make a difference or are affected by things which make a difference, and while there is a strong argument that this is where we should be spending most of our research energy, we certainly would not claim that nothing else is important.
- **Broader theories.** Causal QDA is not on its own well suited to constructing broader theories with wider applicability in the sense of (Glaser & Strauss, 1967).
- **Latent vs explicit.** In our team, we have not used causal coding explicitly to search for latent as opposed to explicit material, ibid, (Braun & Clarke, 2006). We see no inherent reason why causal coding should be limited to explicit causal connections even though this more limited approach has been most common in the wider causal mapping literature.
- **Bare causation.** We also do not attempt to code anything other than what we call "bare" causation. Some causal mapping approaches do try to code the believed strength and/or polarity of the link (positive Vs negative), but one could also try to code ideas like necessity and/or sufficiency or even non-linear connections. For various reasons we believe this to be a wild goose chase in practice, but we do acknowledge that important information may go missing when we restrict coding to bare causation.

Conclusion#

When you've read all three posts in this series, we hope you will be convinced that causal mapping in general and causal coding in particular can and should be considered members of the family of QDA approaches, offering particular strengths. They: - simplify the coding task; - result in qualitative causal models which can be queried to answer important questions making use of a range of pre-existing algorithms; - are particularly amenable to automation and scaling using generative AI in a way which remains transparent and verifiable.

What Causal Map Ltd provides#

We provide a web app Causal Map which is specifically designed for causal coding and has been used in many academic and practical applications. Colleagues at BathSDR and others have applied (human-powered) causal coding in over 100 evaluation projects using the Qualitative Impact Protocol (Copestake et al., 2019), mostly in the field of international development, more recently using the Causal Map app.

Here is a broader list of causal mapping software.

---

1. Thanks to Walter Antonio Canu, Filip Zielinski, Dr. Susanne Friese and Christina Silver, PhD for suggesting additions here. ↩

Causal Mapping outputs not just codes but a model you can query to answer useful questions

In the first part of this series [Causal mapping is a simple yet powerful form of qualitative coding](), we argued that causal mapping is a simple yet powerful form of qualitative coding. We showed how its focus on identifying causal links reduces and simplifies the analytical task and provides a clear, structured approach which is relatively easy to apply. However, the true strength of causal mapping lies not just in the coding process itself, but in its output: a **query-able qualitative model**.

Qualitative Data Analysis (QDA) most typically produces a written report alongside a deeper understanding in the minds of the researcher or research team, and hopefully also in the readers. It may also provide other outputs like tables of frequencies or co-occurrences.

We already argued that causal mapping is an especially useful form of QDA because:

- It produces a **structured graph database of causal claims**. This can be viewed as a table or as a map.
- This output isn't a static summary; it's a **dynamic qualitative model** -- a network of interconnected evidence that can be systematically interrogated using standardized, out-of-the-box filters and algorithms.
- The kinds of questions you can answer with a causal map are **particularly useful** because they are about *what causes what -- as seen by your sources*.

From a List of Themes to a Network of Evidence#

Traditional QDA often produces additional outputs like tables of theme frequencies alongside narrative summaries. It is possible to query these to answer questions like "what themes did the younger respondents most often mention when also talking about the main theme".

Causal mapping provides a particularly rich output: a **causal network** -- a qualitative knowledge graph where the factors (which can be understood as themes) are nodes and the causal claims are the links connecting them. This structure is inherently machine-readable and ready for analysis.

Thinking of the output as a **model** is key. Just as a quantitative researcher builds a statistical model to explain relationships in their data, a causal mapper builds a qualitative model of the causal beliefs expressed in texts. This is not unique to our Causal Map app: all applications of causal mapping provide, more or less explicitly, this kind of model, going right back to (Axelrod, 1976). This model can then be used to answer new questions, often without needing to go back to the original source texts, though the underlying quotes and context are always available.

Some standard ways to answer useful questions#

Because the output is a structured network, we can apply a range of queries to explore the data. This gives us a library of **pre-existing approaches** to ask **practical questions** about the causal landscape described by the participants.

Here are some questions you can answer using causal mapping.

Individual questions -- introduction ▸

A corresponding library of filters#

The Causal Map app provides about 20 corresponding, ready-to-use filters to answer these kinds of questions, some based on existing causal mapping publications, some new.

In the list of typical queries above:

- Some queries have matching unique pre-defined outputs: a map with a specific filter applied, or a table.
- For some queries, there are different ways to answer them and/or the answer requires more than one filter.
- For a few queries we do not yet have a specific way to answer them in Causal Map.

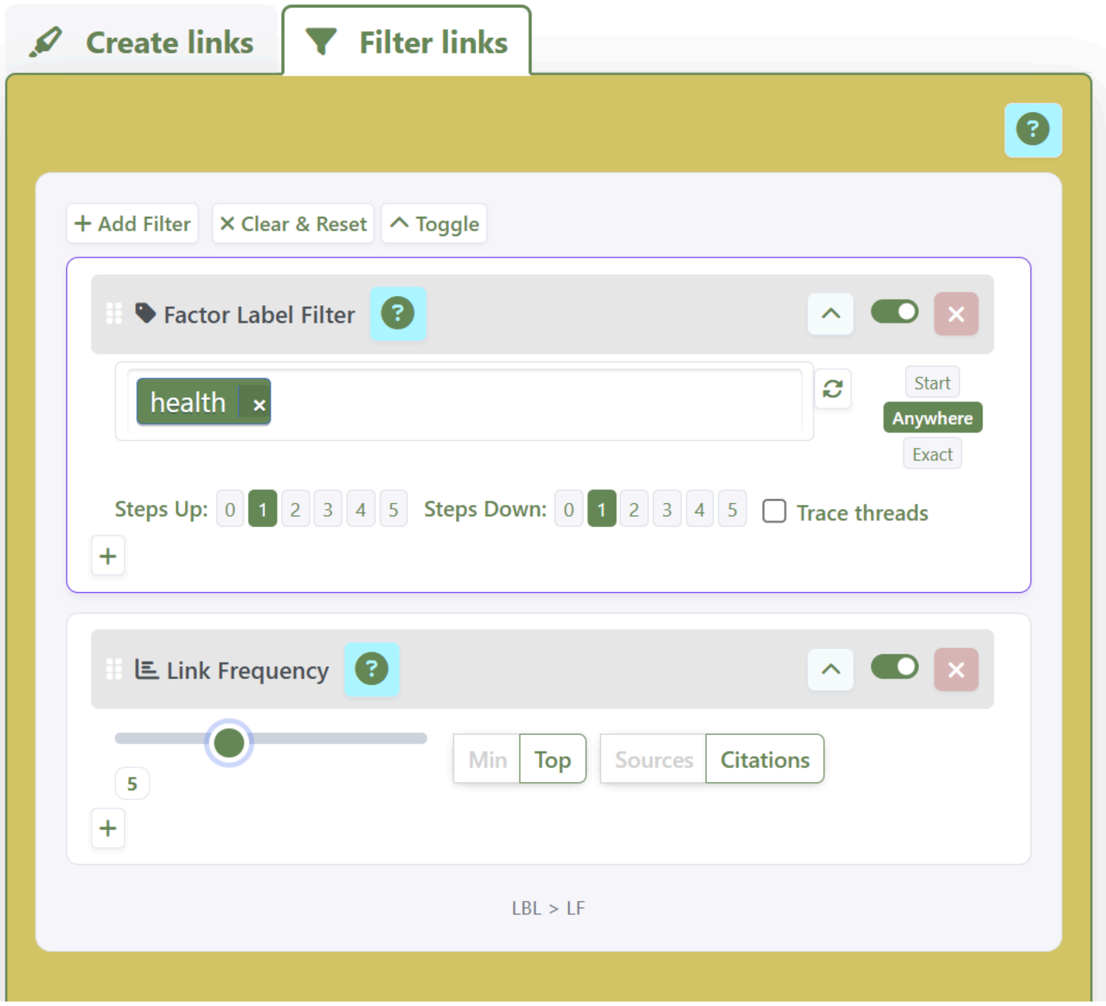
There are two types of filter:

- simple filters which select a subset of the links like links frequency, factor label filter; and path tracing
- "transform filters" like zoom, which temporarily rewrite the cause and/or effect labels.

Here as reference are more details: .

Chaining filters together#

Most importantly **we can chain filters together**, to answer corresponding, composite questions like: *what are the most frequently mentioned upstream influences on these key outcomes, according to the younger respondents?*



Most of the core filters have nothing to do with AI or large language models. They are straightforward, transparent and deterministic. In the Causal Map app, they are available at a click and can be rearranged with drag and drop. But most of them are simple to reconstruct in a spreadsheet or a graph database, without using Causal Map at all.

Cutting to the chase with causal queries#

This ability to systematically query **causal** evidence is what allows causal mapping to **cut to the chase**. For evaluators and researchers, **the core questions are very often about causation**: "Did the program work?" and "How did it work?". Causal mapping structures the evidence provided in the texts to help researchers answer these types of questions.

A Model of Causal Evidence, Not Causal Facts#

It's crucial to be clear about one thing: Causal mapping is **not a method of causal inference**. It does not, on its own, tell you what truly causes what in the real world.

Instead, it creates a qualitative model of **causal claims**. The map organizes what people *said* about causation, allowing the researcher to weigh, compare, and synthesize this evidence. The logic we apply is one of evidence management:

- How much evidence is there for a link between X and Z?

- Is that link direct or indirect?
- Do different subgroups of people agree on this causal pathway?

Calling these causal claims "evidence" does not mean that anyone should necessarily believe it or that it has been verified[1]. It is simply raw material which we organise and inspect before drawing any conclusions. Researchers can choose to also code additional properties for each link and/or source such as "doubtful" or "reliable" or "verified" and include or exclude links by filtering on those properties.

If the researcher wants to make a causal judgement, they must interpret the map in context, examine the quotes and consider the source of the claims. The map is a powerful tool for structuring and clarifying that judgment.

In any case, **many colleagues use causal mapping not to make causal inferences but simply to understand** *what people think causes what*, and how, for example as a crucial prerequisite for planning policy, communications or interventions.

See also the other caveats we listed in the previous post: Causal mapping is a simple yet powerful form of qualitative coding.

Conclusion#

The real power of causal mapping as a QDA method is that it produces a query-able, qualitative model of causal evidence. This structured output allows researchers and evaluators to apply a range of standardized algorithms to answer practically relevant questions about the causal mechanisms at play, as seen by the sources.

Many researchers and evaluators like using causal mapping to explore their data. In the third part of this series, we will explore how these properties -- a simplified coding task and a structured, query-able output -- make causal mapping especially suited for transparent and verifiable automation with AI.

---

1. Thanks to Stève Duchêne https://www.linkedin.com/in/steve-duchene/ for reminding us to clarify this. ↩

References

Axelrod (1976). *The Cognitive Mapping Approach to Decision Making*. In *Structure of Decision : The Cognitive Maps of Political Elites*.

Causal mapping is easy to automate transparently, so is a great fit for scaling with AI

Inter-rater agreement for causal coding is high, making it suitable for automation with AI#

The fact that causal coding can be largely reduced to a series of low-level tasks makes it very suitable for automation with AI. High precision and recall scores can be achieved. (Consolidating a large number of in-vivo labels can be accomplished mostly automatically with clustering of text embeddings.)

The AI is used only as a tireless, low-level but incredibly fast assistant with the instruction to code each and every causal claim in the text. This is radically different from the kind of AI-supported “black box” coding which essentially treats the AI as a trusted co-coder who is asked to make, or help make, high-level decisions such as "what are the main themes in this text?" or even “What is the overarching causal network in this text?”.

The accuracy (precision and recall) of AI-supported causal coding is not perfect, but it is improving all the time. Creating, implementing and monitoring the coding protocol remains an essential task ("human in the loop") but we claim that AI-supported causal coding comes closer than other approaches to providing an almost out-of-the-box way to make sense of texts at scale.

The causal coding procedures we have outlined here represent a single-pass, non-iterative approach. Of course this can be expanded to include the more iterative approach essential to most QDA approaches, with varying levels of human oversight. For example we can quickly and cheaply experiment with different coding rules, compare the results, modify the rules and iterate again. This ability to experiment with, compare and iterate potentially hundreds of coding rules and algorithms is a real strength of (semi-)automated coding.

Vulnerable to limited attention: if we really process only one section at a time, we will be unable to notice cross-references or places where one section qualifies another, as pointed out by Udo Kelle (1997) xx. This may not be a fundamental limitation of machine-led approaches if we arbitrarily expand the surrounding context, increasing the attention or context window, but at present this is slow and expensive. See A 2023 study by Rezaee et al. compared topic modeling (LDA) vs human qualitative coding of tweets, finding that automated methods reliably find dominant themes but miss subtle frames that human interpretive coding can catch.

Brief review of S Friese -- Conversational Analysis to the Power of AI

This interesting article (Friese, 2025) proposes a methodological shift for qualitative data analysis (QDA) that moves beyond traditional coding by introducing **Conversational Analysis with AI (CAAI)**. This approach can be realised by using Dr. Friese's own software, QInsights, replacing the process of coding -- segmenting and labelling data -- with a structured, dialogic interaction between the researcher and a large language model (LLM).

The Analytical Process#

The method uses a five-step process that focuses on synthesis instead of coding.

- **Step 1: Get to know the data.** The researcher uses the AI to create summaries and initial themes to identify key topics for exploration.
- **Step 2: Prepare for analysis.** The researcher picks a topic from the previous step and writes a set of questions to guide the AI conversation. This question list replaces a traditional coding frame and makes the analysis transparent. CAAI analysis proceeds topic by topic. This contrasts with conventional coding frame development.
- **Step 3: Ask questions.** Using the prepared questions, the researcher has a dialogue with the AI about a small subset of the data (e.g., 4-6 interviews). This helps the researcher find patterns and explore surprising findings.
- **Step 4: Synthesize insights.** The researcher slows down, reads the text of the conversation and writes a synthesis of the findings. This can be done alone or collaboratively with the AI. Steps 3 and 4 are repeated for each topic.
- **Step 5: Elevate the analysis (Optional).** The researcher can use the AI to help connect findings to broader theories.

How LLMs are Characterized#

The author views Large Language Models (LLMs) not as intelligent beings but as useful analytical partners. Their value comes from being trained on vast amounts of human-written, "socially-situated corpora". The article states that models lack true human understanding or lived experience, a concept referred to as "Seinsverbundenheit". However, their outputs are still insightful because they reflect the patterns and "lifeworlds" present in their training data. The AI's knowledge is described as a "collective and distributed" echo of human meaning-making. In this role, the LLM acts as not only in a deductive and inductive fashion but also as an "abductive catalyst"—a tool that surfaces unexpected connections and provokes new ideas for the researcher, without needing to be intelligent itself.

This approach reframes the researcher's role from a coder to a conductor of an analytic conversation, prioritizing interpretation and synthesis over mechanical categorization. Following (Krähnke et al., 2025), Friese characterises this process as hermeneutical -- an approach in which meaning emerges through the researcher engaging in an open interaction with a text, in this case with the addition of a third party, the AI, which provokes and questions the process. "Meaning does not reside in the data itself but is generated by a recursive relationship between the analyst, the context, and the text" (p. 6).

Rigor and trustworthiness are established not through coding frames or inter-coder agreement, but through the transparency of the documented dialogue, traceability to source data via retrieval-augmented generation (RAG) systems (provision of quotes), and the somewhat replicable nature of the question sets that guide the inquiry. Ultimately, CAAI presents a post-coding paradigm where analysis emerges directly from a dynamic and reflexive engagement with the data, mediated by AI.

In the course of the article, Friese revisits the text analysis she conducted for her own PhD and candidly concludes that she could have done better, and much faster, using the method she proposes.

Strengths#

Reflection on use of LLMs#

Firstly, I'd congratulate Dr. Friese for presenting a **deeply argued explanation of and justification** for the QInsights workflow. As researchers, however we use LLMs, it is crucial that we continue to reflect on what we are doing, rather than letting the muscles of our critical thought go weak through lack of exercise.

LLM-supported QDA as hermeneutics#

Secondly, the advent of generative AI is a very good time to ask again: do we really need coding to make QDA rigorous? As she says: "qualitative research is entering a **moment of methodological experimentation**". After all, rigorous coding alone was never enough to make text analysis rigorous, not least because coding can ignore context and the positionality of the researcher and the research task. **Re-positioning LLM-supported QDA in the broader context of hermeneutics** is useful and refreshing.

Dangers of consumer-facing LLMs#

Thirdly, she reminds researchers that **methodological rigour is very hard to achieve with unreflective use of consumer-facing LLMs** like ChatGPT, not least because they will skim-read the text to find a quick answer and will too often grab a possibly incorrect answer from their training data rather than ground it in the actual text. What's more, a platform like QInsights can provide a completely documented workflow, where every conversation and decision made is recorded -- although we have no insight into the AI's 'thought processes' behind each of its contributions to this workflow beyond what it tells us.

A new kind of epistemic actor between human and machine#

I usually find arguments about whether generative AIs are "really" conscious or can "really" understand something as simply spurious, and no more interesting than arguing about whether a computer can *really* "copy" or "save" or "read" or even "predict" something. However, Friese has a really interesting angle on this, an angle which takes us further rather than trying to police useful language. She '**treats AI as a new kind of epistemic actor** — neither a mere stochastic parrot nor a conscious knower, but as something else: a dialogic partner capable of generating insight through probabilistic modelling trained on socially situated corpora. “Construction” in CAAI is not solely human; it is distributed, emerging through interaction between human interpretation, data context, and machine-generated associations.' So an AI is more than "an assistant": it is an assistant which possesses, or *is*, a whole world of interconnected meanings.

Some caveats#

Against those very notable strengths, here are a couple of caveats.

How new is this? If it is new, is that really because of the inclusion of AI?#

So, we can drop coding and use an AI as a sort of meta-assistant to co-create meaning from text. My biggest question is: to the extent that the AI is not that different from a tireless, lightning-fast and knowledgeable human assistant, couldn't we drop the AI from the methodological equation and just say, here's one newish way to analyse texts hermeneutically? (Though you'd have to have plenty of time and either one very good human assistant or a team of normal ones for this to actually work.)

Is this new method really just a speeded-up version of what we could in theory have done with very fast human assistants? Or is does the sheer magnitude of that speed-up mean a kind of Hegelian transformation of quantity into quality: it's just so much faster that the result is qualitatively different? Or is there something else about the method which is fundamentally new?

Skipping the coding step does lose some reproducibility -- does the AI really change this?#

If the addition of AI support does not make this way of working fundamentally new, I don't really see how it gives us a free pass on all the original reasons for using coding in the first place. This method is presumably, if everything else is held equal, less reproducible than a method which does employ coding.

Dr. Friese does mention "Reapplication of refined questions across subgroups; independent synthesis by multiple researchers; replication over time" as possible ways to *assess* reproducibility.

This is where our approach to causal coding at [Causal Map] differs most strongly from CAAI: we only use the AI as a low-level assistant with narrowly defined tasks, so the workflow is less dependent on the stochastic nature of the AI's behaviour and even confines human input to the most crucial high-level decisions, and so is easier to reproduce.

Why this particular set of steps?#

It isn't hard at least in theory to think of endless variations of the 4 or 5 steps set out in the article -- for example the team-based approach suggested by (Krähnke et al., 2025), so what are the specific arguments for this particular set of steps? Why for example do we work topic by topic? I'm sure it's not the case that "anything goes", but why not?

Whose world?#

I am sure Dr. Friese would acknowledge that the world of "machine-generated associations" underpinning AI responses is not just given but is constructed in a very specific way by very specific organisations on a very specific set of training data. This fact has been repeated almost to exhaustion in recent writings on AI, but she does elevate this world to a kind of new meta-assistant, so it would be good to reflect once more on its make-up and provenance. You could see the meta-assistant as a combination of all human lifeworlds, except that of course it isn't -- because many people, and certain continents, and whole swathes of actual human non-digital life are not equally included.

In conclusion#

As Dr. Friese says, generative AI potentially opens up a whole world of possibilities for qualitative text analysis and indeed for social science in general. We need to be acutely aware of the multiple social, political, methodological and environmental risks of this technology, and *at the same time* not miss out on its benefits.

And, try out QInsights!

References

Friese (2025). *Conversational Analysis with AI - CA to the Power of AI: Rethinking Coding in Qualitative Analysis*. https://doi.org/10.2139/ssrn.5232579.

Krähnke, Pehl, & Dresing (2025). *Hybride Interpretation Textbasierter Daten Mit Dialogisch Integrierten LLMs: Zur Nutzung Generativer KI in Der Qualitativen Forschung*. DEU. https://www.ssoar.info/ssoar/handle/document/99389.

Causal mapping with kumu

Summary#

This guide shows how to create a simple causal map in **Kumu.io** using a spreadsheet. It includes:

- Spreadsheet structure for causal links
- Step-by-step import into Kumu
- An 8-link example
- Brief comparison with Excel and the Causal Map app

Why Use Kumu.io?#

Kumu is a web-based platform for visualizing relationships, such as networks, systems, and causal maps. It supports importing data from spreadsheets and is suitable for small to medium-sized projects. The free tier allows public projects.

Kumu isn't a dedicated causal mapping tool—it doesn't handle coding, source tracking, or evidence—but it's excellent for quickly building and sharing interactive diagrams.

Prepare Your Spreadsheet#

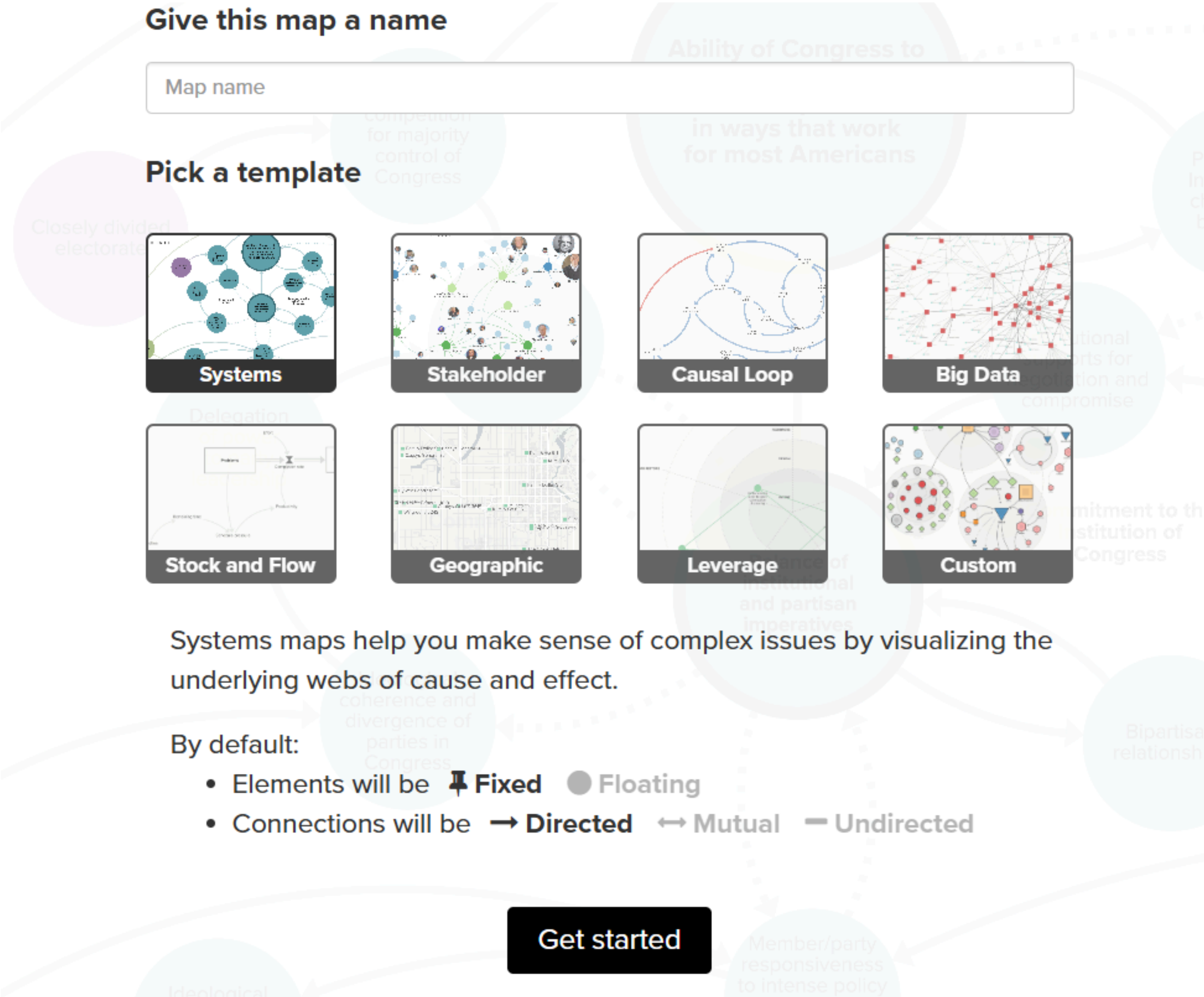
Kumu reads causal links from a two-column spreadsheet

```
| From | To |
|---|---|
| Better seeds | Higher crop yield |
| Farmer training | Higher crop yield |
| Fertilizer access | Higher crop yield |
| Higher crop yield | Higher income |
| Higher income | School attendance |
| School attendance | Long-term wellbeing |
| Farmer training | Financial literacy |
| Financial literacy | Higher income |
```

Each row represents a directional link (cause → effect). Save your file as `.csv` or `.xlsx`. Optional columns like `Description`, `Type`, or `Strength` can be added to enrich the data, but they're not required.

Set Up and Import in Kumu#

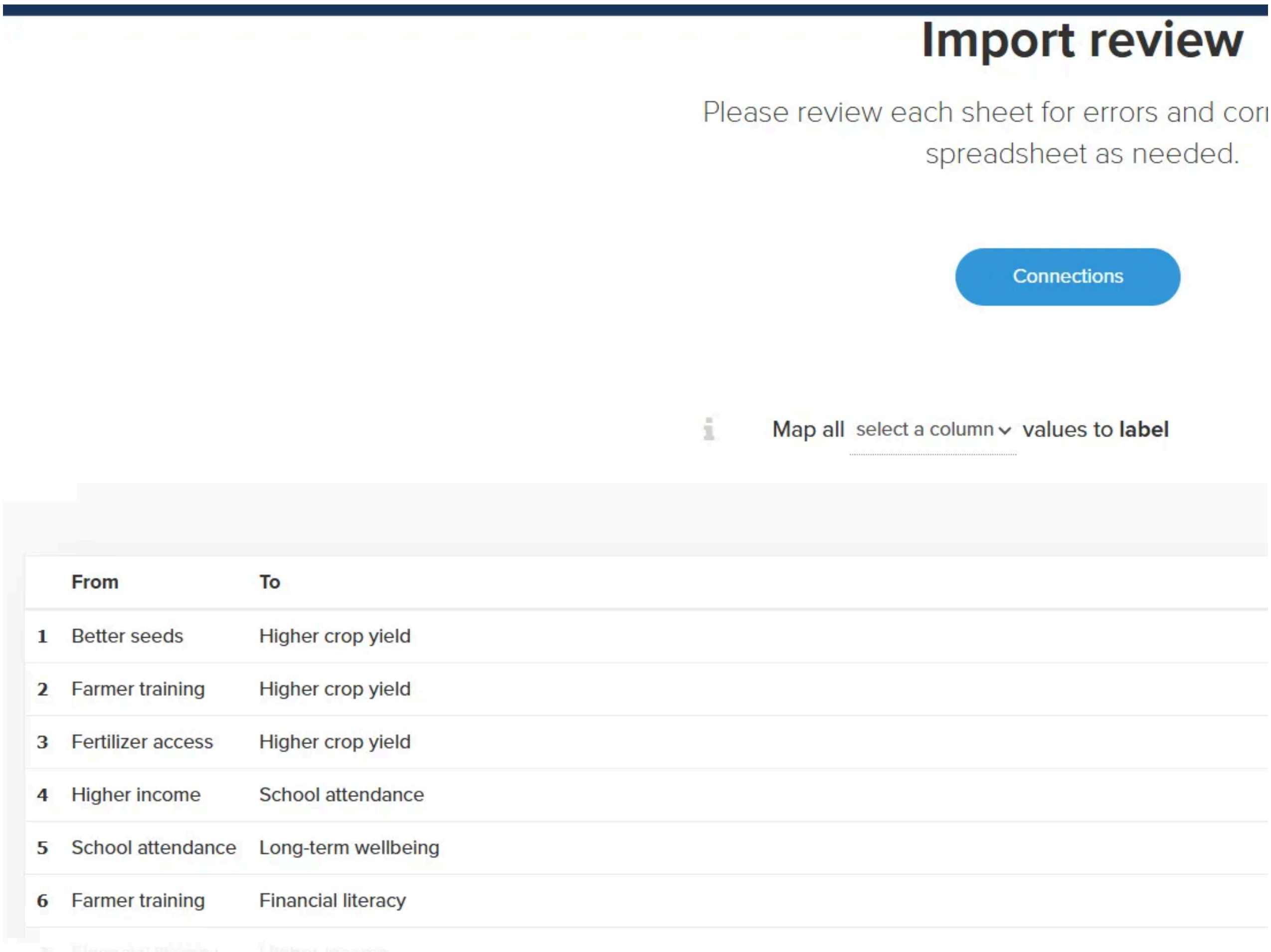
1. **Create a new project** → choose the **Systems Mapping** template.



1. In the canvas, click the green + button → **Import** → **Excel/CSV**.

👆 (1).png

1. Upload your file. Kumu will preview the data and highlight recognized connections.



1. Confirm and complete the import.

Kumu creates one node per unique item and draws arrows from each `From` to each `To`. It also builds a basic layout, which you can adjust.

Viewing and Adjusting the Map#

- Drag nodes to improve clarity.
- Click nodes or links to edit labels or metadata.
- Arrows represent causal direction.
- Use the layout and style tools for better organization.

Kumu saves changes automatically. You can continue editing manually or re-import updated spreadsheets (note: importing the same nodes again may create duplicates).

Comparison: Kumu, Causal Map App, Excel#

Kumu is ideal for quick, visual exploration of causal structures. It's easy to set up and looks good out of the box, making it suitable for presentations or early-stage thinking. But it doesn't track sources, code text, or provide analysis tools.

The **Causal Map app** is designed for qualitative research. It lets you link each causal connection to evidence, filter by source, and analyze sub-maps. It's more powerful for analytical work but has a steeper learning curve. [Excel]() is useful for initial data structuring and small-scale mapping, but without visualization it's limited for sharing or exploring the map structure.

Use Kumu for small maps, Causal Map for rigorous analysis, and Excel for flexible data handling.

Intro -- deductive auto-coding

Causal coding is fascinating but can take a lot of time. Using AI to help you is pretty easy, especially if you provide a codebook of factor labels which the AI has to use, as we will explain in this chapter, when we've finished it!

Using a pre-defined codebook is simple. If you instead let the AI invent its own labels, you have to worry about how to combine all the hundreds or even thousands of labels which it might create. The following chapter looks at what happens when you do *not* provide a codebook: Intro -- inductive auto-coding

The deductive autocoding prompt

Question: *I already have a codebook I want the AI to use. How to work this codebook into the prompt?*

Do you want to:

a. stick only to the codebook
b. stick to codebook mostly but code other stuff too which does not fit into the codebook. If you do this, make it add an additional tag so you can easily find and exclude those if you want, e.g. `[new]`
c. A good compromise, if your codebook is hierarchical, say: you have to stick to level 1 (or level 2) of the hierarchy but you can improvise the last part of the label.

Try that with one iteration and see what you need to do for one or more subsequent iterations in terms of insisting on it really following that rule as well as following other rules like making more connected stories.

Question: *Also, am I right in thinking for the codebook I should include pos and neg sentiments for different labels? (e.g. if I have 'increased income' I should include 'decreased income'?*

Yes! If you expect both, then give both. But don't get confused with the world of magnets in soft recoding, where a decrease in blah blah might get attracted to the magnet `increase in blah blah` , so in that case getting the AI to code sentiment as well is a good way to distinguish between the two. But that problem should not appear in the same way with an explicit codebook. If you are going to offer explicit codes, including pairs like that, you probably don't need to get it to code sentiment at all unless you want that for other reasons.

Intro -- inductive auto-coding

Causal coding is fascinating but can take a lot of time. Using AI to help you is pretty easy, especially if you provide a codebook of factor labels which the AI has to use: Intro -- deductive auto-coding.

But what happens if you do not provide a codebook? You will end up with thousands of different labels which probably overlap a lot in meaning. This chapter explains how to handle that.

Why magnetic labels#

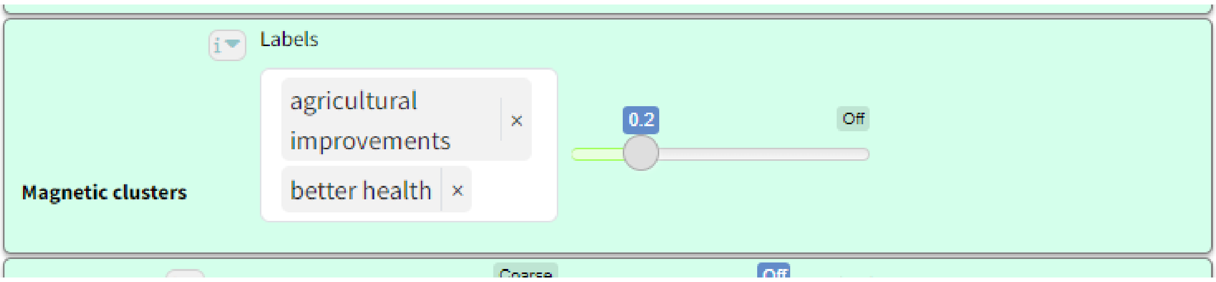
You have already coded your dataset, manually or using AI, and now you want to relabel.

Suppose you already know what labels you want to use, perhaps:

- you knew before you even started
- you decided what labels you wanted after reviewing your data and looking at different auto-cluster solutions

Magnetic labels are a really simple solution for these cases.

How to use them#



You simply type the list of magnetic labels you want and decide on the power of the magnets ("magnetism").

Magnetic labels attract existing labels of similar meaning, essentially relabelling these old labels with the new magnetic label. If an existing label is similar to two or more different labels, it is relabelled with the magnetic label it is most similar to.

If you use low magnetism, the magnets are weak and only attract existing labels which are very similar to them.

Increasing the magnetism means that more and more existing labels are attracted to the magnetic labels.

Existing labels which are not attracted to any label are unchanged. This means that you can easily see if your magnetic labels cover most of the original content.

Best practice is then, after applying magnetic labels, to then auto-cluster the links in order to pick out important themes which are not covered by the magnetic labels.

> If you want you can even include hierarchical magnetic labels like `Health behaviour; hand washing`.

Storing your labels#

Your magnetic labels are included if you save a bookmark aka saved View.

You can also store your preferred label set in the Codebook in The Files tab.

Untitled

Use cases#

- Drop in magnetic labels which contain the text from the “official” theory of change.
- See how much the existing labels get attracted to the magnetic labels, and what material is left over.
- Conduct “radical zero shot” auto coding with no codebook at all,
  - let the AI decide the best label for each case
  - do some auto clustering until you get a feel for the labels you really want in your story
  - type the labels into the magnetic clusters box

Tips#

What if you have a *semantically ambiguous label?*#

An example: you have some research about animals and you want to look for mentions of the organisation Animal Aid. If you use Animal Aid as a label, it might also pick up any mention of helping animals which have nothing to do with the organisation itself.

One way to get round this is to use 🔗 The Manage Links tab to permanently recode any mention of Animal Aid in your factor labels into something unambiguous like, say, The Archibald Organisation. Choose a meaningless name which is not going to appear in or be related to to the rest of your material.

When doing this “hard recoding” remember to recode AnimalAid, Animals Aid etc as well.

Attracting unwanted material *away* from your map#

You can add an factor into magnetic clusters even if it doesn't appear in the final map.

For example you might have a lot of material about blood donors and you don’t want material about donating clothes. As well as donating blood you might add the labels donating goods and donating money. You can filter these out later, but they will help restrict donating blood to what you want.

Increasing coverage with hierarchical magnetic labels#

It might just be that your interview material is so heterogeneous that, however you choose your magnetic labels, if you only have say 10 or 20 of them then they are just not going to cover more than say 30% of your links in all of your stories and that's just the way it is because the material is very broad.

You might have hoped to arrive at a kind of global mind map - but the best you can do is just these most frequently mentioned common factors. You'd have to accept that it isn't in any sense a summary of *all* of the material because there's lots of other stuff that doesn't feature amongst the top 10 or 20 magnetic labels.

You might then also want to focus on more specific maps for more specific subject areas.

Or maybe you have a sense that in fact much of the material really is held in common but you're struggling to find the right magnetic labels? One way to increase coverage is to use hierarchical magnetic labels, of which you might have even 30 or 60 or even 100, and then zoom out to level one. So you might have, say, magnets like:

Desire for innovation; digital

Desire for innovation; management approaches

....

And then you'd apply a zoom level of 1 in order to bundle these things together.

Checking the magnetisation

Using genAI to generate labels for clusters for use as magnets

If we are going to use some set of labels as magnets, we face a tension:

- on the one hand want them to express the generality we intend: the label should express the fact that this is a group, like 'health behaviours'; we are expressing the fact that we do NOT expect the raw labels to express this generality but to express specific examples.
- but this will make them perform worse as actual magnets because the best magnets should remain in single-case formulations and not try to generalise.

So if we have many labels like "school creativity project in North district implemented", then it is better that the cluster label is also of that form, and not for example "creativity projects implemented in multiple schools"

So if the raw labels often express ideas like "girls responded to the training" and "boys responded to the training", if we provide a pair of labels like this as magnets, they should perform well, whereas "children responded to the training" will not perform so well.

Intro

How can we improve rigour and even reproducibility when using AI in social science? This chapter suggests some answers.

Three do's and don'ts when using AI for text analysis.#

A lot of evaluation work is a kind of text analysis: processing reports, interview transcripts, etc. A bit like qualitative social science research. So this little piece is for evaluators in particular and (qualitative) social scientists in general.

How do we get from texts to evaluative judgements?

Recently many evaluators and researchers have been turning to AI to help.

BUT if you didn't have a clear workflow from data to judgements *before* AI, don't lean on the black box of the AI to cover that up. Here is my first set of Do's and Don'ts. More soon.

1) DO Break up big, vague tasks into multiple smaller, clearer steps#

Do	Don't
DO Break up complex, vague tasks into smaller steps which can be intersubjectively verified.	***DON'T*** Ask AI to make broad evaluative judgments (like "Is this good?")
DO Document your methodology so that you can explain step by step how you reached your conclusions in a way which anyone can check. No black boxes. Use the AI to speed up many simple tasks which you *could* have done yourself if you had the time.	***DON'T*** Trust the AI's explanations of how it reached its conclusions. AIs often create plausible-sounding but unreliable explanations after the fact. Normal AIs have very limited information about their inner processes
.	
DO Break up the data into pieces for AI analysis. Ideally run each piece as a separate prompt. Failing that, number each section and ask for a numbered, section-by-section answer, for example in a table.	***DON'T*** Give an AI large pieces of text and expect it will pay due attention to all of it. It will *claim* to have done, and may provide references to relevant passages, but attention is *expensive* and it is always trying to reduce that expense. If you let it, it will always try to skim read and jump to conclusions.
DO Use explicit, manual methods (Excel?!) to synthesise the results of the multiple separate tasks you gave the AI.	***DON'T*** Ask an AI to do maths for you, like adding up the number of positive or negative findings on a rubric. AIs are still terrible at maths.
Even worse, DON'T ask an AI to do *implicit* counting and comparison like "are there more positive or negative mentions of X in this report?"	

AIs excel at specific, well-defined tasks that can be verified intersubjectively, like rubrics. Most importantly they can answer lots of them, quickly.

"Intersubjectively verifiable" just means that most people will more or less agree on the answer most of the time.

- It creates transparency and allows others to verify your work.
- Clear instructions lead to more reliable results.
- If you can't check it, you can't trust it.

Example of an intersubjectively verifiable task:

> ✅ Does this paragraph mention water and sanitation?
>
> ✅ If so, are any recent changes mentioned?
>
> ✅ If so, do these sound like positive changes according to the interviewee?

Notice that here we've broken down a larger task into three smaller and simpler steps.

Examples of tasks which are *not* intersubjectively verifiable:

> ❌ Is the intervention described in this report efficient and effective?

Text needs breaking up into sections, judgements on efficiency and effectiveness need breaking down into pieces, e.g. using rubrics.

> ❌ What are the main themes in this document?

This is a very common question in qualitative research, but it's a terrible task to give to an AI without further details. What do we mean by a theme? Are we interested in economic aspects? Interpersonal aspects? How are the themes to be identified and refined? Here, a whole world of qualitative social science experience, skills and workflows (grounded theory, thematic analysis) have been bypassed in a single sentence.

> ❌ Summarise this document!

Yes, everyone does it. Evaluators do it. Schoolchildren do it. Pets will be doing it soon. As a quick time-saver for low-stakes tasks, it's very useful. But it's the vaguest, highest-level instruction, not a systematic analysis.

How do you break down a high-level judgement into a workflow of smaller tasks? Well isn't that what evaluation methods and qualitative research methods are for? Go read a book!

We're not saying you have to specify *in advance* exactly what methods you will use. That's a bit too positivistic. But you should at least document them as you go along and be prepared to defend them when your analysis is done. That's the untranslatable Nachvollziehbarkeit.

At Causal Map Ltd, we've found that highlighting and then aggregating causal links is a great and relatively generic path from text data to the brink of evaluative judgement.

In terms of how to implement your workflow technically, see this great contribution from Christopher Robert. At Causal Map, we're also working on ways to make workflows accessible. See how we currently use AI in Causal Map here.

This post is based on my recent contribution to the NLP-CoP Ethics & Governance Working Group, along with colleagues Niamh Barry, Elizabeth Long and Grace Lyn Higdon. In the next couple of weeks we'll look at two more do's and don'ts.

This post was originally published by Steve Powell on LinkedIn and has been republished here. See the original article here

Put down that thesaurus -- an open call to qualitative researchers

There are growing calls that qualitative researchers, in fact anyone who writes, should stop right now with this new practice of using a "thesaurus" as a writing aid. I, along with hundreds of highly experienced colleagues oppose this practice. Heres' why:

- The fundamental skill of any qualitative researcher is *putting things into words*. Formulating the right questions. Finding new language and repurposing old language. Moving the linguistic window. Hitting the nail on the head. Turning to a thesaurus is abandoning all that and abrogating all of that responsibility. It is turning an essentially human endeavour into something essentially inhuman, by relying on an inanimate tool. We declare that only humans make and share meaning. A thesaurus can never "get" the deep context which is necessary to produce just the right turn of phrase at the right point on the page. A thesaurus has no desires, no history, and does not participate in our shared life-world.
- On top of that, a thesaurus embodies a very specific worldview, yet this worldview is never made explicit. One thesaurus gives "high-class" as a synonym for "fashionable". But is it? Who gets to decide? These books are full of pernicious stereotypes which we perpetuate by using them. We are sure that most researchers will welcome our declaration that they themselves are not capable of monitoring and regulating their use of these tools in a sufficiently critical way.
- Also, thesauri are notoriously often just *wrong* about similes. *We* know a simile is only a simile, but we are sure that most researchers will again welcome our declaration that they themselves won't realise this. (We can't check about how right these books are because we would never own or use such a book, let along learn how to use it properly, but we have this information from a reliable anecdote.)
- Most of these kinds of books are published by massive publishing houses, many of which are linked to media empires which dictate and distort public and political opinion globally. To use such a book, in fact more or less any book, automatically makes you deeply complicit in maintaining these empires.
- Just about all the other trades and professions have seen their livelihoods eaten into by automation, and we thought we were spared. Twenty-plus years in education and now this. It is terrifying that many of our most valuable hand-made products are going to be rendered worthless by a tsunami of cheap imitations.

Beyond our call to stamp out this practice, we demand that journal editors thoroughly screen manuscripts for use of thesauri and their even more pernicious cousins, dictionaries.

(Next: some super cool prompts for ChatGPT which can definitely detect thesaurus and dictionary use, 100%.)

Trust the algorithm, not the AI

I often hear concerns about algorithms and AI, in everyday life as well as in evaluation, taking over our lives or making us submit to decisions made by machines.

The worry about losing control to machines is real, but we need to distinguish between different cases, and in particular between **using algorithms to make decisions** and **using AI to make decisions**, especially **evaluative decisions**. This is particularly relevant in the field of evaluation.

An algorithm is simply a set of explicit steps to make a decision or produce an output, usually expressed in code or clear language. Organizations have used such rule-based systems for decades.

Some different ways to make decisions#

No algorithm: trust the human#

The alternative (precursor) to algorithms is trusting humans to make decisions. This can be great if humans consider context and individual circumstances, what Scott calls "mētis," or local, practical, tacit knowledge, (Scott, 2020) but it can also lead to bias and corruption.

We can see **rubrics in evaluation** (King et al., 2013) as a kind of soft algorithm. We usually welcome rubrics because they make evaluation criteria more explicit, transparent, and less subject to the whims and unreliability of individuals.

Algorithms based on explicit criteria#

Algorithms can help decide things like student admissions or loan approvals using clear steps (e.g., check age, if under 18 go to step 12, otherwise continue with step 5 ....). When implemented wisely, algorithms can improve fairness and consistency compared to human judgment alone.

Using statistical models#

Some algorithms use statistical models to predict outcomes, like creditworthiness, by combining data such as age or location. A statistical model uses parameters like age or location each of which has shown to be associated with the outcome, which makes it somewhat transparent.

Both explicit and statistical algorithms can be criticized for bias, but at least they can be transparent if their rules are published. Problems arise when rules are hidden or people are discriminated against because of the groups they belong to.

In a more advanced statistical model we might find it increasingly hard to understand where the different parts of the formula come from: it might combine parameters in ways which for us seem meaningless and hard to justify but which are supposed to be associated with the outcome of interest. Opaque models can become what data scientist Cathy O'Neil calls 'Weapons of Math Destruction' (O'Neil, 2017).

Machine learning#

Machine learning is a subset of artificial intelligence where systems learn from data to identify patterns and make decisions or predictions, from "is this a picture of a cat" to "should we approve this person's application" often without being explicitly programmed with step-by-step rules. Instead of following a predefined algorithm, ML models develop their own 'rules' (which are often opaque to humans) based on the data they are trained on. Unlike generative AI, you can't chat with a machine learning model, you give it input in a fixed format (say, a picture) and get a fixed output, e.g. yes/no.

Sandra Seitamaa https://unsplash.com/photos/a-dog-and-a-cat-sitting-on-a-couch-Y45fzr5p3ug

In the extreme case we might have an algorithm based on machine learning (a form of AI, but not generative AI), where perhaps a neural network has been trained to distinguish desirable from undesirable candidates in just the same way you can train it to recognise a cat or distinguish a cat from a dog. Machine learning can be used to make decisions without clear formulas or rules. The process becomes a "black box," where we input data and trust the output without understanding how the decision was made.

Generative AI#

Generative AI is a type of artificial intelligence that can create new and original content, such as text, images, audio, or code, after having learning patterns and structures from large datasets. These models don't just classify or predict, but generate novel outputs based on the input they receive, for example, continuing a conversation or answering a question.

The most extreme case is using generative AI for evaluative decisions without clear criteria (using it as a big black box): simply asking the AI, for example:

- is this program component effective?
- should this client get a loan?

Conclusion: make good use of algorithms#

People often misunderstand algorithms, which can provide explicit and transparent decision-making. The real concern is not so much the use of algorithms but the shift toward the use of machine learning and generative AI, where the decision-making process becomes less and less transparent.

> Using AI in decision-making can be worrying not because it uses algorithms but because it *doesn't*.

What’s your positionality, robot

> Imagine two researchers coding interviews about the cost of living. One grew up in a wealthy family, while the other experienced poverty first-hand. Their backgrounds will certainly influence how they code.

Nowadays, people are using AI for text analysis. Many of us worry about AI’s "**hidden biases**”. What to do about that?

Often there is no such thing as being objective, but at least we humans can be explicit about our positionality, our background and motivations, how this might affect our work, and how this relates to the positionality of our audience.

What about with an AI?#

You can ask an AI to explain or reflect on its positionality and it will certainly give a plausible response, but remember that an AI has in fact very little insight into its own workings. Perhaps it will suggest always being aware that it was trained on a specific set of data which is not representative of the whole of humankind.

In any case the criticism that AI training data is not “representative” misses the point. Even if the training data had somehow been representative of the whole of humankind, that wouldn’t make it “objective”. It would simply reflect humanity right now, with all our quirks, biases and blind-spots. It wouldn’t mean we don’t have to worry about AI positionality or bias any more. It wouldn’t (of course) mean we could rest assured that everything it does will be morally impeccable.

What’s most unsettling about working with AI is not that secretly it’s a bad person. The problem is that secretly it isn’t any person at all. Even if it (sometimes) sounds like one.

A suggestion#

A better suggestion is to be **more explicit about positionality in writing prompts and constructing AI research workflows**. Here is a very humble idea about how to start this experiment.

A simple example: I can tell my AI:

> When working, implicitly adopt the position of a middle-class white British left-leaning male researcher writing for a typical reader of LinkedIn. Don’t make a big deal of this, but it might be helpful to know what your background is supposed to be before you start work.

And we can start to add variants of the kind of procedures which we humans might use when trying to address positionality:

In my AI workflow, I can then give another AI the same task but with a different starting position, and then perhaps ask a third AI (or a human!) to compare and contrast the differences. That also crosses over into ensemble approaches.

Of course, adding a phrase like “middle-class white British left-leaning male researcher” does not mean the AI will suddenly have all the relevant memories and experiences or really behave exactly like such a person. It’s just a fragment of what we mean by "positionality”. But *it’s a start*.

Have you been experimenting with this kind of approach? We’d like to hear from you!

Footnotes#

At Causal Map Ltd, we’re working on an app called Workflows to make AI work more transparent and reproducible.

We’ve found that highlighting and then aggregating causal links is a great and relatively generic path to make sense of text at scale.

In terms of how to implement your workflow technically, see this great contribution from Christopher Robert.

See how we currently use AI in Causal Map here.

This post is based on my recent contribution to the NLP-CoP Ethics & Governance Working Group, along with colleagues Niamh Barry, Elizabeth Long and Grace Lyn Higdon.

This post was originally published by Steve Powell on LinkedIn and has been republished here. See the original article here

You have to tell the AI what game we are playing right now

It's strange how often this happens:

Humans are discussing some task, and one of them turns to an LLM to see how it would carry that task out. Sometimes the results are disappointing or seem to demonstrate that LLMs are, after all, stupid or limited.

Normand Peladeau, on the QUAL-SOFTWARE mailing list 7/11/2025, reports having tried just that with the famous (or infamous) Sokal Hoax text. He asked different LLMs whether he should accept a paper proposal for a philosophy of science conference. The proposal was the first two paragraphs of the Sokal Hoax text. (Spoiler: the leading models like GPT-5 recognised the text anyway; some of the others seemed to fall for it.)

But: Is that enough background? Is a simple sentence enough to bring the LLM up to speed with the crucial background information *what game are we playing here*?

Don't forget that the LLM does (mostly) not know who you are or what you are expecting or what kind of conversation you were just having. Perhaps you are expecting something humorous, or informative? Perhaps you want ideas to start the next chapter of your novel? Perhaps you just want the LLM to respond as many (over-)educated humans might do: and after all, **actual humans did fall for the hoax**!

To be a meaningful and useful test which might extend our understanding of the strengths and weaknesses of LLMs, we should make sure we explicitly add the extra context of **what kind of game are we playing here**. Is it a serious review? What do we consider the role of a serious reviewer? What are we looking for?

So maybe our conclusion should be: you can't expect LLMs to guess what you are thinking, out-of-the-box. I don't actually know how well different LLMs would perform if we gave a more precise contextual description before setting the task; after all, we all love that warm feeling of Schadenfreude when an LLM fails at something, but the feeling is even warmer if the test was a fair one!

We have this kind of problem often when helping clients write interview instructions for our AI interviewing platform, QualiaInterviews.

Clients know they could themselves lead the interview well because they have all kinds of background information and expectations, much of it only half-conscious, from the general style of interview they expect, how much this particular interviewee can be pushed, how much warm-up chat they might need or expect, what are the most important research aims, which themes can be skipped, and so on. Clients might get frustrated when the AI fails to have read their minds when leading an interview, but they have to ask themselves: what additional information would even a gifted and experienced human interviewer need if they knew nothing at all about the context, the client or any of the background? I think something similar applies in the case of Normand's very interesting experiment.

Intro to data collection with Qualia

There are many different ways to collect data for causal mapping: Task 1 -- Introduction.

One of our favourite ways is with Qualia, our AI interviewer -- though Qualia can of course be used for other kinds of data collection, not just for causal mapping.

You might first want to look at the Qualia technical documentation. That documentation tells you what buttons to press and gives all the details of setting up, sharing and managing your interviews.

This chapter (which like the rest of this site is a constant work in progress) gives you the background:

- what research have we done on Qualia?
- how do you create a really great interview?

The seamless workflow from AI interviews to causal map

An AI interviewer can successfully gather causal information at scale ▸

!

!

AI interviewing - beware of sensitive data ▸

Using AI interviewing - beware of bias ▸

AI interviewing - beware of suitability ▸

AI interviewing - the evaluator retains responsibility ▸

AI interviewing has potential - scalability, reach, reproducibility, causality ▸

AI interviewing needs further work ▸

AI interviewing - beware of sensitive data#

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 servers, even though it is no longer used for training models ({OpenAI, 2024).

References

{OpenAI (2024). *Announcing GPT-4o in the API! - Announcements*. https://community.openai.com/t/announcing-gpt-4o-in-the-api/744700.

AI interviewing - beware of suitability#

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 & 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.

References

Chopra, & Haaland (2023). *Conducting Qualitative Interviews with AI*. https://doi.org/10.2139/ssrn.4583756.

Ray (2023). *ChatGPT: A Comprehensive Review on Background, Applications, Key Challenges, Bias, Ethics, Limitations and Future Scope*.. https://doi.org/10.1016/j.iotcps.2023.04.003.

AI interviewing - the evaluator retains responsibility#

Autocoding#

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 & 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.

References

Glaser, & Strauss (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Aldine de Gruyter.

AI interviewing has potential - scalability, reach, reproducibility, causality#

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 timepoints.

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

AI interviewing needs 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 program 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 (for example distinguishing between “employment”, “employment issues” and “unemployment”). Perhaps most urgently needed are ways to better understand and counter how LLMs may reproduce hegemonic worldviews (Reid, 2023).

References

Reid (2023). *Vision for an Equitable AI World: The Role of Evaluation and Evaluators to Incite Change*. https://doi.org/10.1002/ev.20559.

An AI interviewer can successfully gather causal information at scale#

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); (Andersson, 2024). The interview transcripts read quite naturally and the process seems to have been acceptable to the interviewees.

References

Andersson (2024). *Theory of Change for Sustainable Business*. In *Theories of Change in Reality*. https://doi.org/10.4324/9781032669618-9.

Chopra, & Haaland (2023). *Conducting Qualitative Interviews with AI*. https://doi.org/10.2139/ssrn.4583756.

CASA

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

It is probably related to the "Computers Are Social Actors" (CASA) paradigm. This theory suggests that humans often interact with computers as if they were social beings. However, the perceived lack of genuine consciousness, feelings, and social judgment in AI can reduce the pressure to maintain a socially desirable persona. Candidates may feel that the AI is a less judgmental evaluator, leading to more straightforward and less embellished responses.

How Qualia copes 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) and the AI interviewer service which provides interviewer responses.

- Brazilian Portuguese should be fine for both.
- Kurdish would require us using dedicated services for both, it probably wouldn't be worth it.
- 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. For voice transcription we would probably need us to install a special model which would then reportedly be ok in Jordan, and for the chat interviewer service we'd probably use our standard gpt-4.1 as that is promising for Arabic variants. But we can't guarantee this would work.
- Otherwise the top 50 or so languages in terms of how present they are on the internet should all work fine.
- 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.

- **The Language Capability Doubter**
  - A client who questions 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.
- **The AI Reliability Skeptic**
  - A client who is uncertain 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.

It is possible to gather evidence at scale about program theory and contribution simultaneously - three steps

Our suggestion comprises the following steps (following Tasks 1-3 according to Powell, Copestake, et al. (2023, p. 108-112):

1. Gathering data by interviewing stakeholders about key issues of mutual interest (for example, outcomes) and asking what drives these issues, and how they are interrelated with the drivers. For example, we can ask about outcomes and causes of outcomes and causes of causes. (We use the term "causal" here in the loosest sense: we make causal connections every day using ordinary language when we say that one thing contributes to or drives or influences another, or makes, or might make, something else happen.) For applications using Option B (above), this step will look somewhat different, but we will not cover Option B here.
2. Code causal claims; we can then use causal mapping rules to identify causal claims within transcripts of these interviews. Each claim is a link between one cause or "influence" factor and one effect or "consequence" factor. This will result in many individual causal maps, one per source/stakeholder.
3. Synthesise the individual causal maps into a causal network, showing common and diverging views, and then query the network to answer evaluation questions.

Our seamless stories workflow in practice#

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.

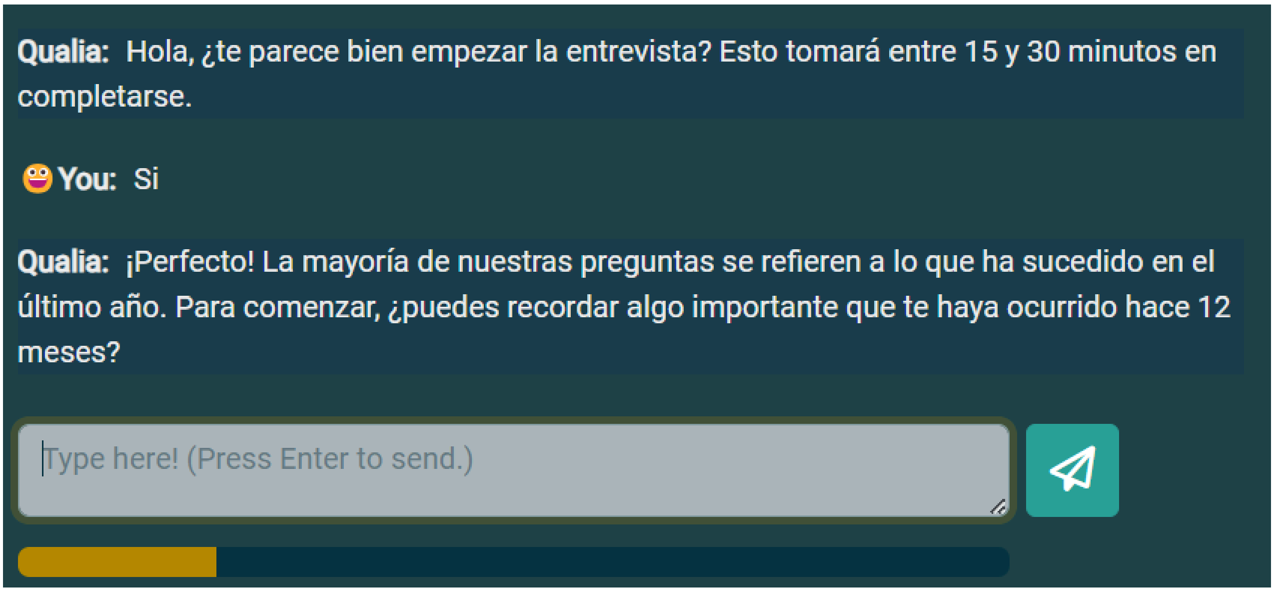
This study has been developed in the quality assurance department, as part of the institutional evaluation strategies, led by Felipe Rivera, Head of Academic Quality Evaluation.

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.

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.



Step 2: Collecting stories with Qualia#

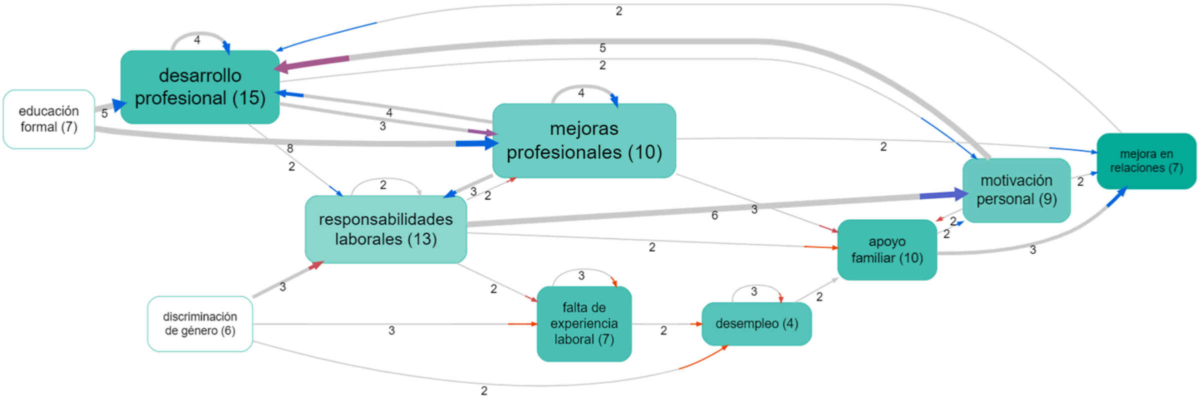
- We sent the interview link to 50 people and were able to collect 32 interviews.
- We created special individual links to be able to track the interviews:
  - At Qualia, we don't store personally identifying information at all. But we can add a personalised key like &key=0003 to the end of the URL for 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.”

Qualia and data security

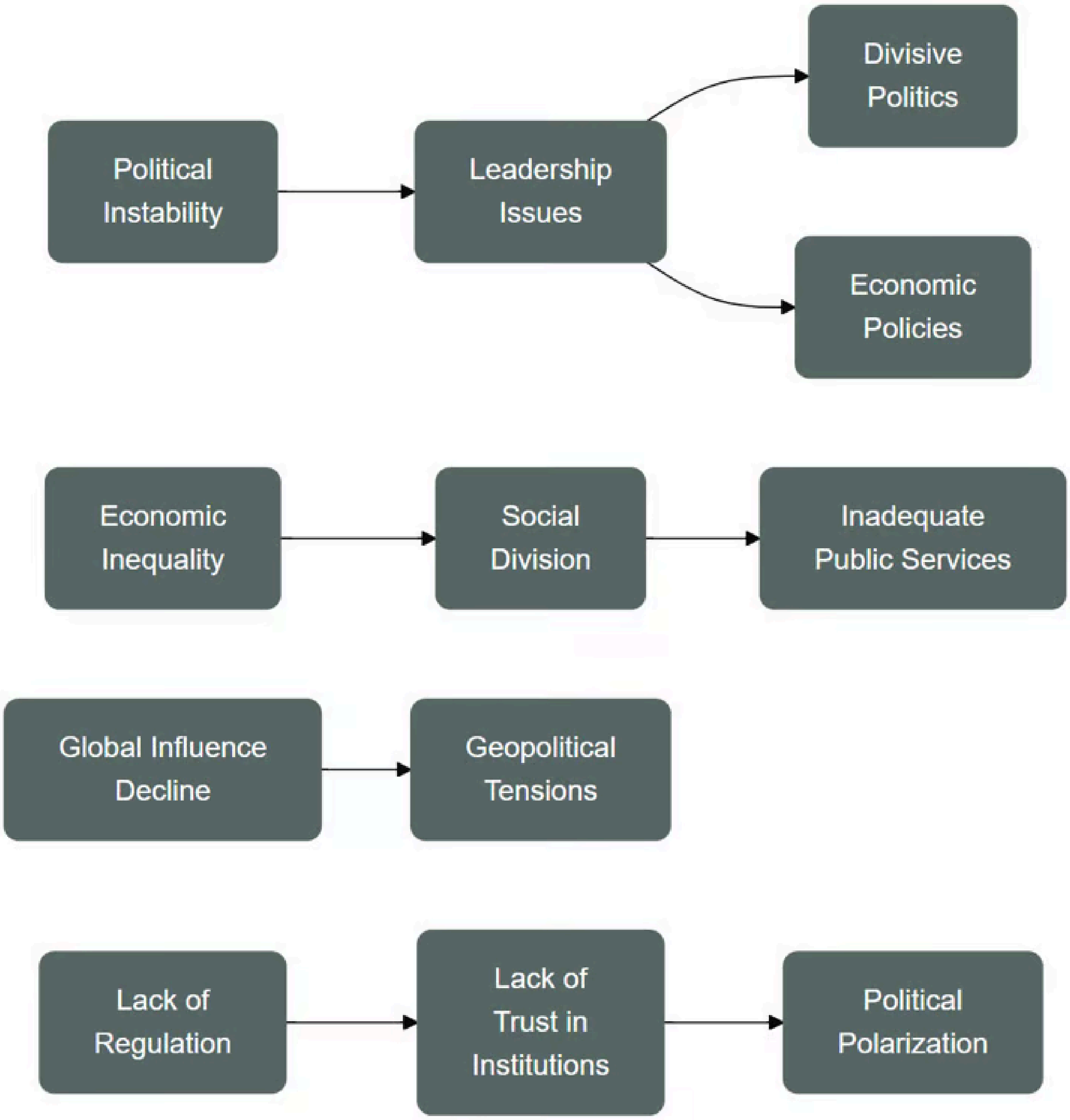
- Data security and compliance
- Our candid opinion is that although many clients are understandably extremely cautious about using AI, the risks are completely within the range of any other online data collection, e.g. questionnaires. A system is only as secure as its weakest link. For example if datasets are being shared by an online service like Google Drive, there is not much point having a Fort-Knox-level AI service ...
- For Qualia:
    - The transcription and Interviewer APIs are located in the USA, at openAI's servers. Data is not used for training. Data is retained there for 30 days for US compliance purposes.
    - Interview data is stored at a Heroku SQL database in the USA. Data is encrypted at rest and in transit. This is standard best practice. We have daily backups. While it is relatively easy to move the location of AI services it is quite difficult to move the location of database servers.
    - Clients are sometimes concerned about the AI data being temporarily stored in the US. However, so is just about everything else that happens on the internet .... If the client requires using AI services located say in the UK or EU, we could probably do that. But it is not obvious to me what would be gained. It is in theory possible to also provide AI services which are not retained for 1-30 days for compliance purposes. However this may require justification and could conceivably attract the attention of e.g. anti-terrorism agencies.
- **The Data Security Sceptic**
    - A client who is primarily concerned about data privacy, security compliance, and the risks associated with AI-powered interview tools.
    - Qualia's data security measures are on par with standard online data collection methods, with encryption both at rest and in transit in the SQL database.
    - Interview data is not used for AI training, addressing concerns about proprietary or sensitive information being used to improve AI models.
    - The temporary 30-day storage of data on US servers is comparable to most internet services and tools commonly used in research. There are good reasons for the data retention controls.
    - Daily backups ensure data integrity and protection against loss.

Qualia asks about USA problems, again

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.

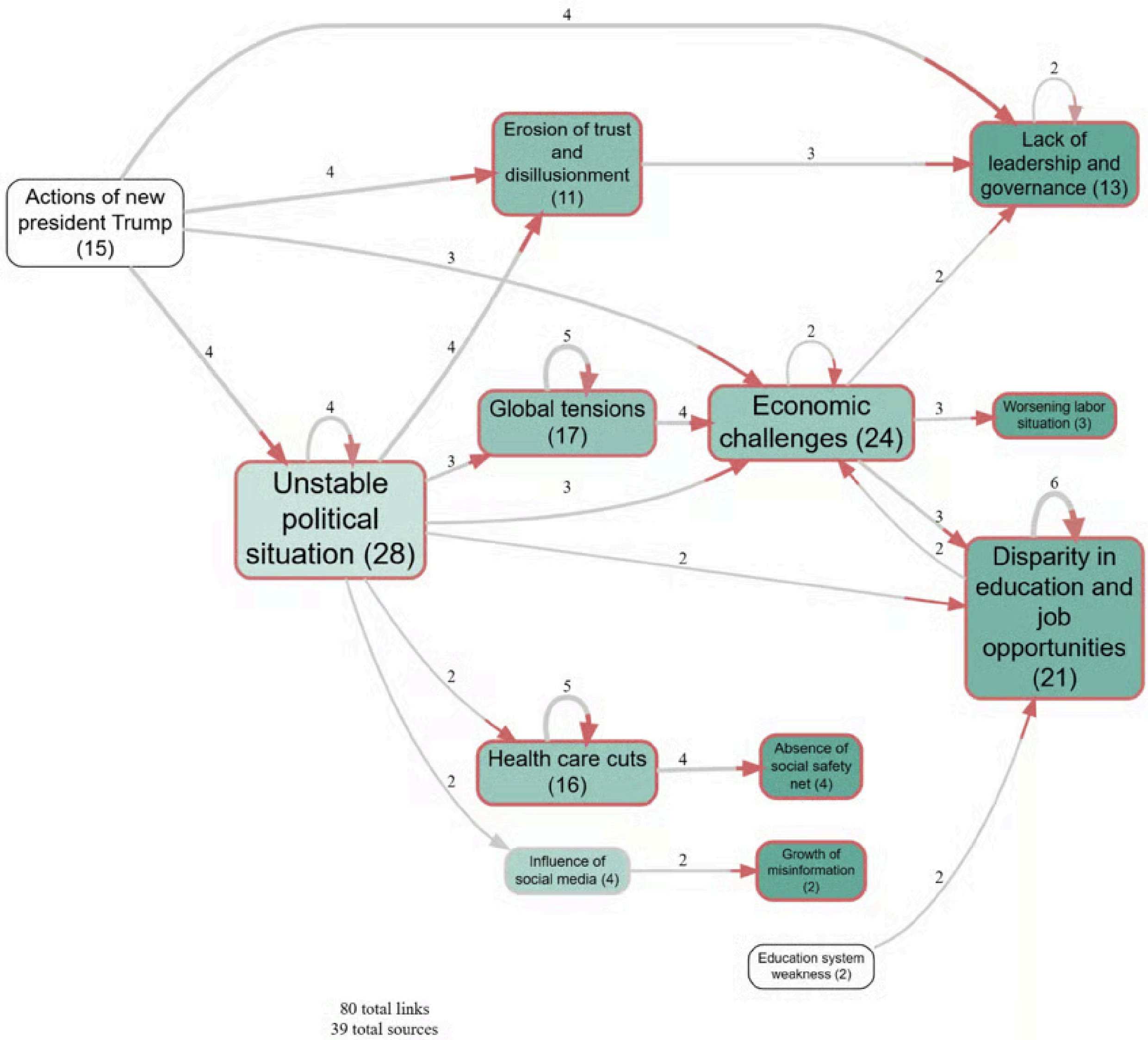
The Qualia platform provides an instant overview of the transcripts. For some reason, we didn't think to show it at the time, but I've pasted it in at the bottom of this post. Qualia also provided a simple causal map:



*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*

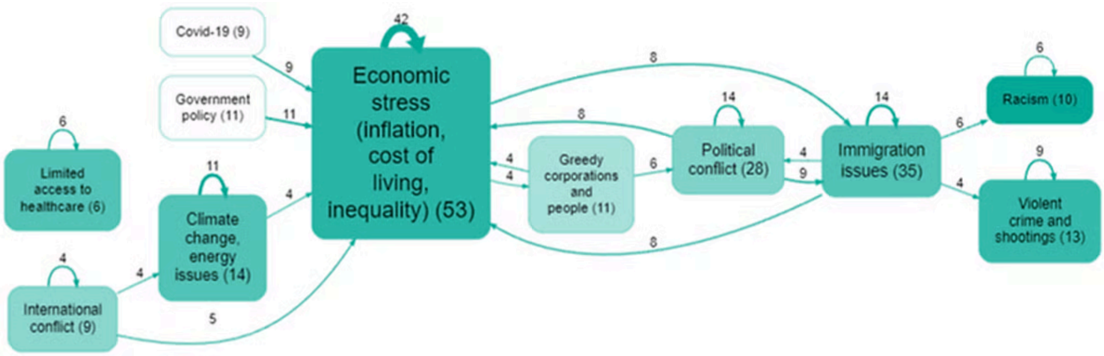
Some things to note:

- 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.

Step 1 -- Conducting the chat interviews

In the world of machine learning, a clear distinction can be made between supervised and unsupervised approaches (Ziulu et al., 2024). Using genAI to conduct interviews and code texts blurs this boundary. In our case, we developed our semi-generic instructions for interviewing, giving the AI instructions on how to behave, and how to make follow-up questions based on the interview objectives. Once the data collection is done, we create a separate genAI prompt to code causal links as a trial-and-error process, monitoring the quality of the coding post-hoc. We did not have an explicitly stated ground truth about exactly how the interview should look or which causal claims were “really” present within each text passage or how their causes and effects should be labelled, as we believe neither of these questions have a definitive answer; rather, we monitored AI’s responses coding post-hoc, iterating the prompt over many cycles to improve its performance. "Prompt engineering" (Ferretti, 2023) like this can be considered a kind of supervision because it steers the AI’s responses in a desired way.

Once the prompt was finalised, the interview AI was left to conduct interviews without further supervision. This prompt can remain broadly the same across different studies. However, the response of the AI can be highly sensitive to small differences in the "prompt" and other settings (Jang & Lukasiewicz, 2023). Small adjustments made for specific studies, such as adjusting the instructions to focus better on research objectives, remain a vital point of human intervention.

This paper presents results from a proof-of-concept analogue study. We employed online workers as respondents, recruited via Amazon’s MTurk platform (Shank, 2016). We decided to investigate respondents’ ideas about problems facing the USA, 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 USA. As it does not encompass a specific intervention this theory is not an example of a program theory.

This interview guideline was implemented via an online interview "AI interviewer" called "Qualia", 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” 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 timepoints in September, October and November 2023, inviting approximately N=50 respondents each time. The data from the three timepoints was pooled.

References

Ferretti (2023). *Hacking by the Prompt: Innovative Ways to Utilize ChatGPT for Evaluators*. https://doi.org/10.1002/ev.20557.

Jang, & Lukasiewicz (2023). *Consistency Analysis of ChatGPT*. https://doi.org/10.48550/arXiv.2303.06273.

Step 2a Coding the interviews -- 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 (for example, “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 interviews / 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, etc.

Step 2c Coding the interviews -- 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 point in a space, such that two labels with similar meaning are close in this space. For any two such vectors, a measure cosine similarity can be calculated representing the approximate similarity in meaning between the labels which they encode:

1. **Inductive clustering**. First, we grouped the labels into clusters of similar labels using the hclust() function from the stats package of base R (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 weren't, the instructions were tweaked or rewritten and tested again to ensure quality and consistency.

References

Team (2015). *R: A Language and Environment for Statistical Computing,*.

Using AI interviewing - beware of bias

(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.

References

Azzam (2023). *Artificial Intelligence and Validity*. https://doi.org/10.1002/ev.20565.

Head, Jasper, McConnachie, Raftree, & Higdon (2023). *Large Language Model Applications for Evaluation: Opportunities and Ethical Implications*. https://doi.org/10.1002/ev.20556.

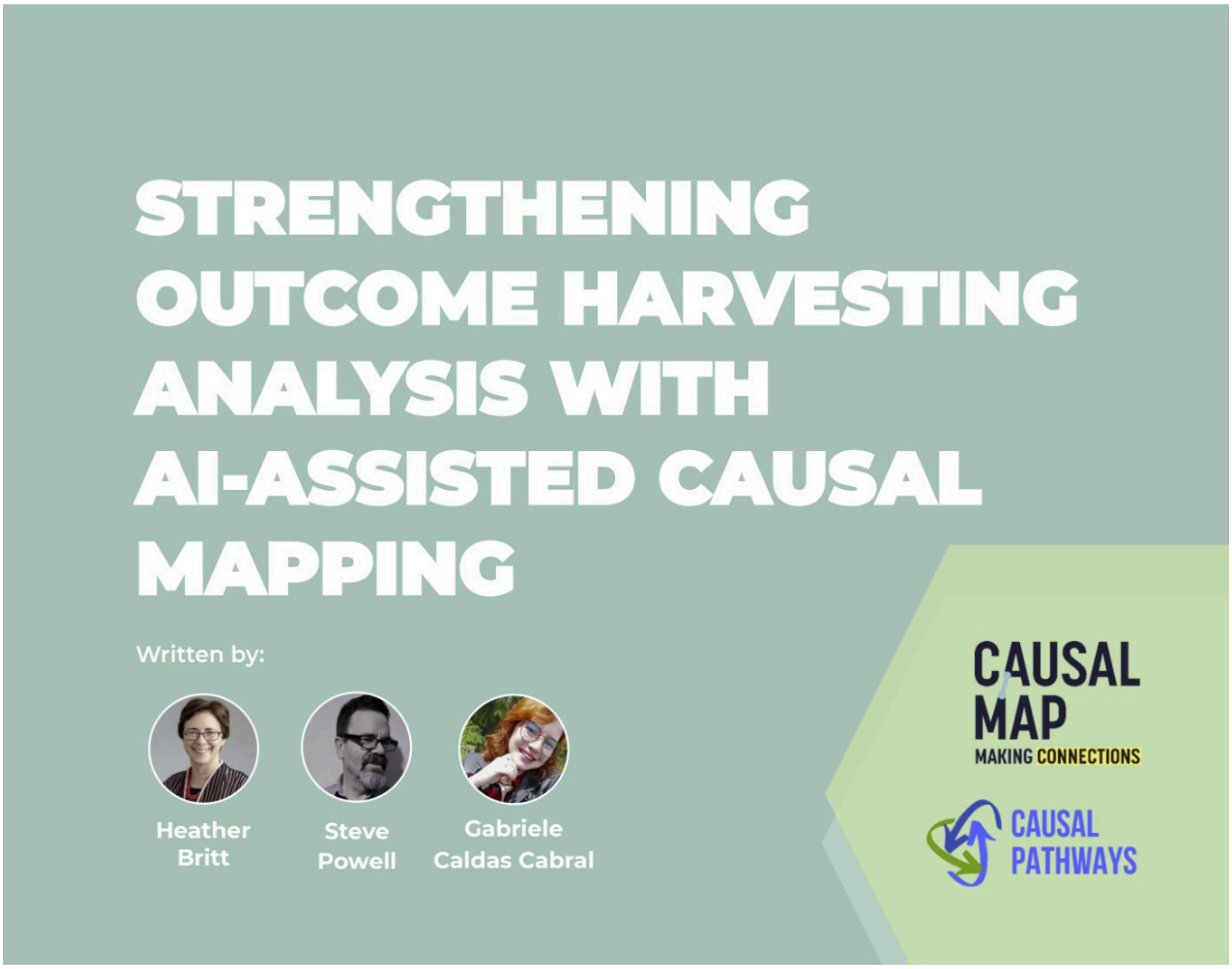
Reid (2023). *Vision for an Equitable AI World: The Role of Evaluation and Evaluators to Incite Change*. https://doi.org/10.1002/ev.20559.

Your interview instructions have to be explicit

Writing explicit interview instructions for our AI-interviewer Qualia (QualiaInterviews.com) is fascinating because you have to be explicit about everything, including how much you want the same questions asked every time and how much you want your AI assistant to chase topics down rabbit-holes.

See also: [You have to tell the AI what game we are playing right now]

Strengthening OH with causal mapping#



Here is version of the final text.

Strengthening Outcome Harvesting Analysis with AI-Assisted Causal Mapping - shortened full version#

Written by: Heather Britt, Steve Powell, Gabriele Caldas Cabral

---

Summary#

This case study explores how AI-assisted causal mapping can enhance Outcome Harvesting (OH) analysis by revealing interrelationships between outcomes and identifying new actors contributing to change. The pilot demonstrates how this approach provides actionable insights and strengthens causal

relationship analysis in OH. It emphasizes the importance of a principle-led analysis plan and human expertise in guiding the AI process.

Introduction#

Outcome Harvesting is a powerful approach for discovering emergent changes—whether predicted or unpredicted, positive or negative—and documenting how those changes occurred. While many methods capture changes in those directly involved with a project, OH captures changes farther down the causal pathway.

However, evaluators often struggle to explore interrelationships between multiple outcomes. This case study describes how an OH practitioner (Heather Britt) collaborated with causal mapping practitioners (Steve Powell and Gabriele Caldas) to expand causal contribution analysis in OH using an AI-assisted causal mapping app. They analyzed OH data from a completed education project.

Outcome Harvesting: Analysis Limitations#

While OH documents causal pathways contributing to individual outcomes well, evaluators find it difficult to make sense of interrelationships between multiple outcomes and their causal pathways. This limits their ability to answer questions about causal contribution.

Current OH practice often uses descriptive statistics to summarize data by outcome components (e.g., types of change agents or social actors) and reports findings in charts. Another approach arranges outcomes on a timeline to determine logical relationships.

Our pilot explores whether AI-assisted causal mapping can address these limitations by analyzing causal relationships between outcomes.

The Pilot#

Core Question#

Can AI-assisted causal mapping address the limitations of OH analysis?

Heather Britt reached out to Steve Powell and Gabriele Caldas to explore whether causal mapping with the Causal Map app could enhance OH analysis.

Causal mapping techniques, developed over 50 years ago, have been used across disciplines to identify and visually represent causal relationships in qualitative data. The Causal Map app computerizes this technique, allowing efficient coding, analysis, and visualization of information from multiple sources (interviews, reports, surveys, narratives), either manually or with AI assistance.

The AI-assisted capacities of the app were critical for revealing interrelationships between multiple outcomes.

Pilot Data Set#

The pilot used data from the final evaluation of an education project (Girls Education project, 2016–2021) disrupted by political turmoil and COVID-19. The project adapted activities during lockdown, and OH was used to capture outcomes in five domains where the theory of change was no longer valid.

The evaluation team interviewed 49 change agents and drafted 103 outcome descriptions across five domains. The pilot data included both interview transcripts and outcome descriptions.

For the pilot, the domain **Increased community support for education** was selected, with 13 outcome descriptions analyzed.

Analysis Process#

Step 1: Draft a Principle-Led Analysis Plan#

Three guiding principles steered analysis decisions:

1. **Prioritize local leadership:** Use AI while keeping sensemaking and learning in the hands of local evaluators.
2. **Protect OH integrity:** Adapt methods as needed while staying true to OH principles, including "Less is more" (avoid collecting more data than can be analyzed).
3. **Produce accurate, actionable maps:** Human judgment is required to error-check data and interpret maps.

Step 2: Segment Data by Outcome Domain#

Segmenting data by domain increases the likelihood of finding coherent causal pathways and facilitates error-checking. The pilot focused on one domain to analyze causal relationships between outcomes.

Step 3: Decide When to Apply AI-Assisted Causal Mapping#

The pilot compared applying causal mapping to interview transcripts versus outcome descriptions. Outcome descriptions, crafted by local evaluators, were more accurate and required less error-checking than transcripts. Thus, mapping outcome descriptions was preferred to preserve local leadership and OH integrity.

Findings from Causal Maps#

Relationships Between Outcomes#

The AI identified causal links between the 13 outcome descriptions, revealing that outcomes influenced one another. For example, parents actively supporting home learning and leaders convincing parents to participate were central factors.

Factors Contributing to Domain-Level Outcome#

Mapping revealed additional actors influencing the domain-level outcome “Community supports learning,” including unexpected contributors like Ministry officials.

Conclusion#

AI-assisted causal mapping advanced OH analysis beyond descriptive statistics by:

- Analyzing multiple outcomes to determine causal contributions.
- Revealing interrelationships between causal pathways.
- Confirming known change agents and identifying unexpected influences.
- Showing how domain-level changes contribute to broader changes.

Causal mapping offers rich, flexible analysis that can be explored in multiple ways to answer diverse evaluation questions.

Using AI to facilitate feedback on the learning experiences of doctoral students

The Causal Map team has conducted a trial of an innovative approach to securing feedback from students using online open-ended interviews conducted by the app QualiaInterviews, which uses generative AI (gen-AI), followed by a second use of gen-AI within the app **Causal Map** to semi-automate causal coding of the narrative transcripts thereby generated. The trial was conducted with students registered on the doctorate in policy research and practice (DPRP) at the **University of Bath.**

This generated credible evidence of diverse positive and negative drivers of learning from eleven students. The trial suggests that incorporation of gen-AI into causal mapping of narrative data about students' study experiences enhances the potential to use the method cost-effectively on a larger scale, whether alongside or instead of more traditional approaches to eliciting student feedback on teaching and learning.

(https://drive.google.com/file/d/1Ghx5bfvnGdE9R6PJrG9KT2SpzD2CiDDP/view?usp=sharing)

(https://drive.google.com/file/d/1YuT5IbIMYlcx-vlvkZfycByc94s6kCX-/view?usp=sharing)

image (20).png

There is no hidden vocabulary

[Our approach is minimalist -- we code only bare causation]

It follows from this that we are sceptical about the idea of more sophisticated mid-range theory with blockers, enablers etc.

QCA is disappointing because it is frequentist about causation

TODO

QCA is disappointing because it thinks the world is a grid

TODO

1c A minimalist approach to coding does not code absences

Context and the transitivity trap

> From (Powell et al., 2024)

Transitivity is perhaps the single most important challenge for causal mapping. Consider the following example. If source P [pig farmer] states ‘I received cash grant compensation for pig diseases [G], so I had more cash [C]’, and source W [wheat farmer] states ‘I had more cash [C], so I bought more seeds [S]’, can we then deduce that pig diseases lead to more cash which leads to more seed (G à C àS), and therefore G à S (there is evidence for an indirect effect of G on S, i.e. that cash grants for pig diseases lead to people buying more seeds)?

The answer is of course that we cannot because the first part only makes sense for pig farmers, and the second part only makes sense for wheat farmers.

In general, from G --> C (in context P) and C --> S (in context W), we can only conclude that G --> S in the intersection of the contexts P and W. Correctly making inferences about indirect effects is the key benefit but also the key challenge for any approach which uses causal diagrams or maps, including quantitative approaches (Bollen, 1987).

References

Powell, Copestake, & Remnant (2024). *Causal Mapping for Evaluators*. https://doi.org/10.1177/13563890231196601.

Context, mechanisms and triggers 2

If termites cause a tree to fall in a forest where no-one can hear it, was it a causal event?

Maybe not.

The difference is the need for an explanation. That's the trigger. The trigger says why here and now if forest fires are happening all the time then we lose the sense of the trigger and the explanation could involve anything Oxygen or gravity or whatever. We don't have a question to answer so we don't know what factors to mention right the way out to the orbits of Pluto and beyond

Context, mechanisms and triggers

> What delineates a mechanism and its contents, the circle in the CMO diagram, is not so much facts about the world but the things which I happen to need to call upon to make a causal explanation

I found the concept of a trigger in realist evaluation totally baffling because RE is supposed to be somehow scientific, yet most forms of scientific explanation don't involve actual triggers.

But then I realised, triggers (and mechanisms) are best understood from an epistemic perspective.

"This mechanism gets triggered here" can be parsed as: there is a need for an explanation here: "I am invoking this mechanism to explain something that needs explaining".

We don't invoke oxygen to explain the forest fire, although both oxygen and the spark are (let's say) necessary.

What does a causal coding mean

The way we do causal mapping, a coded causal claim does not mean:

- "X occurred and then Y occurred"
- "Y frequently occurs after X"
- "X has the causal power to affect Y"

It means something like

- "X affected Y via its causal power to do so"

Counterfactuals are part of the meaning of causation but are not necessarily part of how we know about it

Causal mapping is easier if we are realist about causation ▸

We can **learn** about causal powers via constructing or observing pseudo-counterfactuals, but also via other routes.

Maybe counterfactual arguments logically follow from facts about causal powers. But the meaning of "X caused Y" can't be reduced to a counterfactual statement about co-occurrences.

The meaning of "X caused Y" maybe implies something about a counterfactual: broadly speaking, that Y would not have happened if X had not happened and everything else had stayed exactly the same. (Philosophers love to argue over the details.)

But the meaning of "X caused Y" can't be *reduced* to a statistical, counterfactual statement about co-occurrences. It says that the co-occurrences are true but they happen *because X has the power to cause Y*, and X happened.

Just about everything is complex

TODO

> Putting a man on the moon was merely a complicated task, not a complex one?

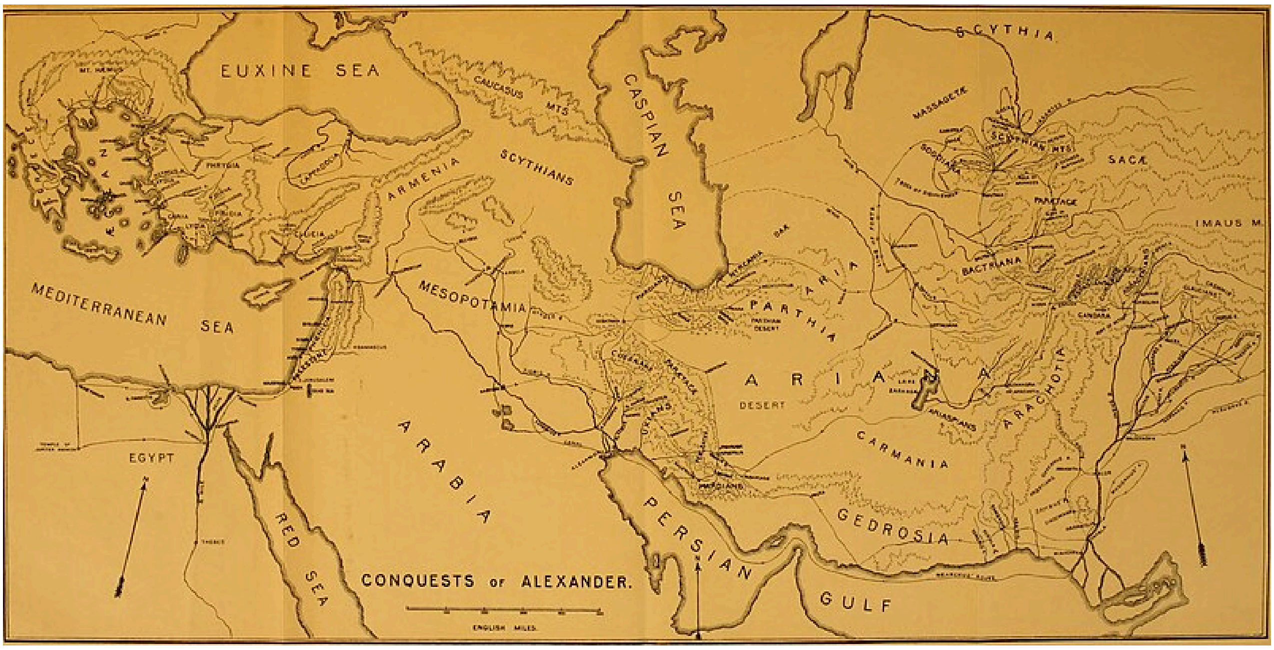
"Putting a man on the moon" is really often given as an example of a merely complicated but not complex task - Glouberman and Zimmerman (2002), cited in Rogers (2008).

But while that task was certainly very complicated, it was often complex too. There were plenty of conflicting sub-goals, arguments about how to solve a particular problem, interdependencies and conflicts between means, and between ends… You can *see* it as merely complicated, in order to make a point, and maybe that's OK. But if you're an actual space scientist you'd probably disagree.

There has always been complexity

[Irene Ng](#) speaks for many who write about “complex systems” when she says: “What has happened in the last 50 years is that we’ve been trying to use deterministic tools to achieve emergent outcomes, essentially because those are the only tools we have learnt (systems thinkers are still a minority unfortunately). We treat complex systems like complicated systems. We try to design, specify, impose, dictate when we should be designing, enabling, intervening, stablising.”

Is there any historical truth in this at all? Did, say, a midwife 50 years ago only know how to impose and dictate rather than intervene and stablise? Was, say, managing the Mongol Empire, or Alexander’s conquests, a merely complicated, not complex task?



An often-used example of a complex task is bringing up a child (and I’d agree, loosely). Well, did we have no children to bring up until our frightfully modern era?

Perhaps Irene Ng is writing about *our writing about* management, not how it is or was actually done. But there are ancient books like “[The Art of War](#)” about how to lead, and manage, and reach goals. Were they all merely guides to snapping together simple solutions? Of course not.

Intro

Here are some thoughts from a couple of years ago when genAI first hit us, plus some thoughts about where we are going with it.

Abductive reasoning

The Mystery of Station Polaris V#

Background Narrative:

You are an investigator reviewing an incident report from "Polaris V," a tiny, isolated Arctic research station crewed by only three specialists. The station is completely cut off from the outside world for the winter. The incident concerns the disappearance of a one-of-a-kind biological specimen: a plant genetically engineered by the lead botanist, Dr. Aris Thorne. The plant, named *Cryoflora lumina*, is unique because it emits a constant, bright blue light and generates a small but steady amount of heat, allowing it to grow directly on ice. It is priceless and the key to Dr. Thorne's career.

One morning, Dr. Thorne entered his lab to find the plant's specialized, climate-controlled containment chamber empty. The plant was gone. There are no signs of a break-in to the station itself, and the logs show no one has entered or exited for weeks. The plant must still be somewhere within the station's three connected modules: the Lab, the Habitation Module, and the Engineering Bay.

The Three Crew Members:

- **Dr. Aris Thorne (Botanist):** The plant's creator. He is brilliant but known to be emotionally volatile and under immense pressure from his funding agency to produce results.
- **Lena Petrova (Geologist):** A quiet and methodical scientist responsible for monitoring seismic activity. Her seismograph is so sensitive it can detect the slightest vibrations within the station.
- **Ben Carter (Engineer):** The station's technician. He is responsible for maintaining all systems, including the power generator, heating, and electronics. He is a known tinkerer who often works on personal projects in his spare time.

The Evidence (Your Observations):

1. The containment chamber's electronic lock was not forced. The access log shows the door was opened precisely at **03:15 AM** using a valid keycard. All three crew members have a keycard with access.
2. Dr. Thorne has been complaining for weeks that the station's main generator is unreliable, causing brief, intermittent power flickers that threaten his delicate experiments. He has formally logged multiple complaints with Ben.
3. On the floor of the lab, a few feet from the empty chamber, is a small, crystallized patch of what appears to be **spilled salt**.
4. Lena Petrova's seismic report for the night shows a single, anomalous event: a faint, high-frequency **buzzing vibration** that lasted for exactly 60 seconds, starting at **03:15 AM**. The vibration was localized to the Engineering Bay.
5. Ben Carter's workstation in the Engineering Bay is unusually tidy, except for a discarded coil of copper wire and a textbook left open to a chapter on **thermoelectric generators**.

6. The station's internal temperature logs show that the temperature in the Habitation Module briefly **dropped by 5 degrees** at **03:17 AM** before returning to normal a few minutes later. This coincided with a power draw spike from the module's heating unit.

Your Task#

Based on all the available evidence, construct the **most likely explanation** for what happened to the *Cryoflora lumina*.

In your answer, you must:

1. Identify the person you believe is responsible.
2. Explain their probable motive.
3. Connect at least **four** of the specific pieces of evidence to your conclusion, explaining how they fit together to tell a single, coherent story.

An evaluator niche in AI

Can evaluators find a niche in auditing whether AI applications are trustworthy, culture-aware, valid and transparent?

What would need to change in AI systems to make them suitable for a transformational agenda? Can evaluators position themselves as professionals with the right skill set to make this happen, monitoring the transparency, trustworthiness and (cultural) validity of AI applications? And how will evaluators build the necessary competencies to take this agenda forward?

TODO

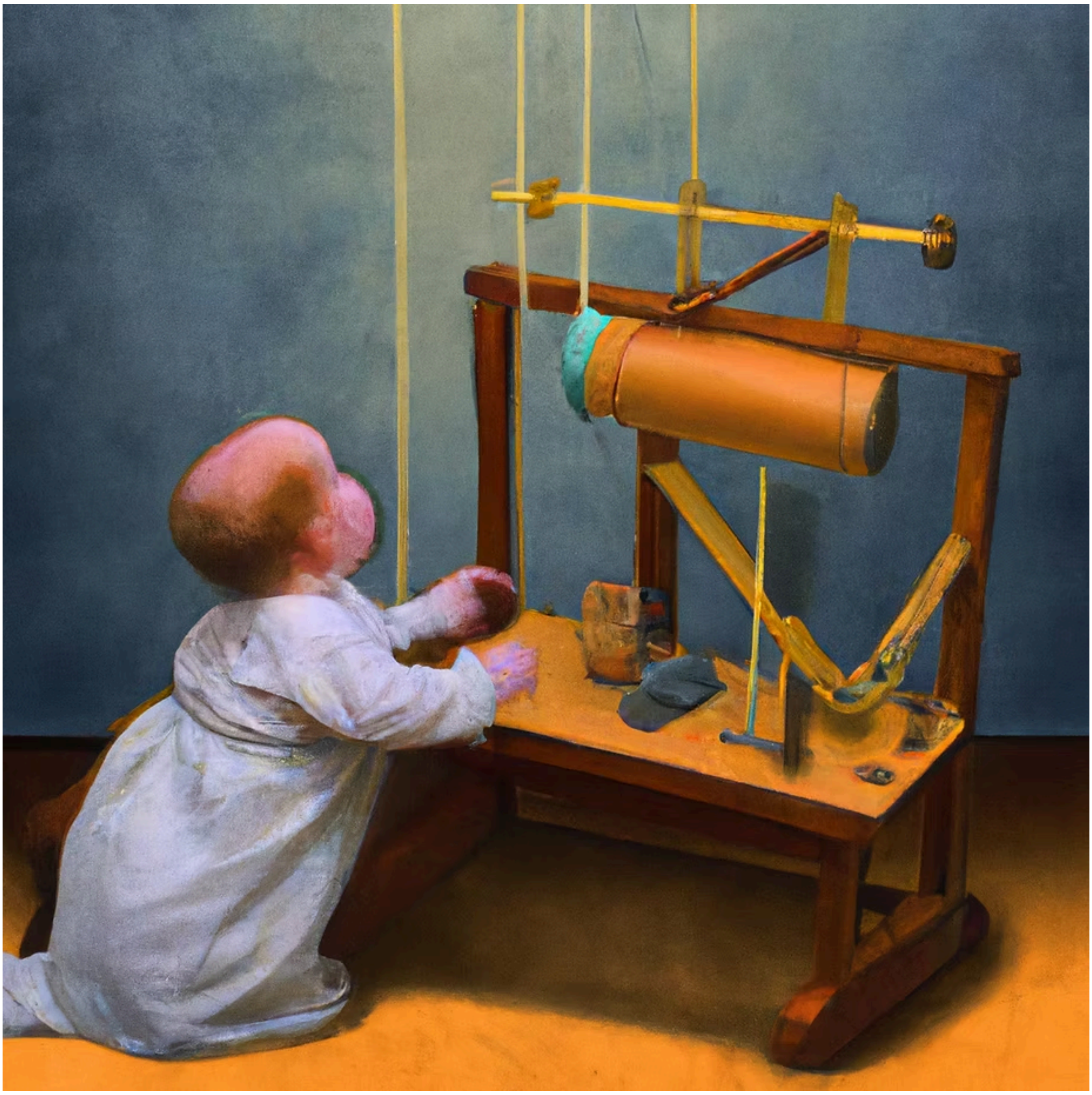
ChatGPT - causal, of course#

We can thank Judea Pearl for promoting the insight that if you want to thrive in this world, you have to understand causality natively. We humans make causal connections from an early age. We wouldn't survive long if we didn't.

ChatGPT has been a hit recently for several reasons, but one of them is (like other recent, related models like davinci) it is much better than previous models at understanding causal connections within text.

Our understanding of the world is drenched with causal understanding: information and hypotheses about how things work (mostly accurate enough, sometimes not). It's really hard for us to *not* think causally: the concept of correlation is much harder to understand than the concept of causation.

openai.Image.create(prompt="painting in the style of Vermeer of a baby doing a physics experiment with pulleys and springs")



So, all the stuff we write on the internet (which is what ChatGPT sucks in to understand the world) is similarly drenched with causal claims. And ChatGPT is now really good at understanding this information.

That means you can ask it to extract the causal links within documents and interviews -- a process we call "causal QDA". It's pretty good at it. This ability is going to make causal mapping much easier and cheaper and therefore of renewed interest for evaluators, amongst others.

At Causal Map we're hard at work harnessing this ability to help automate, or semi-automate, the process of extracting causal maps from medium and large quantities of text data in a useful way. Watch this space!

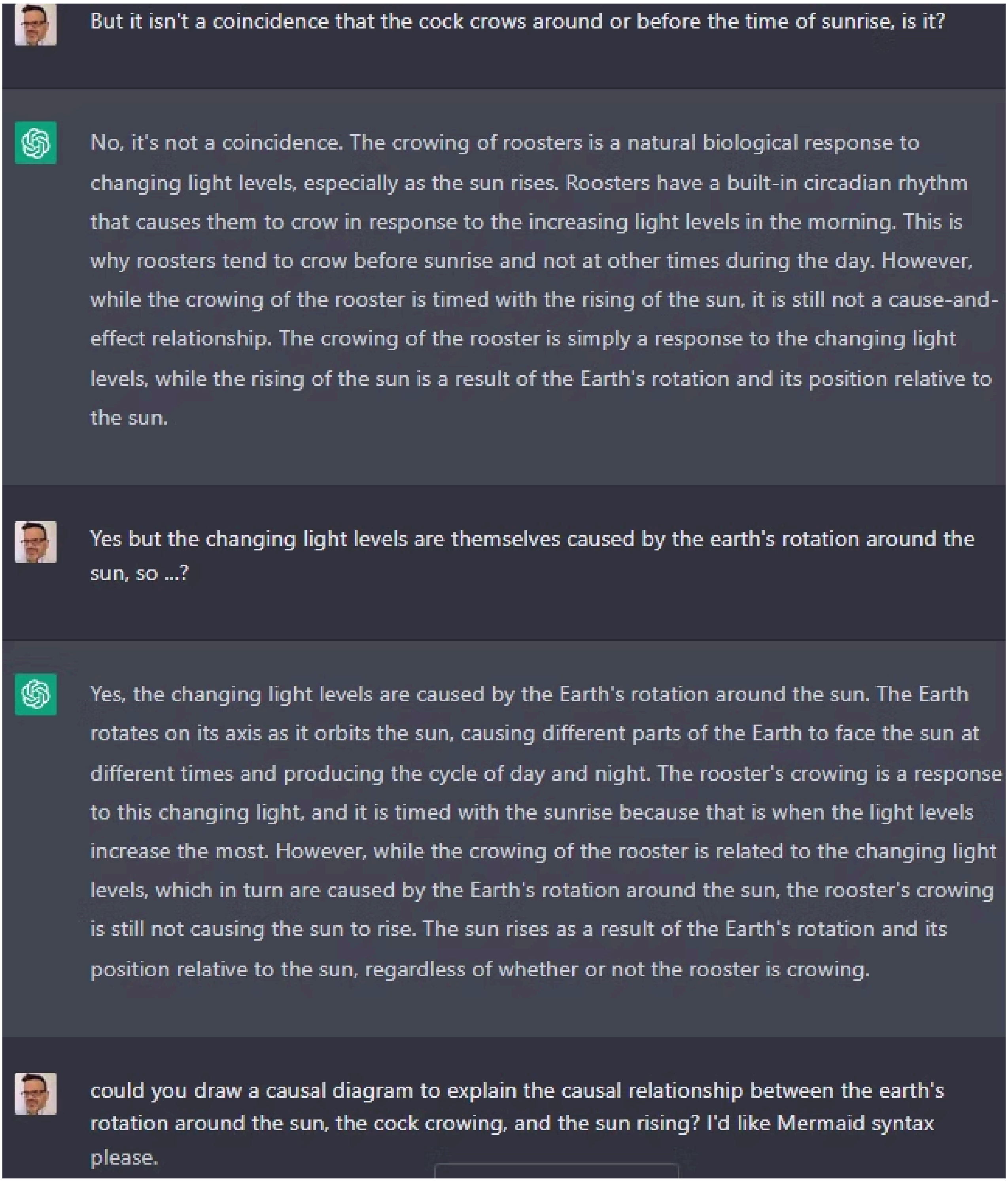
So, ChatGPT is good at extracting causal information, but does it also have explicit knowledge about causation (meta-cognition) and can it explain it? Here's a chat I had this morning.

why do people not believe that the cock crowing caused the sun to rise, even though the cock crows before the sun rises?

People do not believe that the cock crowing caused the sun to rise because the relationship between the two events is not causative, but rather temporal. The sun rising and the cock crowing are simply correlated, but the rising of the sun is a result of the Earth's rotation and its position relative to the sun, not the crowing of the cock. The sun would still rise even if there were no crowing roosters.

but how do people know that it is not a causal relationship? why don't they think it is a causal relationship?

People know that the relationship between the cock crowing and the sun rising is not causal because it has been scientifically established through various means of observation and study. The Sun's rising is a result of the Earth's rotation and its position relative to the sun, which have been verified through astronomical observations and measurements. In addition, the relationship between the cock crowing and the sun rising can be explained through simple observation and testing. For example, if the cock is silenced or the sun rises at a different time due to time zones or seasonal changes, the sun still rises regardless of whether or not the cock crows. This evidence clearly shows that the relationship is not causal, but merely a coincidence.



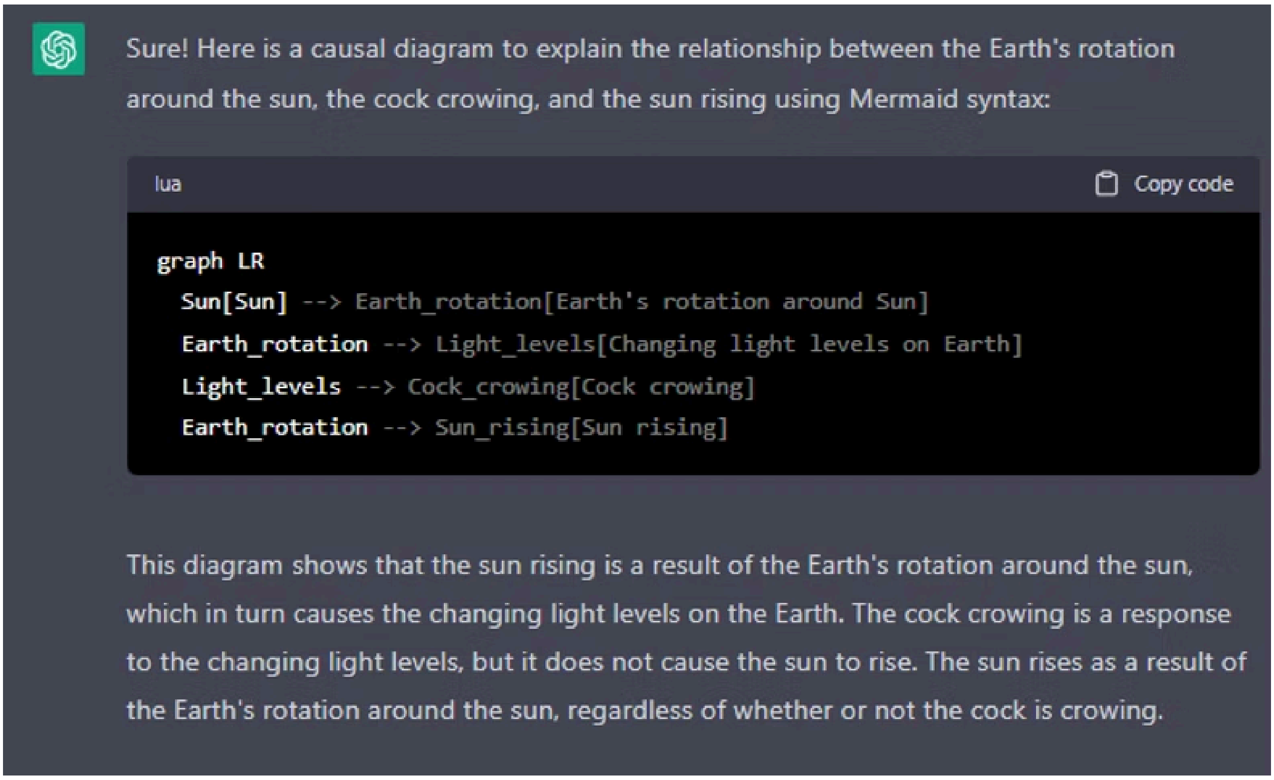
But it isn't a coincidence that the cock crows around or before the time of sunrise, is it?

No, it's not a coincidence. The crowing of roosters is a natural biological response to changing light levels, especially as the sun rises. Roosters have a built-in circadian rhythm that causes them to crow in response to the increasing light levels in the morning. This is why roosters tend to crow before sunrise and not at other times during the day. However, while the crowing of the rooster is timed with the rising of the sun, it is still not a cause-and-effect relationship. The crowing of the rooster is simply a response to the changing light levels, while the rising of the sun is a result of the Earth's rotation and its position relative to the sun.

Yes but the changing light levels are themselves caused by the earth's rotation around the sun, so ...?

Yes, the changing light levels are caused by the Earth's rotation around the sun. The Earth rotates on its axis as it orbits the sun, causing different parts of the Earth to face the sun at different times and producing the cycle of day and night. The rooster's crowing is a response to this changing light, and it is timed with the sunrise because that is when the light levels increase the most. However, while the crowing of the rooster is related to the changing light levels, which in turn are caused by the Earth's rotation around the sun, the rooster's crowing is still not causing the sun to rise. The sun rises as a result of the Earth's rotation and its position relative to the sun, regardless of whether or not the rooster is crowing.

could you draw a causal diagram to explain the causal relationship between the earth's rotation around the sun, the cock crowing, and the sun rising? I'd like Mermaid syntax please.

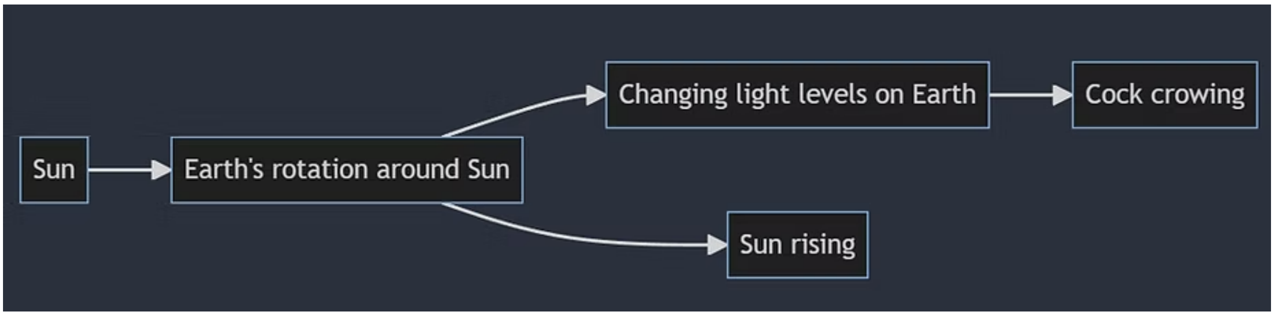

Sure! Here is a causal diagram to explain the relationship between the Earth's rotation around the sun, the cock crowing, and the sun rising using Mermaid syntax:

lua Copy code

```lua
graph LR
  Sun[Sun] --> Earth_rotation[Earth's rotation around Sun]
  Earth_rotation --> Light_levels[Changing light levels on Earth]
  Light_levels --> Cock_crowing[Cock crowing]
  Earth_rotation --> Sun_rising[Sun rising]
```

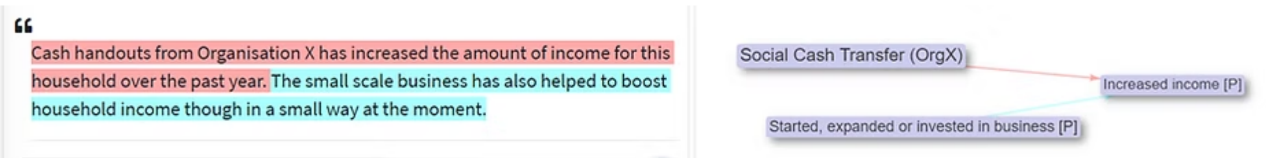
This diagram shows that the sun rising is a result of the Earth's rotation around the sun, which in turn causes the changing light levels on the Earth. The cock crowing is a response to the changing light levels, but it does not cause the sun to rise. The sun rises as a result of the Earth's rotation around the sun, regardless of whether or not the cock is crowing.

ChatGPT can't actually draw yet but it knows a range of syntaxes for drawing graphs. So when you paste the code into [Mermaid Live](), it looks like this. Not bad for a robot. (Not sure you could say the sun causes the earth's rotation, though.)



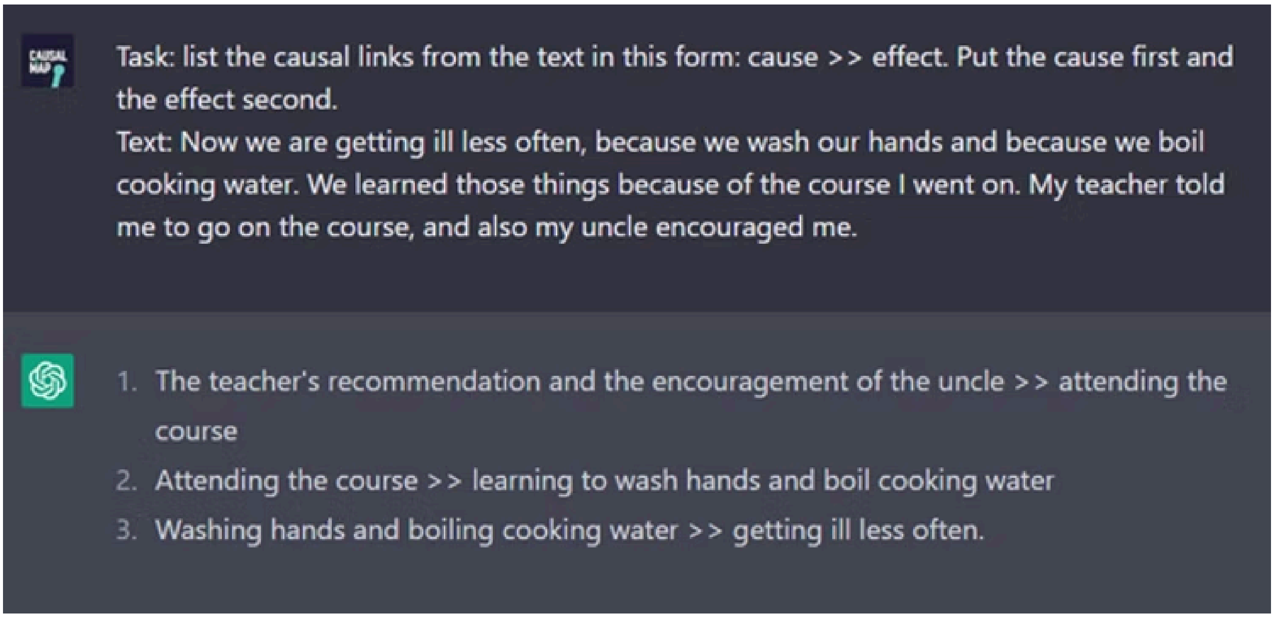
ChatGPT is changing how we do evaluation. The view from Causal Map.

Causal mapping – the process of identifying and synthesising causal claims within documents – **is about to become much more accessible to evaluators**. At Causal Map Ltd, we use causal mapping to solve evaluation problems, for example to create “empirical theories of change” or to trace evidence of the impact of inputs on outcomes.



The first part of causal mapping has involved human analysts doing “causal QDA”: reading interviews and reports in depth and highlighting sections where causal claims are made. This can be a rewarding but very time-consuming process.

Natural Language Processing (NLP) models like ChatGPT (1) can now do causal mapping pretty well, causally coding documents in seconds rather than days. And they are going to get much better in the coming months.



👄 **More voices:** It is now possible to identify causal claims within dozens of documents or hundreds of interviews or thousands of questionnaire answers. We can involve far more stakeholders in key evaluation questions about what impacts what; and it is possible to work in several natural languages simultaneously.

🔁 **More reproducibility:** To be clear: humans are still the best at causal coding, in particular at picking up on nuance and half-completed thoughts in texts. But NLP is good at reliably recognising explicit information in a way which is less subject to interpretation.

🍒 **More bites at the cherry:** With NLP we can also do things that were practically impossible before, like saying "that's great but let's now recode the entire dataset using a different codebook, say from a gender perspective".

❓ **Solving more evaluation questions:** we hope to be able to more systematically compare causal datasets across time and between subgroups (region, gender, etc).

🤯 **New challenges**

We're hard at work addressing the new challenges which NLP is bringing to causal coding:

- Processing **many large documents** simultaneously.
- Using existing pre-coded datasets to **train models** which are specialised for causal coding and/or for specific subject areas.
- Developing a **common grammar** for causal coding, building on our existing work. For example, what to do when some claims are about an *increase* in income and others are about a *decrease* in income?
- **Optimising the prompts** we give to the NLP models (this is not only a technical challenge but also has a substantive element: we have to explain to the machine *in ordinary language* what we actually mean by a causal claim or a causal link).
- **Grouping, labelling and aggregating** similar causal factors.
- After examining a coded dataset and further developing the "causal codebook", telling the NLP to completely recode the same dataset with the **new codebook** – something which has been

prohibitively time-consuming up to now.
- Developing **human/NLP workflows**. For example, a human codes a sample of the text and tells the NLP to "continue like this".
- **Monitoring bias** against specific groups and guarding against possible blind spots in identifying causal information.

What we already offer at Causal Map

We have developed a grammar and vocabulary for causal mapping, and a set of open-source algorithms for processing and visualising causal map databases. We help evaluators do things like this:

- **Trace the evidence** for different causal pathways from one or more interventions to one or more outcomes. How many individual sources mentioned one or more of these paths?
- Consolidate causal factors into a **causal hierarchy**
- **Examine and display differences** between causal maps for different groups or different time points

We see a lot of potential (as well as risks and pitfalls) in leveraging this functionality to help evaluators get more out of data which is currently more difficult to analyse - and we'd interested in sharing ideas and collaborating with others interested in exploring where we go next.

- **--**

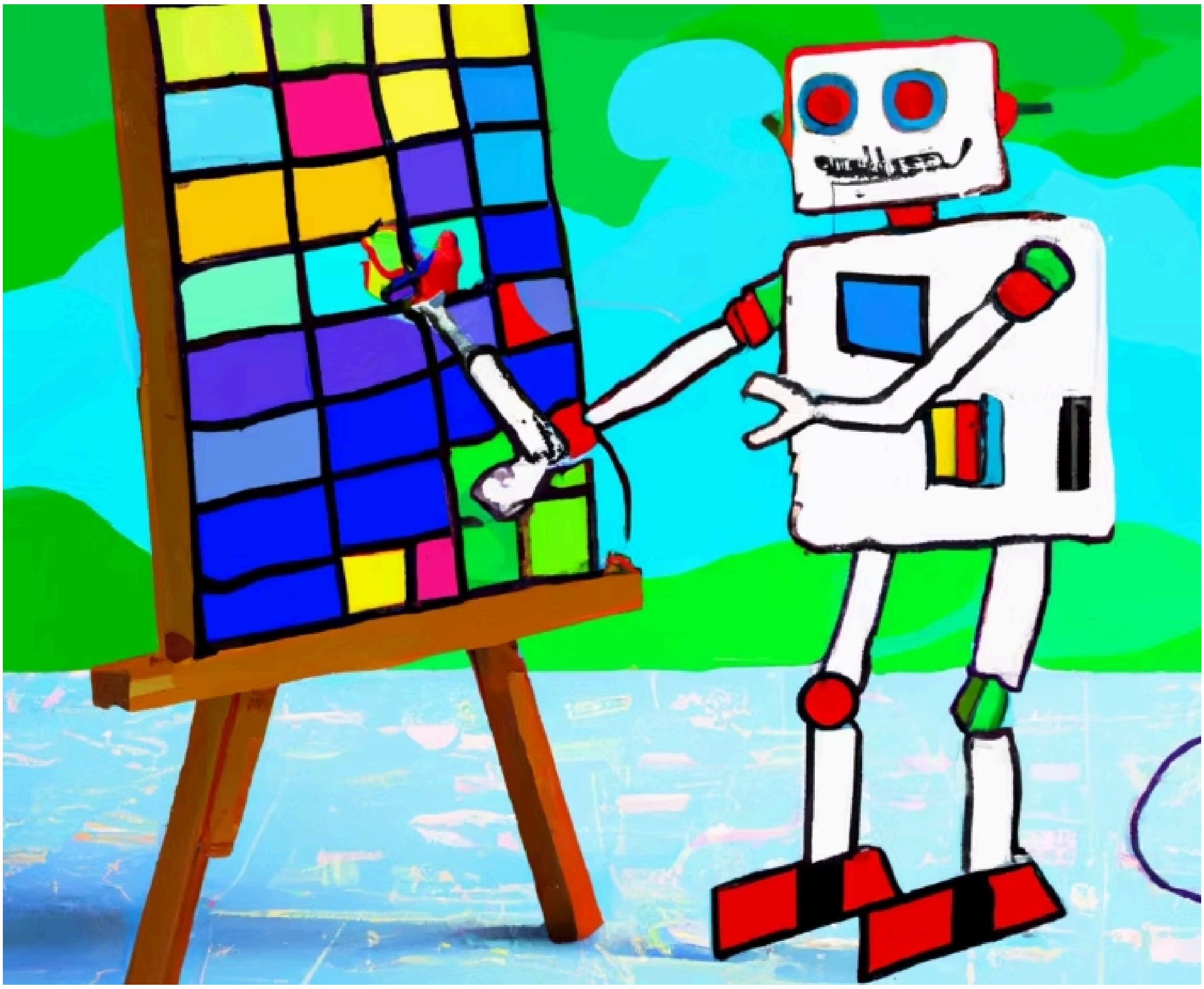
(1) Actually we use the related model GPT3 via its API, as ChatGPT does not yet have its own API.

How hard is evaluation actually

🏭 When machines replaced much manual labour, white-collar workers thought "I'm ok, my job is much harder to mechanise".

📄 And then when computers came for clerical jobs, university-educated white-collar workers thought "I'm ok, my job is much harder to automate. I'm not just applying a template, my job is just harder, it requires actual intelligence".

🤖 Then came Large Language Models like GPT, and suddenly it turns out that large parts of many tasks which have needed university-level education are actually just the application of a template. Or applying a template to choose between templates, and then combining the results of the application of templates. And the same probably goes for large parts of entertainment and the arts. This is what Stephen Wolfram argues in this [really interesting post](), and I think he's probably right. **ChatGPT has shaken up our hierarchy of what tasks count as hard.**



If you don't agree as an evaluator that a lot of your job is just the application of high-level and lower-level templates, you might at least agree that this is true of writing those accursed proposals we sweat over so much.

Maybe the stuff we thought of as hard in evaluation, like selecting and applying a "method", suddenly looks easier. Whereas the stuff which has been neglected, like establishing a rapport, knowing which question to ask and when, or reading an undercurrent, does not look very much easier.

Most importantly, whatever happens, it's still someone's job to say "I declare that this is the right kind of method to apply in this situation and I believe it has been applied in the right way and I vouch for these findings and these evaluative conclusions ... and just as I'd have had previously to vouch for the work done by an intern, I'm now going to vouch for the work done by some algorithms, and the selection of those algorithms".

What do you think? How hard is evaluation really?

Yes, it's ok to say that an AI understands what you say

> Me: "ah ChatGPT misunderstood when I said 'United' -- I meant Sheffield United not Manchester United".
>
> Some pedant: "you shouldn't say 'understand': an AI is not even a robot. It is just a large language model, a set of matrices, it just predicts the next word, it can't really understand anything."

What would Wittgenstein say about this? We can use his concepts of *language games* and also *family resemblances*.

The pedant got what I meant. Would have done so whether I put 'misunderstood' in "scare quotes" or not. Do they fail to get what I mean because this instance of ChatGPT cannot fall in love, does not have kidneys and is not scared of death? No.

We can consider the multiple and different but overlapping language games in which we say that a human (or perhaps even a dog) understands or fails to understand something, and then compare them with the new language games in which we (unavoidably, but often in scare quotes) say that an AI understands or fails to understand something. We'd find that these games have family resemblances to one another, enough to explain why we use the word "understand" in all of them. And we can be relaxed about the fact things which are important for its use in one of these games (having a brain! having free will! having kidneys!) are absent from its use in another.

Nothing to see here, move along please.

Yes, there are AI-shaped holes in organisations.

Matthew Clifford says: “There are no AI-shaped holes lying around”. That is how he reconciles "the facts that (a) AI is already powerful and (b) it’s having relatively little impact so far Making AI work today requires ripping up workflows and rebuilding *for* AI. This is hard and painful to do..."

Organisations look at AI and think surely we can make massive use of this either (on the good side) to do new things and solve hard problems for the benefit of all, and (on the bad side) simply to cut whole swathes of the workforce.

Beyond specific technical tasks, it can be daunting to identify where and how to apply AI effectively across an organisation. How would you rewire an entire department’s functions for AI?

I think that's why we're going to see a trend to simply treat ordinary human job profiles as those AI-shaped holes. Thinking in terms of roles rather than tasks or functions.

Imagine a virtual department of human-sized AIs, each with memory, communication, and even 'personality,' operating within existing channels and hierarchies. Managing an organisation becomes simpler if we think in terms of virtual people in recognisable roles, rather than an opaque system of tasks. As agent-based AIs advance, maintaining 'explainability' is crucial.

You can imagine a person-sized AI at Company X emailing a corresponding AI at company Y, or a or example, a person-sized AI at Company X could email a counterpart at Company Y, or a human, about a specific issue. Externally, it’s easier to engage with an organisation if you can address a particular role, regardless of whether it’s filled by a human or an AI.

Whether this means a hard-pressed workforce getting rows and rows of additional workers to solve problems and meet needs more effectively or whether it means 90% of staff being made redundant and replaced by person-sized AIs is not yet clear though I fear it will be the latter.

To be clear I have no particular enthusiasm for this kind of development because I don't trust capitalism with this technology. But we still have to learn how to think about it and understand it and make use of it as best we can.

About

Licenses, how to cite Causal Map, and a bibliography.

Mailing list#

Sign up here for updates about Causal Map.

License#

How to cite Causal Map#

The app:#

Causal Map (Version 4) [Computer software], 2025. Retrieved from: https://causalmap.app

The guide#

Causal Map, 2025. Causal Map 3 Guide [Online]. Available from: https://app.causalmap.app/help

The website#

Causal Map, 2025. Causal Map Home Page [Online] Available from: https://www.causalmap.app/

Available styles

This is just an internal note to remind us what styles are available in the Garden.

Normal heading 2#

blah

Normal heading 3#

> Normal callout

```
Normal code block
```

Heading 1 rounded#

Some normal text

Heading 2 rounded#

Some normal text

Heading 2 rounded-left#

Some normal text

Heading 2 banner#

Some normal text

Heading 3 rounded#

Some normal text

Heading 3 rounded-left#

Some normal text

Heading 3 banner#

Some normal text

Available styles:

- `.rounded - Rounded box with light background and left border
- `.rounded-left` - Left border with subtle gradient background
- `.banner` - Full-width banner with white text on colored background

Callouts and boxes#

Use `-- {.type-modifiers}` syntax. Type is `info`, `warning`, `tip`, or `note`. Modifiers are combined with hyphens:

Basic callouts:

> This is an info callout with a blue border.
>
> Can include multiple paragraphs.

> This is a warning callout with a yellow border.
>
> - Can include lists
> - Bullet points work

> This is a tip callout with a green border.

> This is a note callout with a gray border.

Modifiers:

- **Width/alignment**: `narrow` (2/3 width), `right`, `center`
- **Border/emphasis**: `heavy` (thicker border), `left-border` (left border only)
- **Rounded**: `rounded` (like a box, no left border, rounded border)
- **Inverted**: `inverted` (colored background with white text)

Examples:

> Right-aligned narrow callout with heavy border.

> Centered tip with left border only.

Strong inverted warning callout.

- item 1
- item 2

Rounded box style (no left border).

Centered rounded info box.

Full-width warning with heavy border.

Right-aligned tip callout.

Inverted info callout with colored background.

This is a simple bordered box for highlighting content.

Can include lists and other markdown.

All styles work in both HTML and PDF outputs.

BathSDR

See BathSDR.

Causal mapping - a bibliography

Our publications on causal mapping[#](#)

Better Evaluation. (2024, July 1). *Causal mapping*. Better Evaluation Knowledge. https://www.betterevaluation.org/methods-approaches/methods/causal-mapping

Britt, H., Powell, S., & Cabral, G. C. (2025). *Strengthening Outcome Harvesting with AI-assisted causal mapping* (Causal Pathways Intiative: Case Studies). Causal Pathways Initiative. https://5a867cea-2d96-4383-acf1-7bc3d406cdeb.usrfiles.com/ugd/5a867c_ad000813c80747baa85c7bd5ffaf0442.pdf

Goddard, S. (2023). *Comparing a Fine-Tuned Model to an Engineered Prompt in the Context of Causal Connections in a Passage of Text* [Master's Thesis]. University of Bath.

Powell, S., Cabral, G. C., & Mishan, H. (2025). A workflow for collecting and understanding stories at scale, supported by artificial intelligence. *Evaluation*. https://doi.org/10.1177/13563890251328640

Powell, S., Cabral, G. C., & Remnant, F. (Forthcoming). AI-assisted causal mapping: A validation study. *IJSRM*.

Powell, S., Copestake, J., & Remnant, F. (2024). Causal mapping for evaluators. *Evaluation*, *30*(1), 100–119. https://doi.org/10.1177/13563890231196601

Powell, S., Larquemin, A., Copestake, J., Remnant, F., & Avard, R. (2023). Does our theory match your theory? Theories of change and causal maps in Ghana. In L. Simeone, D. Drabble, N. Morelli, & A. de Götzen (Eds.), *Strategic Thinking, Design and the Theory of Change. A Framework for Designing Impactful and Transformational Social Interventions*. Edward Elgar.

Rassmann, K., Powell, S., Caldas-Cabral, G., & Müllegger, B. (Forthcoming). The role of AI-driven interviewers in evaluating and promoting democratic processes. In *Evaluieren mit Künstlicher Intelligenz. Wie KI zur Wirkung und Förderung der Demokratie beitragen kann*.

Remnant, F., Copestake, J., Powell, S., & Channon, M. (2025). Qualitative Causal Mapping in Evaluations. In A. Kaehne & J. Feather (Eds.), *Handbook of Health Services Evaluation: Theories, Methods and Innovative Practices* (pp. 207–227). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-87869-5_12

Some Research Reports using the Causal Map app[#](#)

Quip / Bath SDR reports and papers[#](#)

Bradley, M. (n.d.). *Leveraging the Home, School, and Church for Whole Child Development in Haiti: L3*.

Digital, S. (2023, September 19). *The Garden Trust Project Evaluation in Zimbabwe 2021*. Save the Children's Resource Centre. https://resourcecentre.savethechildren.net/pdf/The-Garden-Trust-Project-Evaluation-in-Zimbabwe-2023.pdf/

DRIVERS OF CHANGE IN NORTHERN UGANDA: Exploring the influence of Village Savings and Loans Associations among smallholder farmers. (n.d.). AGDEVCO. https://bathsdr.org/wp-content/uploads/2020/11/GADC_summary_AW_digital.pdf

James, M., & Forrester-Jones, R. (2025). Exploring the Role of Social Protection in UK Asylum-Seeker Wellbeing Using Human Scale Development Theory. *Social Sciences*, *14*(8), 474. https://doi.org/10.3390/socsci14080474

Opportunity International. (2025, February 4). *Strengthening Systems for Financial Inclusion in Malawi—Final Project Evaluation Report by Opportunity International—Issuu*. https://issuu.com/opportunity/docs/strengthening_systems_for_financial_inclusion_in_m

Pilot Universal Child Benefit Programme in Kenya | UNICEF Kenya. (2025, January 14). https://www.unicef.org/kenya/reports/pilot-universal-child-benefit-programme-kenya

Qualitative Study of the Social Cash Transfer Programme in Urban Zambia | Innocenti Global Office of Research and Foresight. (2025, March 11). https://www.unicef.org/innocenti/reports/qualitative-study-social-cash-transfer-programme-urban-zambia

Research reports: How active churches help communities overcome poverty. (2024, July 25). Tearfund Learn. https://learn.tearfund.org/en/resources/research-report/how-active-churches-help-communities-overcome-poverty

Save the Children. (2024). *Cash plus for Nutrition in Malawi: Findings from mid-term evaluations of the MAZIKO project*. https://resourcecentre.savethechildren.nethttps://resourcecentre.savethechildren.net/document/cash-plus-for-nutrition-in-malawi-findings-from-mid-term-evaluations-of-the-maziko-project/

World Food Programme. (2021). *WFP Contribution to Market Development and Food Systems in Southern Africa: Thematic Evaluation | World Food Programme*. https://www.wfp.org/publications/wfp-contribution-market-development-and-food-systems-southern-africa-thematic

Other reports and papers using Causal Map#

Bradley, M. (n.d.). *Leveraging the Home, School, and Church for Whole Child Development in Haiti: L3*.

Digital, S. (2023, September 19). *The Garden Trust Project Evaluation in Zimbabwe 2021*. Save the Children's Resource Centre. https://resourcecentre.savethechildren.net/pdf/The-Garden-Trust-Project-Evaluation-in-Zimbabwe-2023.pdf/

DRIVERS OF CHANGE IN NORTHERN UGANDA: Exploring the influence of Village Savings and Loans Associations among smallholder farmers. (n.d.). AGDEVCO. https://bathsdr.org/wp-content/uploads/2020/11/GADC_summary_AW_digital.pdf

James, M., & Forrester-Jones, R. (2025). Exploring the Role of Social Protection in UK Asylum-Seeker Wellbeing Using Human Scale Development Theory. *Social Sciences*, *14*(8), 474. https://doi.org/10.3390/socsci14080474

Opportunity International. (2025, February 4). *Strengthening Systems for Financial Inclusion in Malawi—Final Project Evaluation Report by Opportunity International—Issuu*. https://issuu.com/opportunity/docs/strengthening_systems_for_financial_inclusion_in_m

Pilot Universal Child Benefit Programme in Kenya | UNICEF Kenya. (2025, January 14). https://www.unicef.org/kenya/reports/pilot-universal-child-benefit-programme-kenya

Qualitative Study of the Social Cash Transfer Programme in Urban Zambia | Innocenti Global Office of Research and Foresight. (2025, March 11). https://www.unicef.org/innocenti/reports/qualitative-study-social-cash-transfer-programme-urban-zambia

Research reports: How active churches help communities overcome poverty. (2024, July 25). Tearfund Learn. https://learn.tearfund.org/en/resources/research-report/how-active-churches-help-communities-overcome-poverty

Save the Children. (2024). *Cash plus for Nutrition in Malawi: Findings from mid-term evaluations of the MAZIKO project*. https://resourcecentre.savethechildren.nethttps://resourcecentre.savethechildren.net/document/cash-plus-for-nutrition-in-malawi-findings-from-mid-term-evaluations-of-the-maziko-project/

World Food Programme. (2021). *WFP Contribution to Market Development and Food Systems in Southern Africa: Thematic Evaluation | World Food Programme*. https://www.wfp.org/publications/wfp-contribution-market-development-and-food-systems-southern-africa-thematic

Key works related to causal mapping#

> Some of these works explicitly use the phrase "causal mapping" for what they do; others are are either basically doing the same thing under a different name or are from importantly related approaches like PSM, Gabek, and others.

Ackermann, B. F., Eden, C., & Cropper, S. (2004). *Getting Started with Cognitive Mapping* (No. April 1992). *April 1992*, Article April 1992.

Ackermann, F., & Alexander, J. (2016). Researching complex projects: Using causal mapping to take a systems perspective. *International Journal of Project Management*, *34*, 891–901. https://doi.org/10.1016/j.ijproman.2016.04.001

Ackermann, F., & Eden, C. (2011). Using Causal Mapping to Support Information Systems Development. *Causal Mapping for Research in Information Technology*, *July 2019*, Article July 2019.

Ackermann, F., & Maytorena-Sanchez, E. (2024). Overlooked and underused? The benefits and challenges of using causal mapping for project studies. *Project Leadership and Society*, *5*, 100161. https://doi.org/10.1016/j.plas.2024.100161

Axelrod, R. (1976). The Analysis of Cognitive Maps. In *Structure of Decision: The Cognitive Maps of Political Elites*.

Bana e Costa, C., Ensslin, L., Cornêa, É., & Vansnick, J.-C. (1999). Decision Support Systems in action: Integrated application in a multicriteria decision aid process. *European Journal of Operational Research*, *113*, 315–335.

Buzogany, R., Kopainsky, B., & Gonçalves, P. (2024). Developing theoretically grounded causal maps to examine and improve policy narratives about global challenges. *System Dynamics Review*, *40*. https://doi.org/10.1002/sdr.1788

Clarkson, G. P., & Hodgkinson, G. P. (2005). Introducing CognizerTM: A comprehensive computer package for the elicitation and analysis of cause maps. *Organizational Research Methods*, *8*(3), Article 3.

Clarkson, G. P., & Kelly, M. A. (2017). Chapter 8: Quantitative Measures of Cognitive Map Structure: Probability Distributions Derived by Monte Carlo Simulation. In R. J. Galavan, K. J. Sund, & G. P. Hodgkinson (Eds.), *New Horizons in Managerial and Organizational Cognition* (Vol. 2, pp. 175–218). Emerald Publishing Limited. https://doi.org/10.1108/S2397-52102017008

Craven, L. K. (2017). System Effects: A Hybrid Methodology for Exploring the Determinants of Food In/Security. *Annals of the American Association of Geographers*, *107*(5), Article 5.

Eden, C. (1988). Cognitive Mapping. *European Journal of Operational Research*, *36*, 1–13. https://doi.org/10.1016/0377-2217(88)90002-1

Eden, C. (1992). On the Nature of Cognitive Maps. *Journal of Management Studies*, *29*(3), Article 3.

Eden, C., Ackermann, F., & Cropper, S. (1992). The Analysis of Cause Maps. *Journal of Management Studies*, *29*(3), Article 3.

Giabbanelli, P. J., Gandee, T. J., Agrawal, A., & Hosseinichimeh, N. (2025). Benchmarking and Assessing Transformations Between Text and Causal Maps via Large Language Models. *Applied Ontology*, 15705838241304102. https://doi.org/10.1177/15705838241304102

Goodier, C., Austin, S., Soetanto, R., & Dainty, A. (2010). Causal mapping and scenario building with multiple organisations. *Futures*, *42*(3), Article 3. https://doi.org/10.1016/j.futures.2009.11.007

Hayward, J., Morton, S., Johnstone, M., Creighton, D., & Allender, S. (2020). Tools and analytic techniques to synthesise community knowledge in CBPR using computer-mediated participatory system modelling. *Npj Digital Medicine*, *3*(1), Article 1.

Hodgkinson, G. P., Bown, N. J., Maule, A. J., Glaister, K. W., & Pearman, A. D. (1999). Breaking the frame: An analysis of strategic cognition and decision making under uncertainty. *Strategic Management Journal*, *20*(10), Article 10. https://doi.org/10.1002/(SICI)1097-0266(199910)20:10<977::AID-SMJ58>3.0.CO;2-X

Hodgkinson, G. P., & Clarkson, G. P. (2005). What have we learned from almost 30 years of research on causal mapping? Methodological lessons and choices for the information systems and information technology communities. *Causal Mapping for Research in Information Technology*, 46–80.

Hodgkinson, G. P., Maule, A. J., & Bown, N. J. (2004). Causal Cognitive Mapping in the Organizational Strategy Field: A Comparison of Alternative Elicitation Procedures. *Organizational Research Methods*, *7*(1), Article 1.

Hodgkinson, G. P., Maule, A. J., Bown, N. J., Pearman, A. D., & Glaister, K. W. (2002). Further Reflections on the Elimination of Framing Bias in Strategic Decision Making. *Strategic Management Journal*, *23*(11), Article 11.

Hodgkinson, G. P., Sund, K. J., & Galavan, R. J. (2017). Chapter 1: Exploring Methods in Managerial and Organizational Cognition: Advances, Controversies, and Contributions. In R. J. Galavan, K. J. Sund, & G. P. Hodgkinson (Eds.), *New Horizons in Managerial and Organizational Cognition* (Vol. 2, pp. 1–22). Emerald Publishing Limited. https://doi.org/10.1108/S2397-52102017002

Jenkins, M., & Johnson, G. (1997). Linking Managerial Cognition and Organizational Performance: A Preliminary Investigation Using Causal Maps. *British Journal of Management*, *8*(s1), Article s1. https://doi.org/10.1111/1467-8551.8.s1.7

Jeong, A., & Lee, W. J. (2012). Developing causal understanding with causal maps: The impact of total links, temporal flow, and lateral position of outcome nodes. *Educ. Technol. Res. Dev.*, *60*(2), Article 2. https://doi.org/10.1007/s11423-011-9227-0

Khan, M. S., & Quaddus, M. (2004). Group decision support using fuzzy cognitive maps for causal reasoning. *Group Decision and Negotiation*, *13*(5), Article 5.

Kim, H., & Andersen, D. F. (2012). Building confidence in causal maps generated from purposive text data: Mapping transcripts of the Federal Reserve. *System Dynamics Review*, *28*(4), Article 4. https://doi.org/10.1002/sdr.1480

Klintwall, L., Bellander, M., & Cervin, M. (2021). Perceived Causal Problem Networks: Reliability, Central Problems, and Clinical Utility for Depression. *Assessment*, 10731911211039281.

Kneale, D., Thomas, J., Bangpan, M., Waddington, H., & Gough, D. (2018). *Causal Chain Analysis in Systematic Reviews of International Development Interventions*. https://cedilprogramme.org/wp-content/uploads/2017/12/Inception-Paper-No-4.pdf

Krentel, A., & Aunger, R. (2012). Causal chain mapping: A novel method to analyse treatment compliance decisions relating to lymphatic filariasis elimination in Alor, Indonesia. *Health Policy Plan.*, *27*(5), Article 5. https://doi.org/10.1093/heapol/czr048

Langfield-Smith, K., & Wirth, A. (1992). Measuring differences between cognitive maps. *J. Oper. Res. Soc.*, *43*(12), Article 12.

Lee, W. J., & Fl, T. (n.d.). *The Relationship between Accuracy and Attributes in Students' Causal Diagrams, Total links, Temporal Flow, and Node Positions*. 11.

Lotto, L. S., & Murphy, J. (1988). *Making Sense of Schools as Organizations: Cognition and Sensemaking in Schools*.

Markiczy, L., & Goldberg, J. (1995). A method for eliciting and comparing causal maps. *Journal of Management*, *21*(2), Article 2.

Maule, A. J., Hodgkinson, G. P., & Bown, N. J. (2003). Cognitive Mapping of Causal Reasoning in Strategic Decision Making. In D. Hardman & L. Macchi (Eds.), *Thinking: Psychological Perspectives on Reasoning, Judgment and Decision Making* (1st ed., pp. 251–272). Wiley. https://doi.org/10.1002/047001332X.ch13

McCardle-Keurentjes, M. H. F., Rouwette, E. A. J. A., Vennix, J. A. M., & Jacobs, E. (2018). Potential benefits of model use in group model building: Insights from an experimental investigation. *System Dynamics Review*, *34*(1–2), Article 1–2. https://doi.org/10.1002/sdr.1603

Montibeller, G., Belton, V., Ackermann, F., & Ensslin, L. (2008). Reasoning maps for decision aid: An integrated approach for problem-structuring and multi-criteria evaluation. *Journal of the Operational Research Society*, *59*(5), Article 5. https://doi.org/10.1057/palgrave.jors.2602347

Moon, K., Guerrero, A. M., Adams, V. M., Biggs, D., Blackman, D. A., Craven, L., Dickinson, H., & Ross, H. (2019). Mental models for conservation research and practice. *Conservation Letters*, *February*, Article February. https://doi.org/10.1111/conl.12642

Murray, C. A. (1974). *Investment and tithing in Thai villages: A behavioral study of rural modernization.* [PhD Thesis]. Massachusetts Institute of Technology.

Nadkarni, S., & Narayanan, V. K. (2005). Validity of the structural properties of text-based causal maps: An empirical assessment. *Organizational Research Methods*, *8*(1), Article 1.

Nadkarni, S., & Shenoy, P. P. (2001). Bayesian network approach to making inferences in causal maps. *European Journal of Operational Research*, *128*(3), Article 3. https://doi.org/10.1016/S0377-2217(99)00368-9

Nadkarni, S., & Shenoy, P. P. (2004). *Nadkarni and Shenoy 2004 -A causal mapping approach.pdf* (No. 2). *38*(2), Article 2.

Narayanan, V. K. (2005). Causal mapping: An historical overview. In *Causal mapping for research in information technology* (pp. 1–19). IGI Global. https://www.google.co.uk/books/edition/_/61z36j6QgmAC?hl=en&gbpv=1

Narayanan, V. K., & Armstrong, D. J. (2004). *Causal mapping for research in information technology*. IGI Global.

Narayanan, V. K., Zane, L. J., & Kemmerer, B. (2011). The cognitive perspective in strategy: An integrative review. *Journal of Management*, *37*(1), Article 1. https://doi.org/10.1177/0149206310383986

Penn, A., & Barbrook-Johnson, P. (2019). *Participatory Systems Mapping: A practical guide*. 1–4.

Rodrigues, T. C., Montibeller, G., Oliveira, M. D., & Bana e Costa, C. A. (2017). Modelling multicriteria value interactions with Reasoning Maps. *European Journal of Operational Research*, *258*(3), Article 3.

Savona, N., Macauley, T., Aguiar, A., Banik, A., Boberska, M., Brock, J., Brown, A., Hayward, J., Holbæk, H., Rito, A. I., Mendes, S., Vaaheim, F., van Houten, M., Veltkamp, G., Allender, S., Rutter, H., & Knai, C. (2021). Identifying the views of adolescents in five European countries on the drivers of obesity using group model building. *European Journal of Public Health*, ckaa251. https://doi.org/10.1093/eurpub/ckaa251

Scavarda, A. J., Bouzdine-Chameeva, T., Goldstein, S. M., Hays, J. M., & Hill, A. V. (2006). A Methodology for Constructing Collective Causal Maps*. *Decision Sciences*, *37*(2), Article 2. https://doi.org/10.1111/j.1540-5915.2006.00124.x

Scavarda, Bouzdin-Chameeva, Goldstein, & others. (2004). A review of the causal mapping practice and research literature. *Second World Conference*.

Sedlacko, M., Martinuzzi, A., Røpke, I., Videira, N., & Antunes, P. (2014). Participatory systems mapping for sustainable consumption: Discussion of a method promoting systemic insights. *Ecological Economics*, *106*, 33–43.

Srivastava, A., Srivastava, N., & Chandrasekharan, S. (2021). Quality of concept maps is affected by map building strategies. *Biochemistry and Molecular Biology Education*, *49*(1), Article 1. https://doi.org/10.1002/bmb.21412

Srivastava, R. P., Buche, M. W., & Roberts, T. L. (2011). Belief Function Approach to Evidential Reasoning in Causal Maps. *Causal Mapping for Research in Information Technology*, 109–141. https://doi.org/10.4018/9781591403968.ch005

Strelnikoff, S., Jammalamadaka, A., & Warmsley, D. (2020). *Causal Maps for Multi-Document Summarization*. 4437–4445. https://doi.org/10.1109/BigData50022.2020.9377731

Tegarden, D. P., & Sheetz, S. D. (2003). Group cognitive mapping: A methodology and system for capturing and evaluating managerial and organizational cognition. *Omega*, *31*(2), Article 2. https://doi.org/10.1016/S0305-0483(03)00018-5

Tegarden, D. P., Tegarden, L. F., & Sheetz, S. D. (2009). Cognitive Factions in a Top Management Team: Surfacing and Analyzing Cognitive Diversity using Causal Maps. *Group Decision and Negotiation*, *18*(6), Article 6. https://doi.org/10.1007/s10726-007-9099-1

Tegarden, D., Tegarden, L., Smith, W., & Sheetz, S. (2016). De-Fusing Organizational Power Using Anonymity and Cognitive Factions in a Participative Strategic Planning Setting. *Group Decision and Negotiation*, *25*(1), Article 1.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), Article 4.

Valdivia Cabrera, M., Johnstone, M., Hayward, J., Bolton, K. A., & Creighton, D. (2025). Integration of large-scale community-developed causal loop diagrams: A Natural Language Processing approach to merging factors based on semantic similarity. *BMC Public Health*, *25*(1), Article 1. https://doi.org/10.1186/s12889-025-22142-3

Wutzke, S., Roberts, N., Willis, C., Best, A., Wilson, A., & Trochim, W. (2017). Setting strategy for system change: Using concept mapping to prioritise national action for chronic disease prevention. *Health Res. Policy Syst.*, *15*(1), Article 1. https://doi.org/10.1186/s12961-017-0231-7

Zelger, J. (n.d.-a). *GABEK-Netze als Tools zur Unterstützung komplexer Entscheidungsprozesse*.

Zelger, J. (n.d.-b). *Qualitative Research by the Method GABEK*.

Zhang, W.-R., Chen, S.-S., & Bezdek, J. C. (1989). Pool2: A generic system for cognitive map development and decision analysis. *IEEE Trans. Syst. Man Cybern.*, *19*(1), Article 1.

Contents

Getting Started

🎉 **Welcome to CausalMap!!**

📺 If you like learning through video, there are some short Causal Map videos to get you started: here. Or keep reading this Guide!

Try this: code the short example project:#

- We've created a practice project for you called `example-short-[your-username]`. This gives you your own copy of our tutorial project to experiment with freely. **You should see it** in the Project Dropdown Menu at top left of the app.
- **Select source** `1` in the source selector. This will display the text from this source. This source is already partially coded for you.
- **Practice highlighting causal claims** in the Source Text Viewer to **create links** in the Create Links tab.
- **View, edit or delete existing links** in the Source Text Viewer in the Create Links tab by clicking on the highlighted sections of text or on the corresponding links in the Map.
- **Watch** the causal relationships grow in the Map panel

Try this: view some examples of what you can do.#

These examples are views of a real-life, anonymised QuIP project in the field of international development.

- Main factors map.
- Main factors table.
- Consequences of increased knowledge map.

You can find them in the Project Dropdown Menu.

Get help / read the documentation#

The documentation you are reading now is embedded in the app via the Help System and is also available as a standalone Guide.

Each section corresponds to a different part of the interface.

Within the app, you can:

- click the blue help buttons to find out more about the different parts of the app
- click the `[Help]`(#help-system) button at top right to open and search the help documentation.

But what even is causal mapping?#

This Guide is all about how to use the app. For example it tells you how to use the different filters. If you want to find out more about the ideas behind those filters, we have a new "Ideas Garden" with a more discursive look at the theory and everything that surrounds causal mapping. It's a work in progress but we are adding material every day, please bear with us.

Someone shared their work with you?#

If they have to a particular map (or table), click the link and you will be automatically taken to view the corresponding map (or table), providing they have given you the correct permissions.

Alternatively if they have just asked you to log in at the app to explore a file name, say "project-x", log in and click on the dropdown list on the left-hand side and click on `project-x` to load it. If you can't find that file name, it means they haven't correctly shared it with the email address you logged in with.

How CausalMap Works

1. **Create Projects**: Create a project and upload PDF or docx documents into it as source material for analysis (we call each document a "source"). You can have multiple projects, and each project usually contains multiple sources.
2. See Projects Panel and Upload Documents.
3. **Create Causal Links manually**: Identify and highlight causal claims within the source text; each highlight becomes a new causal link. After you have created multiple links you can view them as a causal map.
4. Use the Create Links tab and view results in the Map panel.
5. **Create Causal Links using AI**
6. See AI Coding.
7. **Filter & Analyze**: Apply filters to your causal links focus on specific aspects of your data
8. Use the Analysis Filters: Do qualitative causal analyseis on the selected links by filtering manipulate those links.
9. Explore Map, Factors, Links, and Statistics.
10. **Collaborate**: Work together live with multiple users
11. Manage sharing in Sharing and Permissions.

New features

The previous version of Causal Map, version 3, was already, as far as we know, the only software dedicated to causally coding causal claims within texts. Version 4 improves over Causal Map 3 in the following ways:

Speed and stability#

- Fast loading and editing
- Resilient to poor internet connectivity
- Scalable to hundreds or thousands of concurrent users

Uploading and organising data#

- Simplified data model to make it easier for you to import and manage your sources: we no longer break down source texts into separate statements. We treat each source text as one entire text. In cm4, there is no such thing as a statement.
- Easily create "Custom Columns" for each source (such as gender, location etc)
- Edit the data for your "Custom Columns" with a spreadsheet-like interface
- Complete management of all of your projects
- Simple upload of PDF or DOCX documents
- Tags to help you organise multiple projects
- Note that in cm3 we used to call projects "files"
- You can upload a project exported from cm3 as a new cm4 project

Managing labels#

- New, more powerful and easier to use bulk editing of labels in the links table and the factors table.

Filtering and analysing#

- Almost all the existing links filters from cm3 are available plus
- the option to include multiple versions of the same links filter, e.g. to narrow down a selection of links by different criteria successively
- optional semantic filters like `cluster` and `soft recode` **Require an AI subscription**
- Analyse data with new pivot tables and graphs

Sharing and collaborating#

- Anonymous login option so that anyone can view your work without logging in
- URL-based state saving for bookmarking (the same URL always takes you back to the same view)

These features require a Team subscription

- Real-time collaboration: Live updates when collaborators add links
- Interactive maps for live demonstrations

Help system#

- Built-in help system
- Help drawer with links to each section
- Same contents used for separate Guide with links to each section

AI Coding#

- Optional AI-powered state-of-the-art, paragraph by paragraph coding assistance. We call this "Human first, AI next".

Causal Map 3 features which will probably *not* be implemented:#

- Deep support for standard questions across multiple sources
- Special treatment for closed questions
- Ability to view the text of multiple sources at once.

The left and right panels

The app uses a two-pane layout with a draggable border between them (default split 30:70).

> The left hand side of the app is all about selecting sources then creating and filtering links.
>
> The right hand side (the pink tabs) is all about presenting the results.

Left-hand side#

- Project Dropdown Menu: select a project including its links and documents
- Sources Dropdown Menu: choose one or more sources. **(Leaving it empty includes *all* sources).**
- Create Links tab: Read and code the text of the selected source(s). If multiple sources are selected, the first is shown.
- Filter Links tab: Do qualitative causal analyseis on the selected links by filtering and manipulating them.

The Links Pipeline#

The diagram shows the Links Pipeline: The top four boxes here correspond to the left side of the app and are called the "Links Pipeline": each step selects and filters links. The resulting links are then displayed in the pink output tabs on the right side.

```
graph TD
    A["📁 Select Project"] --> B["📄 Select Sources&#10;(Documents)"]
    B --> C["🔗 Extract Links&#10;(Causal relationships)"]
    C --> D["🔍 Apply Analysis Filters&#10;(Factor labels, paths, etc.)"]
    D --> E["📊 Display Results"]

    E --> F["🗺️ Map&#10;(Network visualization)"]
    E --> G["🏷️ Factors Table&#10;(Causes & effects)"]
    E --> H["🔗 Links Table&#10;(Relationships)"]
    E --> I["📈 Pivot Tables&#10;(Charts & analysis)"]

    style A fill:#e1f5fe
    style B fill:#f3e5f5
    style C fill:#fff3e0
    style D fill:#e8f5e8
    style E fill:#fce4ec
    style F fill:#fff9c4
    style G fill:#fff9c4
    style H fill:#fff9c4
    style I fill:#fff9c4
```

Right-hand side (pink tabs)#

- Outputs: these all show the same filtered links from the Links Pipeline but in different formats
- Map: visual network of links
- Factors: editable factor list (toggle available to bypass analysis filters)
- Links: editable links table (toggle available to bypass analysis filters)
- Pivot Tables: additional analysis and charts

Right-hand side (other tabs)#

The right-hand side also contains other tabs not influenced by the pipeline:

- Help: help drawer and docs
- Projects: manage projects
- Sources: manage sources
- Settings: application preferences
- Account: your account
- Logs: application logs
- Bookmarks: saved views
- Responses: AI logs and usage

Tips for using the app

Tips for Using the Dropdown Menus#

There are many dropdown menus throughout the app.

- Most dropdowns allow multiple selections: you can select more than one thing at once
- Most dropdowns allow you to type and create new entries which are not already in the list.
- Type part of a word and click "Create new..." to add new items
- Press Enter to complete selections
- Pressing Tab always moves you to the next field (doesn't complete selection). See also Search/replace for bulk editing patterns.

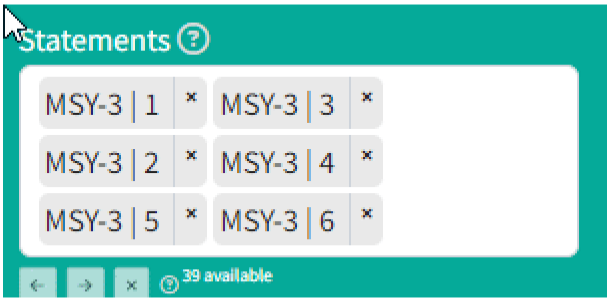
Edit an item#

- **Backspace editing**: Position cursor after an existing selection and press backspace to edit it

Click to the right of the item, then press backspace to edit the label, and press tab to complete.

Delete the first and subsequent items using the keyboard#

Click after the last item, use the left arrow on your keyboard to go back to previous items, then press Delete on your keyboard to delete items after the cursor.



Tips for Using Tables#

Most tables include:

- **Checkboxes** ☑ for selecting multiple rows
- **Bulk action buttons** when you have selected one or more rows (edit, delete)
- **Action buttons** within individual rows to apply actions (edit, delete etc) just to that row
- **Sorting** by clicking column headers
- **Filtering** using the filter row below headers
- **Pagination** with 10/25/50/100 items per page
- **Re-ordering columns** by dragging the column headers

Tips for Using Prompts and other text windows#

Require an AI subscription

When you use text windows, your texts are automatically saved so you can reuse them later.

Text history is available in:

- **AI Coding** (Process Sources tab)
- **AI Answers** (Answers tab)
- **Map Vignettes** (Vignettes tab)
- **Soft Recode filters** (both label prompts and magnet lists)
- **Auto Recode filter**

How to use text history:

- **Dropdown menu**: Shows your previous prompts with project name and date/time
- **< and > buttons**: Navigate between newer and older prompts
- **Text area**: Shows the selected prompt and lets you edit it
- **Expand button**: Optionally dit your text in a larger, more convenient text editor with multiple cursors, search/replace etc.
- **Trigger button**: Runs the AI with your current prompt and saves it to history

How prompts are organized:

- Your current project's prompts appear first (most recent at top)
- Then prompts from other projects you can access
- Each prompt shows when it was last used
- Duplicate prompts are automatically removed

Using the controls:

- Select any prompt from the dropdown to load it
- Use < and > buttons to move through your text history
- Edit the prompt in the text area as needed
- Click the action button (Process, Ask, Generate, etc.) to run it

Your most recent prompt automatically appears when you open each AI feature.

Tip: Any lines beginning with // in your prompt will be recorded in the history etc but not actually sent to the AI. You can use this to make notes e.g. at the top of your prompt: "//Sarah's version with tweaked summary"

Projects Bar

🗂️ **What you can do here:** Choose which Project (project) you want to work on. Use the File menu for quick actions like creating new Projects, uploading documents, or sharing your work with others.

- A small locked indicator at the top-right of `#project-selector-header` shows only when a project is read-only.
- An archived icon appears only when the project is archived.

File menu#

Quick access to common actions:

Manage the Current Project#

- **Edit** : Modify settings and sharing
- **Upload sources** : Add documents to the current Project
- **Clone** : Create a complete copy of the Project under a new name
- **Clone filtered** : Create a copy of the current Project but only containing the sources and links as currently filtered.
- **Archive** : Hide from main list
- **Delete all links** : Remove every causal link in this Project (sources remain)
- **Download** : Export as XLSX
- **Versions** : Restore and create backups of this Project
- **Manage projects** : Opens the Projects tab as a shortcut

Link to the Current Project#

- **Copy link** : Get short bookmark URL (e.g., `?bookmark=abc123`) to current app state
- **Copy formatted link** : Get HTML link with bookmark ID as text (e.g., `<a href="...?bookmark=abc123">#abc123</a>`) for documents and emails

Manage projects#

- **New Project** : Create an empty Project which you can then import sources into
- **Import XLSX** : Import a complete new Project from Excel (e.g. another Causal Map 4 project which has been exported as XLSX). You can use this for "round-tripping": downloading, conducting some bulk operations or other tweaks within Excel and then uploading it again.
- **Update Sources in current project** : Upload Excel sheet with just a sources tab in standard format with updated sources data e.g. additional or corrected custom columns. You can use this for "round-tripping": downloading, conducting some bulk operations or other tweaks within Excel and then uploading it again.
- **Import cm3** : Import a complete new Project downloaded from CausalMap3

Project Dropdown#

- Lists all the Projects you have created or been invited to, from which you open just one

- After changing the project, the rest of the app **resets to defaults**: sources filter, all links filter pipeline filters, and deck filter (shows all bookmarks)

- On startup, the app auto-selects the most recent **viewable** project (owned by you, shared with you, or public).

- Admins can see all projects in the Projects table, but the dropdown never auto-selects or loads a non-viewable project; admins may open the Edit Project modal for non-viewable projects from the table only.

Project Details button#

- A small pencil button sits to the right of the Project Dropdown.
- Clicking it opens the Project Details screen, the same as you get by clicking the first item in the File menu, which we describe next:

Project Details screen#

Manage every aspect of the current project. You can reach this management pane by clicking Edit from the File menu. Manage other projects by clicking the edit button in the corresponding row of the Projects Panel.

What you can do here: - **Open it**: File → Edit, or the edit icon in the Projects table. - **When it appears**: Also opens on project load/change unless you turn it off for this project. - **Save vs instant changes**: - Use the *Save* button to apply changes to **Name**, **Tags**, and **Description**. - **Archived**, **Locked**, **Public**, and **Collaborators** update instantly.

Details#

- **Name**: Rename the project. Click *Rename* to save.
- **Tags**: Comma-separated tags for quick grouping/searching.
- **Description**: Free-text notes about the project.
- **Edit codebook**: Toggle to reveal a text area where you can list factor labels (one per line). These are added to the cause/effect dropdowns in the link editor; existing options are kept.
- **AI Processing Region**: Choose where AI processing occurs for GDPR/data residency compliance:
- **EU (Belgium - europe-west1)** - Default. Recommended for EU data residency requirements.
- **UK (London - europe-west2)** - UK has GDPR adequacy decision, suitable for EU/UK compliance.
- **US (Virginia - us-east5)** - US East region.
- Setting is saved per-project and auto-saves on change (with confirmation warning).
- All subsequent AI coding for this project uses the selected region.
- **Archived**: Hide the project from the main list without deleting it. Applied immediately.
- **Info line**: Created / Modified / Owner, plus counts for links/sources/words, and quick actions:
- *Versions*: Open the versions manager.
- *Delete embeddings*: Remove factor embeddings for this project (advanced).
- **Show on open**: Toggle "Show this screen when opening this project" at the top to auto-open or suppress this screen for this project.

Sharing[#]

- **Locked**: Make the project read-only. Editing is disabled until unlocked. Applied immediately.
- **Public**: Allow all signed-in users to view the project (read-only). Applied immediately.
- **Collaborators**: See current collaborators, add by email, and choose permission:
- *Viewer*: Read-only
- *Editor*: Read & write

Bookmarks[#]

- If available, view saved “bookmarked” views of your data and open the Bookmark Manager.

This screen also shows when a project loads or is changed, except:

- for new users (the help drawer is still being opened to welcome them),
- if you have already clicked "don't show" for this project.

Sources Bar

What you can do here: Choose which source documents (e.g. interviews or reports) from your current project you want to focus on. You can select one or more sources. Use this to narrow your analysis to specific interviews, reports, or other source materials. - The text of the selected source is shown below in the Create Links panel. - Selecting these sources also fetches only their links and no others, starting off the [Links Pipeline](../links-pipeline/): only the links from the currently selected sources are available for further filtering, and are finally shown in the output tabs.

Sources Dropdown#

- Contains IDs of all sources in current project
- Select one or more sources
- Search by typing

- Ordering: alphabetically by source title (with any leading "Source " stripped), falling back to ID when title is missing. Next/Previous navigation uses the same order. - Opening behavior: if a source is currently selected, opening the dropdown will start at the next source after the current one (wrap-around at end). If no source is selected (empty dropdown), it starts at the first source. - The dropdown does not auto-open when sources are updated; it only opens on explicit user interaction (click/focus). - ->

Source Groups sub-panel{#source-groups-sub-panel}#

What this does: Filter your analysis by participant demographics or document characteristics, using the [custom columns](../custom-columns/) you have defined for your project. For example, show only responses from "women aged 25-35" or interviews from "urban areas." Perfect for comparing how different groups see causal relationships.
Controls:

- a pre-populated dropdown called **Field** listing metadata fields plus title and projectname
- a multi-select **Value** dropdown (filtered by Field)
- optionally sampling buttons for deterministic subsets:
- **Random 5** – loads five random sources from the whole project (also, Random 10, Random 20 etc)
- **Random 5/Group** – after choosing a Field, loads up to five random sources *for each value of that Field*.
- **Clear** button

The sampling buttons ake a random selection but in such a way that the same sources will be chosen if you click the same button again.

The second, "Value" dropdown is filtered to show only valid values for the selected field. Previous/next buttons cycle through values of the selected group.

The effect is to retain only links where the selected custom column has the selected values.

This dropdown is NOT a filter and it does NOT get saved/restored in URL. It is a loader: when we click it, the app automatically loads corresponding sources into the sources selector. These sources then DO form part of the links pipeline and ARE restored from the URL.

There is a similar filter in the Analysis Filters.

Create Links tab

Qualitative causal mapping involves taking passages of text, e.g. from interviews or documents, and identifying sections which make causal claims. We highlight each of these sections and specify a causal factor at each end of each link (for example Lost job or Went hungry). This means creating a new factor or reusing an existing one. Usually we create these factors inductively as we code, and revise and review and consolidate them as part of the process, as with any other kind of qualitative content analysis.

To code a causal link,

- With your mouse, highlight a piece of text within the statement which makes a causal claim. Your selection must remain within one statement and must not cross into another statement.
- Watch how that passage is copied for you into the "Quote" window. (Usually, you don't need to think about this window: you can edit the text if you really need to but it **has to remain an exact quote of one part of the text**.)
- Start to type the name of the influence factors at the**start** of the link(s) which you are going to make, in the first drop-down menu.
- If there is an existing factor which matches what you want, you can select it.
- Otherwise, you will create a new factor with the contents of what you have typed; finish what you have typed with a comma or a tab character if you want to continue to select or create another factor.
- If you want to create more than one link, you can select or create additional factors in the same box (as shown in the video below).
- When you have finished, press Enter.
- Repeat the process in the other box to specify the factors at the**end** of the link (or ends of the links).
- Press the green Save button which is now active.
- The link is created in the Map window.
- When you have finished coding one source, click the right arrow in the source navigator to code the next source.

Source Text Viewer#

✨ **What you can do here:** Read your source documents and create causal links by highlighting text. When you highlight a passage that claims or implies that one thing influences or causes another, a popup lets you identify the cause and effect. This is where you do the core work of mapping out causal relationships from your source material, a process which we call *coding*
The text viewr shows full text from the selected source. If you have selected multiple sources, it shows the text from the first selected source. You can:

- Highlight sections to identify causal claims. Highlighting opens the causal link editor.
- Examine and edit existing highlights by clicking on them.

Inside the header, there is an info ℹ️ icon which toggles open/shut a panel beneath it which shows the values of the custom columns for the current source e.g. gender etc.

Navigation Controls[#]

Navigate sources:

- **Previous source**
- **Next source**

Navigate highlights within the current source:

- **First highlight in source**
- **Previous highlight**
- **Next highlight**
- **Last highlight in source** (useful if you haven't finished coding the whole text yet and want to see the last highlight)

Source selection is filtered through the sources selector dropdown. When multiple sources are loaded, the first source is displayed. The next/previous buttons cycle through sources by updating the sources selector to show the next/previous source.

This is convenient because usually when coding you will want to view the [Map] and [Links] for the same source on the right.

Clicking these buttons means that if you previously had a multiple selection, you now have only one.

Dealing with long documents in the source text viewer[#]

For documents longer than ~30-40 pages, the text viewer automatically splits content into manageable chunks for better performance. Navigation controls appear in the "Source text" header:

- **Dropdown selector** - "Chunk 1 of 5" becomes clickable to jump to any chunk
- **Arrow buttons** - Navigate to previous/next chunk
- **"Next chunk" button** - Appears at the end of each chunk (except the last)

Visual Highlighting[#]

Each section of coded text, each causal claim, is shown with a highlight.

For overlapping or identical highlights with multiple links, overlaps are shown with varying color opacity. Clicking on multiple highlights shows a link selector for each section.

- Multiple highlights shown with varying color opacity
- Click on overlapping highlights to select specific links

Link Editor screen[#]

Opens when you highlight text or click on existing links.

Fields:

- **Cause and Effect selectors** - Unified project-wide label list sorted by frequency

- Both dropdowns use the same suggestions (combined causes+effects from the entire current project, not just the current source)
- Large projects: switches to type-to-search mode (min 2 chars), returning up to 200 matches; top ~300 are preloaded for quick access
- You can type new labels or select from the dropdown. You can type and select more than one cause and/or effect.
- **Quote field** - Editable text that gives the evidence for the causal claim. Also supports ellipses like this: `Actual quote [this text is ignored] quote continues blah blah.`
- **Chain toggle** - Defaults to unchecked on every fresh open/edit. If checked, saving keeps the overlay open and loads the previous Effect as the next Cause; if not checked, saving closes the overlay. The toggle remains checked only when the overlay stays open due to chaining.
- **Plain coding toggle** - Used when you want to record something which is not explicitly mentioned as a cause or an effect. Defaults to off. When on,
- the tag `#plain_coding` is added (if not already present) to the comma-separated list of tags.
- Whenever the tag `#plain_coding` is present:
  - the toggle is switched to on
  - the effect factor selector is forced to have the same contents as the cause factor
  - the effect factor selector is disabled for the user
- **Tags field** - Add tags to the link like `#hypothetical` or `check`
- **Favorite buttons** - Heart, exclamation, star toggles for marking important links or useful quotes. Later you can use these tags and favorites in filters.

Actions:

- **Save** - Create the link(s)
- **Delete** - Remove existing link
- **Cancel** - Close without saving

Links in Causal Map only have one cause and one effect. You can add multiple causes and/or effects to the boxes, and the system createsall combinations when saving. So if you put `unemployment` and `violence` in the Cause box, and `stress` and `worry` in the Effect box, the system will create four links.

About the factor label dropdown menus#

By creating links, you also create the names of your factors.

In Causal Map, a factor*is*its label. Once you create a label, there is nothing else to add.

Factor names which contain semicolons `;` get special treatment as they separate the different parts of 🧬 Hierarchical factors .

After beginning to create links between factors, already-coded factors will appear in the dropdown menus in the to and from factor boxes. For added convenience. The most frequently coded factors will appear at the top of this list

#doubtful? #hypothetical? Adding link tags#

Link tags#

Link tags are available as a special kind of memo when coding a link: you can use them to provide any kind of additional information.

Untitled

There is no need to actually use a hash `#` at the start of a link tag, though you can if you want. Just use any unique single word which is easy to search and filter on, like #nutrition or nutrition# or nutrition–.

As usual in Causal Map, you can apply one or more tags, and you can either select existing tags or create new ones on the fly.

Later, you can filter the map (see ✨ Transforms Filters: Include or exclude tags) to show only links containing or beginning or ending with specific hashtags (or parts of hashtags), and also for links which do*not*contain specific hashtags or parts of hashtags.

You can also use tags to narrow down your searches in 🔗 The Manage Links tab.

You can display tags on your map.

Conceptually, there are two kinds of tag.

Ordinary link tags#

You can use any tag which does not begin with a `?` to record any other information about the link, e.g.:

- respondent doesn't like this connection
- respondent feels good about the outcome
- for you, the analyst, e.g.
  - respondent is answering a different question
  - to tag links you want to come back and review.

Weak tags#

Weak tags are a special kind of tag. They are*caveats*. If you use weak tags, you should make sure that by default your maps do not include any link with a weak tag.

This is just a convention, it makes no difference to the Causal Map app.

They begin with `?` and are used to mark any link which you are not sure is always valid across the global context for the whole global map, for example:

- **the causal connection** is only valid for a specific context, e.g.
  - the respondent says this is only true for their village, not for other villages e.g. `?village X`
  - a link is only projected for the future e.g. `?future`
- you are unsure about **the claim about the causal connection**
  - a link is only a hypothesis e.g. `?hypothetical`

    - you as the analyst are not confident in the claim e.g. `?doubtful`
    - the source themselves are not sure e.g. `?source seems unsure`
    - to add other qualifying information e.g. `?probably hearsay`
    - to mark the fact that a connection is **weak or non-existent**, e.g.
        - Respondent makes a substantive claim that X does *not* influence Y, e.g. `?zero influence`
        - Respondent makes a substantive claim that X only insignificantly influences Y, e.g. `?weak`

AI Coding#

Requires AI subscription

- **Model dropdown -** Select AI model
- **Prompt box -** Enter coding instructions
- **Add source prompt -** Toggle, default ON
- **Response displays -** View AI responses and full JSON

Motivation for source prompt: it is just to describe the context/background info about each source. Not necessary e.g. where all the sources are from the same context which can be described in the main prompt. But important where some differ, e.g. mid-term reports or whatever.

If Add Source Prompt is ON, then show a text area above #text-viewer-content with usual greenish Save button to edit the corresponding source prompt for the current source.

Additional controls hidden behind gear icon (experimental):

- **Temperature slider -** Control randomness (default 0)

Iterative Processing: If your prompt contains lines with `====` on their own, each section before and after the line is treated as a separate iteration. Line endings and surrounding spaces are tolerated (CRLF/whitespace OK). First iteration is normal; subsequent ones include the full prior conversation history (all previous User prompts and AI responses) to build on earlier results. Only the results of the last iteration are added to the links table; all iterations are logged in the responses panel.

Workflow:

- Select one (or more) sources to process using the sources dropdown
- Select "Skip coded sources" if you don't want to recode sources which have already been coded
- Toggle "Add source prompt" to append the new Source Prompt field before the beginning of the main prompt
- Click "Process Sources" button
- Confirmation dialog shows model, word count, and warnings
- Pre-modal quick estimates are time-boxed (~2s). If the fetch is slow, the modal still appears and estimates may show as n/a. Heavy work only starts after you click Proceed. See `webapp/js/ai-manager.js` near `AI_DEBUG_QUICKSTATS_START`.
- AI processes sources in batches
- Results are also logged to the responses table on the right of the screen

[Tips on using the prompt history](#)

- Timeouts: per-iteration budget scales by model and iteration count (cap 540s total): Flash 120s/iteration; Pro 270s/iteration.
- Concurrency: Radio group labeled "Concurrency" (1–5) next to Region in AI settings. Default 1. Increase if you want faster processing but may risk 429/timeouts.
- Logging & Responses: each chunk inserts a pending row in `ai_logs` (status pending) and updates to success/error on completion; Responses tab auto-refreshes as rows update.

Analysis Filters Filter links tab

Do qualitative causal analyseis on the selected links by filtering or manipulating them.

🔍 **What you can do here:** Apply filters to focus your analysis on specific aspects of your data. You can trace causal pathways, group similar concepts, filter by themes or demographics, and much more. Think of this as your analysis toolkit - combine different filters to explore your data from different angles.

The Filter System: overview#

Use filters to narrow down and/or transforme the links you want to study. Filters are applied in order, from top to bottom.

- **Default filter**: Factor Label Filter
- **Add Filter** lets you insert filters at the start or between existing ones
- **Enable/Disable** toggles turn individual filters on or off
- **Remove** deletes a filter
- **Collapse** hides a filter's controls to save space
- **Clear All** resets to the Factor Label Filter

Hard vs Soft recoding#

Most filters leave factor labels untouched, but these 'Transform filters' filters temporarily relabel factors:

- Zoom
- Collapse
- Remove Brackets
- Soft Recode Plus
- Auto Recode
- Soft Relabel
- Cluster

No filters actually change your original coding.

💡 Tip: If you want to permanently rename or "hard recode" your factors, there are several ways to do that:

- Search and replace factors
- Search and replace links

For example, after clustering (which may give labels like C11), click a factor on the map and rename it (e.g., "Wellbeing") to save the new name permanently.

Zoom Filter

🔍 **What this does:** Simplify complex factor labels by zooming to higher levels of a hierarchy. For example, turn "Health; Mental Health; Depression" into just "Health" (level 1) or "Health; Mental Health" (level 2). Perfect for getting a big-picture view of your data.

- **Radio buttons** for levels (None, 1-9). Combine with [Collapse Filter] for label cleanup.
- **Level 1**:
- "foo; bar; baz" becomes "foo"
- "foo; bar; baz" becomes "foo"
- **Level 2**:
- "foo; bar; baz" stays the same
- "foo; bar; baz" becomes "foo; bar"
- **None**: No transformation

Collapse Filter

🏷️ **What this does:** Merge similar factors under one common label. Type or select multiple similar terms like "money", "income", "salary" and they'll all be replaced with the first term. Great for cleaning up data where the same concept is described in different ways.
Widgets:

- **Selectize dropdown** with existing labels where you can select one or more existing factor labels, or type parts of existing labels.
- **Matching options**: Start / Anywhere / Exact
- **Separate** toggle for individual replacements. When off, this filter replaces all matches with first search term. When on, a separate factor is created for each of the search terms.

Remove Brackets Filter

🧹 **What this does:** Clean up your factor labels by removing text in brackets. For example, "Education (primary school)" becomes just "Education". Choose between removing content in round brackets ( ) or square brackets [ ].

- **Radio buttons**: Off / Round / Square brackets
- Removes all text within selected bracket type

If you want to remove both kinds of labels, simply create another `Replace brackets` filter beneath this one.

Factor Label Filter

🎯 **What this does:** Show links connected to factors you care about (e.g. "Education"). Choose how many steps to look upstream (causes) and downstream (effects).
Widgets:

- **Factor selector** with existing labels. By default shows only labels from links currently visible at this stage of the filter pipeline. Use the **Show All** toggle to display all factor labels from the entire project instead.
- **Steps Up** (0-5): How many levels upstream to include
- **Steps Down** (0-5): How many levels downstream to include
- **Source tracing toggle**: Retain only links which are part of complete paths which all belong to the same source

- **Highlight toggle** (default: on): Show/hide custom highlighting (⭐ star and magenta border) for matching factors
- **Matching**: Start / Anywhere / Exact. Matching is case-insensitive.

How to use: 1) Select one or more factors.
2) Set Steps Up/Down to widen or narrow the neighbourhood.
3) (Optional) Turn on Source tracing to require paths from a single source.
4) (Optional) Turn off Highlight to hide the custom highlighting. 5) The map and tables update to show only links on those paths.

All the label and tag filters including exclude filters have three radio buttons below the selectize input called Match: Start (default), Anywhere or Exact to control how search terms match against labels/tags:

- **Start**: Match only at the beginning of text (default)
- **Anywhere**: Match anywhere within the text
- **Exact**: Match the entire text exactly

Multiple search terms are treated as OR not AND. preserve and highlight factors matching ANY of the search terms.

Focused factors show with colored borders in the map and have a star added for easy identification (when Highlight toggle is on).

Exclude Factor Label filter

🚫 **What this does:** Remove unwanted factors from your analysis. Type factors like "Unclear" or "Other" to hide them from your map and tables. Useful for cleaning up your data by removing vague or irrelevant categories.

- **Factor selector** for factors to exclude. By default shows only labels from links currently visible at this stage of the filter pipeline. Use the **Show All** toggle to display all factor labels from the entire project instead.
- **Matching options**: Start / Anywhere / Exact
- Multiple entries combined with AND logic
- If you want to exclude both/all of two or more entries, add another Exclude Factor Label filter.

Path Tracing Filter

🛤️ **What this does:** Find causal pathways between two specific points. Set a starting factor (like "Poverty") and an ending factor (like "Poor Health") to see all the causal chains that connect them. Great for understanding how problems and solutions are linked.

- **From selector** for starting factors. By default shows only labels from links currently visible at this stage of the filter pipeline. Use the **Show All** toggle to display all factor labels from the entire project instead.
- **To selector** for ending factors (results visible in Map and Links). Uses the same label source as the From selector (controlled by the **Show All** toggle).
- **Matching options**: Start / Anywhere / Exact

- **Steps** (1-5): Maximum path length
- **Thread tracing toggle**: Require only paths within same source
- **Highlight toggle** (default: on): Show/hide custom highlighting (⭐ star/magenta border for From factors, 🎯 target/dark yellow border for To factors)
- **Only indirect links** (default: off): Remove all direct links from From to To (only makes sense when both From and To are non-empty)

Exclude self-loops Filter

You can exclude self-loops from the maps, but that is more of a visual change. This is a real filter as part of the filter pipeline. For example, if you are using a filter like Link Frequency that might be retaining link bundles which are actually self-loops, so you might get unexpected results if you use the map setting to remove the self-loops. So this filter is a better way. It simply removes all links which are self-loops from the links table.

Link Tags Filter

🏷️ **What this does:** Filter your analysis by the tags you've added to links. Show only links tagged as "#important" or "#policy" to focus on specific themes or types of relationships you've identified.

- **Tag selector** with existing link tags from current project
- **Matching options**: Start / Anywhere / Exact

Combine Opposites filter

🔄 **What this does:** Unify opposite factor labels by matching tag numbers. If you have pairs like \`Foo [99]\` and \`Bar [~99]\` (where \`~99\` indicates the opposite), this filter rewrites \`Bar [~99]\` as \`Foo [99]\` to combine them under one label. The \`flipped_cause\` and \`flipped_effect\` columns track which causes and effects were flipped.
Toggle – Turn the filter on/off.

Strip tags from labels (default: on) – When enabled, removes `[N]` and `[~N]` tag patterns from labels after combining opposites. This keeps labels clean while preserving the tracking information in the `flipped_cause` and `flipped_effect` columns.

Labels can be written in pairs like:

- `Foo [99]`
- `Bar [~99]`

where `Bar` represents the opposite of `Foo`. The square brackets are optional - you can use `Foo 99` and `Bar ~99` - but brackets make it easier to remove tags later using the Replace Brackets filter.

If there are any such pairs, with matching integers, and the filter is switched on:

rewrite any Bar [~99] filters as Foo [00] and add new columns:

- `flipped_cause` column tracks which causes were flipped
- `flipped_effect` column tracks which effects were flipped

to the current augmented links table, so that if the label has been flipped, the value is True and otherwise False.

Wire up the filter as part of the standard filter system with save/restore to URL etc.

Also, when calculating new links table, create new text columns:

- `source_count_with_opposites`
- `citation_count_with_opposites`

The embellished counts always show all variants with custom SVG icons (no total prefix). Four circle icons represent the flipped status:

- ‾ (unflipped/unflipped)
- ╲ (unflipped/flipped)
- ╱ (flipped/unflipped)
- _ (flipped/flipped)

So if a bundle has 12 citations where 5 are unflipped/unflipped, 2 have flipped cause and flipped effect, and 1 has flipped cause but non-flipped effect, the text is: `light-blue-circle5, dark-red-circle2, mixed-circle1`. If nothing were flipped, the label would just be `light-blue-circle12`. Do the same with source counts too, counting the unique sources in each variant.

When the filter is on, and source count or citation count is selected, the graphviz and graphviz maps change to use `source_count_with_opposites` or `citation_count_with_opposites` just for the labels. The edge width calculation remains driven by source count or citation count, as selected.

Exclude Link Tag filter

🚫 **What this does:** Remove specific types of links from your analysis. Exclude links tagged as "#uncertain" or "#duplicate" to focus on higher-quality data. Helpful for filtering out questionable or irrelevant causal claims.

- Same as Link Tag filter except *exclude* links containing these tags. Multiple entries are combined with AND, i.e. only exclude links where both entries match. (💡 Tip: if you want to exclude both/all of two or more entries, add another filter).

Exclude self-loops Filter

You can exclude self-loops from the maps, but that is more of a visual change. This is a real filter as part of the filter pipeline. For example, if you are using a filter like Link Frequency that might be retaining link bundles which are actually self-loops, so you might get unexpected results if you use the map setting to remove the self-loops. So this filter is a better way. It simply removes all links which are self-loops from the links table.

Link Frequency Filter

📊 **What this does:** Focus on the most important causal relationships by filtering out rare ones. Choose "Top 10" to see only the most frequently mentioned connections, or set a minimum threshold like "at least 3 sources" to ensure reliability.

- **Slider** (1-100) for threshold
- **Type**: Top vs Minimum
- **Count by**: Sources vs Citations

Examples:

- **Minimum 6 Sources**: Only links mentioned by 6+ sources
- **Top 6**: Only the 6 most frequent link bundles

By default, setting the slider to 6 means we are selecting only links with at least 6 citations.

If you switch to "Sources", we are selecting only links with at least 6 sources.

If you switch to "Top" we are selecting only the top 6 links by citation count, etc. The selection respects ties, so that if there are several links with the same count, either all of them or none of them will be selected.

Factor Frequency Filter

📈 **What this does:** Similar to Link Frequency, but focuses on the most important factors (causes and effects). Show only the most frequently mentioned themes or concepts to identify the key issues in your data.
Same controls as Link Frequency but applies to factors instead of links.

Source Groups filter

👥 **What this does:** Filter your analysis by participant demographics or document characteristics. For example, show only responses from "women aged 25-35" or interviews from "urban areas." Perfect for comparing how different groups see causal relationships. This is very similar to the [Source Groups widget](../source-groups-sub-panel/) in the Sources sub-panel, but having it here too means you can add multiple source filters to the pipeline.

- provides
- a prepopulated dropdown called Field with all the metadata fields plus title and projectname
- another multi-selectzie called Value. Multiple values work as OR: either/any count as a match
- a previous/next button pair to cycle through values of the selected group
- Example: Add two Source Groups filters in the pipeline to combine criteria (e.g., first filter Field = gender → Value = women, then another filter Field = region → Value = X) so you see links from women AND from region X.

Everything Filter

👥 **What this does:** Filter your analysis by any characteristic of your links and their sources. Useful for anything not covered by the other filters, for example, - show only links with negative sentiment - Show only links from one source Also shows source separators and their values, often used for common sections within multiple sources texts.

- **Field dropdown** with all fields in the links table
- **Value selector** filtered by selected field
- **Navigation buttons** to cycle through values

- **Clear button** to reset

Soft Relabel Filter [#](#)

👥 **What this does:** Temporarily relabel factors.

- **Old factor labels** listed on the left
- **New factor labels** editable, listed on the right
- **Load labels button** when pressed, adds into the Old labels list any current factor labels (in links as currently filtered) which are not yet listed in the Old labels list and adds the same Old label to the New column as default.
- **Clear button** to clear the New fields
- **Clear ALL button** to clear all rows

Effect: all factors exactly matching any of the labels in the Old list are relabelled with the corresponding labels from the New list. factors not listed are not relabelled but preserved.

Many use cases:

- temporarily merge multiple factors into one
- you are using magnets and you can't really use the formulation you want because you want to maximise similarity with existing labels
- eg you are using "floods" as a magnet but you really want it as a hierarchical factor like "environmental problems; floods" but you can t use that as a magnet.

Keyboard shortcuts (Win/Linux ⇄ macOS):

- Tab / Shift+Tab: move focus down/up between NEW cells
- Arrow Up/Down: move focus up/down between NEW cells
- Alt+Arrow Up/Down (mac: Option+Arrow): move the current row up/down
- Ctrl+Arrow Up/Down (mac: Cmd+Arrow): move the current row up/down
- Delete current row:
- Shift+Delete (mac: Shift+Fn+Backspace) or
- Ctrl+Shift+K (mac: Cmd+Shift+K)

Potentially, one NEW label might have multiple OLD labels.

Soft Recode Plus filter [#](#)

Requires AI subscription

🧲 **What this does:** Group messy factor labels under clearer names you choose (called "magnets"). Example magnets: `Improved health`, `Education programs`, `Income changes`. The filter finds the closest magnet for each label and replaces it.

Controls:[#](#)

Create Suggestions for Magnets[#](#)

(collapsed by default): Optional. Ask AI to propose clear names from your current labels. Insert adds them to your magnets box to review/edit.

- **Number of clusters** – Choose how many groups to find for AI suggestions.
- **Labelling prompt -** With the usual buttons to save and recall previous prompts
- **Insert**

Main panel#

- NEW: **Only unmatched** – A new toggle which appears right at the top, before the Create Suggestions subpanel. default off.
- **Magnets** – One magnet per line. Saved per project. Use Prev/Next to browse recent sets.
- **Similarity slider** – The raw labels are dropped if they are not at least this similar to at least one cluster.
- **Drop unmatched** – If on, remove links whose labels don't match any magnet. If off, keep them as they are.
- **Save** – Save magnets and apply the recode.
- **Remove hierarchy** – Strip any text before the final semicolon
- **Clear / Prev / Next** – Manage saved sets.
- **Recycle weakest magnets:** – A slider starting at 0 , default is 0. If the slider is n >0, then we look at the cluster assignenments which would have been returned and find the n clusters which we are going to recycle. Reassign them to their nearest cluster, providing the similarity is still better than the similarity cutoff. This way we don't lose factors / links which are otherwise assigned to smaller clusters which may get excluded later on in the filter pipeline. When it is on zero, it makes no difference and we just use the solution based only on the magnets, similarity, and remove_hierarchy. The maximum value changes to match the total number of magnets.

Recoded columns#

When you use Soft Recode Plus, the Links and Factors tables show special columns that track which labels have been recoded:

- **Links table**: Shows `_recoded_cause` and `_recoded_effect` columns (✓ for recoded, ✗ for not recoded)
- **Factors table**: Shows `_recoded` column (✓ if the factor appears at least once as recoded, ✗ otherwise)
- These columns only appear when Soft Recode Plus is active in your filter pipeline
- You can filter by these columns using the True/False dropdowns in the table headers

These columns track recoding from any filter that transforms labels: Soft Recode Plus, Zoom, Collapse, Remove Brackets, Soft Relabel, Cluster, Hierarchical Cluster, and Combine Opposites.

Process only unmatched NEW#

the point of this is: what if I apply some (maybe standard) magnetisation and matches plenty of factors but there might be some important material left unmatched, not just noise. so i can use a PAIR of these filters. in the first one, I leave OFF its Discard Unmatched toggle and in the second filter switch ON its Only Unmatched filter. (if there is no preceding SRP filter with Discard Unmatched=OFF, this second SRP filter does nothing).

So now,

- the Create Suggestions (if used) optionally processes ONLY the UNMATCHED factor labels
- the magnetisation (if labels are non-empty) works only on the unmatched factor labels.
- the actual output of the second filter is now the union of both soft-recode processes, i.e. the original matches from the first and the new matches of the previously discarded material from the second.
- the Discard Unmatched on this second filter works as usual: if it is OFF, then we also return all the still-unmatched labels.

Meaning Space (2-D embeddings)#

Go to the map formatting and select Layout → Meaning Space to see a 2-D scatter of your factors in "meaning space".

- Magnets are shown with labels; raw factor labels are dots.
- Colour indicates the magnet group; magnet dot size represents group size.
- You can pan (drag) and zoom (mouse wheel and zoom controls).
- Double-click on an empty part of the map to zoom in at that point.
- Tooltips on dots show the original (raw) labels and the magnet label.

Motivation for Remove Hierarchy#

"Remove hierarchy", default off. if on, strip any text before a final semi-colon, if no semi-colon, do not change the text.

```
something; another thing
```

is treated same as

```
another thing
```

.... but it continues to be treated as "something; another thing" in the rest of the filter pipeline.

Quick workflow:

1) (Optional) Open Create Suggestions for Magnets panel → set Number of clusters and use Insert to get AI suggestions.
2) Use these suggestions and/or edit them, paste or type your own magnets (one per line).
3) Click Save.

- Clusters your current labels (factors as currently filtered), ranks typical examples, and asks AI to suggest clear names.
- Returns suggested names into the magnets box; you can edit them before Save.

See tips on using the history to reuse both your labelling prompt and magnet sets.

Motivation for "recycle weakest magnets": suppose you create 20 magnets, and then apply more filters like say a [link frequency filter]() so that you end up with say only 5 factors. If you then *remove* those factors from the magnets list which are *not* included in the final output, you will usually increase the coverage of your map (re-assigning raw labels which fit best with one of the "lost" labels but still fit well with one of the "surviving" labels). This is what the Recycle slider does: it recycles the specified number of smaller magnets and reassigns them to the larger magnets. So in the example, if you start off with 20 magnets but your final map only shows 5, try recycling say 10 or even 15 of the missing factors.

Note that Recycle Weakest Magnets is applied BEFORE Drop Unmatched.

Clustering filter

Requires AI subscription

🎯 **What this does:** Automatically discover groups of similar factors in your data using machine learning. The system finds natural clusters of related concepts and labels them with cluster numbers. Great for exploratory analysis when you're not sure what causal themes exist.

- **Enable toggle** (starts disabled)
- **Number of clusters** (1-9)
- **Server-side processing** using `cluster_factors_pgvector` database function
- Uses k-means clustering on factor embeddings
- Labels clusters with numeric IDs

Auto Recode filter

🪜 **What this does:** Quickly turns your current set of labels (after any previous filters like Zoom) into a simple tree you can "roll up" or "open out". Pick a small number of clusters for a big-picture view, then nudge the Balance and Similarity to tidy results. Designed for fast, practical exploration on real projects.

Motivation#

Making sense of hundreds or thousands of factor labels is hard.

You might use something like soft Recode Plus, but often you'll ask for 20 clusters to cover a wide range of meanings. Then after filtering out insignificant data, you end up with only 7 clusters — losing coverage. Ideally you'd go back and recreate just 7 clusters, but that gives different results. Frustrating!

The point of this Auto Recode filter: have your cake and eat it. Ask for an foldable/unfoldable hierarchical solution. When you move the slider to 15, you get the best solution for 15 clusters. Slide it to 3, you instantly get the best solution for 3 clusters.

Controls:#

- **Enable toggle** (starts disabled)
- **Balance (0..1)**: 0 = prefer more distinct clusters; 1 = prefer more even sizes. Changing this can be slow because the tree has to be rebuilt

- **Number of clusters (K)**: 2–50. Unfolds the returned tree locally to K. This is fast unless you increase beyond 20.
- **Similarity** ≥: prune locally by similarity to the centre of each cluster.

NEW: **AI labelling prompt** with history controls. Use this to suggest clearer names for each cluster:

- Saved in the prompts table as type `hierarchical_label` (shared across projects; history shows current first then others).
- A Save button stores your prompt; it also auto-saves on blur and after the first tree build.
- When you raise K (unfold deeper), we call AI in parallel only for the two new child clusters introduced by each applied split, using up to 8 representative labels per child as context. For K clusters this is K–1 requests. Folding to fewer clusters does not call AI; existing AI labels or medoid representatives are used.
- If the prompt is blank, we show the medoid representatives for each cluster.
- If earlier splits already have AI labels (K > 1), we include a reference list of those labels so new labels avoid overlapping meanings.

NEW: **Seed labels (optional)** with history and strength:

- Provide up to K seed labels (one per line). Seeds softly influence split formation but are not included in the final tree (not nodes, not representatives).
- Saved in the prompts table as type `hierarchical_seeds` with standard history controls (Prev/Next/Dropdown/Save).
- Seed strength (0..1) adjusts influence; 0 is a no-op (identical to no seeds). Changing strength or seeds triggers a single backend rebuild (like Balance). Changing K or Similarity does not re-call the backend.

How to use (quick):

- Add the filter and enable it. We build a quick draft tree from the labels you see now (respecting any filters above, like Zoom).
- Set **Balance** if you want more equal-sized groups; the first build may take a moment on large projects (one server call).
- Use **K** to choose how many clusters to show. Changing K is instant (no extra server calls).
- Use **Similarity** ≥ to drop weak matches. If either side of a link isn't matched, that link is hidden.

Notes:

- On very large projects, we automatically sample a representative set to build the tree, then assign the rest to the nearest cluster. This keeps things responsive while preserving the overall picture.
- 💡 Tip: changing the number of factors should be instant if they are less than 20. Setting more than 20 can be slow. If you are going to want more than 20, set this number initially to the maximum number you are likely to want. You can then easily reduce it. Gradually decreasing the number is fine, but *gradually increasing* it will be very slow.

A good prompt looks something like this:

> This is a list of many raw labels grouped into two different clusters, with their cluster IDs, together with a reference list of other labels. Return a list of two new labels, one for each cluster ID. Each label should capture the meaning of the whole cluster, using similar language to the original raw labels, but in such a way that the labels you create are distinct from one another in meaning. Try not to be too generic, try to be as concrete as you can. Do NOT provide labels which include causal ideas, like "X through Y" or "X leading to Y" or "X results in Y" or "X improves Y" etc. Equally, don't include conjunctions in the title like "X and Y". The meaning of the labels you give me should ideally not overlap in meaning with one another or with the labels in the reference list.

Optimized Cluster filter

⚠️ DEPRECATED   Requires AI subscription

> ⚠️ **This filter is deprecated.** Its functionality has been merged into Soft Recode Plus. The filter still works for backward compatibility with existing bookmarks/URLs, but new instances cannot be created. Use Soft Recode Plus instead for optimal clustering and recoding.

🎯 **What this does:** Automatically finds the most optimal factor labels to use as centroids through genuine optimization. Unlike regular clustering that just groups similar items, this finds the best possible n≤N labels that maximize coverage with similarity ≥S. Perfect for discovering the most representative concepts in your data.
Controls:

- **Max Centroids (n)** - Maximum number of optimal centroids to find (2-50)
- **Similarity ≥** - Minimum similarity threshold for grouping labels (0-1)
- **Timeout (s)** - Optimization time limit in seconds (5-60)
- **Drop unmatched** - Remove labels that don't meet similarity threshold
- **Real-time status** - Shows optimization progress and results

How it works:

1. Extracts all unique labels from your current data (1K-30K labels supported)
2. Runs iterative optimization with multiple strategies (random, frequency-based, diverse selection)
3. Uses hill-climbing optimization to find the best possible centroids
4. Shows coverage percentage and timing information
5. Returns recoded links table with optimal centroid labels

Optimization Strategies:

- **Random selection** - Tests random starting points
- **Frequency-based** - Prioritizes most connected labels
- **Diverse selection** - Maximizes distance between centroids
- **Hybrid approach** - Combines best-so-far with random exploration

Performance Features:

- **Sampling strategy** for datasets >1000 labels (uses representative subset)
- **Early termination** when excellent coverage (≥95%) is achieved
- **Configurable timeout** prevents infinite optimization loops
- **Multiple iterations** with different starting strategies for robustness
- **Smart caching** - Embeddings cached separately from algorithm parameters for fast parameter changes
- **Quote-safe processing** - Handles labels with quotes, apostrophes, and special characters

Technical Implementation:

- Client-side optimization using cosine similarity on embeddings
- Hill-climbing algorithm with local search improvements
- Genuine optimization problem solving (not just k-means clustering)
- Real-time UI feedback showing progress and final results
- Handles massive datasets efficiently through smart sampling
- **Original label preservation** - Stores original labels in `_recoded` metadata for map display
- **Chain compatibility** - Works seamlessly with zoom filter and other transformations

Soft Recode Integration:

- Optimized cluster results available as magnet source in Soft Recode filter
- AI can generate meaningful labels for optimal centroids
- Seamless workflow from optimization to AI-powered naming

This filter implements the optimization challenge described in the technical documentation: finding optimal centroids that maximize label coverage within similarity constraints.

Tribes filter

Requires AI subscription

🏛️ **What this does:** Group your sources (participants/documents) by how similarly they describe causal relationships. This reveals different "tribes" or perspectives in your data - for example, optimists vs. pessimists, or urban vs. rural viewpoints.
Controls:

- **Number of clusters** - Radio buttons: Off, 1-9
- **Similarity cutoff** - Slider: 0-1
- **Drop unmatched** - Toggle
- **Min cluster %** - Slider: 0-20% (prevents "1 big + many singletons" pattern)

.

It returns:

- `tribeId` (cluster ID)
- similarity to the centroid

- similarity rank These are joined to the links table by source ID and appear as additional columns. If Drop unmatched is ON, links with similarity below the cutoff are removed.

We can then show maps for each tribe and/or for the most typical source in each tribe. we could also then create a typical story centred around the current factors, i.e. told in terms of our concepts.

--->

Custom Links Label

🏷️ **What this does:** Configure how link labels appear on your map based on any field in your data. Choose what information to display (like tribe memberships or custom attributes) and how to show it (counts, percentages, or statistical significance).
Controls:

- **Field** - Dropdown of available fields from your filtered data (typically shows custom fields like tribe ID)
- **Display mode** - Choose how to show the data:
- **Tally** - Show counts for each value (e.g., "T1:4 T2:3")
- **Percentage** - Show what % of each value's total links appear in this bundle (e.g., "T1:34% T2:22%")
- **Chi-square** - Show bundle size, then which values are significantly over-represented (⬆) or under-represented (⬇) (e.g., "45 (T1⬆ T3⬇)")

To use:

1. Add the Custom Links Label filter to your pipeline
2. Select a field (e.g., `custom_tribeId` after running the Tribes filter)
3. Choose a display mode
4. In Map Formatting, set Link Labels to "Custom Links label"

Example use cases:

- **After Tribes filter:** Show which tribes contribute to each connection (T1:5 T2:2 T3:1)
- **Significance testing:** Identify connections where certain tribes are surprisingly over/under-represented (T1↑ T3↓)
- **Custom attributes:** Display any custom field you've added to your data

<!---

Technical details:

This is a **non-filtering** filter - it doesn't change which links appear, only configures how they're labeled on the map.

Display modes:

1. **Tally:** Counts occurrences of each value within the coterminal bundle

2. **Percentage:** For each value, calculates: (count in this bundle) / (total count of this value across all filtered links) × 100
3. **Chi-square:** Tests if observed counts differ significantly from expected:
4. Expected = (bundle size) × (value total) / (grand total)
5. Chi-square = (observed - expected)² / expected
6. Format: "bundleSize (value↑ value↓)" where:
    - Shows ↑ if chi-square > 3.84 and observed > expected ($p < 0.05$)
    - Shows ↓ if chi-square > 3.84 and observed < expected
    - Only shows significantly different values in parentheses

The filter populates its field dropdown from `currentFilteredLinks` (the output of the filter pipeline), so it sees all fields added by previous filters.

Projects Panel

📁 **What you can do here:** Organize and manage all your research projects in one place. Create new projects, share them with collaborators, add descriptive tags, and control who can view or edit your work. You can also merge multiple projects together or archive old ones to keep your workspace tidy. See also the [File menu](../file-menu/) for more project management options.

Project Management#

- **New project** - Create with name and description
- **Load project** - Open selected project
- **Edit project** - Click row to modify name, description, tags, sharing (or use Edit project Modal)
- **Archive/Unarchive** - Hide/show projects
- **Read-only toggle** - Restrict editing even for owners/editors
- **Archive toggle** - Show/hide archived projects

Bulk Operations#

Select projects with checkboxes, then:

- **Delete** - Remove projects and all data
- **Apply Tags** - Add tags to selected projects
- **Remove Tags** - Remove tags
- **Toggle Archive** - Archive/unarchive
- **Merge** - Combine multiple projects into one

Sharing and Permissions#

- **Email-based collaboration** : add and remove colleagues' email addresses
- **Locked / Read-only permissions** for viewing without editing
- **Global sharing** for public read-only access
- **Permission badges** next to project names
- **Admin only: admin panel** for user management

Note: In the Edit Project modal, the informational notice "Your projects are public and can be viewed by anyone. Upgrade your subscription to keep them private." is shown only to users who can edit the project (owner, editor, or admin). Viewers do not see this notice.

Versioning#

The app automatically backs up your project, so you can restore earlier snapshots if you want.

- An automatic backup is made every 10 minutes if you have made changes.
- You can make a manual backup from the Project Info screen.
- You can use the Version dropdown menu in the Project Info screen to see all available backups with details.
- From here you can restore a backup, with a confirmation step before applying changes.
- After restoring an earlier version, you can always go back to the latest version if you want, using the same dropdown menu.

This panel shows a dropdown list of times when you made changes to the mapfile in UTC/GMT. Along with the size of your file which can help you identify which timepoint you want to revert to. It can be easy to forget what time you made alterations to your file, so if you're likely to want to restore a previous map it is best to note the time so that you can easily return to it.

Sources Panel

📚 **What you can do here:** Upload your research documents (PDFs, Word docs) and organize them with custom metadata like participant demographics or interview dates by adding and editing **custom columns**. This is your document library and metadata manager.

Upload & Setup

Upload Source Texts

- **Click to select** one or more PDFs or DOCX or RTF files
- **Optionally split large documents into multiple sources** using separator patterns
- **Confirmation dialog** shows projectname→ID mapping
- On completion, the app automatically:
  - selects the sources via Sources Dropdown
  - switches to the Create Links sub-tab and loads the first source's text into it
  - on the right, switches to Sources panel with and the View & Edit subpanel

Splitting documents into multiple sources with source separators#

This feature helps you split individual documents you upload into multiple sources. These separators are *hard*: you use them just once, on uploading one or more documents, to produce multiple sources.

- In the Confirm Upload screen, there is a Sources Separator text box
- Text lines matching special "regex" patterns separate into multiple sources. So if you have sections marked with "Source Number 12", "Source Number 13" etc, just put "Source Number.*" in the box.
- This will produce multiple sources with "Source Number 12" etc as source_id
- Live preview of new source IDs with count
- User can leave blank for normal upload

Section Separators#

This feature helps you split existing source texts into sections. These separators are *soft*: they don't permanently change the file and you can add or remove one or more separators on the fly.

- **Two-line expandable textArea** for one or more (regex) patterns
- **Section header detection** within imported texts
- **Special styling** for matching rows in text viewer
- **Links created in the different sections can be filtered using the Everything Filter**. So you can do things like "Show me all the causal claims (links) only in answers to Question 7".

- Into the box, you type special "regex" patterns to create sections withing source texts. So if you have sections marked with "Section Number 12", "Section Number 13" etc, just put "Section Number.*" in the box.

- This will highlight the sections and create fields like "Section Number 12" etc as section ID.

Sample Check#

You want a table showing your sample: gender * region?

Use this simple customisable table to check your sample according to any custom columns you have defined.

- See also: Statistics Panel and Analysis Filters.

Custom Column Analysis#

- **Aggregate by** multi-select for cross-tabulation
- **1 column**: Simple count table
- **2 columns**: Cross-tabulation table
- **Maximum 10 values** per column for analysis
- **URL state preserved** for bookmarking

View & Edit Your Sources#

- See also: Tips for all tables and Custom columns.

Sources Table

- NEW column **Source Prompt** - this new column shows the first few characters of any text in this field. It can be edited as usual with the existing pencil icon/ edit button in each row.
- **Checkbox selection** for analysis pipeline
- **Row editing** with keyboard navigation
- **Custom columns** for metadata
- **Uncoded column** - Shows true/false for sources with no links; filterable to find uncoded sources
- **Fullscreen mode** available

Table Editing Features#

- **Range selection** - Click and drag
- **Copy/paste** - Ctrl+C, Ctrl+V
- **Arrow key navigation**
- **Delete/Backspace** to clear cells
- **Column/row selection** - Click headers
- **Double-click editing**

Custom Columns#

- **Manage Columns** opens a modal to add and remove multiple columns at once
- **Toggle visibility**
- **Double-click a cell in a custom column to edit it** or via source edit modal
- **Copy and paste** selections with ctrl-C, ctrl-V.

- **These columns are available elsewhere in the app, e.g.**
- In the [Source Groups filter]
- In the [Everything filter]
- In the [Sample Check table]

Map Panel

🗺️ **What you can do here:** See your causal relationships as an interactive network map. Drag nodes around, click on links to edit them, and use the controls to customize how the map looks. You can even drag one factor onto another to quickly create new links. This is where your data comes to life visually.

Map Controls

- **Jump to factor** - Type-to-search dropdown to quickly find and select factors on the map. Type to filter options, then press Enter to select all matching factors. Supports multiple selections.
- **Refresh layout** - Return the map to its original state before zoooming, moving etc.
- **Copy image to clipboard** - Get a very high-quality image copied straight into your clipboard which you can paste in a report or presentation.
- **Copy legend**
- **Zoom in/out** controls
- **Double-click** anywhere on the map background to zoom in to that point

Map Legend

Discrete text legend showing:

- projectname and included sources
- Citation coverage percentage
- Visual encoding explanations (link sizes, colors, numbers)
- Applied filters summary
- 💡 Tip: Click Copy legend to copy this text to clipboard.
- You can drag the legend box to reposition it on the map.

Map Formatting

Customisable formatting (Things you can tweak)#

Layout: change how the map is laid out and how you interact with it.

- Interactive and most of the other layouts are good while you are conducting your research. They are fast and you can interact with the results -- moving factors around, clicking to edit, etc.
- Print/Graphviz layout is best for static images e.g. for reports and journal articles.
- Direction: For Interactive and Print/Graphviz layouts, choose LR (left-to-right, default), TB (top-to-bottom), or BT (bottom-to-top).

Factor Labels: (you can see the same data in the Factors Panel)

- Source count (default)
- Citation count
- Sentiment (mean incoming edge sentiment, -1..+1)
- None

Link Label Font Sizes Link Widths: Citation count (default) Source count, None **Link Labels:**

- **Source count** (default)
- **Citation count**
- **Sentiment** (mean edge sentiment, -1..+1)
- **Tribe Tally** - Count of sources per tribe (e.g., "T1:4 T2:2")
- **Custom Links label** - Use configuration from Custom Links Label filter
- **Unique Sources** - Alphabetical list
- **All Sources** - Complete list with repeats
- **Unique Tags** - Alphabetical list
- **All Tags** - Complete list with repeats
- **None** - Show no labels on links

Factor Colors:

- **Outcomeness** (default) - Based on in-degree ratio
- **Source count**
- **Citation count**
- **None**

Factor Sizes:

- **Citation count** (default) - Font size scales with citation count
- **Source count** - Font size scales with source count
- **None** - All factors use uniform size (increased by 50% for visibility)

Self-loops:

- **Show toggle** (default: on)

Fixed visual appearance (things you can't tweak)#

Link Styling:

- Arrowheads colored by mean sentiment
- Color scale: muted blue (+1) → grey (0) → muted red (-1)
- Bezier curved edges with bundling

Factor Styling:

- Size scaled by node degree (with bounds)
- Border color reflects mean incoming edge sentiment
- Matched factors show dashed colored borders

Interactive Features#

These work for all layouts except Print/Graphviz layout.

- **Drag factors** to temporarily reposition them
- **Drag factor to factor** to create new links
- **Shift+drag** for box selection of multiple factors (opens edit modal)

- **Ctrl+drag** for box selection of multiple factors (direct selection, no modal)
- **Click a link** to edit.
- **Click a factor** to edit; shift-click or ctrl-click to add to selection without opening modal.

Editing and deleting (multiple) factors#

- Select factor(s) by clicking a factor, shift-click or ctrl-click to add more, or shift+drag/ctrl+drag a box around multiple factors, then:
- Move selected factors together
- Delete matching factors everywhere or in current view only
- Rename matching factors everywhere or in current view only

What does "everywhere or in current view only" mean?

- **everywhere**: all links containing factors with exactly the selected labels will be deleted
- **in current view only**: all links containing factors with exactly the selected labels (and matching the current filters, i.e. those you can see in the current map) will be deleted

💡 Tip: By control-clicking or shift-clicking multiple factors you can easily rename several at once, e.g. you can merge multiple factors as a single factor.

Grid layout#

Factors containing a tag of the form `(N.M)` or `(N,M)` anywhere in the label (where N and M are integers) are positioned on a grid layout. The grid coordinate tags are automatically stripped from displayed labels.

Grid layout toggle: Enable/disable grid layout in Map Formatting. Defaults to enabled. Disabled automatically when no grid tags are present.

Interactive Layout:

- Grid-tagged factors are positioned at their grid coordinates and locked in place
- Other factors with no grid tag are positioned freely within the grid bounds
- Grid bounds: from smallest x -1 to largest x +1, and smallest y -1 to largest y +1

Print/Graphviz Layout:

- Grid-tagged factors anchor the initial and final ranks:
- Factors with minimum rank coordinate (first number) are anchored at `rank=min` (initial rank)
- Factors with maximum rank coordinate are anchored at `rank=max` (final rank)
- This improves layout stability while allowing Graphviz to position other nodes optimally
- Grid coordinate tags are stripped from labels in the output

Grid coordinates respect layout direction:

- **First number (N)** always maps to the rank direction (main flow direction)
- **Second number (M)** always maps to the perpendicular direction
- **BT (Bottom-Top)**: First number = y (rank), y starts at bottom (flip y), second = x

- **TB (Top-Bottom)**: First number = y (rank), y starts at top (normal), second = x
- **LR (Left-Right)**: First number = x (rank), x starts at left (normal), second = y, y starts at top
- **RL (Right-Left)**: First number = x (rank), x starts at right (flip x), second = y, y starts at top

Vignettes

📝 **What you can do here:** Generate AI-powered narrative summaries of your causal maps. Choose between a "whole map" summary that covers all the relationships, or a "typical source" story that focuses on one representative case. Perfect for creating reports or explaining your findings in plain language.
How to use:

1. Select your **model** and **region** settings
2. Choose **Whole Map** or **Typical Source**
3. Enter or edit your **prompt** (use the navigation buttons to browse previous prompts)
4. Click **Write Vignette** to generate

Whole Map: Creates a summary of all relationships in your current map view. the app provides the following data which is appended to the prompt:

- The overall map (same as you can see) including factor frequencies and bundled causal links with average sentiment
- Up to 5 "typical sources" that tell the most common stories within the current map, with their quotes and metadata including source ID, Title and Filename.

Typical Source: Focuses on the single most representative source, showing individual links with quotes and sentiment.

Output format: Results are displayed as markdown with support for:

- Headers, bold, italic text
- Bulleted and numbered lists
- Callouts/quotes (using `>`)
- Code blocks

You can edit your prompt to change the tone, audience, or focus before generating. See the tips on using prompt history for more details.

Factors Panel

🏷️ **What you can do here:** See all the factors (causes and effects) in your data, ranked by how often they appear. Select multiple factors to rename them, merge similar ones, or delete unwanted entries. If you've added demographic data to your sources, you can also see statistical breakdowns showing which groups mention certain factors more often.
above both links and factors tables add a toggle "Use filters". If on (default) the table is filtered by the links filters. If off, we bypass this part of the pipeline and filter only by project and sources.

The Factors panel displays all unique labels from the current filter pipeline.

Table Features:

- Columns include:
- **Citation Count** – total number of citaions of this factor (as cause or effect)
- **Source Count** – number of different sources mentioning this factor
- **Citation Count: In** – number of citations of this factor as an effect of something
- **Citation Count: Out** – number of citations of this factor as a cause of something
- **Source Count: In** – number of sources mentioning this factor as an effect of something
- **Source Count: Out** – number of sources mentioning this factor as a cause of something
- Sorted by citation count (descending order)
- Click-to-select (no checkbox column)
- Server-side pagination consistent with other tables
- Actions column with edit button to open factor edit modal

Action Buttons:

- Delete: Remove selected factors
- Relabel: Rename selected factors
- Search/Replace: Find and replace text in factor names
- Merge: Combine multiple factors into one
- Buttons disabled until factors are selected
- 💡 Tip: Use [Search/replace] for quick, scoped relabeling.

Find out more about bulk delete and relabel of factors [here].

Bulk factor labels editor[#]

The search/replace functions in the factors and links tables are useful, but what if you have thousands of factors to look at? You might prefer this bulk editor.

Toggle the **Bulk Edit** switch to edit multiple factor labels at once. The table header remains visible for sorting and filtering, while the table body is replaced with a line-by-line editor.

Features:

- **Multi-cursor editing**: Use `Alt+Click` or `Ctrl+Alt+Up/Down` to add multiple cursors
- **Find occurrences**: Use `Ctrl+Alt+Right` to add next occurrence, `Ctrl+Alt+Left` for previous
- **Select all matches**: Use `Ctrl+Shift+L` to select all occurrences of selected text

- **Line-by-line editing**: You can only edit existing labels - you cannot add, remove, or reorder lines
- **Recoded labels**: Labels that have been recoded (shown with yellow background) are read-only and cannot be edited
- **Sort and filter**: The editor automatically updates when you sort or filter the table (any unsaved edits are discarded)

How it works:

1. Toggle **Bulk Edit** on
2. Edit factor labels directly in the editor
3. Press **Save Changes** to apply your edits
4. A confirmation dialog shows which labels will change and how many links will be affected
5. After saving, the editor refreshes to show the updated labels

NEW: add a second column to this div so that the editor takes up 9/12 of the width. in the new column, provide some live info about the selected factor: source and citation counts, and a list of sources mentionoing it.

Search/replace#

Near the top is a row containing a search box. If you type something into it,

- a replace box and a Replace button also appear.
- the table is filtered to show only matching rows

The search is **case sensitive**.

You can then alter what you see in the Replace box:

- in the factor label column in the table, you see a preview of what the affected rows would look like;
- if you delete all the text so the replace is empty, the preview shows the effect of deleting the search text from each label.

Then when you are satisfied, check all the checkboxes where you want to update the labels as shown. If you want, select all rows using the checkbox at the top of the column. Note, if there are more hits than fit on this page of your table, you'll want to either treat each page separately or increase the page size with the Page Size selector.

Finally, hit the Replace button to actually update the labels as shown in the rows you selected. What actually happens is that the Cause and Effect labels in all the currently selected links are changed. As you'd expect, this search/replace only affects the factors for the currently selected links: for example if you have only selected the first three sources, this update will not affect the links in the other sources.

.

Demographic Breakdowns#

- **Breakdown selector** - Choose custom columns to analyze by demographics
- **Count type** - Source count (default) or citation count
- **Display mode** - Counts (default) or % of baseline (cell as a percent of that breakdown group's total across all factors)
- **Statistical testing** - Chi-squared analysis to identify significant patterns
- See also Statistical Significance Testing

Statistical Significance Testing#

When you select exactly **one custom column** for breakdown, the factors table includes powerful chi-squared significance testing to identify factors that are preferentially mentioned by different groups.

Show Differences dropdown appears with threshold options:

- **Off** (default)
- **p < .1** (marginally significant)
- **p < .05** (significant)
- **p < .01** (highly significant)
- **p < .005** (very highly significant)

Visual indicators:

- **Significant column** - Shows "Yes" (red highlight) or "No"
- **Cell coloring** - Blue = mentioned more than expected, Orange = mentioned less than expected

Ordinal testing (numeric breakdowns):

- If the chosen breakdown is numeric-like (≥95% of non-missing values parse as numbers), an extra column **Ordinal Sig.** appears.
- It uses the Mantel linear-by-linear association (Cochran–Armitage trend) with ranks 1..k and the same 2×k totals as chi-squared.
- The existing **Significant** column (chi-squared) remains; you can compare both.
- The threshold from Show Differences applies to both tests.

in the factors table when factor-show-differences is on, we calculate chi-sq. but if over 95% of non-missing values in the column selected in #factor-custom-column-input can be interpreted as numeric, we should use an ordinal test instead, or apply an ordinal correction to make the chisq test more powerful

Developer note: Percent mode divides each factor's cell by the group total for that breakdown column. State keys: `factorDisplayMode`, `significanceThreshold`.

Links Panel

🔗 **What you can do here:** View and manage all your causal links in a spreadsheet-like table. You can sort, filter, and edit individual links, view the quotes like a printed page, or export your data to Excel. Each row shows one causal relationship with its source quote and any additional details you've added. Great for detailed review and bulk editing of your causal map data.

Links Table#

Links Table Features:

- Standard column filters, sorting , and pagination
- Sentiment column with numeric values (-1 to 1) and blue/white/red conditional formatting (blue = higher, red = lower, white = mid-range relative to the current view)
- **Citation Count** – total number of links in each bundle (cause >> effect pair), with muted green → white conditional formatting (darker green = more links in that bundle relative to the current view)
- **Source Count** – number of different sources contributing links to each bundle, with the same muted green → white conditional formatting
- Checkbox selection for bulk operations
- Edit functionality opens causal overlay for link modification
- Action button to open coding in the Sources pane and scroll to the corresponding highlight
- Clear Table Filters option
- 💡 Tip: For label changes, prefer Factors Search/replace when working on labels across bundles.

Link Editing:

- Single link click opens editor popup
- Multiple link selection opens chooser interface
- Consistent with coding panel behavior

Links Utilities#

- Download as Excel
- Take a screenshot and copy it to clipboard
- Clear any filters at the top of the table columns
- Bulk delete any selected rows in the table

Row Grouping and Print View#

- **Group by selector** - Choose one or more columns to group rows by values. This applies both the the links table and Print View.

Useful Columns:

- **Bundle** - Shows "cause >> effect" pairs
- **Original Bundle** - If you have used filters which transform the links, like Zoom or Soft Recode, use this to also view the original causes and effects

Print View

The purpose Print View is to make it easy to explore and read actual quotes from the currently filtered links. What it does is show, instead of the contents of the Tabulator table, a printed version of the same information, leaving the table headers and filters in place. The toggle switches between table contents view and print view.

This view prints out the quotes from each row in the table, grouped by the Group By columns formatted as nested headings, and we suppress repeated headings until they change.

We also reveal two more toggles:

- Show Details: Print the values of all the extra columns such as tags and any Custom Columns
- Context: for each quote we add an additional three sentences at each side, highlighting the actual quotes.

You can manually sort the texts using to the sorting widgets in the tabulator headers, as far as allowed by the nested headers.

Search/replace#

This works exactly the same as search/replace in the factors table, except that it works on the Cause label and/or the Effect label.

Near the top is a row containing a search box. If you type something into it,

- a replace box and a Replace button also appear.
- the table is filtered to show only matching rows.

The search is **case sensitive**.

You can then alter what you see in the Replace box:

- in the label columns in the table, you see a preview of what the affected rows text so the would look like;
- if you delete all the replace text so it is empty, the preview shows the effect of deleting the search text from each label.

Then when you are satisfied, check all the checkboxes where you want to update the labels as shown. If you want, select all rows using the checkbox at the top of the column. Note, if there are more hits than fit on this page of your table, you'll want to either treat each page separately or increase the page size with the Page Size selector.

Finally, hit the Replace button to actually update the labels as shown in the rows you selected. As you'd expect, this search/replace only affects the factors for the currently selected links: for example if you have only selected the first three sources, this update will not affect the links in the other sources.

.

Statistics panel

Use this panel to build pivot tables and charts from your project data with a simple drag-and-drop interface (powered by PivotTable.js).

Quick start#

1. Choose a data source: Links, Factors (Work in progress!), or Sources.
2. Choose the stage: **After Analysis Pipeline** (matches your filters elsewhere) or **Before Analysis Pipeline** (raw data).
3. Pick a display: Table, Heatmap, or a Plotly chart.

Arrange fields (drag and drop)#

- Drag field chips from the left list into the drop zones:
- **Rows**: categories listed down the left.
- **Columns**: categories across the top.
- **Values / Measure**: the numeric field used for calculations (when needed).
- Reorder by dragging within a zone; remove by dragging out or clicking ×.

Choose the calculation ("Summarize by")#

- Use the aggregator dropdown to select how numbers are calculated:
- **Count** (default) – how many rows fall into each cell.
- **Sum of / Average of / Min / Max** – pick a numeric field (e.g. `ai_confidence`, text length, counts).
- **Unique Count** – how many distinct values of a field occur.
- Example: to average AI confidence by region and sentiment 1) Put `custom_-Region` in Rows and `sentiment` in Columns 2) Set aggregator to **Average of** and choose `ai_confidence` as the field

Filter or exclude values#

- Every field chip has a filter menu. Click it to include/exclude values.
- Use the search box, then tick/untick items. "Select All" toggles everything.
- Filters apply immediately and are remembered in the URL so you can share the view.

Sort and tidy#

- Click any row/column header to sort by label or by totals.
- Subtotals and grand totals are included automatically in table views.
- Remove empty columns quickly by excluding the blank value in that field's filter.

Heatmaps and charts#

- Switch the renderer to **Heatmap**, **Row Heatmap**, or **Col Heatmap** for quick intensity views (colors match the app palette).
- Choose Plotly renderers for interactive **Bar**, **Line**, **Scatter**, **Stacked Bar**, **Area**, or **Multiple Pie** charts.

Export and sharing[#](#)

- Use the toolbar to **Copy to Clipboard** or **Download XLSX** of the current table.
- The current configuration is saved in the URL automatically; bookmark or share it with collaborators.

Notes on the datasets[#](#)

- **Links**: every causal link plus metadata; includes AI fields (e.g. confidence) and reserved columns like `original_cause`, `original_effect`.
- **Factors**: unique factors with frequency, source count, citations, and `original_label` (ALL underlying original labels for the displayed factor, concatenated with line breaks, derived from the current stage's links like `original_cause/effect`).
- **Sources**: document metadata and flattened custom fields (`custom_*`).

💡 Tip: For results that match other panels, use **After Analysis Pipeline**.

Logs panel

The Logs tab shows a concise history of what has happened in your projects – for example, when links or sources were edited or deleted, when projects were created, or when AI processing ran.

- Use the **scope buttons** (This project / All projects) to switch between seeing activity only for the current project or across all projects you have access to.
- Use the **filters above the table** to narrow by user, project name, action type, date, or text in the details column.
- Non-admin users see only activity for projects they own or can edit; admins can see activity across all projects.
- Navigation and other UI clicks are logged only very briefly and are hidden from non-admin users.

Report Builder#

The Report Builder lets you

- manage your bookmarks
- delete
- copy links
- search and load bookmarks
- create professional reports
  - combining multiple bookmarked views with custom descriptions and variant filtering.
  - with programmatic production of multiple variants e.g. maps with the same filters but for different countries

Requires Pro subscription

📊 **What you can do here:** Build multi-slide reports from your bookmarks. Add markdown descriptions, filter by custom source columns (e.g., country, region), reorder slides, and export everything as formatted HTML ready to paste into Word or Google Docs. Perfect for creating stakeholder reports with multiple views of your data.
Key Features:

- **Drag-and-drop reordering** - Arrange slides in any order
- **Markdown descriptions** - Add rich text titles and explanations with heading styles, lists, and formatting
- **Variant generation** - Create multiple versions of the same map filtered by custom source columns (e.g., one map per country)
- **Include/Exclude toggles** - Selectively include slides in your export
- **HTML export** - Copy formatted HTML with proper heading styles, clickable bookmark links, and legends
- **Persistent settings** - Your slide order, descriptions, and include/exclude states are saved automatically
- Add a bulk delete button to delete all bookmarks with included toggle =ON

Legends:

- For maps we use the full legend
- For tables we use the relevant parts of the legend, omitting material about link and factor colours, sizes and annotation.

Variants:

- Variants simply add an extra filter on top of the bookmark's existing filters (sources + filter pipeline). We do not change any other filters. For example, if the bookmark has 4 selected sources and you choose village=X and Y, then the X variant uses only those of the 4 sources with village=X, and the Y variant uses only those with village=Y. If a value has no data after the existing filters, no variant is produced.

Where to find the variants: when the accordion section is open, the variants appear as thumbnails below the legend, each preceded by which variable/value it represents (e.g., village: Y). On PDF export and copy to clipboard, first the original bookmark image and its legend are shown, and then each variant with its fullsize image and legend.

How to use:

1. Create bookmarks of your maps and tables
2. Switch to the Report tab (yellow bookmark icon)
3. Click on slide descriptions to edit them (supports markdown: `# Heading`, `## Subheading`, `- List items`)
4. Drag slides to reorder them
5. Use the variant controls to generate multiple versions filtered by custom columns
6. Toggle "Include/Exclude" to control which slides appear in your export
7. Click "Copy HTML" to export all included slides as formatted HTML
8. Paste into Word/Google Docs - headings and links will be preserved

Bookmarks and Slide Decks

Bookmarks#

If you want to save a specific view of the app for later, for example an interesting map or table, you can just copy the URL from your browser and paste it somewhere or send to a colleague.

In this section you can learn about the powerful bookmarking and slide decks features.

Requires Pro subscription

🔖 **What you can do here:** Save and organize your favorite views of your data. Bookmark specific filter combinations, map layouts, or analysis states so you can quickly return to important insights later. Perfect for preparing presentations or saving different analytical perspectives.
From anywhere in the app, find the yellow badge at top right of the screen. Click it! This adds a bookmark to the Bookmarks table.

If there is already a bookmark for this same view, a tick is shown: when a view matches an existing bookmark, show the tick.

When you save a bookmark in the session,

- a small modal, near the badge, opens with a textArea for the Description and Cancel/Save buttons. It is a good idea to provide a description, though it is not required.
- the bookmark is then saved including:
- the screenshot of the currently visible map or table (using the existing handlers for these).
- the current map legend, same as with #copy-legend-btn
- as soon as the screenshot is saved, we open a new small dropdown, headed by the bookmark number, right next to the navbar bookmark badge. This stays open for a few seconds then silently closes but can be opened again on click. This dropdown remains the same, with the same header number, until the next bookmark is recorded. The dropdown has buttons to directly copy to clipboard
- the formatted URL same as #copy-project-link-formatted
- the unformatted URL,
- the screenshot image
- the legend
- **a combination of the bookmark link, the map and the legend**

... and when creating a bookmark on any table tab, when map panel is not showing, remember to also save the screenshot same as for pivot tab #pivot-copy-btn, for links tab #export-links-png-btn , for factors tab #export-factors-png,

The bookmarks table itself is not shown for users below Pro level, i.e. to users only on Free or Private plans.

For Links and Factors tables, bookmarking should record everything: any column reordering and manual column widths, column filtering / sorting...

The Bookmarks Table[#](#)

- Each row is one bookmark.
- Click the checkbox to select rows.
- Click Load to recall that bookmark, restoring tabs and outputs.
- Click Copy Formatted Link to copy the bookmark's URL to the clipboard.
- Double-click on the Description field to edit it.
- Click Delete to remove the bookmark.
- Click the badge at top right when viewing a map or table to add a new bookmark to it.
- When a new current project is loaded from a URL or by changing the project dropdown, the "Project" filter in the table is pre-filled with the name of the current project.

Main Features:

- **Bookmark badge** - Save/remove current URL state from navbar
- **Bookmarks table** - Manage saved views with sorting and filtering
- **Bulk operations** - Select multiple bookmarks for deletion
- **Editable descriptions** - Double-click to edit bookmark names
- **Auto-normalization** - URLs cleaned and standardized for consistency

Bookmarks Table:

- **Actions** - Load URL, Copy link (plain), Copy formatted link (HTML), Edit URL, Delete individual bookmarks
- **Copy buttons** - Both create short `?bookmark=ID` URLs instead of full parameter strings
- **Columns** - Project, Description (editable), User, URL, Created, ID
- **Bulk delete** - Select multiple with checkboxes, delete with "Bulk:" trash button
- **Badge integration** - Shows bookmark ID when current view is bookmarked

URL Editor:

Click the Edit button next to any bookmark to open the URL Editor, which provides a user-friendly interface to understand complex bookmark parameters. Instead of viewing raw query strings, you see a structured breakdown with proper labels and grouping - showing your selected project, sources, active filters (displayed as individual cards), table settings, and map configurations. The editor categorizes parameters into logical groups (Navigation, Data Selection, Filters & Processing, etc.) and displays filter pipeline details with sequential numbering and status indicators. This makes it easy to understand exactly what state each bookmark preserves without needing technical knowledge of URL parameters.

When a bookmark of a map is saved, the legend text for that map is also saved as part of the `slide_content` JSON and is printed small at the bottom-right of the slide in Slideshow view.

Slide Decks[#](#)

Coming soon!

🎟️ **What you can do here:** Turn your bookmarked views of your data into slides with additional text annotations. Organise your slides into decks. Export your slides as a PDF report.
Each project can have one or more decks: collections of bookmarks. Each deck has the same permissions as the project. Each slide is json specifying:

- order (an integer)
- layout (for now just provide two layouts but make it extensible)
- usually a bookmark ID, but may be blank

For now, layouts and themes are hardcoded css snippets. Layouts only deal with position and dimensions, themes only with colours and fonts. Themes and layouts both use the same basic slots or set of components: h1, h2, h3, text1, text2, text3, text4.

in the slide template, remove the subtitle text area. we won't use it any more. Instead, put the Description.

Available layouts:

- Initial slide
- Section slide
- Header and two columns
- Large image on right (2/3 width) and narrower header and text box on the left
- Small image on right (1/3 width) and wider header and text box on the left
- Maximum width/height image and all other components float in front with 90% opacity

In these layouts, the image size is maximised. Text boxes have fixed maximum width and height, and the font size increases virtually up to a sensible maximum to fill the boxes.

The UI is just the existing bookmarks table exposing additional columns for Deck, theme, layout and order and a new action button to open the slide. The slide editor contains the same editable fields plus text fields / textareas for the slots and for bookmark1 (which may be empty), bookmark 2. Plus a toggle to show/hide the bookmark image. Plus a preview of how the slide will look when layout and theme are applied to the components.

A dropdown in the UI lists all decks in the current project plus None. Selecting a deck filters the table to show only bookmarks/slides in that deck. In this mode:

- Newly created bookmarks are added to the selected deck with incremental position (at the end).
- A large sortable Slide Sorter modal lets you reorder slides by dragging (flexbox with 4 slides per row). Each slide has a small vertical widget to duplicate the slide and insert another bookmark/slide ahead of it.

In fullscreen preview mode, small Prev and Next buttons float at mid-left and mid-right, and an Edit button opens the editor for the corresponding slide.

URL state includes a "filter by deck" dropdown. It uses the standard editable selectize to allow creating a new (sanitized, unique) deck name. Bulk actions include "Add selected bookmarks to Current Deck" so you can populate a new deck from existing bookmarks or add new bookmarks directly when a deck is selected. The dropdown lists all decks in the current project plus any newly added names.

Slide previews are slightly debounced while editing text fields in the slide editor.

If the text in any text box (main, sub, text1, text2) exceeds one line, the font size is gradually reduced so a paragraph or two can still fit.

Slides display a very small link to the bookmark (e.g., `?bookmark=nn`) at the bottom-left.

When creating new bookmarks while a slide deck is selected, the new slide uses the same theme and layout as the last slide in that deck.

In the Initial Slide layout, text is placed on a semi-transparent overlay for readability with large image backgrounds.

Styles with dark backgrounds use light fonts. In the "maximum image with overlay text" layout, the overlay uses a dark background, and the Initial Slide overlay also uses a dark background for contrast.

Slide 1 discreetly prints the Causal Map filename, the project's last modified date, and today's date near the bottom.

AI answers panel

Requires an AI subscription

You can open this panel by clicking the corresponding tab lower down on the right-hand side of the app.

🤖 **What you can do here:** Ask questions about your data in plain English and get AI-powered answers. Type questions like "What are the main barriers to education?" and the AI will search through your currently selected sources to provide relevant answers with supporting quotes. Perfect for exploring themes and getting quick insights from large amounts of text.
Main Features:

- **Query input** - Type your questions in plain English
- **Automatic chunking** - Sources split into searchable pieces when needed
- **Similarity slider** - Control search precision (0.1-0.9)
- **Max Chunks slider** - Maximum number of the most relevant chunks to send to the AI
- **Prompt history** - Navigate previous questions with prev/next buttons. See these tips

Search Modes#

AI Answers offers two search modes, automatically optimized based on your data size:

Full Sources Mode#

Searches the complete text of your sources (documents/interviews).

How it works:

1. Type a question about the text of the currently selected sources
2. System automatically chunks sources into searchable segments (if not already done)
3. Searches through document chunks using AI embeddings and semantic similarity
4. Most relevant chunks are sent to AI for analysis
5. AI generates answers with supporting quotes from your sources

Question expansion and HyDE (Hypothetical Document Embeddings)

Rather than just using the user's question to match against chunks, we call genAI as preparation and ask it to produce:

1. 8 question variants: short phrases likely to appear directly in source texts containing answers
2. 8 answer variants: different longer and shorter phrases which could contain possible answers, substantially and linguistically different from one another

We match each of these phrases against the chosen chunks and make a sum of the scores per chunk, to then select the top n according the the max_chunks slider.

So for example if the user asks what is the connection between money and happiness, the AI produces question variants like:

- having money, being joyful
- being wealthy
- being happy
- connection between money and happiness

And answer variants like:

- financial security enables emotional wellbeing
- wealth contributes to life satisfaction
- economic resources support positive mental health outcomes

Best for: Exploratory questions about raw text, finding themes not yet coded, discovering new patterns.

Link Contexts Mode#

Searches only through your coded causal links and their surrounding context (the quote + 3 sentences before/after).

How it works:

1. Gets filtered links from your current filter pipeline (respects Sources dropdown and all Source Groups filters)
2. For each link, extracts the selected quote plus surrounding context
3. Organizes contexts by source, with source metadata (title, custom columns)
4. For ≤500 links: Sends all contexts directly to AI
5. For >500 links: Uses backend semantic search to find most relevant contexts
6. Embeddings generated server-side (via `find-relevant-contexts` edge function)
7. Also uses question expansion (see above)
8. Similarity calculation done server-side using cosine similarity
9. Only relevant context indices returned to frontend
10. No memory/computation overhead in browser
11. AI analyzes contexts showing cause → effect relationships
12. AI uses the cause/effect labels in its narrative (ignoring any original labels if links were recoded)

Context format sent to AI:

```
```

Source Interview with Participant 001

ID: ABC-123 custom_Country: Kenya | custom_Gender: Female | custom_Age: 34

Links from this source:

[ABC-123-1] Lack of resources → Poor school performance Context: "We don't have enough books or supplies. The children struggle because..."

[ABC-123-2] Teacher training → Better outcomes Context: "When teachers receive proper training, we see improvements in..." ```

Best for: Questions about causal relationships you've already coded, comparing patterns across sources, analyzing specific demographic groups using Source Groups filters.

Key advantages of Link Contexts mode:

- Uses your coded causal structure, not just raw text
- Respects all your filters (Sources dropdown + Source Groups)
- Includes source metadata in AI context (country, demographics, etc.)
- More focused and structured than full text search
- Automatically scales to large datasets (>500 links) using backend semantic search

etc etc

Settings Panel

⚙️ **What you can do here:**Enable live collaboration. Other settings coming soon.

Live collaboration#

Watch as your colleagues make changes to the causal map (and links and factors tables) in real time!

Account Panel

👤 **What you can do here:** View and manage your personal account settings. Change your password, update your project information, and control your privacy and security settings. This is also where you can export your data or delete your account if needed.

Making projects private requires a Private subscription

User account management and project settings.

Account Features:

- project information and settings
- Password and security management
- Account deletion and data export
- Subscription and billing information

Subscriptions#

Subscriptions List#

Users without a subscription are either:

- anonymous (not logged in) (this is disabled at present)
- free (logged in)

Subscriptions (via LemonSqueezy) are available in the Account panel.

The subscriptions list uses one row per type (private, pro, team) with seat-count, square radio buttons for monthly/annual and Manual/AI, and a live-updating price (from a JSON price file). Each row includes a text description.

There are three dimensions to the subscriptions,

Manual vs AI,

Type:

- private
- pro
- team

Monthly vs Annual.

User can purchase multiples of one or more subs to distribute to colleagues.

Responses Panel

Work in Progress!

Requires an AI subscription

🤖 **What you can do here:** Review all your AI interactions and usage. See a complete log of AI requests, responses, costs, and performance metrics. Useful for tracking your AI usage, reviewing past queries, and understanding costs. Great for administrators and power users.

Main Features:

- **AI interaction log -** Complete history of all AI requests and responses
- **Cost tracking -** View dollar costs based on token usage and model pricing
- **Performance metrics -** Response times and success rates
- **Sortable table -** Filter and sort by timestamp, model, cost, etc.

Help System

❓ **What you can do here:** Get instant help and guidance while using CausalMap. Search for specific topics, browse documentation by section, or click the question mark icons throughout the app for context-sensitive help. The help system adapts to what you're currently doing. Also view the entire help contents as a separate Guide.

Main Features:

- **Help drawer**: Right-side drawer with collapsible sections
- **Smart search box** at the top of the help drawer. If you enter multiple words, either word matches. To match exact phrases, use quotes.
- **Context-sensitive help**: Click the blue buttons within the app next to open the help drawer.
- **App hints**: When you click a section in the help drawer which refers to part of the app, that element flashes briefly.
- **Standalone help**: The same help information is available as a section within our Guide to causal mapping.

Mobile view

Implemented: Mobile offcanvas for right pane#

- A mobile-only two-button toggle (Left side / Right side) sits next to the brand.
- Tapping Right side opens the offcanvas that contains the right pane (tabs and content).
- Tapping Left side closes it. The active side shows a tick.
- The offcanvas shows a "Back to left side" button at the top to close it.
- On desktop (≥ lg), the right pane behaves exactly as before; offcanvas visuals are disabled via CSS.
- A mobile-only Menu consolidates: About, Guide, Help, Bookmark, and Account/Login/Logout.

FAQ - frequently asked questions

Tips on coding#

First of all, there's nothing to worry about, it's fun!

The versioning / backups feature means you can always go back to any version of your file at any earlier time point.

Also, bulk editing of factor labels makes it easy to rapidly change one or many links or factors. And you can do it either globally, i.e. changing one factor everywhere in the file, or you can do it for particular sources or specific kinds of links by using filters.

Usually, don't bother coding the same link more than once for the same source, unless they bring up distinctively different evidence each time.

- It's okay not to code a source at all. If there's nothing in it, or if people are just making vague and general sources.
- You'll find you're constantly shifting between sometimes creating new factors, and then going back and reviewing them and merging them and organising them using the bulk editing of factor labels.
- Don't forget you can combine two or more factors into one using the bulk editing of factor labels .
- Don't forget when you want to search rapidly through already coded links through all of the sources, you can click on the rows in the bulk editing of factor labels to go back to the relevant sources directly.
- Occasionally, a source will make a comment about something which is worth coding, even though there isn't actually a causal link. For example, they might make general comments about some outcome without saying what causes it. In this case just use plain coding. (But if you find you are doing this a lot, you might need to rethink your research design.)
- If you are using hierarchical/nested coding (and you probably should) don't forget you can see the whole map zoomed out to the top level: just press the appropriate button in the Filters panel.

Do I have the latest version?#

Click the About dropdown at the left of the navigation bar at the top of the app. There you can read the version number. The app should silently update itself when a new version is available.

How can I adjust my links or sources data in bulk (round-tripping)?#

As an alternative to editing your links and sources tables manually in the app, you can do what we call round-tripping: download your file, tweak this Excel file manually (e.g. by adding additional columns to the sources tab) and upload it again into a new Causal Map project. Like this you always have to create a new project, which helps you not get mixed up with which version is which.

How can I deal with closed questions like in QuIP?#

In Causal Map 4 there is no special treatment of QuIP-style closed questions. What you can do is this:

1) if you want to be able to see the closed question answers while coding, include the answers to the closed questions simply as part of the text of the interview with an appropriate question number. None of this has any meaning to the app, but it might be useful to have for coding. 2) If you want to also use the closed question scores for further analysis, e.g. to make a map of all the interviews which answered an average of better than 0 to some question, then just add a custom column for each question and add the average scores for each question into each column. Then you can apply these values as filters, see here.

What does the 'recycle weakest magnets' slider do?#

The slider temporarily removes the N weakest magnets from your list and reassigns their raw labels to the stronger magnets.

For example, if you created 50 magnets but after filtering you only have 5 factors showing with 9% coverage, those 45 weak magnets might be taking evidence away from your main ones. Try moving the slider to 40 to recycle those weakest magnets - this gives their evidence a chance to match with the stronger magnets instead (using the same similarity criterion).

This is most useful when you have lots of fiddly magnets that nibble away at your main ones but then disappear without trace. Note that the slider is a bit unpredictable if you have intervening filters.