

CHAPTER

# Causal Mapping as QDA

## CONTENTS

[Intro](#)

[Causal mapping is easy to automate transparently, so is a great fit for scaling with AI](#)

[Causal Mapping outputs not just codes but a model you can query to answer useful questions](#)

[Causal mapping is easy to automate transparently, so is a great fit for scaling with AI](#)

[Brief review of S Friese -- Conversational Analysis to the Power of AI](#)

## Intro

---

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

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

**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 it. 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  $B \rightarrow C$  is not the same as  $C \rightarrow B$ ). 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 by "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 Frieze](#) and [Christina Silver, PhD](#) for suggesting additions here. ↩

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

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, Frieze 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. Frieze 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, Frieze 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>.