Guide to Causal Mapping

Steve Powell & Causal Map Ltd

2021-09-14

Contents

Chapter 1

Overview

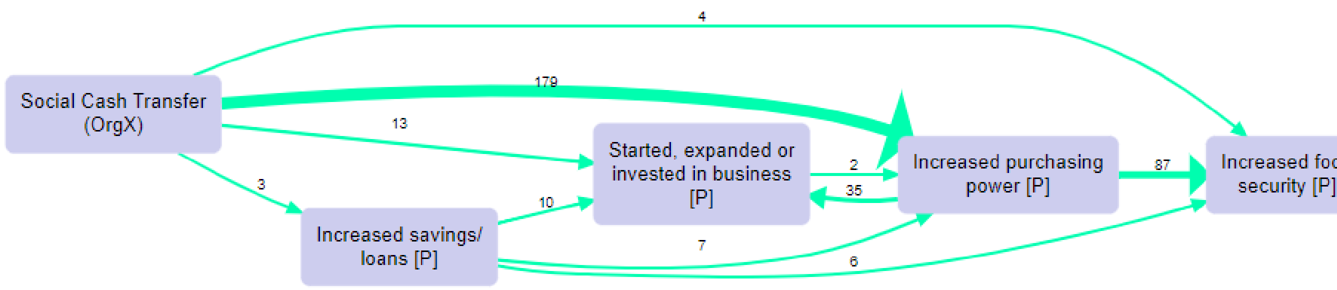


Figure 1.1: img

This Guide is for you if:

- you want to find out more about causal mapping as an approach (look for this symbol)
- you want to find out more about the Causal Map app (look for this symbol)
- you are on, or have completed, the QuIP Lead Evaluator training

---

This Guide accompanies Causal Map 2, the new version of Causal Map, and is valid from September 2021.

The old version of this Guide covers the old version of Causal Map (the one hosted at go.causalmap.app).

Part I

The new Causal Map platform

Chapter 2

Welcome to Causal Map 2!

Read this especially if you have been using the legacy version of Causal Map (CM1) at go.causalmap.app

We will continue to maintain the legacy version until all users have happily transitioned to the new platform, Causal Map 2 (CM2).

You may also have been using the Causal Map **Viewer** at (https://causalmap.shinyapps.io/CausalMapViewer/). This has been superseded by Causal Map 2.

All your files (and other files you had access to) from CM1 are accessible from within CM2. You will find them in the file chooser dropdown menu as before, but preceded by `cm1/...`

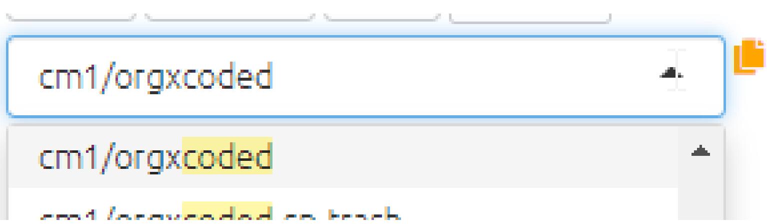


Figure 2.1: image-20210914103154772

You can view and analyse these files in CM2 without needing to do anything.

You can't make changes to them from within CM2.

If you are still coding a file in CM1, changes you make will be visible from CM2 as soon as you refresh the page.

If you want, you can save a CM1 file in the new CM2 format so that you can continue to edit it in CM2 (as long as you have *edit* or c*opy*, not just *view*, authorisation to the original file).

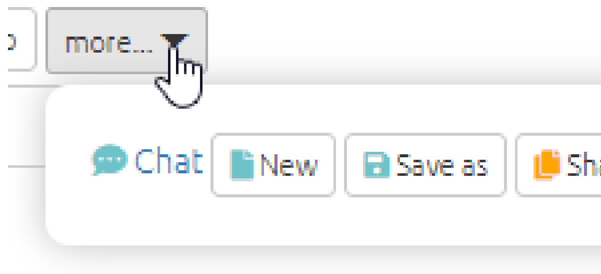


Figure 2.2: image-20210914103228048

This creates a new file which is no longer connected to the original file.

The following sections will take you on a quick tour of the app.

Chapter 3

Quick tour of the app

3.1 The left-hand side

3.1.1 Getting data into the app

- Upload a map in the form of an Excel file (.xls or .xlsx).
- in the file chooser, select any files to which you have access, which include an anonymised dataset from a QuIP study
- follow a direct link sent to you by someone else to view a file with pre-defined filters applied and which you can then explore for yourself.
- view maps created in the legacy version of Causal Map.

3.1.2 Commands and buttons to apply filters to your map

The left-hand side of the app really only contains the text in the Advanced Editor (which you can view if you want, but close the window if it scares you).

The text window uses a simple syntax for filtering and manipulating the maps and tables.

Nearly all the other buttons on the left-hand side are just ways to manipulate this text. Each line in this window is a filter which manipulates the existing map in some way. The lines in the windows are applied successively to the original map to produce the final map which is then displayed.

The way these buttons work is still evolving, we are open to your suggestions!

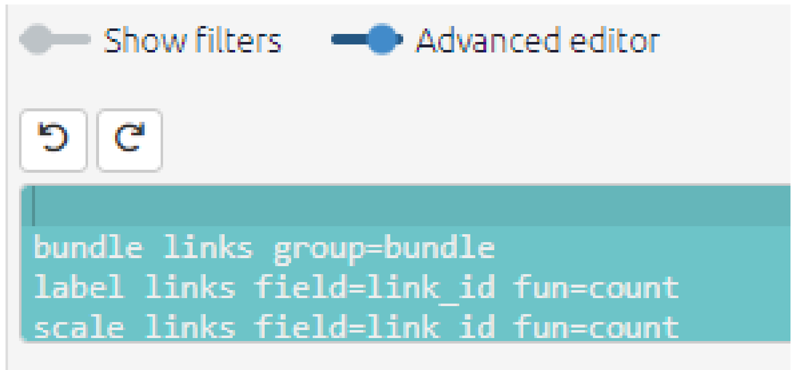
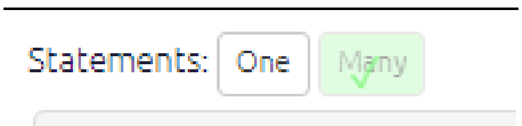


Figure 3.1: image-20210914103354673

3.1.3 Do some coding (view one statement) or do some analysis (view many statements)?



These buttons add the correct filters to either view just one statement or many statements.

When you are coding, you will want to view just one statement, and when you want to explore and analyse the entire causal map or sections of it, you will want to view many statements.

(Even when you are viewing one statement, it is still possible to apply filters for example if the map associated with this one statement is quite large.)

3.1.4 Top row

Most people like to hide these filters when they are coding, so they hide the filters using the toggle.

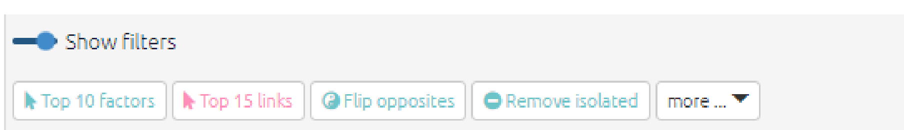


Figure 3.2: image-20210914103505559

But for analysing your map, you will want to open the filters panel.

3.1.4.1 Shortcut buttons

These buttons offer quick ways to add filters without going through a dialog panel.

3.1.5 The filter buttons

are in three sections: analysis, conditional formats and simple formats. See the section on filters.

3.2 The right hand side

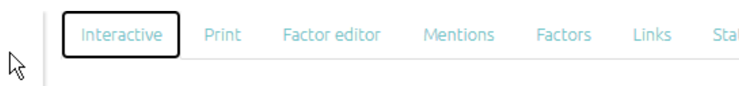


Figure 3.3: image-20210914103654314

3.2.1 Interactive View

An interactive version of the map in which the elements can be moved around and also the upstream and downstream factors are highlighted when the user mouses over them.

- Drag the factors around
- Copy a map as an image by right-clicking on it.
- Save a PNG image by pressing the button at bottom right.
- Hover over factors to highlight the connected links and factors.
- Hover over factors to display basic information about them and delete them from the entire map.
- Hover over links to see the associated quote and other information.
- Click on links to edit or delete them them.

3.2.2 Print View

A print-quality version of the map with advanced layout.

3.2.3 Factor editor

This is the same as the old Multi-edit Panel

3.2.4 Tables

Several tables showing lists of factors, links and various metrics and which can be filtered and reorganised in marvellous ways.

Chapter 4

Is Causal Map for me?

So, you've taken a look at the **features of the app**, and you're getting excited about creating maps – but is your project right for it?

Use Causal Map if you:

- have a relatively large amount of narrative data (enough to provide at least 20-30 causal links)
- need help to organise a large number of links and summarise them into an overview or synthesis
- have information from more than one source (for example different respondents, different documents, or different places in one document) and the information about the source is important to you: they aren't all interchangeable
- are interested in possible differences between the sources and groups of sources – and/or you don't necessarily have a preconceived idea of the contents or boundaries of the map.
- want to capture what your sources actually say, systematically and transparently

Causal Map map is not suitable if you:

- 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 sketch out a plan (e.g. Theory of Change or similar) without much reference to the different sources underpinning each link

4.1 Prerequisites for using the app

The app is tested on Chrome and Firefox on Mac and Windows, and Safari on Mac, and the new Edge on Windows, and it should work fine on Chrome, Chromium and Firefox on Linux.

You need a screen with a large resolution, preferably HD or better. It won't work well on tablet or phone. You will probably want to zoom out to about 90% (in Chrome, press Ctrl – once or twice) to make sure everything fits on the screen. You may also like to view the app "full screen" – how to do this depends on your browser and operating system; on Chrome on Windows, press F11.

As Causal Map is a web application, you will also need a reasonably reliable internet connection.

Chapter 5

Features of Causal Map

5.1 Main features of Causal Map

- Code causal claims in the form A –> B
- Code multiple claims at once
- Code additional claims as a continuation of previous claim (“chaining”)
- Import data (and meta-data about your sources) in different formats
- Add memos and hashtags
- See a summary of all statements from a specific source
- View additional information about each statement (e.g. question and respondent characteristics)
- Create simple or hierarchical factor labels
- Edit many factors and links at once in a powerful bulk editor
- Filter the global map by: Current statement; Factor label; Link (quote, hashtag, memo etc); Statement (text, additional data e.g. respondent characteristics)
- Trace paths from one set of factors to another
- Use the interactive viewer to edit your map
- View and export print-quality maps
- View all quotes for a specific view, and get a smart summary
- Create detailed tables: Factors; Sample; Questions; Closed questions
- Calculate **Robustness of Argument** for causal paths between sets of factors
- Download/export your data in different formats
- Share files with others for viewing or editing
- Code additional properties of links like strength and certainty, or mark pairs of factors which are opposite to one another

5.2 Additional new features in Causal Map 2

- Anyone can use it, and it is free for small projects
- Make your findings more visible with conditional formatting
- Bundle links by stakeholder, by district, by question …
- Really nice infinite versioning, so you can restore any single change you ever made to a file from any timepoint. This isn't just backups, it is every single change.
- Much easier to import and manage your data
- You can import a standard-format Excel file into cm2, and download it again, and edit it, and upload it again ("round-tripping"). You can also include any additional fields aka columns you want and they will be imported and available for searching, conditional formatting etc.
- Build a map by importing any combination of tables or just some of them - factors, links, sources, questions & statements, in any order at any time and it should all just work
- You can edit the tables directly

/- The URL in your browser always points to the map and filters you are using right now, so a) if you refresh your browser or go away for a while, you can always come back to where you left off, and b) you can share or save the link for later/

- A Google Chat room for support
- Bundle factors
- Bundle links by field
- Clickable legends and ordinary-language explanations of filters
- All outputs recreatable from the filters
- merging maps:
  - You can use the merge_map filter to temporarily merge other files into the current file if you wish. You can share a link to that merge and revisit it. Viewing a merge of file A and file B will take longer, so you will probably want to save the merged file as a new file.
  - The tables (factors, links etc) have a new field called factor_map_id etc, which you can use to visualise the merge e.g. by presenting links in a different colour according to their source.

- manage all your files at once with a comprehensive file manager from which you can restore, lock, delete, archive and share any file (not just the current file).
- log in to the same or different maps from the same or different accounts, including from multiple browser tabs.
- collaborate on a file in almost real time.
- view, filter, sort etc and directly edit all the tables - factors, links, sources, questions & statements - making up a map. These tables provide all the existing tables functionality plus more. So you will be able to do things like, say, find all the links with a certain hashtag and delete them.
- delete factors and edit and delete links directly from the interactive map
- create a new blank file
- clone an existing file under a new name, containing all the current factors or only including the factors currently visible if you wish.
- If border colours are not set, or all the same, borders are superfluous so the border size is set to zero in Print view
- New colour scales which are colour-blind friendly and make sense when printed on black & white printers
- Like Print View, if Interactive View is given a very big map with lots of links, it now bundles up the links so that it loads faster (you can still view a random quote from the bundle when you hover over a link)
- When you edit a link, you are able to replace it with more than one link if you wish, e.g. you can edit a link A→B and replace it with A→B and A→C by typing both B and C in the consequence factor box. Actually, when you edit a link, CM2 loads up the details from the old link into the coding panel, and if you press , the old link will be deleted and the new one added. You can cancel this process, as in cm1
- Less than, greater than: The find factors, find links, find statements filters allow comparisons like greater than, so it is possible to do things like find statements with id greater than 100.

Chapter 6

Signing up and signing in at Causal Map

Start your journey at https://causalmap.shinyapps.io/CausalMap2/.

The first time you use the app you will see a screen like this:

You can sign in with a Google account.

If you have a paid subscription, however you sign in, make sure the associated email is the same as the one for which you have a subscription.

If someone has shared files with you, however you sign in, make sure the associated email is the one the files were shared to.

Chapter 7

Coding: creating factors and links in the app

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. This section is about how to create factors and their labels.

In Causal Map, a factor *is* its label. Once you create a label, there is nothing else to add.

7.1 Create and edit links in the app

To code a causal link, - With your mouse, highlight a piece of text within the statement which makes a causal claim.

- Watch how that passage is copied for you into the "Quote" window below. (Usually you don't need to think about his 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 or you will get a warning.) - 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 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.
- 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, colour-coded with the quote which is now highlighted on the left. If you mouse over the highlighted quote, the link in the map is activated.

To edit an existing link, - Click on it in the Interactive Map - Make any adjustments in the left panel, e.g. change the influence factor (in the first box), and/or the consequence factor (in the second box). You can change the quoted text just by re-highlighting the correct passage in the statement panel above in the same way that you made the original highlight.

- Press the green Save button.

NOTE: your factor names should not contain semicolons **;**. Semicolons are special, they are used for hierarchical coding.

After beginning to create links between factors, one will notice 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.

7.2 Using memos and hashtags

Qualitative coding usually involves making notes and memos, and you can do this in Causal Map too.

Hashtags are available as a special kind of memo when coding a link: you can use them to provide any kind of additional information:

- where a link is only relevant in a particular context
- where a link is only a hypothesis
- where a link is only projected for the future
- to tag links you want to come back and review

- to add other qualifying information like "source seems unsure."

As usual in Causal Map, you can apply one or more hashtags, and you can either select existing hashtags or create new ones on the fly.

Later, you can filter the map to show only links containing specific hashtags (or parts of hashtags), and also for links which do *not* contain specific hashtags or parts of hashtags.

Chapter 8

All the filters

8.1 Analysis

Top row: Active analysis buttons with explanation. Click to edit, move, delete

Bottom row: All analysis buttons. Click to activate another instance.

These filters apply powerful filters which change the structure of the unfiltered map. Each filter can be applied more than once and they can (to some extent) be moved up and down.

For example suppose you want to show only links from statements in which the word "women" appears. So you apply a `find statements` filter, you click the button in the bottom row. A dialog appears,

and a new button with a human-readable explanation appears in the top row. The button in the bottom row gets a green tick to show it is in use, but you can click it again to apply another, different statements filter, for example to find links from all statements in which "women" appears and which were mentioned by *men*.

8.1.1 Find factors

8.1.2 Searching and filtering factors

Type to select factors which you want to find.

- Select one or more pre-existing factors. Press `enter` to accept
- And/or just type fragments of text which might match several factors; type a comma to complete

The map changes:

- All factors matched by the search are, by default, *highlighted*.
- Only the factors matched by the search and those upstream and downstream (left and right) of them are shown, as many steps as set by the `Upstream steps` and `Downstream steps` sliders.

Factors are only matched if the search term uses the same case, e.g. "Farming" will not match "farming" or "FARMING".

For a more visual representation: The following examples demonstrate how applying different upstream and downstream step filters affect the causal pathways presented in the map.

Searching for factor 'B' The default: one step upstream – one step downstream

One step upstream – zero steps downstream

One step upstream – two steps downstream

Why would I use this filter? This analysis function is particularly useful for searching for important factors such as intervention activities or intended programme outcomes. This filter is best applied once you know which factors are of importance or interest.

8.1.3 Find statements, find links

These filters enable you to filter the causal map to factor in additional data or respondent/source characteristics.

How do I use this filter?

Select an additional data field from the dropdown menu Find Statements. The options available will vary depend on the data you have collected, a typical additional data field might be respondent ID, age, or sex, for example.

Once you've selected your field, then type/select the specific criteria you want to filter by (i.e. individual respondent, age category, male/female) into the Filter by… search box. Use the arrows to the right of the search box to toggle between the different options. You can add more than one value (e.g. several age groups or several respondents).

8.1.4 Filtering statements

- The second box is prepopulated with all the values for that field, for example education levels. You can select one or more values (e.g. several age groups or several respondents).

- **TIP**: As well as selecting pre-existing values for your search, you can also just type fragments of text which might match several values. So for example, if you are searching question numbers, and you have the questions e1, e3 and e5, instead of selecting all of them you can just type "e": all the values where this text appears are included in the filter. But be careful that this does not match other values which you did not intend.
- **TIP**: You can use multiple filters all at the same time.

8.1.5 Remove brackets

This filter hides any tags or other words written between different kinds of brackets. So instead of this:

you see this:

8.1.6 Trace paths / trace robustness

Why would I use this filter?

This analysis function is a powerful tool which enables you to view full causal pathways and to interrogate the relationships between specific causal factors.

Using our OrgX example, there is a clear causal path from the 'Social Cash Transfer' to 'Increased Purchasing Power.' The complexities of this causal path are best seen and shown by using path tracing, as it simplifies the map and highlights the intervening factors between the two factors.

Also, path tracing is the prerequisite for calculating Robustness.

A path length of 1 will only show the one step in the causal chain from/to your chosen factor, i.e. A B. A path length of 2 will also show the next step in the causal chain (if there is one!), i.e. A B C.

8.1.7 Bundle factors

8.1.8 Combine opposites

8.2 Conditional formats

Conditional formats calculate and visualise information in your map.

The buttons in this section apply conditional formats to the map after it has (optionally) been transformed in the analysis section. Each filter can only be applied once, so when you click an inactive button in the bottom row it becomes active and moves to the top row, and when you delete it from the top row it appears again in the bottom row.

8.2.1 Colours

- if you specify a fixed colour, everything will be set to that colour.
- otherwise, if the field is numerical, the values of that field will be assigned to a colour gradient with the given low, medium and high colours. If you specify white or grey as the low point, the mid point will be ignored.
- otherwise, if the field is not numerical, the values of that field will be assigned random colours up to a maximum of eight.

8.3 Simple formats

There is a filter for "cluster factors" which groups your factors into arbitrary groups.

So if you type Intervention, Firms and Impact, three boxes appear just grouping together factors which begin with those phrases, and the factors are regrouped to fit into the boxes.

The setting "layout" makes a huge difference. Usually you will want "dot" layout which lays out your map in a left-to-right direction.

However, other layouts can be useful, for example when you are looking just at the ego network for a single factor, i.e. just the factors immediately adjacent to it, searching just for that one factor and one step up or down, the "circo" layout can be very helpful:

Chapter 9

All the tables

From simply creating a table showing the total links to and from each factor to a visual representation of each respondent's response to a closed question. The possibilities are large - and increase with the amount of additional data your project has.

9.1 Features common to all the tables

Numerical tables are presented as *heatmap tables*. The higher the number, the darker the colour of the cell will be.

9.1.1 Presets

If you want to keep things simple, try creating a table using one of the quick presets:

… and the filters relevant to each table are also accessible from within that table:

9.1.2 Main controls

Each table has a set of controls, which are almost the same across all the tables.

When the `Which version` toggle is set as “Filtered”, the tables respond to any filters you have applied in the left-hand panel of the app, just as the interactive maps do. The table shows data corresponding to the map as it is currently displayed.

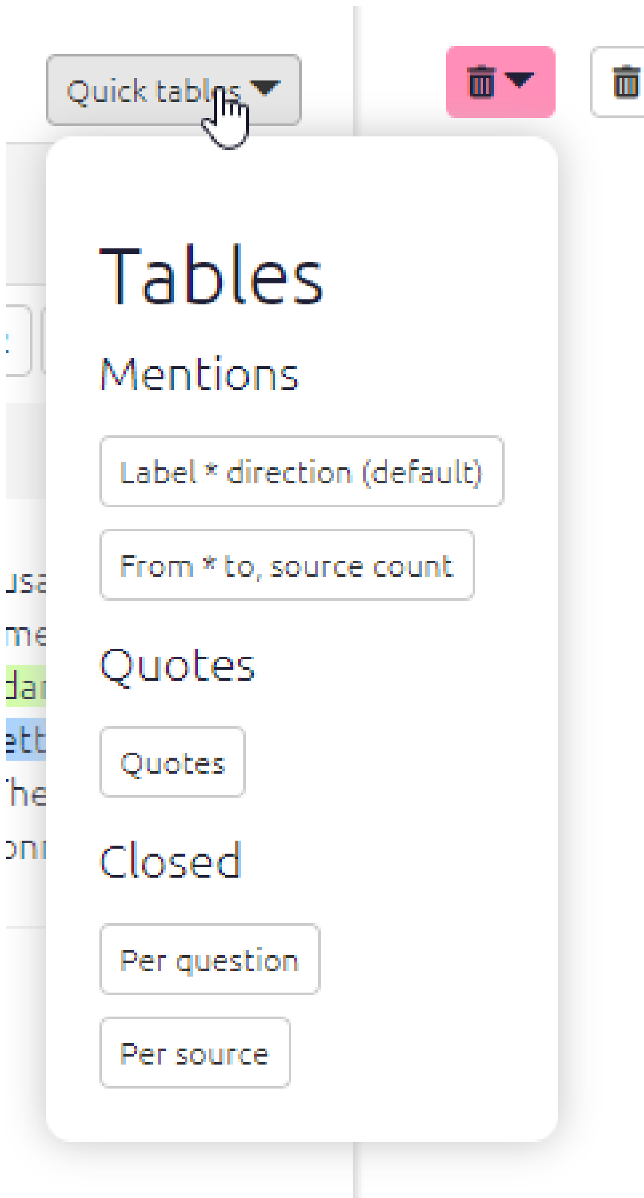


Figure 9.1: image-20210914081424695

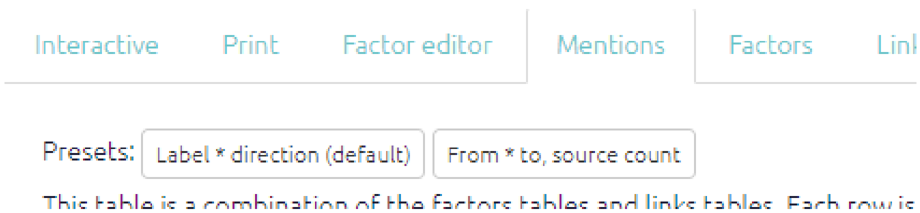


Figure 9.2: image-20210914081506509

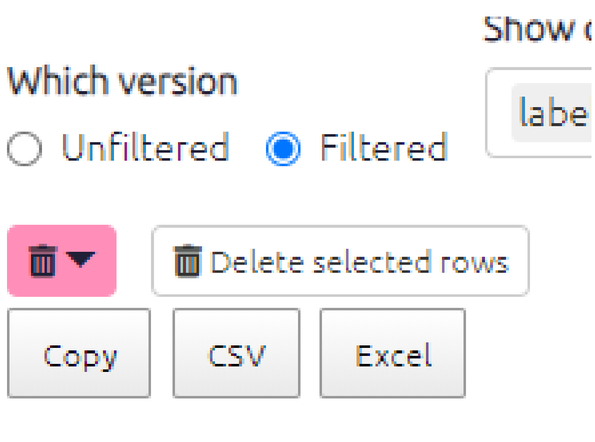


Figure 9.3: image-20210914081641240

If you want to see all the data in one table for the unfiltered map, switch this toggle to Unfiltered.

- You can copy the data from the tables by clicking Copy table to csv or Excel. You can then paste the data in Word or Excel to create your own tables, graphs, or visualisations. You could also screenshot the table if you prefer!

9.1.3 Adding columns, grouping, counting

You can also group the rows in the factor tables to show how the data presented differs between various respondent characteristics such as age, education, and sex. Simply select the desired filter from the Group rows filter.

When you put a field in the `count` box, your table will get an extra final column called `total`:

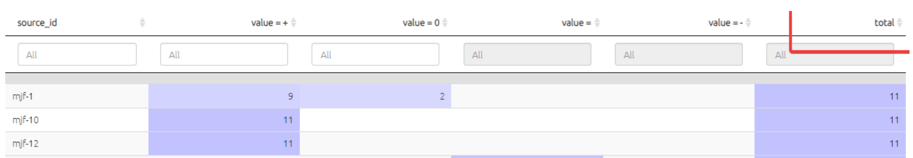


Figure 9.4: image-20210914092140216

9.1.4 Search

You can search / filter the whole table using the box at top-right. And you can search / filter individual columns using the individual boxes. (These boxes are greyed out if all the values in the field are the same so there is nothing to search.)

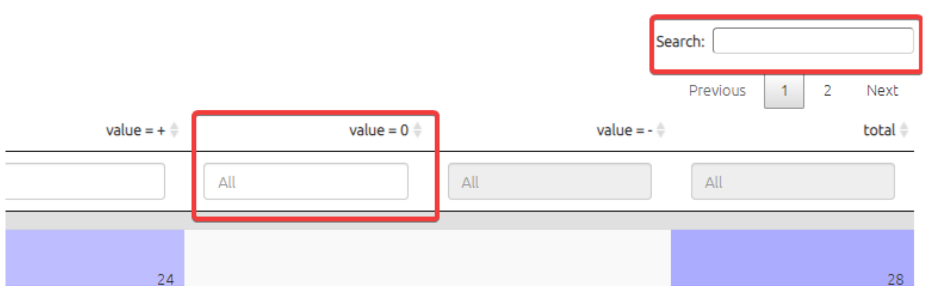


Figure 9.5: image-20210914101220510

You can search a column of numbers by using the slider, or by typing an equivalent range:

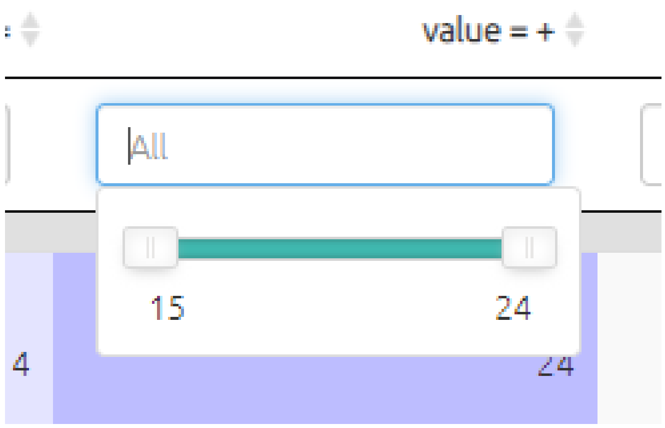


Figure 9.6: image-20210914101522120

… so if you type "15…15" you will search just for the number 15:

9.1.5 Sorting

You can sort the whole table by any column by clicking on the appropriate header:

9.2 The individual tables

9.2.1 Factors

9.2.2 Mentions

What does this table show me?

The factors table presents the factors applied during coding (and which are relevant for the current filters) and the number of times they were reported as an influence factor and/or a consequence factor.

- From = how many times the factor was applied as an **influence** factor, i.e. leading to another factor.

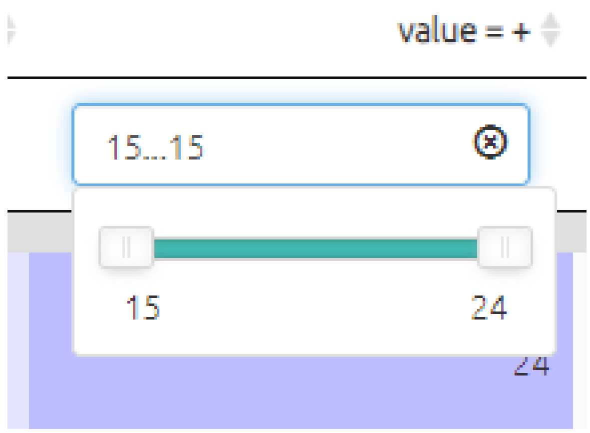


Figure 9.7: image-20210914101553467

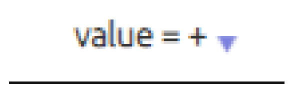


Figure 9.8: image-20210914101749080

- To = how many times the factor was applied as a **consequence** factor, i.e. as a result of another factor.
- The factors are listed with the most frequent first.

Why would I use this table?

This table is merely an overview which can help us to understand which factors are reported most frequently, and whether specific factors were more often cited as an influence or a consequence. To fully understand what the factors mean, they need to be seen in the context of the causal stories they appear in. This table can be useful for initial communication with the commissioner about which factor labels have been created and how often they have been applied, but do exercise caution when presenting this data as it only shows the factor in isolation, whereas QuIP is most interested in the relationships between factors.

9.2.3 Links

9.2.4 Statements

9.2.5 Sources

What does this table show me?

The sample table [without any filters applied] simply lists all of the respondents and any additional data collected about them (sex, age, location, education.) It also shows the interview type (individual or focus group) for each source. When filters are applied, the table will only present the relevant sources/respondents. If you are searching and filtering for a specific factor, this table will update and only show sources who reported that factor.

Why would I use this table?

In the first instance, this table can be useful to check that all the sources have been imported into the app correctly. The table also provides a summary of the respondents which can be useful for presenting respondent demographics in the report, either in the sampling section or as an appendix.## Questions

9.2.6 Questions

What does this table show me?

This table displays the text and ID for every question included in the questionnaire. It may also show any other additional data fields which are the same for each question, e.g. questionnaire subsection or question group.

Why would I use this table?

You may wish to include this table as an appendix in your report or as a reference point for looking up particular questions.

9.2.7 Closed question blocks

There are two tables which provide a summary of the responses to the **closed questions** asked at the end of each QuIP questionnaire domain. The following symbols are used in the tables to represent the direction of change indicated by the response to the closed question:

Symbol	Direction of change	Example responses
0	No change	“No change” “Stayed the same”
+	Positive change	“Better” “Improved” “Increased”
-	Negative change	“Worse” “Decreased”

The closed questions table gives an overview of how each **individual** respondent answered each closed question. The closed question summary table presents the **total respondent counts** for each direction of change. For every closed question you can see how many respondents reported positive, negative, or no change in that domain. You may wish to search and filter the statements to view only the closed question responses from a particular respondent group.

Why would I use this table?

These tables provide a snapshot of the overall trends of change across the domains, so they can be a helpful introduction and “easing in” to the findings - before diving deeper into the causal stories! The closed question responses can also provide interesting insights when compared to the open-ended responses, especially in cases where they might differ.

A full list of closed question responses can be found here

Part II

Intro to Causal Mapping

Chapter 10

Causal Mapping: Definitions

10.1 What is a causal map? What is causal mapping?

A **causal map** is a diagram, a 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, in some sense, C causally influences E. Causal maps are used by many research and practitioner teams around the world in a range of disciplines, who employ a variety of methods to construct and interpret them. While one group of such methods is actually called "causal mapping", there are many similar methods which go by a wide variety of names.

At the same time, the term "**causal mapping**" is often used as a name for a specific kind of *data collection* method, along with suggestions for analysis. (However, it is possible to have a causal *map* without anyone having intentionally done any causal *mapping*, using information gathered for other purposes.) There are a vast variety of possibilities, with seemingly every author having their own suggestions, from individual interview (Ackermann & Eden, 2004) to reusing open-ended questionnaire questions (Jackson & Trochim, 2002). We distinguish two main kinds of causal mapping:

- Individual respondents are deliberately asked for information about causal links, for example via open questions at the end of a questionnaire or via a series of interviews in which people are directly asked questions of the form "what causes what?" or "what contributed to this event?" For example in QuIP, respondents are asked for causes of changes, and then for causes of the causes, etc. We call this **unmerged** causal mapping because the information from different sources is not yet merged into a single map. When analysing the data, we try to read what the different sources tell us, and bit by bit ("inductively") try to identify the common elements in their narratives, such as "Health" and "Amount of exercise". Different respondents will, of course, not always use exactly the same phrases and it is a really exciting and creative challenge to create and curate this list of causal factors: a special kind of *causal* qualitative data analysis. This is your job as causal mapping analyst. For example, if Mo says "Feeling good about the future is one thing that increases your wellbeing", is this element "feeling good about the future" the same as "confidence in the future" which Sara mentioned? Should we encode them both as the same thing, and if so, what shall we call it? Positive view of future? Does that cover both cases? The Causal Map app is designed for this kind of causal mapping.

- A group of people are deliberately asked about causal links and this information is merged straight away into an overall picture, as a participatory process with the group (Penn & Barbrook-Johnson, 2019). We call this **merged** causal mapping because information from different sources is already combined into one map. The Causal Map app can be useful here too, but it is not its primary use case.

10.2 Summary

Causal maps are used by a wide range of research and practitioner teams around the world in a variety of disciplines, from management science to ecology and programme evaluation, who employ a number of methods to construct and interpret them. While one group of such methods is actually called "causal mapping", there are many similar methods which go by a wide variety of names. The aim of this section is to point out the similarities and highlight some of the differences, in the hope that bringing these various approaches into one big tent could increase mutual learning.

This definition *could* cover diagrams representing causal connections between variables which are measured in a strictly quantitative way and would therefore also include a wide variety of important approaches from Structural Equation Models (Bollen & Long, 1993) to the modern and general approaches to causation in statistics centred around Directed Acyclic Graphs (DAGs) and the work of Judea Pearl (**?**). However the phrase "causal mapping" is usually reserved for qualitative or merely semi-quantitative approaches and we will follow that restriction here, while noting that many ideas from these quantitative approaches have implications for their qualitative cousins (Powell, 2018).

This definition is still very wide, covering as it does applications as diverse as "Theories of Change" in programme evaluation and management, and systems modelling in ecology (Moon et al., 2019). The phrase "causal mapping" goes back at least to (**?**)), based in turn on Kelly's personal construct theory (**?**). The idea of wanting to understand the behaviour of actors in terms of internal 'maps' of the word which they carry around with them goes back further, to Kurt Lewin (Lewin, 1982) and the field theorists. Causal mapping in this sense, used widely in project management and political science, is loosely based on "concept mapping" and "cognitive mapping", and sometimes the three terms are used interchangeably, though the latter two are usually understood to be broader, including maps in which the links between factors are not necessarily causal and are therefore not causal maps.

Literature on the theory and practice of causal mapping includes a few canonical works like Axelrod (**?**). Already by 1990 Anne Huff had edited a book presenting some of the wide variety of concept mapping approaches in use in the U.S. at the time (Huff, 1990), including causal maps and argument maps. However, literature on causal mapping tends to be dispersed between disciplines, such that causal mapping is "invented" once again in one discipline or another every few years.

Chapter 11

Glossary

11.1 Generic terminology for qualitative causal mapping

We collect qualitative data in the form of **text**.

Coding that textual data means **highlighting** pieces of text which express **causal links** (or **causal claims**) between two (or more) **causal factors**. To code a causal link we specify a causal factor at the beginning of the link ("**influence factor**") and a causal factor at the end of the link ("**consequence factor**"). You create labels for the causal factors, e.g. Increased crop yields. Usually you will start to reuse these same factors to create new causal links as you continue coding.

Some causal claims can be coded as just a single link between two factors like this.

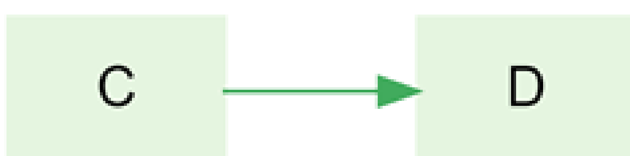


Figure 11.1: image-20210109103241094

In the most basic case, a link means simply that C has some kind of causal influence over D; we could say it makes a **causal contribution** to D, or it **causally influences** D.

Often we will code a causal claim as a chain like this. These chains are particularly important in QuIP, where we try to trace outcomes back to their drivers:

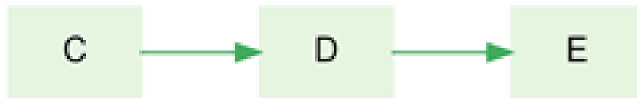
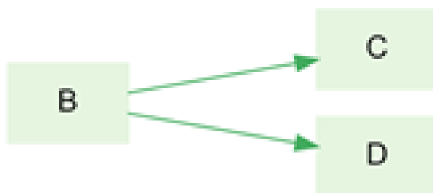


Figure 11.2: image-20210109103251196

Sometimes claims are more complex; in general we can call each claim a **causal map** (or a "**sub-map**" to make clear that it is only part of a larger story).



or

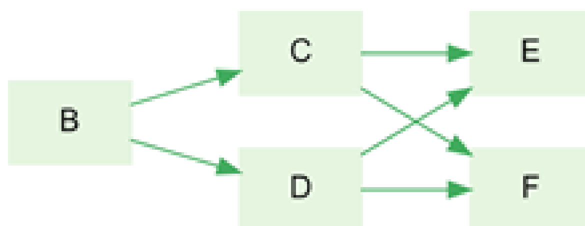


Figure 11.3: image-20210109103305430

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 "less of this" or "this happened" or "this didn't happen": we call these semi-quantitative factors.

These factor names can optionally also include flags: fragments of text, for example, *(Farming).* These should be chosen so they are easy to search for and wouldn't come up in another factor label to mean something different. General terms, such as *(Farming),* can be used to organise the factors into **themes**. Analysts can also use flags to mark factors associated with specific interventions or project inputs, for example the word "*Intervention*", as well as to mark factors which are regarded (usually in conversation with the evaluation commissioner) as desirable outcomes, e.g. by using the word "*Outcome*".

Causal factors can also optionally be **nested** within one another. This means for example that it is possible to view **zoom out** to see just the higher-level factors, with their subsidiary factors 'rolled up' into them.

Additional coding possibilities for advanced users (more about these later in the guide):

Hashtag: you can use these to make any notes about the *links*, e.g. to note that a link mentions a particular topic or type of statement (e.g. #future) or is a link that you want to review later. Hashtags can use an actual hashtag, e.g. *#hashtag*, but don't have to, e.g. *hashtag*.

Memo: as is standard in qualitative social research, as an analyst you can make short memos of any type on each statement, as well as about each causal link, each factor and the project globally.

Strength and **actualisation** of the causal link.

A useful concept is that of a **bundle** of links, i.e. sets of links which are coterminal, i.e. between the same influence and consequence factors, usually mentioned in different statements by different sources. These can be shown separately or, with certain caveats, together, with a thicker line and/or marked with a label to show the number of links in the bundle.

Each individual piece of text is called a **statement**, usually this is a sentence, a paragraph or a few paragraphs. Each statement has an ID; a number like 1, 2, 3 etc.

Each statement usually has **additional data** associated with it, e.g. respondent ID number, respondent age, question number. An additional data **field** like "age" usually has various **values**, e.g. 0-20, 21-30.

We call the set of statements (and possibly the set of respondents etc.) which correspond to a particular value of a particular field a group, e.g. the group of women.

Usually, statements come from a set of **sources**; in QuIP and in most other kinds of study these are usually *respondents* – individual people or individual focus groups – but in other contexts the statements come from a set of *documents*. Each source is usually associated with several documents. In the app, you can, for example, search for all the statements made by a particular respondent (in the app, the term "source" is used rather than the term "respondent").

11.2 Formatting in this guide...

We display statements and quotes from them, respondents' actual words, in quote marks:

"Thanks to the financial support, I can now go to school"

We display factor labels with underlines, like this:

Able to go to school

When we display whole coded causal claims which include an arrow, we don't use the underlines:

Financial support   Able to go to school

This first part of the Guide covers all the basic ideas of causal mapping. The second part covers some additional ideas, which correspond to the features enabled for those with “Extra” subscriptions to the Causal Map app.

Bundle:

Any links between any pair of factors can be called the original links; these links can be bundled so they visually look like one. The “simple bundle” is just all of the links bundled into one bundle, but you can also create more than one bundle if you e.g. bundle by source_id. Once you bundle the links, you can point to the bundle and say, look, a bundle. But if someone calls it a link that is ok too, providing we are clear that this link aka bundle may contain several original links.

File: “File” specifically for what we used to call “projects” within the app.

Hybrid format: Where some rows are statements and some are metadata

Chapter 12

Causal mapping for evaluators

> Causal mapping: a way to understand how people think; and perhaps to understand how the world works

12.1 Causal mapping for evaluators

A causal map is a way to organise a heap of claims about causal links between causal factors. Causal maps allow us to ask and answer questions like "what kind of effect will/did tweaking C make to E?", which is one of the central tasks of programme evaluators. We may be asked

- if an intervention had, or could have, any effect on some desired outcome.
- to assess other relevant but unintended consequences of the programme to assess the causal relevance of some intermediate step to some hard-to-measure goal (is this a good step? Are there better steps?)
- whether C was *the* cause of E.
- whether C had *some* causal influence on E.
- what would have happened without some particular assumptions or events taking place
- is this contribution larger than some comparison contribution
- is this contribution more valuable than some comparison contribution
- which are/were the important influences on this outcome

This means that as programme evaluators, we often have to deal with bundles or heaps of claims about causal influences and somehow combine them. For example, we might have some questionnaire data which suggests that an intervention (C) improved teachers' skills in the desired way (D), and we might have some research studies which suggest that this will have a positive impact on student outcomes (E); on the other hand we have an interview with a school director who insists overall that the intervention (C) is useless and did not influence outcomes (E) at all.

Occasionally we may have only one source of information from one method, such as a questionnaire or clinical trial, and we may even believe we have an algorithm which tells us how to make a judgement based on that source, but most often we will have several pieces of information expressed in more or less vague terms, and most often they have to be weighed up and combined based on our own best judgement. To be sure, we can develop some scoring algorithm to help us with that process, but we still need to choose and justify the algorithm. Occasionally an evaluation question can be reduced simply to a question about the *direct* influence of C on E, but most often, as suggested by theory-based evaluation approaches, we have to consider a *network* of causal factors which influence one another, mostly *indirectly* along the paths in the network.

In theory-based evaluation, we may work more deductively, with a pre-existing model or theory; we have to look for evidence for the different links in the theory, revise the theory, and then make evaluation judgements based in part on the revised theory. Or we may work more inductively, developing step by step a theory about relevant causal factors and the links of causal influence between them; this means in particular being able to identify some common causal factors within the different claims in order to combine all this information. Whichever route we take, we will have to make decisions about boundaries (what is part of the model, what is not, and who decides?) and about values (what is good, is this good, is it good enough?).

Most evaluators don't ever physically combine all the fragments of causal evidence at their disposal into one single, composite causal map. But in this documentation we argue that they are still in a sense in possession of a causal map, they just haven't drawn it yet; and perhaps they could, and should. We also argue that it is useful to think of a pile of interconnected causal information as a "map" in the sense of an abstract structure for storing causal information. Such a structure might look like a bewildering hairball if we tried to just print out all of it, but if we know the right rules for doing so, we can print out various *summary* maps and *sub*-maps to help us answer various questions. We also argue that in practice evaluators need software to help them to create, organise and store the different pieces of causal information they collect in the course of piecing together answers to evaluation questions; software which understands the translation rules and can help us with producing the right sub-map or aggregate map to answer a particular question.

The heaps of information which evaluators have to deal with, combined into a causal map, are usually composed largely of number-free judgements like "B makes a considerable contribution to E" or "S believes that E happened entirely because of B". This means that the special set of rules which are available to statisticians when processing entirely quantitative maps are not available to us. So what *do* we do, faced with such a heap of information? Even if we are not conscious of it, we make use of a larger set of generally weaker rules which are still available to us, mostly based in the end on "common sense".

A social scientist might throw up their hands in despair at the vagueness of the information we as evaluators have to deal with: causal factors are not clearly defined, links in maps are ambiguous as to whether they refer to groups or individuals, situations or events; it is not clear whether the claims are eternal or momentary, generalisable or specific, and so on. Yet, decisions have to be made, and information is required, so we do the best we can with what we have.

12.2 Features which causal mapping approaches have in common

12.2.1 Causality

But in our overall definition of causal mapping, what all approaches share is that a causal map's translation rules have to be **explicitly causal**. As Pearl points out, a genuinely causal arrow can be understood as something like "**if you do C, that will make E happen** (perhaps, with probability p)", whereas, for example, Bayesian belief networks are not strictly causal maps because the arrows say "if you observe C, you will observe E (perhaps with the probability p)". Causation is not correlation.

For Pearl and colleagues, the definitive way to check if C has an effect on E is to intervene in the system, if necessary breaking or disabling any incoming links which might determine C, and tweak E to see what difference just that tweaking makes to E.

C and E can be expressed specifically and uniquely or very generally, or anything in between. But as the claim is causal, the link has to express some kind of causal mechanism which is hardly going to make sense restricted to just one single case.

12.2.2 Modularity

All these approaches share the basic idea that causal knowledge can be at least partially captured in small relatively portable nuggets of information (like "drought causes hunger" or "mosquitoes cause mosquito bites", something like the idea of a "mechanism"), and that these nuggets can be assembled into larger models of how things work, or at least of how they worked in one or a few cases.

We are for the most part not interested in total or exclusive causation but in causal influence.

12.3 Causal maps as a summary of qualitative data analysis of textual causal claims for each link

Usually when we do causal mapping, we try to listen to (or read) what people tell us, and bit by bit ("inductively") try to identify the common elements in their narratives, such as "Health" and "Amount of exercise" for example. Different respondents will, of course, not always use exactly the same phrases and it is a really exciting and creative challenge to create and curate this list of causal factors. This is your job as causal mapping analyst. For example, if Mo says "Feeling good about the future is one thing that increases your wellbeing", is this element "feeling good about the future" the same as "confidence in the future" which Sara mentioned? Should we encode them both as the same thing, and if so, what shall we call it? Positive view of future? Does that cover both cases?

The possibility of coding links between concepts is mentioned briefly in a well-known QDA handbook (Saldaña, 2015) as a possibility, and the Axelrod school has its own coding manual describing how to highlight areas of text expressing causal connections and code them as links between causal factors, inspired by evaluative assertion analysis

This challenge is central to the overlapping field of qualitative data analysis (QDA), which often makes use of tools like NVivo, Dedoose and AtlasTI. However those tools are designed to capture general concepts like "Wellbeing" but are not as well suited to coding links *between* concepts, which is what we need for causal mapping. We believe Causal Map is the only app which is dedicated to helping you with this task.

12.4 Causal mapping as a form of data collection

"Causal mapping" is often used as a name for a specific kind of data collection method, along with suggestions for analysis. There are a vast variety of possibilities for gathering data for causal mapping, with seemingly every author having their own suggestions, from individual interview (Ackermann & Eden, 2004) to reusing open-ended questionnaire questions (Jackson & Trochim, 2002). All of these methods can be coded using the Causal Map app.

Some examples of data collection modalities:

Individual respondents are deliberately asked for information about causal links, for example via open questions at the end of a questionnaire or via a series of interviews in which people are directly asked questions of the form "what causes what?" or "what contributed to this event?"

A set of documents is gathered (either strictly comparable documents as in a medical meta-analysis, or complementary as in a broader review often known as "deskwork," and criteria are drawn up for which sections are to be analysed (e.g., just the executive summaries).

A group of people are deliberately asked about causal links and this information is merged straight away into an overall picture, as a participatory process with the group (Penn & Barbrook-Johnson, 2019) (Markiczy & Goldberg, 1995).

When interviews are carried out, there are different ways to elicit causal claims, for example:

Backwards questioning about problems as in the "problem tree" approach: respondents are asked about a problem in their lives, and what causes it, and what causes those causes, etc.

Backwards questioning about changes, as in QuIP (Copestake, Morsink, & Remnant, 2019) and (rather differently) in Most Significant Change technique (Dart & Davies, 2003): respondents are asked about changes in their lives recently, and then for causes of changes, and then for causes of the causes, etc.

Forwards questioning about effects, as in iterative scenario planning, and in particular in ParEvo: e.g. people are asked what might happen next, and then what that would lead to, and so on.

Each of these approaches have their own detailed suggestions for how to gather and analyse data.

One popular (and cost effective!) way to source data for a causal mapping study is to reuse existing text data which was gathered for another purpose, in particular open ended questions in surveys, providing data protection agreements allow this.

Of course all of the usual procedures for ethical review, protection of respondents, gaining assent and data protection apply as for any other comparable piece of research.

12.5 QuIP as a form of causal mapping

The QuIP is a form of data collection with some very special features:

- Interviewers are usually blindfolded to the commissioner and to the specific intervention.

- Respondents are asked about changes in key domains, generating a backwards chain of causal explanations ("and what influenced that? …. And what influenced *that*?").

QuIP analysis also involves qualitative, inductive coding. The above features have implications for the kind of coding we do. The causal factors in QuIP:

- deal with actual events which happened to the respondents rather than with general principles
- usually take the form of changes ("improved" / "decreased" etc).

Finally, in the QuIP we are interested first of all in people's *beliefs* about what causes (or caused) what, constructing a causal evidence map. Only then, as an optional next step, do we consider whether, and how much, we can deduce from that what *actually* causes (or caused) what. In particular, we are interested what we can deduce about causal paths from explicitly and implicitly identified interventions to other specific factors ("Outcomes") which are agreed to be important.

12.5.1 QuIP: explanations of changes do not themselves also have to be changes

One question which comes up a lot is this: In QuIP, we ask people to describe changes in their lives over (say) the last three years, and then generate a backwards chain of causal explanations by asking "and what influenced that? …. And what influenced *that*?". So, do the *explanations* also have to be expressed as changes in the last three years? For example, someone might say this:

I know more about water conservation now because of the radio broadcasts sponsored by Organisation X, which we didn't have three years ago. But I also know more because of the local government's agricultural officers who explain things to me. They've been around for ever, and they haven't changed their activities, it's just that I learn something new from them every time I meet them.

We recommend including the agricultural officers as a causal factor influencing the change in knowledge, even though their activity has not changed.

This means that we also sometimes code factors like "God" and "Unemployment" which are often mentioned as causal influences even though they may not themselves have changed. Of course, we hope that the interviewers have been trained to gently question whether the respondent really means to describe a causal influence and not is not just producing an empty formula out of habit.

If we look at the same question from a perspective of quantitative statistics, we note that if we had data both from three years ago and from now (which we don't), we would indeed observe a correlation between presence of radio broadcasts and knowledge of water conservation. On the other hand, we wouldn't have any correlation between presence of radio broadcasts and input from agricultural officers, because there is no variation in the officers' input; it was the same all the time. This fact might lead us to feel that there is something illegitimate about our recommendation to include the officers' input as a causal factor. We would perhaps like to have data from a parallel world in which there were no agricultural officers over the whole three years, to see whether the knowledge increased, but we don't. But if we think more carefully, we will realise that nor do we in fact have data from three years ago on the radio broadcasts either. What we have in both cases is not a statistical contrast but our respondents' more or less implicit causal claim that it was both the addition of the radio broadcasts *and* the presence (rather than the absence) of the agricultural officers which each made a difference. The validity of people's causal information comes primarily from a whole shared knowledge map gained over time from culture, instruction and experience. For example, it is not usually the case that I think "ooh, my knowledge seems to be going up, what could have caused that?" but rather we are well aware of the causes because we are part of, inside, the whole process, and we know what it is like to gain understanding when and because someone explains something. It is true that this knowledge is sometimes updated using systematic observation of contrasting cases, whether before-against-after, or here-against-there, or even occasionally using experimental manipulation; but this is an important additional option rather than a primary or original source of causal information.

12.6 Advantages of causal mapping

12.6.1 Induct

Causal mapping aims to directly understand and collate the causal claims which people make in narrative (and other) data rather than trying to deduce causal connections using statistics or some other method. 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.

In some forms of causal mapping, the map is drawn as a synthesis of the views of contributors in a participatory process. In other forms, like QuIP, each contributor is helped to produce their own map and these maps are synthesised later by an analyst. In either case, the maps do not need to follow any preconceived conceptual framework; types of causal claims are identified inductively and iteratively. This is a partly creative process, however the decisions made during synthesis are transparent as the underlying text is always available.

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

12.6.2 Discover

Causal maps work on two levels. On one level, they are presentations of individual and shared cognitive structures, the maps “in people’s heads” which are real (social-)psychological things which we need to know about if we want to understand, predict and influence behaviour. On the other level, they are putative, fallible maps of the actual causal world: how things work. Like all other research results, these maps may be wrong, but they usually contain at least some truth. At Causal Map we take a realist stance on both of these levels: the maps in people’s heads are real, and the causal world, made up of many causal links between causal factors, is real too.

12.6.3 Distinguish

It is possible for a causal map to be able to encode both general claims of the form B causes E as well as specific historical claims that B in fact caused E.

12.6.4 Present

The results of ordinary qualitative research on texts is usually just more text, with maybe some tables of frequency of occurrence or co-occurrence of particular themes for particular respondents, with maybe a chart or even a network graph to present these results. Causal maps on the other hand are not additional presentations of additional analyses but are the main product of qualitative causal mapping. They are relatively intuitive and easy to understand.

12.6.5 Query

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. With certain assumptions, it is possible to ask and answer questions like “which is the largest influence” or “which is the most positive effect”.

12.6.6 Quote

The original quote or other evidence on which each causal link is based can be stored within the link itself. That means that at every stage of causal mapping, it is possible to directly to return to the story in the original context.

12.6.7 Reuse

Causal mapping also encourages reanalysis of existing narrative data which is often gathered but left unanalysed. It is highly suited to online use, e.g. gathering narratives via online interview, email, questionnaire etc., reducing airmiles and viral risk. There is no need for a foreign evaluator to travel to gather or collate data.

12.7 When to use Causal Map

So, you've taken a look at the **features of the app**, and you're getting excited about creating maps – but is your project right for it?

Use Causal Map if you:

- have a relatively large amount of narrative data (enough to provide at least 20-30 causal links)
- need help to organise a large number of links and summarise them into an overview or synthesis
- have information from more than one source (for example different respondents, different documents, or different places in one document) and the information about the source is important to you: they aren't all interchangeable
- are interested in possible differences between the sources and groups of sources – and/or you don't necessarily have a preconceived idea of the contents or boundaries of the map.
- want to capture what your sources actually say, systematically and transparently

Causal Map map is not suitable if you:

- 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 sketch out a plan (e.g. Theory of Change or similar) without much reference to the different sources underpinning each link

Part III

Analysing maps

Chapter 13

Analysing maps: Overview

Various kinds of software are available to simplify the task of viewing and filtering a causal map, especially for larger maps. For example, gephi.org is a good choice for very rapidly manipulating even very large maps - much larger than Causal Map can handle.

A causal map can be filtered to show only some parts of it. Some submaps are well-known from network analysis, for example the “ego network” for a particular factor is the part of the map which includes only those factors directly connected to it. This is something like searching for that particular factor.

An related approach is to specify one set of Source factors and another set of Target factors and to display only the links along paths which lead from a Source factor to a Target factor. This is called an “etiograph” in a seminal organisational analysis of Utrecht Jazz Orchestra (Bougon, Weick, & Binkhorst, 1977).

Chapter 14

Analysis: comparing maps between particular groups

14.1 Filtering by respondent group

We code a causal map on the basis of text data. That text data can be usefully broken up into statements, usually of a length between a paragraph and a page. Each statement usually has "**additional data**" associated with it, for example the ID or gender of the respondent, the text of a question to which this statement is an answer, the page and name of the document from which this statement comes, etc. When we code a causal claim within a statement, we can associate the resulting link with the additional data. That means that for every link, we should know the additional data, e.g. the gender of the respondent, etc.

We call the set of statements corresponding to a particular value of a particular additional data field a "**group**". This definition of "group" is quite broad and does not have to refer only to respondents, e.g. the group for "question 3" is the subset of all the data relevant to that question.

It is easy to filter a causal map by this additional data. This idea goes back at least to (Ford & Hegarty, 1984). For example, here is one map filtered to show all and only the links mentioned by with female respondents. We call these the **per-value maps**, e.g. the map consisting of all links mentioned by women. However, often the maps for different groups are quite similar as a large proportion of links are shared. When there are many links as in this example, the resulting filtered maps can be uninformative.

There may still be a bewildering hairball of links. We can apply techniques like hierarchical coding to "zoom out" of the map, or simply show only the most frequent factors. This map shows the top five factors for women:

Figure 14.1: image-20210115220243930

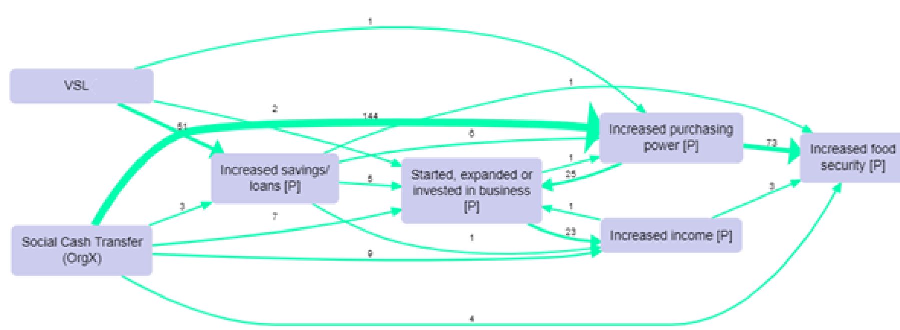


Figure 14.2: Causal Map

And this map shows only the top five factors for men (there were far fewer men in this project).

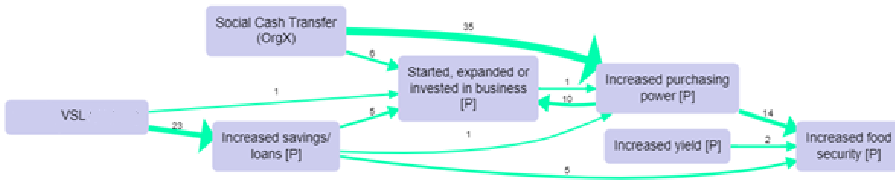


Figure 14.3: Causal Map

This is disappointing. The maps are almost identical, though the frequencies are different. We might have a feeling that the men's answers do indeed differ from the women's, but how can we be sure? Does this pair of maps prove that there are no interesting differences?

These are important problems:

- Which, if any, features of the map for one group are interestingly different from the map for another?
- Are there in fact any relevantly different sub-groups within our data? When we have many additional data fields (gender, age, education level, parental occupation, size of village as well as question number, interviewer ID, etc), how can we focus on the most interesting differences? Is it *worth our time* looking at separate maps for different sub-groups like different villages or genders?

14.2 Showing *distinctive maps* for each group

One solution is to display only features which are distinctive for each group. This means the map for this group, which usually includes all links coded for statements from that group, now excludes any of those links if the factor at each end of the link is not **distinctive** for that group. The factor "increased income" is **distinctive** for men if it is mentioned significantly more often by men than by women taking into account the baseline, i.e. the total number of times men mentioned any factor at all. (For example, maybe there are more men than women, or they talk more, or both.)

A slightly more technical discussion, which can be skipped:

This means that for binary fields like (gender in this particular study), if there are links from factor F to factor G in the distinctive map for women, there cannot be any links between these factors for men. On the other hand, factor

F may appear in both maps even though it is not distinctive for either group, but if it does so this is only because in each sub-map it is connected to a factor which is distinctive for that group.

More generally, we look at whether the proportion of mentions of a given factor from a given group compared to all the mentions (i.e. including those by the other values of the same additional data field, e.g. the other villages, the other gender) is significantly higher than the baseline, i.e. the overall proportion of mentions of all factors by this group compared to the other values. Is (Mentions of factor F by this group / Mentions of factor F by all groups) significantly higher than (Mentions of any factor by this group / Mentions of any factor by any group)?

Given a filtered set of statements, do the usual calculation to get what links would have been shown in the normal view, i.e. which links are coded in the current statements, i.e. create the variable shown vs not-shown for each link. Group all the links into co-terminal bundles (with the same from and to factor). Create the table which has IDs of bundles in the rows and two columns, the number of shown links and the number of not-shown links within each bundle. Run a chi squared test on this table and mark a bundle as "distinctive" if the standardised residual for the "shown" cell is greater than a given cut-off value.

This idea takes a bit of getting used to. A distinctive map for a group filters out all links between pairs of factors which are both not distinctive for that group. That is the same thing as saying that it shows all the factors which are distinctive for that group, plus factors which are linked to those factors.

Why don't we just stop there, and simply display only the factors which are distinctive for a group? Why do we add factors which they are linked to, even if not distinctive? The trouble is that there might be a factor which was very distinctive for the group, but was not linked to any other factors which were distinctive, and would therefore disappear from the map (or we would have to display it floating on its own, which is not appropriate for a map of causal *connections*). In practice this step does not make much difference, and it is usually acceptable to think of a distinctive map as simply a map consisting of factors which are distinctive for that group.

It is also possible, especially with smaller maps, that there are no links which are distinctive for a particular additional data field like gender. This suggests that there are no very interesting differences between the groups. It is also possible that there is a distinctive map for one group, e.g. women but not for men, or vice versa.

To continue the example, we can follow the algorithm and produce the distinctive maps for men and women. These maps even without any further filtering are interesting and readable in their own right, but as there are around 40 factors in each map, here we show the maps filtered further for only the most frequent factors.

Distinctive map for women, showing the top seven factors.

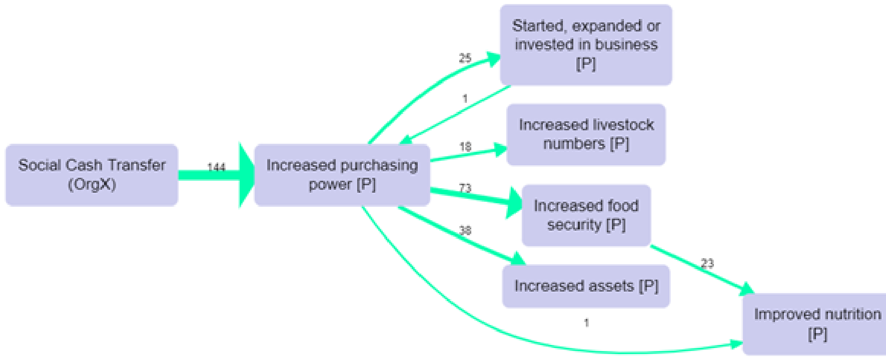


Figure 14.4: image-20210115220338421

Distinctive map for men, showing the top seven factors, also filtering out a few links with just one mention:

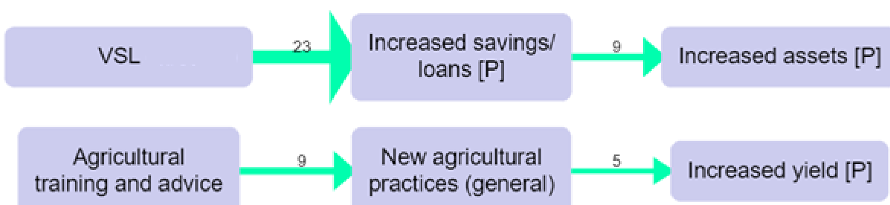


Figure 14.5: image-20210115220351694

Here, some of the paths in the map for men, from VSL to Increased savings and thence to Increased assets, would be certainly worth following up, as well as the paths to and from "Increased purchasing power" in the map for women.

As usual, we should treat the maps just as a gateway into the data, an invitation to read and understand the quotes in more detail and in context.

To repeat, the "distinctive map" for a group is *not* a good summary of everything that was said by that group. It is a good summary of what was said by that group more often than the other groups.

14.3 Filtering statements and groups for comparisons in Causal Map

When filtering statements it is easy, for example, to show the causal map just for specific **groups**, e.g. the female respondents, for which the additional data field “gender” is equal to “female”: you can cycle through the values using the arrow buttons, and you can flip through the statements for each group using the “previous” and “next” buttons in the statement pager.

Chapter 15

Causal coding principles

15.1 Formulating labels (semi-quantitative factors)

Remember that in general, factor names are case sensitive. So funding is **not the same as** Funding.

You can use emojis and other special characters in your factor labels.

Also, semi-colons ';' you can use them if you want, but note that they have a special meaning in the app.

15.1.1 Relax! Use heterogeneous, "in-vivo" factor labels

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.

15.1.2 QuIP specific: back-chaining

In QuIP projects, most of the interview material comes in response to questions about changes in the reference period, usually the last three years. QuIP questioning continues back up the causal chain, asking what was the reason for that? … and the reason for that? That means that most influence factors will also be expressed as changes or differences (a change in F is explained by a change in E which is explained by a change in D….), whereas some are not (e.g. "unemployment"). Some analysts will try to avoid coding this kind of claim (can something which has been around a long time explain a change in the last three years?) but you may decide that this distinction does not matter too much. If something really does describe a *change* (e.g. "became unemployed") then it should be coded.

Most QuIP testimony begins with questions about whether things have got better or worse, so most causal factors, going back up the causal chain, are likely to be semi-quantitative too.

15.1.3 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."

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.

15.2 Opposed pairs of causal factors

What to do when some explanations use a causal factor phrased in a positive way and others use a similar causal factor but phrased the other way around?

"I feel good because my health is good."

"My sister lost her job because her health is bad."

We could code these:

Health improved Feel good

Health got worse Lost job

But we might feel we are missing the fact that the first factor in each case is arguably the same thing, just the other way around.

Solution 1: do nothing

This may mean you find yourself using pairs of opposing factors such as Better health and Worse health to capture the causal claims – things which are in a sense the opposite of one another. You might decide that is fine anyway, because you are happy to have pairs of factors like this, or because you decide that the pair of factors are not really polar opposites at all and therefore you don't want to combine them. For example, illness is arguably not really simply the opposite of health but a quite different state with its own causal rules.

Solution 2: merge them

If you use more advanced coding styles involving "strength," detailed in the "Extra" section of this Guide, you can avoid this by just using one factor like Better health.

Alternatively, 'Health Improved' and 'Health got Worse' could be coded as 'Health Improved' and '~Health Improved' respectively, with '~' indicating the 'opposite' of an improvement in health. The app will then treat links to Health Improved and ~Health Improved separately. More information on this can be found here.

15.3 Formulating factors as desirable

If possible, especially when conducting evaluations or research for policy purposes, try to formulate these semi-quantitative factors so that for each one,