

```
knitr:::opts_chunk$set(cache=T)
library(tidyverse)
```

Overview

What is Causal Map?

[Causal Map](#) is a web app for analysing and visualising causal claims within narrative data from interviews or reports.

```
knitr:::include_url("https://player.vimeo.com/video/689223728")
```

If you are particularly interested in the *theory* of causal mapping, rather than how to use the app, look for the "book" symbol  in this Guide.

This Guide is divided into these Training Levels:

1. [Getting Started](#)
2. [Basic Coding](#)
3. [Basic Analysis](#)
4. [Advanced Coding](#)
5. [Advanced Analysis](#)

Each is followed by a simple [quiz](#).

... and there are also sections on:

- [The rest of the app](#)
- [Technical reference](#)

If you are in a hurry (and who isn't?), just read Level 1: Getting Started.

Features marked  are experimental and are implemented on the [test server](#) only.

Videos

```
knitr:::include_url("https://www.youtube.com/embed/S2acbJJ9t7E")
```

Also check out our growing library of short, focussed training videos, none longer than four minutes!

Support chat

- WhatsApp group for Causal Map
- Whatsapp group for StorySurvey

Quizzes

Test your knowledge of causal mapping and Causal Map with our structured series of quizzes.

Selected projects to date

These are some causal mapping projects completed with Causal Map – either directly by subscribers, or by Bath SDR working for clients, or conducted by Causal Map Ltd for the client.

For more details on selected Casual Map projects please check out our Recent Projects page.

License {unnumbered}

This guide is licensed to you under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

```
readLines("Guide-to-causal-mapping.Rmd") %>%
  str_match(".*vimeo\\\\.com.*|.*youtube\\\\.com.*") %>%
  keep(~!is.na(.)) %>%
  unique %>%
  paste0("```{r, echo=F}\n", , "\n```\n") %>%
```

```

c("# Video list{#xvideo-list}\n\nThese videos are all included at
different points in the Guide. If you love videos, here they are again in
one place. Feel free to binge. Tip: if you play a video, you can press
'Picture in picture' and then switch to the app, and you can follow (and
even resize) the video.",.) %>%
writeLines("83500-video_list.rmd")

readLines("faq.txt") %>%
str_replace_all("//.*", "\\1</summary>\n") %>%
str_replace_all("//", "</details>\n<details><summary>") %>%
c("</details>") %>%
paste0(collapse="\n") %>%
paste0("# FAQs\n\n## Logging in\n\n<!--DO NOT TRY TO EDIT THIS TEXT, EDIT
FAQ.TXT INSTEAD-->",.) %>%
writeLines("80021-faq.rmd")

```

!--chapter:end:index.Rmd--

(PART*) 1: Getting Started{-}

!--chapter:end:10000-PART-1-getting-started.rmd--

Level 1 overview{#xlevel1}

In the pages in this section you will learn how to get started with Causal Map: what kind of device you need, how to access Causal Map for free, and finding your way around the app.

When you've read the pages, try the [quiz](#) to see if you have understood enough or need to review some!

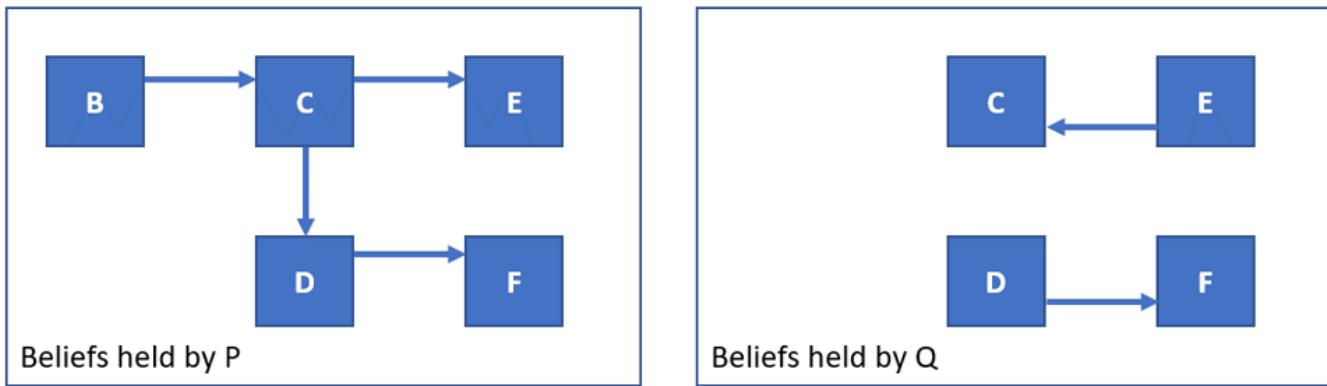
!--chapter:end:10010-level1.rmd--



What is a causal map?{#xcausal-maps}

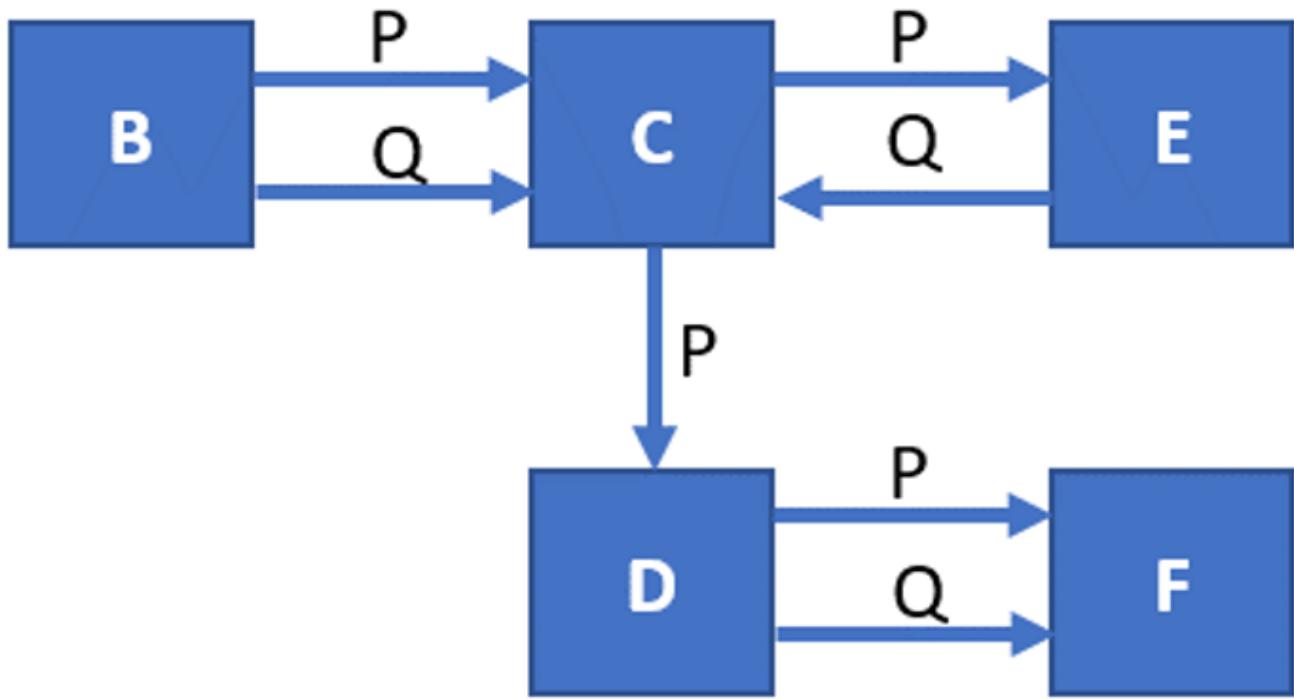
A causal map consists of boxes (which we call "factors") joined by arrows (which we call "links").

A link from factor C to factor E means that someone (P) believes that C is some sense causally influences E. Every link represents one causal claim.



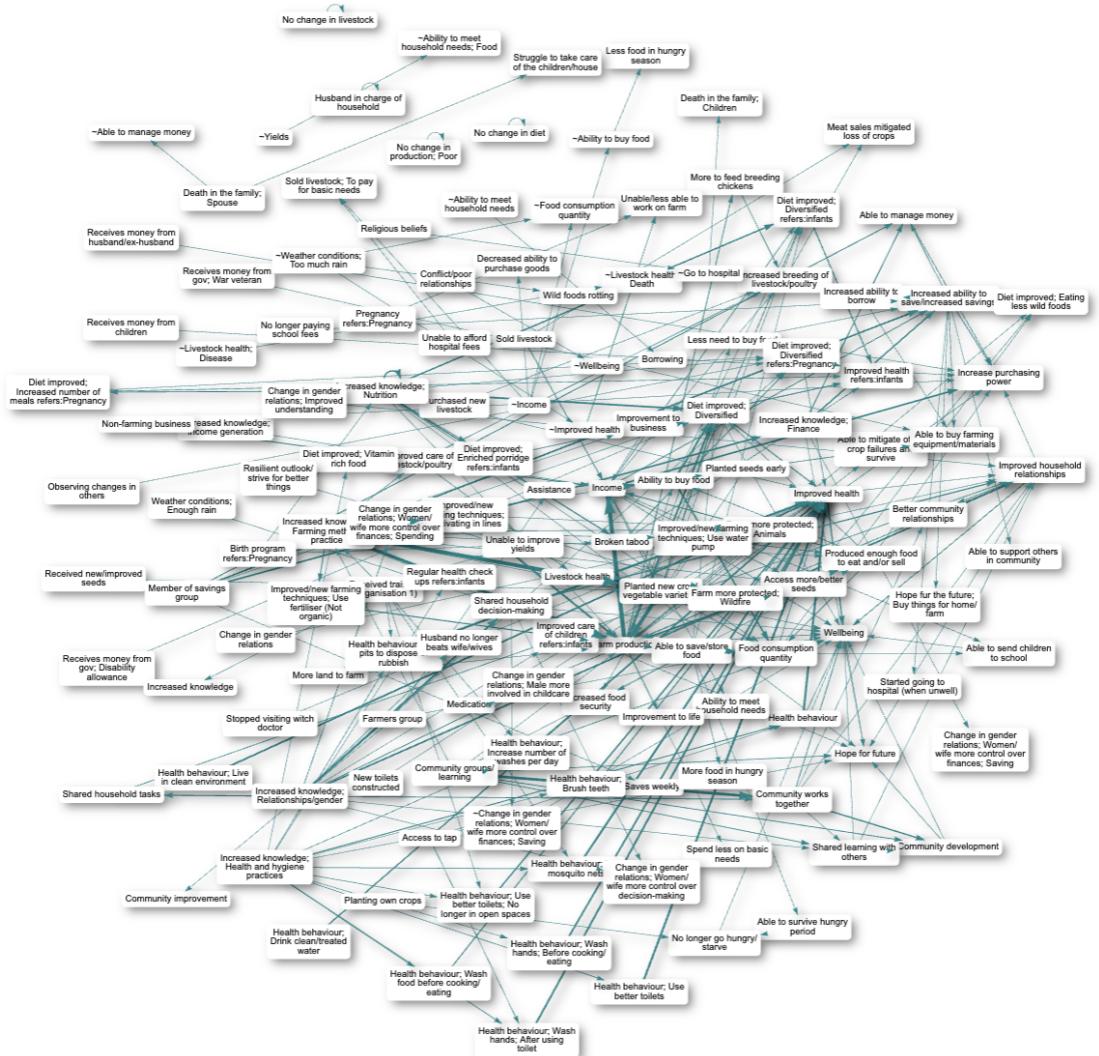
Two separate single-source causal maps: Beliefs held by sources P and Q about causal influences between various factors.

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



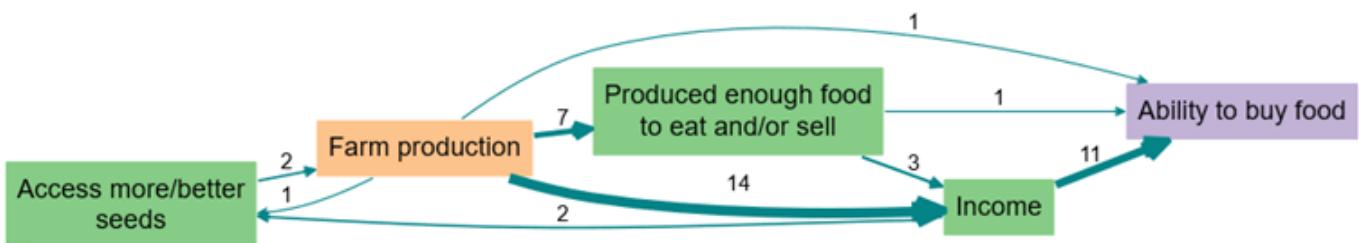
A combined, multi-source causal map about beliefs held by sources P and Q about causal influences between various factors

When we have highlighted many causal claims we can present the entire, global causal map containing all the claims like this:



A global causal map combined from many sources

However this map is bewildering and not very useful. So to answer specific questions, we filter the map in certain 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.



--chapter:end:10020-causal-maps.rmd--



What is causal mapping? {#xcausal-mapping}

This is a shorter version of [this post](#).

Induct

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

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

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

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 we need to know about if we want to understand, predict and influence behaviour.
- On the other level, if we trust our sources we can use this information to tell us something about how things work in the real world.

Present

The results of ordinary qualitative research on texts is usually just more text. 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 (and they look good in presentations 🎉).

Query

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.

Quote

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

Causation - no longer a taboo

Gone are the days when we could think of data or information as primarily about numbers. Many of us who are involved in understanding the social world and evaluating interventions within it spend much of our day understanding, presenting, manipulating and caring about *causal structures* (and even *models of other people's causal structures*).

CEDIL Video on Impact evaluation: causal maps, mechanisms and the elusive quest for useful middle range theory

With Nancy Cartwright, Gary Goertz, James Copestake and Steve Powell

```
knitr:::include_url("https://www.youtube.com/embed/6DBzadDMXew")
```

!--chapter:end:10023-causal-mapping.rmd--

Are Causal Map and causal mapping for you?

So, 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

!--chapter:end:10025-for-me.rmd--

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.

::: {.rmdnote}

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.

::: {.rmdnote}

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

...

Signing up

To use the app, create an account [here](#). It's quickest to just log in with a Google account, or you can use your email address.

It's free to log in and explore the app.

[!--chapter:end:10103-prerequisites.rmd--](#)

Accounts

To use the app, you'll need an account. What account you need depends on what you want to do.

	Free account	Paid subscription account
View and analyse	Unlimited	Unlimited
Code your own data	You can code up to 50 links for free.	Find the subscription option which is right for you.

Free accounts

You create an account just by logging in. You can create a new file and do some coding and analysis.

Paid subscriptions

You log in as usual, but to code more than 50 links you need to have a [subscription](#).

To cancel, renew or manage your subscription see this [section](#).

If you have a paid subscription, make sure you used the email you use to login is the same as the one for which you have a subscription.

Getting help



Toolips



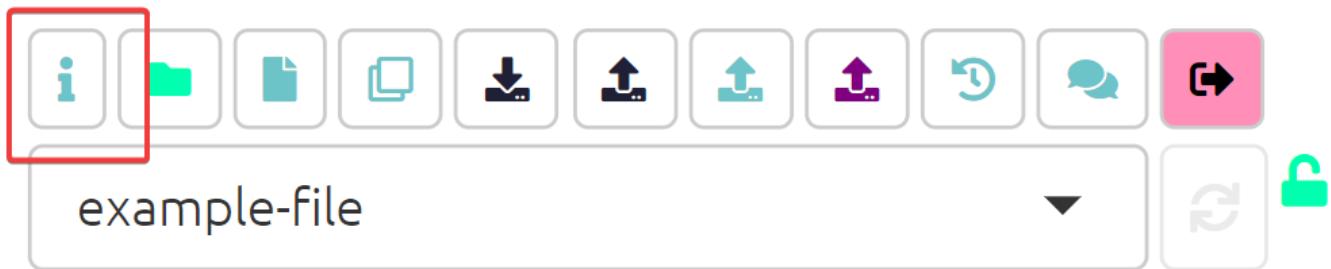
You can get information about most of the buttons in the app by hovering your mouse over the button for a couple of seconds.

Blue info buttons

You can get help at any time by clicking on the little blue info buttons, like below. This will then open a panel displaying the relevant section of this guide (good for when you forget what a feature works for).



Help button



The details of what you see may differ slightly depending on the version of the app you are using.

Support chat

You can also usually get live help on these WhatsApp

- [WhatsApp group for Causal Map](#)
- [WhatsApp group for StorySurvey](#)

!--chapter:end:10104-getting-help.rmd--

Has someone shared a file with you? {#xinvite}

::: {.rmdcaution}

If someone has shared a file with, say, YOU@MAIL.COM, make sure you sign in with YOU@MAIL.COM as well.

:::

If they have sent you a link like <https://causalmap.shinyapps.io/CausalMap2/?s=99> to a particular map (or table), click the link and you will be automatically taken to view the corresponding map (or table), providing they have given you the correct permissions.

Alternatively If they have just asked you to log in at the app to explore a file name, say "project-x", log in and click on the dropdown list on the left-hand side and click on project-x to load it. If you can't find that file name, it means they haven't correctly shared it with the email address you logged in with.

The screenshot shows the Quip interface. At the top, there is a navigation bar with icons for Home, Information, File, Document, Download, Upload, Refresh, and Share. Below the navigation bar is a dropdown menu with the following options:

- example2
- example3-path-tracing
- svukwebinar
- quip-exercise
- quip-coded
- example3-for-testing
- quip-uncoded-short

Below the dropdown menu, there is a preview window showing the content of the selected file 'example2'. The content includes:

```
cm1/advanced quip uncoded
source_id 1
question_id 1
question_textglobal question
```

At the bottom of the screen, there is a footer with two small icons.

If you would like to explore further, you can then apply filters suggested for this file by using the [dashboard](#).

!--chapter:end:10106-just-shared.rmd--

You want to code your own data?

The workflow is usually like this:

1. GATHER THE DATA

- Identify or collect a set of several reports or a set of several interviews from one or more sources.
- Add metadata to this data, e.g. date, name of source, perhaps gender, location, etc.
- Organise this narrative data into an Excel template in the form of paragraph-length chunks called `statements`

2. CODE:

1. UPLOAD: Create a new empty file and upload the narrative data.
2. HIGHLIGHT: View your data statement by statement, and highlight any causal claims within it, telling the app what is the cause ("influence") and what is the effect ("consequence") for each claim. We call this process "coding".
3. ANALYSE: When you've coded a lot of causal claims, you can analyse them to make different kinds of maps and answer different research questions.

!--chapter:end:10127-data.rmd--

Your first 5 minutes with Causal Map{#x5}

Let's have a look around the Causal Map app.

Quick tour of the app{#quick-tour}

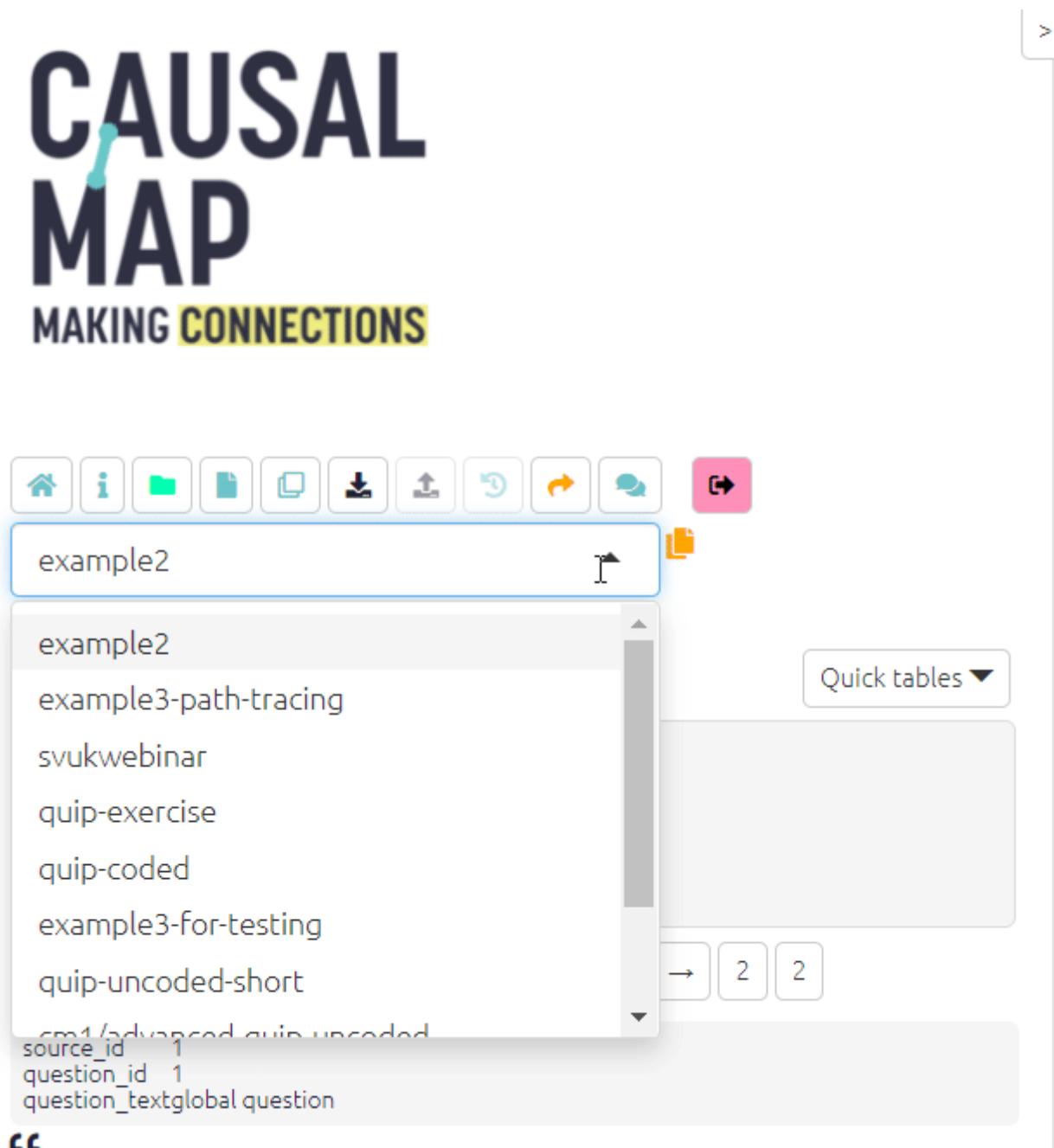
```
knitr:::include_url("https://player.vimeo.com/video/674369121")
```

The app has two parts, the left-hand side and the right-hand side as well as the top menu. Watch the video below to see how these parts work.

Whenever you use the app, you are always working with exactly one file, the "current file".

The top menu contains tools for managing the current file, for instance to restore earlier versions of the file.

Most importantly, the top menu contains the "file dropdown menu" shown here. You use it to select what file you want to view or code.



If you have just loaded the app without clicking on a link, you will be taken to an example file called `example-file`. Other public files or files which have been shared with you are listed in this dropdown.

-  If a file has this symbol, you can edit it, e.g. you can continue to code it.
-  If a file has this symbol, you can view and analyse it but not edit it.
-  If a file has this symbol, you can not edit it but you can make a copy which you can then edit.

Coding a file

If you want to practice coding, you will mostly use the left hand side of the app. You usually start off with an uncoded example file which the app creates for you. You can find it in the dropdown menu, it has a name like `example-file-uncoded-youremailaddress`.

Multiple statement selector{#xstatement-view}

Statements:

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

When you are coding you should click One to view just one statement at a time. Then you will see the navigation bar displayed below.



You can use the arrow buttons to skip through statements. The \rightarrow button takes you to the next statement.

Viewing and analysing a file

Analysing a file means

- **applying different filters** to it, to focus on different parts of the entire causal map in order to answer different questions
- **viewing the results in different ways** by clicking on different tabs in the right hand side to view different kinds of graphical map and various tables.

If you want to view / analyse a file, you will mostly:

- use the right-hand side of the app and the Dashboard
- view many statements at a time by clicking `Many` on the multiple statement selector (above).
- start off by analysing the [example-file](#).

The Dashboard{#xdashboard0}

The Dashboard appears on the right as the app loads, or reveal it by pressing the yellow button at top-right:



Your medals!

Load uncoded example Help

Complete views (replace existing view)

- ★ 10 - Overview 2648
- ★ 11- Overview with smart zoom and combined opposites 2316
- ★ 12 - Overview with combined opposites 2292
- ★ 14 - Contrary factors 2294
- ➔ 28 - Quotes 2296
- 30 - Factor mentions 2755
- 31 - Factors zoomed to level 1 2757
- 34 - Factors - codebook 2299
- ? 39 - Closed questions, per source 2653
- ➔ 50 - Trace from main drivers to main outcomes 2682
- ➔ 51 - Consequences of top 3 drivers 2683
- ➔ 52 - 1 step down from main drivers 2300
- ➔ 53 - Influences on main outcomes 2684
- ➔ 54 - 1 step up from main outcomes 2301
- ➔ 60 - 3 clusters + bundle by cluster + surprise 2302
- ➔ 62 - Maximum betweenness + up 1 + down 1 2303
- ➔ 64 - Top 3 drivers 2305
- ▲ 65 - Top 3 outcomes

Addons (added to existing view)

- ★ Combine opposites ★ Contrary factors get orange text
- ★ Exclude links with only 1 mention ★ Hide plain codings
- ★ Hide quickfields ★ Links with more than one mention
- ★ Pink links - fixed colour ★ Remove isolated ★ Smart Zoom, preserve 5
- ★ Standard format ★ Standard style ★ Top 10 factors ★ Top 15 links

Save the current view

Title for this view Rec? Addon?

Jump to a shortlink

Shortlink nr.

Community development

Each item in the dashboard is a different view of the current file saved by you or other users. You can apply a view by clicking on the corresponding row. You can see more about The Dashboard [here](#)

The other tabs

There are other tabs on the right hand side. You won't need these when you are just getting started.

- [Factor Editor](#) make advanced edits to your factor labels.
- [Files tab](#) manage all your files.
- [Updates tab](#) see activity on your files.
- [Data tab](#) see summary data about the current file such as number of factors
- [Account tab](#) manage your subscriptions and teams.

!--chapter:end:10130-first-5-minutes.rmd--

Interactive View{#xinteractive-view}

The first tab on the right shows 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 moves their mouse over them.

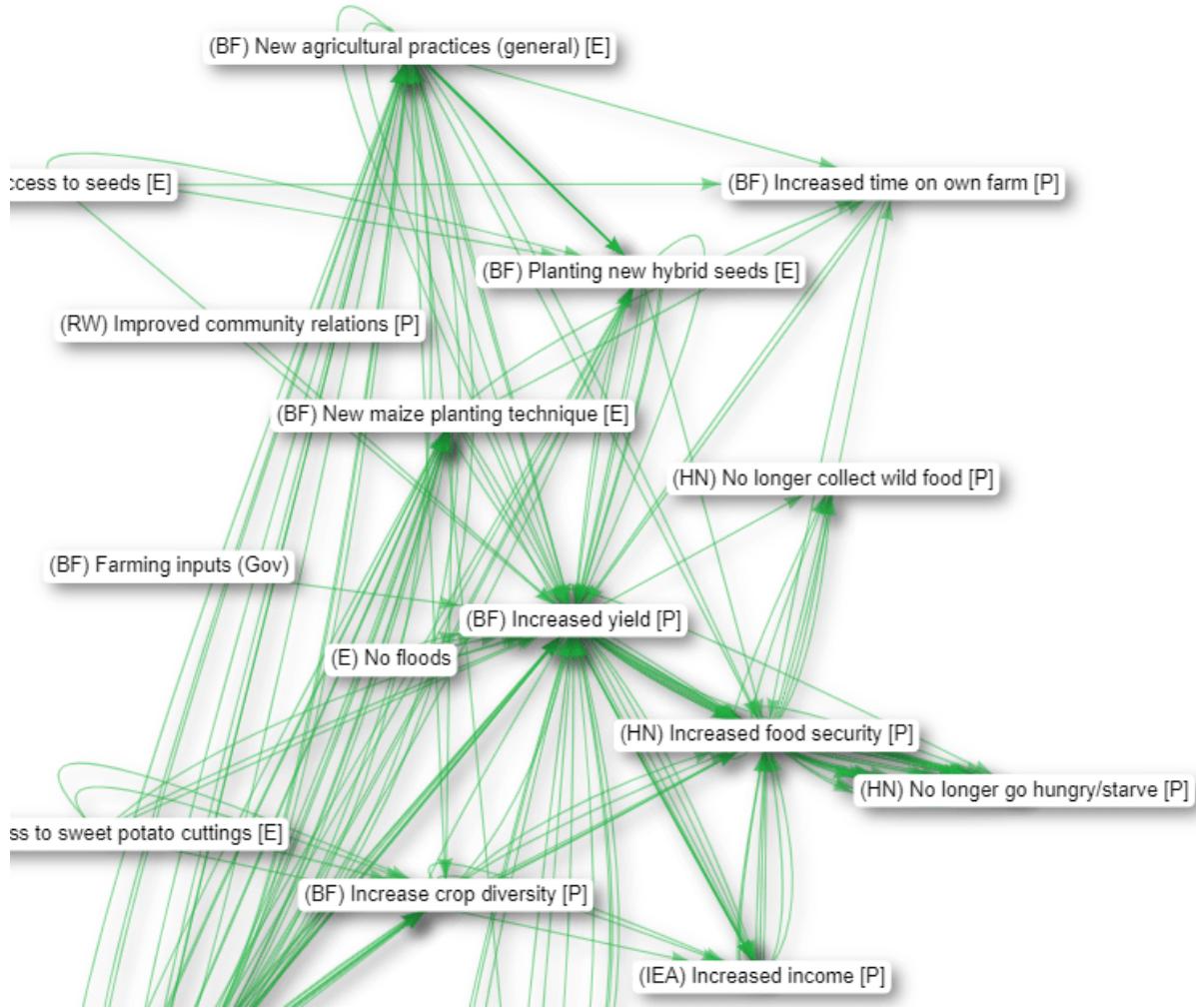
The Interactive View is designed to let you interact with and explore your map.

The video below gives a brief overview of both the interactive and print view tab.

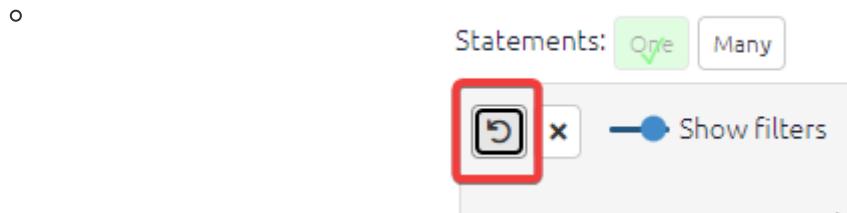
```
knitr:::include_url("https://player.vimeo.com/video/674383670")
```

You can:

- 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.
- Click on factors to focus on them, trace paths from/to them, edit their [memo](#), or delete them.



- As usual, if you want to return to the previous view after focussing on a particular factor, press the Restore Previous Filter button:



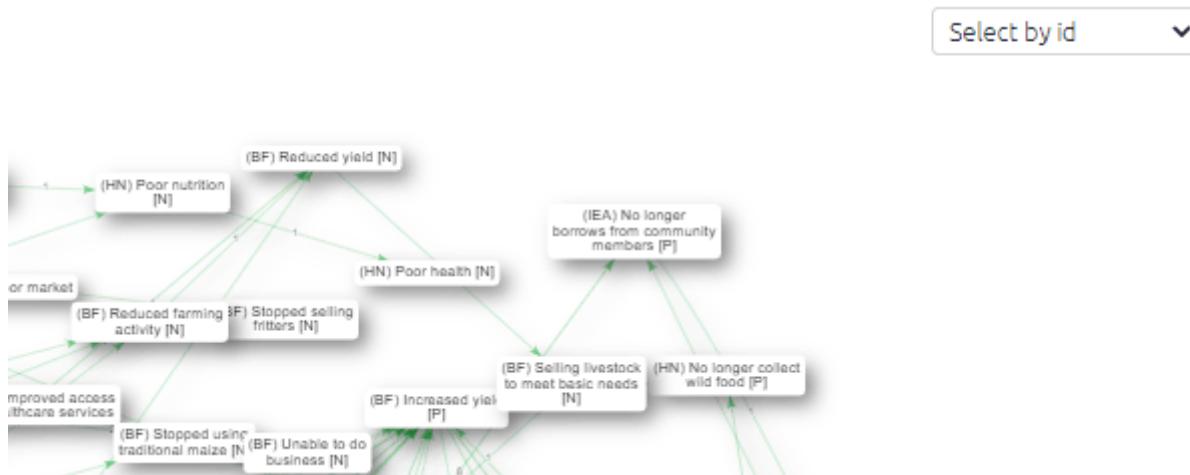
- Click on factors to edit the label or memo

(E) Reduction in mosquito

(HN) Malaria prevention advice [E]

In both Interactive and Print view, if the factor labels have not been wrapped to any specific width, they are automatically wrapped to 22 characters.

When you have a lot of factors in your map, the selector in the top-right corner of the map can help locate them.



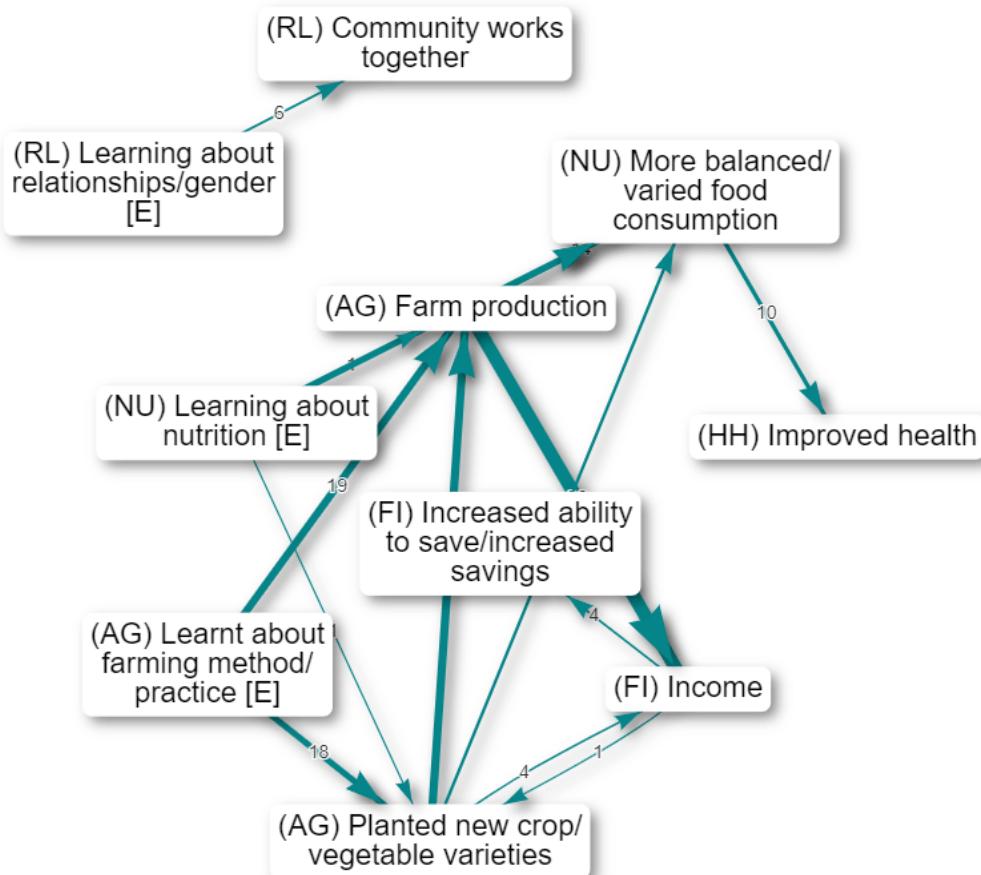
When you are viewing a single statement, the colour of the links corresponds to the colour of the highlighted sections of text.

The screenshot shows the Causal Map app interface. At the top, there are buttons for 'Show filters' and 'Advanced editor'. Below this is an 'Info' panel with fields: source_id: 1, question_id: 1, question_text: global question. A quote from the text is displayed: "Recent periods of high rainfall have led to Flooding. The Flooding has damaged property and businesses. Funds from Organisation X have enabled us to build better flood defences, reducing the flood risk and increasing the investment into the area." Below the quote, it says: "The claims extracted from these statements can be used to build up a network of causal connections." At the bottom are buttons for 'Influence factor(s)' and 'Consequence factor(s)'.



Tracing paths

You can even use the same dialog to [trace paths](#) to or from a particular factor directly from the interactive map

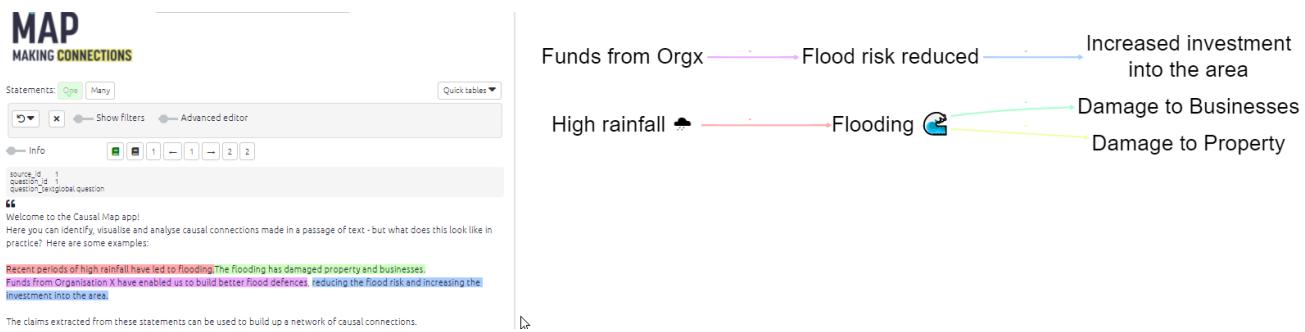


Print view{#print-view}

A print-quality version of the map with advanced layout. You usually use Print View to get high quality static outputs for including in reports.

When you are viewing a single statement, the colour of the links corresponds to the colour of the highlighted sections of text.

- If you have edit permissions, you can edit a factor by clicking on its label.



You can now change the size of the Print map by using the + and - (and Reset) buttons. You can also drag it around.

To change the format of the print view, use the [set print filter](#).

Copying your map as a vector image{#copy-svg}

You can now export your map as SVG using the export SVG button at bottom right.

When you export an SVG image, a shortlink is automatically created for you. The ID of the shortlink is printed in grey at the bottom left of the image, and the link is also copied into your clipboard so you can paste it straight into a report if you want. Anyone who clicks on the shortlink will be taken to the app with the same image with the same filters.

[!--chapter:end:1014.7--print-view.rmd--](#)

Tables{#xall-tables}



You will find all the tables under the table tab at the top of your screen.

A screenshot of a dashboard interface. At the top, there is a toolbar with various icons: a hand cursor, a printer, a document, a list, a pencil, a table (highlighted with a light gray box), a target, a ruler, a folder, a database, a Wi-Fi signal, and a user icon. Below the toolbar, the word "Tables" is written in purple. A mouse cursor is hovering over the "Tables" icon. Underneath, the word "Table" is written in blue. At the bottom, there is a horizontal menu bar with several tabs: "factors" (highlighted in gray), "mentions", "links", "statements", "sources", "questions", "closed_questions", and "metrics".

As you can see there are lots of tables you can create, from a table simply showing the total links to and from each factor to a summary of each respondent's response to a closed question. The possibilities are large – and increase with the amount of additional data your project has.

Presets{#preset-tables}

If you want to keep things simple click the presets button which you will find below the table tabs¹.

The preset button opens a list of views in the dashboard, these links will take you to the most commonly used tables. Once you have clicked on one of these views you will be taken to this table. There is a description of what table you are viewing displayed in the dashboard.



Find out more about tables [here](#).

[!--chapter:end:10149-tables.rmd--](#)

Level 1 quiz

Now do the quiz to see if you've understood about getting started with the app.

[!--chapter:end:10160-quiz1.rmd--](#)

(PART*) 2: Basic Coding{--}

[!--chapter:end:20000-PART2-basic-coding.rmd--](#)

Level 2 overview{#xlevel2}

In the pages in this section you will learn basic coding with Causal Map.

When you've read the pages, try the [quiz](#) to see if you have understood enough or need to review some things!

!--chapter:end:20210-level2.rmd--

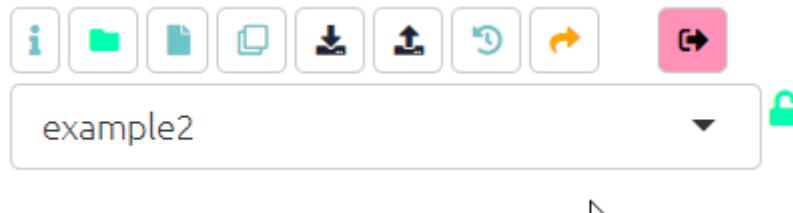
Making your own copy of a file

The app creates an uncoded example file for you. You can find it in the dropdown menu, it has a name like **example-file-uncoded-youremailaddress**.

You can also make your own copy of another file which has been shared with you.

```
knitr:::include_url("https://player.vimeo.com/video/641927229")
```

Copying a file in Causal Map is easily done. Simply load the file you want to copy using the dropdown menu. Then click the 'Save As' button on the top left of your screen.



You can then append something to the end of the filename to make it yours:

Save a version of this map

You don't need to save, your map is being saved constantly in the background.

But you can save a version under a new name.

Name of new map (will be added to existing name)

quip-exercise-

steves-version

Save current map

Click the 'Save' button. This window will then close and your copy of the file will load up. You can now make changes to this new file.

The other buttons help you with these tasks:

- Restoring a previous version of your file
- [Uploading your data](#)
- [Managing the current file](#), for example to share it with others.

!--chapter:end:20230-own-copy.rmd--

Creating a new file

You can create your own file by [copying the example file](#) or by creating a new file. These files will then be added to the dropdown list in the top left-hand side.

If you want to [upload your own data into a new file](#), first press the New File button and give your file a name.



Then you can upload causal mapping data (consisting of at least a set of narrative statements to be coded, e.g. the text of interviews or reports) in the form of an [Excel file](#) (.xls or .xlsx).

!--chapter:end:20252-new-file.rmd--

How to upload and update your data - - Overview{#uploadandupdate}



How to manage your data, whether it is downloading, modifying or appending? These bulletpoints will direct you to the relevant sections later in this Guide.

This page tells you how to get your data into the right format for uploading to Causal Map.

Currently, you can work with 5 different types of formats in the Causal Map App, they are:

1. Text documents: you can upload multiple text files, doc, docx, or pdf files.
2. Standard format Excel: An Excel file with one worksheet called `statements` with one row per statement and optionally other worksheets for `sources` etc.
3. Wide format Excel: An Excel file with one worksheet containing row for each source, one column for each question. e.g. data from Google Forms.
4. Hybrid format Excel: An Excel file with one worksheet containing one row for each answer and includes the statements as one kind of answer but also includes rows for additional data like gender.
5. Links-only Excel: An Excel file with just one worksheet called `links` with just the columns `from_label` and `to_label`. When you want to analyse causal links from elsewhere.

Preparing to code{#xcoding}

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.

Statements: One Many

When you are coding you should switch the multiple statement selector to One, to view just one statement at a time. Then you will see the navigation bar displayed below.



You can use the arrow and numbered button to skip through statements. If you want to skip to a particular statement the easiest way is to use the [find statements](#) filter. This will automatically be applied if you are viewing statements one at a time, so simply click on the filter and edit the number it is searching for.

(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.)

!--chapter:end:20260-prepare-coding.rmd--

The info panel{#xinfo-pan}

This info provides important information about context which you may need to refer to during coding.

During coding, clicking on the `info` toggle opens additional statement info. ¹

source_id TS8
question_id \$A6
question_textSetting recall period (since August 2020)

“

I started the fresh fish business which failed three month later

The memos panel {#xmemos-panel}



When you click the toggle switch, boxes to edit the statement, question and source memos will also appear.

Here you can add useful additional memos on your coding process.

These memos can also be displayed in tables and used to create filters such as "don't show me information from sources with a memo containing the word `doubtful`".

statement_memo	agricultural		
question_memo	open question		
source_memo	global source		

!--chapter:end:20265-info-panel.rmd--

The Coding Panel: creating factors and links{#xcoding-panel}

Creating links in the app

```
knitr:::include_url("https://player.vimeo.com/video/604099226")
```

“

Cash handouts from the social cash transfer has increased the amount of income for this household over the past year. The small scale business has also helped to boost household income though in a small way at the moment.

Social cash transfer

Started, expanded or invested in business

Increased income

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 (as shown in the video below).
- When you have finished, press Enter.
- Repeat the process in the other box to specify the factors at the **end** of the link (or ends of the links).
- Press the green Save button which is now active.
- The link is created in the Map window, 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.
- When you have finished coding one statement, click the right arrow in the statement navigator to code the next statement.

The screenshot shows the 'MAP MAKING CONNECTIONS' software interface. At the top, there's a toolbar with various icons. Below it is a dropdown menu showing 'organisationh22'. The main area has tabs for 'Statements' (set to 'One') and 'Many', with a 'Quick tables' button. There are buttons for 'Show filters' and 'Advanced editor'. A 'Info' section shows details like source_id, question_id, and question_text. The main workspace displays a quote: "In my local area, rising sea levels have led to coastal erosion. Coastal erosion has caused house prices to drop." This quote is highlighted in blue, indicating it's selected. A vertical sidebar on the right has a 'Report a bug' button.

For more information on how to edit an existing link [see here](#).

About the factor label dropdown menus

By creating links, you also create the names of your factors.

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

Factor names which contain semicolons ; get special treatment as they separate the different parts of hierarchical labels.

After beginning to create links between factors, already-coded factors will appear in the dropdown menus in the to and from factor boxes. For added convenience. The most frequently coded factors will appear at the top of this list

Editing factors and links (basic)

{#xedit-factor-and-links1}

It is likely that you will need to regularly review and edit your links, and there are various ways to do this in the app. You can also delete factors and links in the same way, however you cannot delete factors or links when certain filters are applied such as `combine opposites` and `zoom` as these filters represent multiple factors as one factor label.

Editing links in the interactive tab

Click on the link you want to change in the interactive view. 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 to finish editing.

The screenshot shows the 'interactive view' interface. At the top, there is a code block:

```
source_id MNX-1  
question_id2 F3  
question_id SF3  
question_textglobal question
```

Below this is a text area with a red highlight:

“
It has minimized as we do not have a means of earning funds to improve our lives in as much as we want to change a few things.
”

On the left, there is a toolbar with icons for canceling, saving, and other operations. Below it are two input fields: one for the source factor (~Income) and one for the target factor (Decreased ability to purchase goods). At the bottom, there is another text area with the same highlighted text and a toolbar below it.



Editing factors in the interactive tab

To edit your factor labels in the interactive view, just click on the factor and then ‘edit factor’ to open the edit panel. Using this panel you can change the text of the individual factor or add a memo.

What do you want to do with this factor?

 Delete from whole map ▾

Focus on this factor

Edit this factor across the whole map

Label

Decreased ability to purchase goods

Memo

Finance

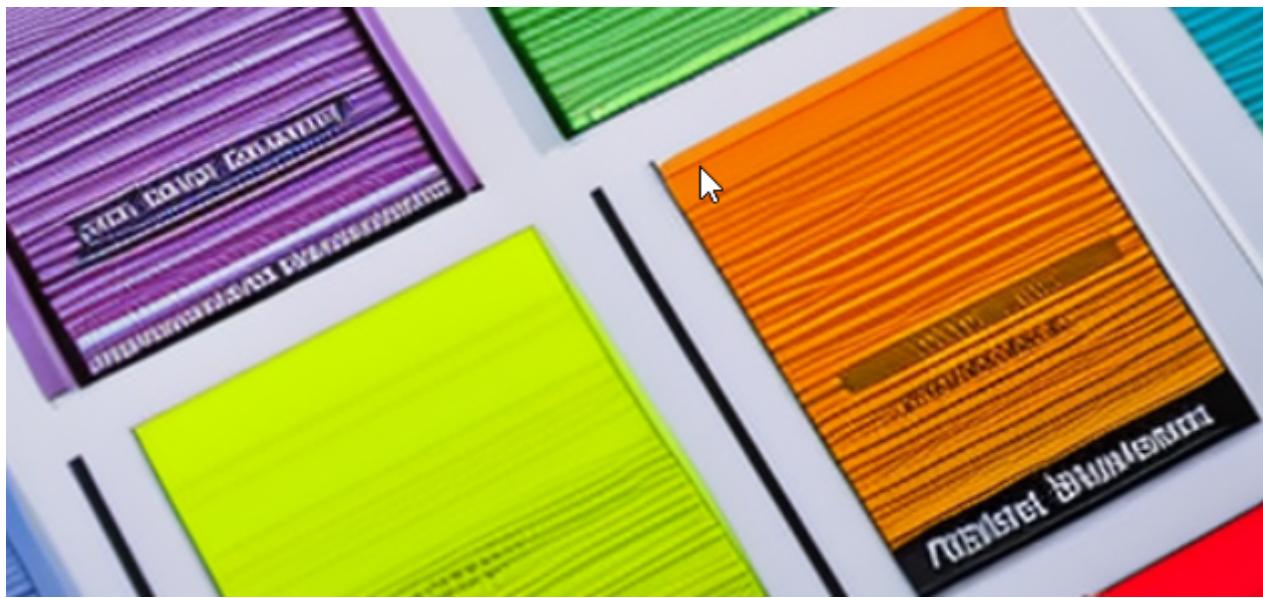


[Dismiss](#)

You can also edit multiple factors using the [tables](#) and the [factor editor tab](#).

!--chapter:end:20300-editing-factors-and-links1.rmd--

Factor labels: introduction{#factor-labels-0}



Many people like to start coding simply by choosing whatever factor labels occur to them, as far as possible using text directly taken from the statements they are coding. That's one good way to start.

```
knitr:::include_url("https://player.vimeo.com/video/580212681")
```

In the following pages we will give you some more tips on making your factor labels work for you.

::: {.rmdcaution}

Remember that in general, factor names are case sensitive. So the app will not treat 'funding' and 'Funding' as the same.

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 the app will then treat this factor label as hierarchical.

:::

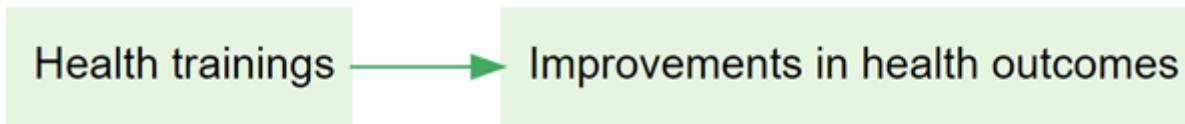
!--chapter:end:20370-factor-labels-intro.rmd--

Factor labels: actual facts?{#facts}

Causal mapping is very flexible and accommodates different philosophies or ways of coding. However, it's good to be aware of this flexibility and perhaps avoid *mixing* philosophies.

For example, you can code in these two different ways, but it's probably better not to mix them:

- To show only causal influences between factors, without recording what did or does happen. Seen in this way, this map simply says that health trainings have the power to improve health outcomes, without any information about what actually happened or is happening.



- Or you can decide that you will only create this kind of link when someone claims that this actually happened: Not only that the first thing is something which has the causal power to influence the second thing but also that in some sense both things actually happened.

!--chapter:end:20375-factor-labels-coding.rmd--

Factor labels: actor-focused{#xfactor-labels-actors}

Often you will find that it is easier and clearer if you formulate your labels to highlight who does what.

So instead of this:

advocacy on renewables

try this:

NGOs conduct advocacy campaign on renewables.

In particular, avoid using passive formulations like "Advocacy is carried out" or "Awareness is raised".

!--chapter:end:20380-factor-labels-actors.rmd--



Factor labels: semi-quantitative{#factor-labels1}

"Semi-quantitative' 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 $B \rightarrow E$ is supposed to mean that B exerts in some sense an “increasing” or “decreasing” influence on E, then both B and E need to be formulated in a corresponding way. In order to ease interpretation, with a few exceptions, factors should be labelled and understood in such a way that it makes sense to say “more of this” or “this happened as opposed to not happening”: we call these semi-quantitative factors.

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

Examples of semi-quantitative factors

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

- Sold cow
- Earthquake happened
- (Had) good harvest
- (Level of) bank account

- (Level of) ethnic tolerance
- Quality of seeds

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

Non-quantitative factors

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

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

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

!--chapter:end:20391-factor-labels-semi-q.rmd--

Factor labels: using hashtags{#xfactor-hashtags}

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

```
knitr:::include_url("https://player.vimeo.com/video/671894620")
```

Hashtags are useful in either of these cases.

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

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

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

More food --> Improved wellbeing

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

More food #nutrition --> Improved wellbeing

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

More food --> Improved wellbeing (Outcome)

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

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

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

!--chapter:end:20400-factor-labels-flags.rmd--

Level 2 quiz

Now do the [quiz](#) to see if you've understood about basic coding.

[!--chapter:end:20490-quiz2.rmd--](#)

(PART*) 3: Basic analysis{-}

[!--chapter:end:30500-PART3-basic-analysis.rmd--](#)

Level 3 overview{#xlevel3}

In the pages in this section you will learn how to do basic analysis in Causal Map: how to produce maps and tables to answer your research questions.

When you've read the pages, try the [quiz](#) to see if you have understood enough or need to review some things!

[!--chapter:end:30510-level3.rmd--](#)

Analysing a file{#xabout-analysis}

Analysing a file means

- **applying different filters** to it, to focus on different parts of the entire causal map in order to answer different questions
- **viewing the results in different ways** by clicking on different tabs in the right hand side to view different kinds of graphical map and various [tables](#).

If you want to view / analyse a file, you will mostly:

- apply filters using [the Dashboard](#) and view the results in the right-hand side of the app.

- view many statements at a time by clicking `Many` on the multiple statement selector (above).
- start off by analysing the [example-file](#).

Background information about filters

::: {.focus}

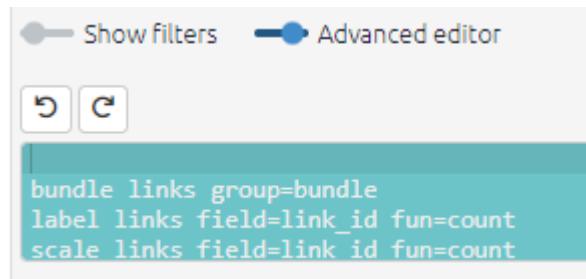
The information in this section is just for orientation, you don't need to remember it because you will probably find it much more convenient to use the dashboard.

:::

Filters simplify your graph to help you answer specific questions. You will often apply more than one of these filters in order, e.g. first select only specific factors and then make the most frequently mentioned factors larger. There are different kinds of filters:

- **Analysis filters**, e.g. to exclude some links and/or factors from the map. So instead of seeing the entire unfiltered causal map, which can be quite bewildering, you see only part of it.
- **Conditional formatting filters**. You can format the links and factors in your map to reflect the data, e.g. you can make factors bigger if they were mentioned more often.
- **Simple formatting filters**. These are simply used to apply overall formatting, e.g. give all the factors a grey border.

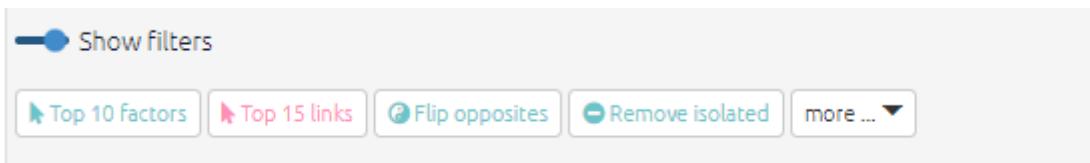
The filters are actually specified by special text in the Advanced Editor (which you can view if you want, but close the window if it scares you). The text window uses a simple syntax for filtering and manipulating the maps and tables. Each line in this window is a filter which manipulates the existing 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.



You won't need this window at this stage.

You can use the other buttons on the left-hand side to manipulate this filter text more easily. Each button corresponds to one line.

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



But for analysing your map, you may want to open the filters panel using the toggle.

If you are just beginning with analysis, you will probably just use the [Dashboard](#) -- it's easier!

!--chapter:end:30530-about-analysis.rmd--

The Dashboard{#xdashboard}

Your medals!

- 1555
- 82
- 5
- 53
- 3
- 10
- 3348
- 38

Load uncoded example Help

Complete views (replace existing view)

- 10 - Overview 2648
- 11 - Overview with smart zoom and combined opposites 2316
- 12 - Overview with combined opposites 2292
- 14 - Contrary factors 2294
- 28 - Quotes 2296
- 30 - Factor mentions 2755
- 31 - Factors zoomed to level 1 2757
- 34 - Factors - codebook 2299
- 39 - Closed questions, per source 2653
- 50 - Trace from main drivers to main outcomes 2682
- 51 - Consequences of top 3 drivers 2683
- 52 - 1 step down from main drivers 2300
- 53 - Influences on main outcomes 2684
- 54 - 1 step up from main outcomes 2301
- 60 - 3 clusters + bundle by cluster + surprise 2302
- 62 - Maximum betweenness + up 1 + down 1 2303
- 64 - Top 3 drivers 2305
- 65 - Top 3 outcomes

Addons (added to existing view)

- Combine opposites Contrary factors get orange text
- Exclude links with only 1 mention Hide plain codings
- Hide quickfields Links with more than one mention
- Pink links - fixed colour Remove isolated Smart Zoom, preserve 5
- Standard format Standard style Top 10 factors Top 15 links

Save the current view

Title for this view Rec? Addon?

Jump to a shortlink

Shortlink nr.

Community development

The Dashboard appears on the right as the app loads. You can also reach the dashboard by pressing the 🏅 icon on the right, or with the key combination Alt+q.

The Dashboard lists views of the current file saved by you or other users.

Each view usually:

- applies a particular **filter**, e.g. "show me only causal claims made by women".
- and/or it may take you to a particular **tab** on the right hand side -- the Print map, one of the Tables, etc., including, in the case of tables, the way the table is set up as well. A view also saves the name of the current file.

To load one of the views, just click on the item.

You can also go to a view if you know its unique number. So to go to view 369, type 369 into this box

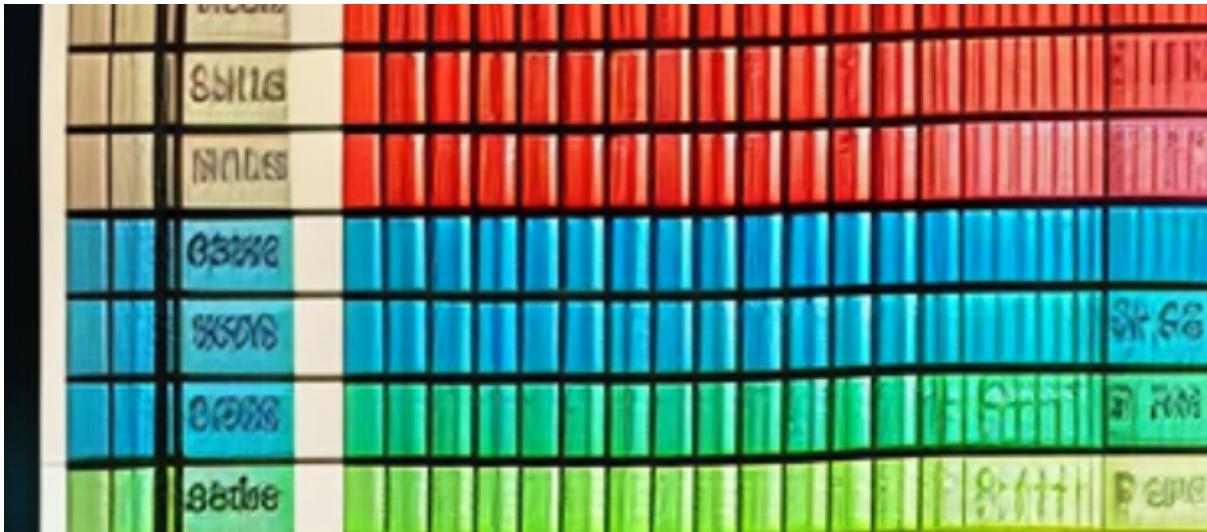
Apply a view

 and click the green arrow button.

(Maps which have been downloaded as an image often have a unique number like this in the bottom left-hand corner.)

!--chapter:end:30561-dashboard.rmd--

Tables{#xtables-basic}



Features common to all the tables{#tables_common}

Main controls

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

unfiltered filtered

When the table is set to "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.

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

Search

You can search / filter the whole table by clicking on the equals sign at the top of a column. To search numerical volumes you can type in the value you are searching for or you can use > and < to search for values less than or greater than a set number.

103 rows 

Click the = signs to filter, click the headers to sort. You can also search 'blanks' or 'notblanks'. Click the pink '...' to edit

=	=	=	=
label	factor_memo	source_count	...
Improved health	Health and hygiene	18	...
Increased knowledge; Farming method/practice	Agriculture	17	...
Diet improved; Diversified	Nutrition	17	...
Farm production	Agriculture	16	...
Income	Finance	16	...
Food consumption quantity	Nutrition	16	...
Planted new crop/vegetable varieties	Agriculture	15	...
Improved/new farming techniques; Cultivating in lines	Agriculture	15	...

Sorting

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



The Factors table{#xthe-factors-table}

The factors table presents the factors applied during coding. There are columns for label, factor memo, and columns for any other fields you have provided, maybe "domain" for instance. The app also calculates some additional fields like:

- `betweenness` (how central is the factor in the map?)
- `in_degree` (how many links come in to this factor?)
- `out_degree` (how many links come out of this factor?)

This table is an overview which can help us to understand which factors are reported most frequently, and to view and understand metadata like `memos`. As usual, this table responds to any filters you have applied on the left-hand side and can be additionally sorted and filtered.

The factors table has two presets:

- The `codebook` preset which lists all factors used in the file without any metrics, this is useful for reports.
- `Top factors` shows you the factor label and source count so you can easily see which factors were most mentioned.

As with all tables, you can create your own table. For example, sometimes it can be useful to break down hierarchical labels into first, second, third levels and so on. By clicking adding levels in the columns box you can view a break down of factors' hierarchies.

Table

factors mentions links statements sources questions closed_questions metrics unfiltered filtered Presets

[About this table](#) [How to use the tables](#)

Columns

label	factor_memo	source_count
level_1_label	level_2_label	
level_3_label		

Rows

label	
-------	--

Pivot column

--

133 rows

Click the = signs to filter, click the headers to sort. You can also search 'blanks' or 'notblanks'. Click the pink '...' to edit

label	factor_memo	source_count	level_1_label	level_2_label	level_3_label
Improved health	Health and hygiene	18	Improved health		...
Health behaviour; Use pits to dispose of rubbish		18	Health behaviour	Use pits to dispose of rubbish	...
Increased knowledge; Farming method/practice	Agriculture	17	Increased knowledge	Farming method/practice	...
Diet improved; Diversified	Nutrition	17	Diet improved	Diversified	...
Community groups/learning		17	Community groups/learning		...
Farm production	Agriculture	16	Farm production		...
Income	Finance	16	Income		...
Health behaviour; Use mosquito nets		16	Health behaviour	Use mosquito nets	...

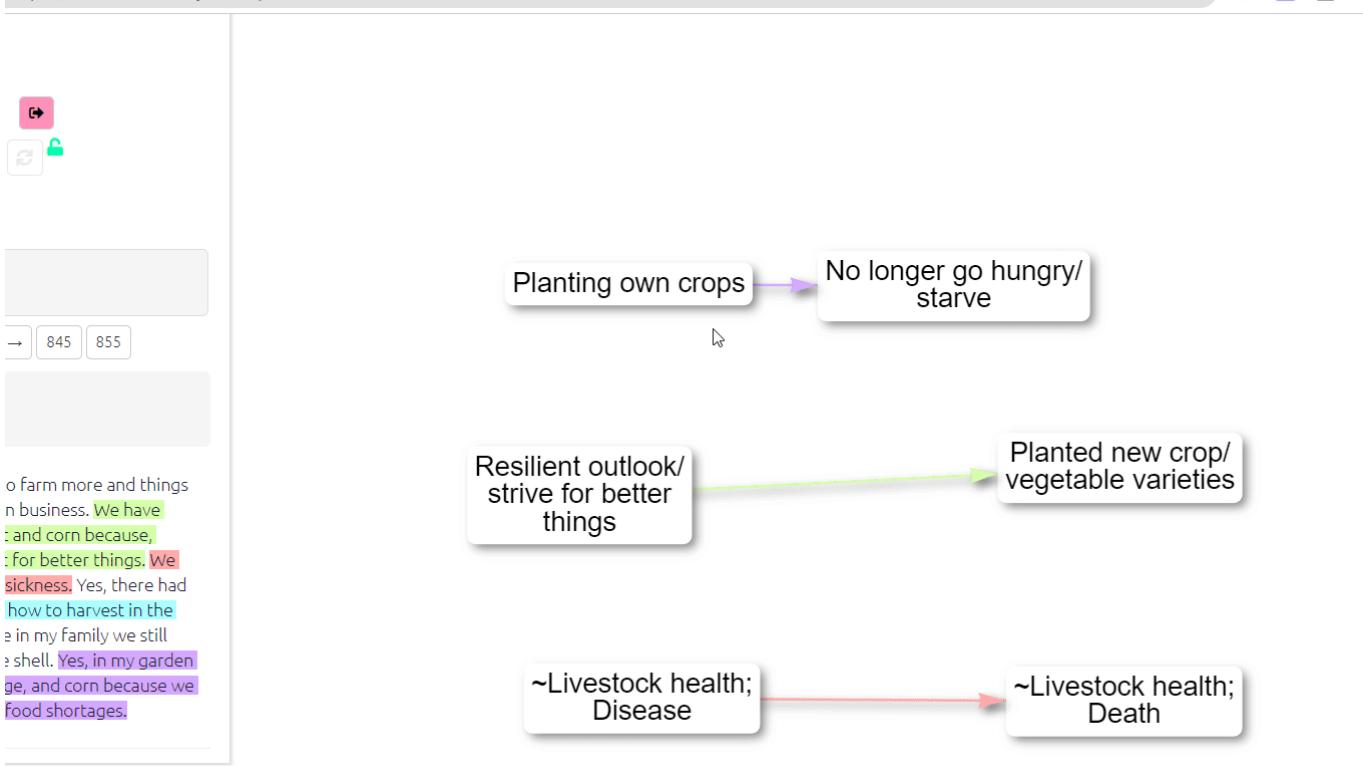
Factor memos{#xfactormemos}

You can include a lot of information about your factors directly in the factor label, especially using hashtags. But sometimes you want to store information separately from the label, perhaps because you are not interested in ever seeing this information displayed in maps. For example, you might want to record who created this factor -- you or a colleague.

Factor memos are an important part of your Codebook. They appear:

- when you hover in interactive view
- in the sidebar in the factor editor
- in the factors table.

You can create and edit the memos from interactive view:



... or from the factors table.

You can find factors with particular memos in the factors table or using filters:

Find factors

[Apply Filter!](#)

Keep this panel open?

Include factors how many steps up?

0 10

Include factors how many steps down?

0 10

What field to search in?

factor_memo

Type to find a value or values for this field

SP

Add SP...

contains

Active

!--chapter:end:30680-factors.rmd--

The Mentions Table{#the-mentions-table}

This table is useful when looking at how many times factors have been mentioned in links. In the [Factors table](#), you can display how many links are coming in and out of each factor overall, but you can't for example ask how many different *sources* mentioned it, or in how many different districts.

It is likely that you will want to view this table using the preset **Mentions (basic)** which can be found in the dashboard.

Table

The screenshot shows a Tableau dashboard with a top navigation bar featuring tabs for factors, mentions, links, statements, sources, questions, closed_questions, and metrics. There are also buttons for unfiltered (unchecked) and filtered (checked), and a Presets button. Below the navigation is a toolbar with links for 'About this table' and 'How to use the tables', and a refresh icon.

The main area is a pivot table interface. The 'Rows' section has a dropdown set to 'label'. The 'Columns' section has a dropdown set to 'label direction', which is further expanded to show 'direction' (selected), '...', and 'Unpivot'. The 'Pivot column' section has a dropdown set to 'direction'. The 'Summary' section has a dropdown set to '---'.

A 'Fill the cells with' section contains two dropdowns: 'the count (number of distinct values)' and 'of this field: source_id'.

The table itself has 133 rows, as indicated by the text '133 rows' below the table. A note says: 'Click the = signs to filter, click the headers to sort. You can also search 'blanks' or 'notblanks'. Click the pink '...' to edit'.

direction			
=	=	=	=
label	consequence	either	influence
Improved health	18	18	5
Health behaviour; Use pits to dispose of rubbish	3	18	16
Increased knowledge; Farming method/practice	1	17	16
Diet improved; Diversified	17	17	5
Community groups/learning		17	17
Farm production	15	16	16
Income	16	16	15
Health behaviour; Use mosquito nets	4	16	14

The Direction field shows you the number of times each factor was reported as an influence factor and/or a consequence factor.

- **Influence** = how many times the factor lead to another factor.
- **Consequence** = how many times the factor was mentioned as a result of another factor.
- **Either** = how many times the factor was mentioned each direction.

Remember, in each case these counts depend on what you put in the Count box. You can count link_id (which gives us the overall count, the highest number) or statement_id, source_id, district etc (which usually give a lower number).

As usual you can change what is counted by selecting a different field in the Count box. You can also change what is displayed by deleting or adding columns and rows.

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 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 when doing causal mapping we are most interested in the relationships between factors.

!--chapter:end:30690-mentions.rmd--

The Links Table{#xthe-links-table}

This table counts factors differently to [the mentions table](#) as it counts links rather than individual factor mentions. Each row represents one link and has columns for that link's influence factor and consequence factor, as well as many other columns. This table is also linked to the [statements](#) and [sources](#) tables. Use this table to show the links between factor labels. Information like the related quote is also available in this table.

This table has three available presets:

- `Links (basic)` as shown below, which gives you information about each individual link
- `Frequent bundles` which shows the `source_frequency` for each link bundle
- `Hashtags` which shows the hashtags used and their `source_frequency`

Table

Factors: mentions, links, statements, sources, questions, closed_questions, metrics

unfiltered (radio button) filtered (radio button) Presets

About this table How to use the tables

Columns

Rows

link_id	from_label	to_label	quote	source_id	hashtags	simple_bundle
1	~Livestock health; Disease	~Livestock health; Death	We have nothing, the chicken we had all died due to sickness.	MNX-1	~Livestock health; Disease / ~Livestock health; Death	...
3	~Improved health	Unable/less able to work on farm	[Quantity of food decreased because...] It has decreased because in my family we do get sick, my husband and I likewise for my husbands other wives. We all get sick and when this happened everything is affected and put on hold. [Income decreased because they were all sick and unable to plant.]	MNX-1	~Improved health / Unable/less able to work on farm	...
4	Member of savings group	Increased knowledge; Finance	My family earns money only because of the savings that Organisation 1 has provided for us. They taught us how to save money. My husband has his own business. Without getting along well nothing has changed we live in the same way. [...]The earnings we obtain from agriculture has not changed at all, and everything we have is due to business that my husband does and also through savings. My husband sells sieves, brooms and much more and it is through this money that we make a living.	MNX-1	Member of savings group / Increased knowledge; Finance	...
5	Member of savings group	Increased ability to save/increased savings	My family earns money only because of the savings that Organisation 1 has provided for us. They taught us how to save money. My husband has his own business. Without getting along well nothing has changed we live in the same way. [...]The earnings we obtain from agriculture has not changed at all, and everything we have is due to business that my husband does and also through savings. My husband sells sieves, brooms and much more and it is through this money that we make a living.	MNX-1	Member of savings group / Increased ability to save/increased savings	...

857 rows
Click the = signs to filter, click the headers to sort. You can also search 'blanks' or 'notblanks'. Click the pink '...' to edit.

Fields

There are many fields you can use as columns or rows, such as:

- `simple_bundle`. Shows the whole link, both the influence and consequence factor.
- `from`. Contains numbers which match the `factor_ids` of the factors.
- `to`. Contains numbers which match the `factor_ids` of the factors.
- `statement_id`. Contains numbers which match the `statement_ids` of the statements.
- `link_memo`. For making notes about a link as you code.
- `hashtags`. Re-usable tags which are usually common across several links, for example `#hypothetical` or `#certain` or `#check`.

!--chapter:end:30710-links.rmd--

The Statements Table {#xthe-statements-table}

This table makes searching for a specific statement easy. By clicking on a row in this table you will jump straight to the statement, meaning you can edit the associated factors in the left-hand-side panel. This table shows you all the statements in your map file. You can use the search bars to look for statements by their id or by a word/phrase in the text.

There are two presets for this table:

- `Statements (basic)` this will automatically load when you click on this tab and displays statement id, text, source_id and question id
- `Statements metadata` this preset gives you more information about who said which statement and shows you household code, name of province , village , age as well as the fields listed above. If your metadata fields are called something different just add them to the columns box to display them.



Table

Factors mentions links statements sources questions closed_questions metrics unfiltered filtered Presets

[About this table](#) [How to use the tables](#)

Columns

Rows

statement_id

text source_id question_id
statement_id

Pivot column

855 rows

Click the = signs to filter, click the headers to sort. You can also search 'blanks' or 'notblanks'. Click the pink '...' to edit

statement_id	text	source_id	question_id
1	This is an example file to showcase the different functionalities in the Causal Map App.	MNX-1	\$A8
2	In my family we are considerably healthy because people get sick from time to time. Even this year we got sick, although we do not visit the hospital we do get better eventually. Instead we go to church as we are prohibited of going to the hospital because of cultural issues, and beliefs (they believe church are more serious and that some diseases are linked to spiritual problems) and also going to the hospital takes longer time to be assisted. We do get sick because it is the way God made things to occur we do not know why. We try to change but we still do get sick even though we make use of mosquito nets. We also dig holes in which we put litter and we also have better latrines/toilets. We take three baths per day, and also wash our hands after visiting the washrooms. We do use mosquito nets. Yes, there are. We previously took a bath per day, and we would go to the bushes to alleviate our selves. We changed our behaviours because we learned from Organisation 1 that the way one behaves contributes to a health status. Caring for our health is considerably important. Men and women in my house hold mostly use the latrines/toilets and we take three baths per day.	MNX-1	\$B1
3	Improved	MNX-1	\$B2
4	This year things are getting worst, we often get sicker. I cannot explain why only God can explain as sickness was created by him. Our churches prohibited us from visiting the hospital and as a result five of my children are dead.	MNX-1	\$B3
5	Not much has changed as we have the strength to farm more and things have not changed regardless of my husbands own business. We have altered because in the last year we planted millet and corn because, regardless of the difficulties in life we try to fight for better things. We have nothing, the chicken we had all died due to sickness. Yes, there had been, Organisation 1 taught us how to plant and how to harvest in the best manner. It has not changed as much because in my family we still arrange the barn and sift corn after removing the shell. Yes, in my garden in general we plant tomatoes, vegetables, cabbage, and corn because we grow crops in our farm and that way we have no food shortages.	MNX-1	\$C1
6	Decreased	MNX-1	\$C2
7	It has decreased because in my family we do get sick, my husband and I likewise for my husbands other wives. We all get sick and when this happened everything is affected and put on hold.	MNX-1	\$C3

Remember the print view button may come in handy to view your table as plain text.



Fields

- statement_id. A running number from 1 upwards.
- source_id. A code which matches the field source_id in the sources table.
- question_id. A code which matches the field question_id in the questions table.
- statement_memo. For making notes about a statement as you code.

!--chapter:end:30720-statements.rmd--

The Sources Table{#xthe-sources-table}

The sources table, when there are no filters applied or when viewing the preset filter `Sources (basic)`, simply lists all of the respondents and any additional data collected about them (sex, age, location, education). As with all the tables 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.

This table can be useful to check that all the sources have been uploaded 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.

Table

Rows: source_id

Columns:

source_id	Household code	#Name of province	#Name of village	#Age of the main respondent
MNX-1	MNX1	A	1	20-45 years
MNX-2	MNX2	A	1	20-45 years
MNX-3	MNX3	A	1	46+ years
MNX-4	MNX4	A	1	46+ years
MNX-5	MNX5	A	1	46+ years
MNX-6	MNX6	A	1	46+ years
MNY-1	MNY1	A	1	46+ years
MNY-2	MNY2	A	1	20-45 years

19 rows
Click the = signs to filter, click the headers to sort. You can also search 'blanks' or 'notblanks'. Click the pink '...' to edit

source_id Household code #Name of province #Name of village #Age of the main respondent

MNX-1 MNX1 A 1 20-45 years ...
MNX-2 MNX2 A 1 20-45 years ...
MNX-3 MNX3 A 1 46+ years ...
MNX-4 MNX4 A 1 46+ years ...
MNX-5 MNX5 A 1 46+ years ...
MNX-6 MNX6 A 1 46+ years ...
MNY-1 MNY1 A 1 46+ years ...
MNY-2 MNY2 A 1 20-45 years ...

11 more rows ...

Fields

There are many fields you can use as columns or rows, such as:

- `source_id`. A code which matches the field `source_id` in the `statements` table.
- `source_memo`. For making notes about a source as you code.
- Any other fields which you have uploaded.

!--chapter:end:30750-sources.rmd--

The Questions Table{#xthe-questions-table}

This table displays the text, ID, `memo` and map id for every question included in the questionnaire.

This table has one preset:

- `All questions (basic)` which displays id and text as shown below



Table

factors mentions links statements sources **questions** closed_questions metrics unfiltered filtered Presets

[About this table](#) [How to use the tables](#)

Columns

Rows **question_id** question_text Pivot column

question_id	Question HH	Pivot column
question_id		

69 rows

Click the = signs to Filter, click the headers to sort. You can also search 'blanks' or 'notblanks'. Click the pink '...' to edit

=	=	=	=
question_id	Question HH	question_text	
#A1	Household code	global question	...
#A2	Name of district	global question	...
#A6a	Age of the main respondent	global question	...
#A6b	Household size (residents > 6months)	global question	...
#A6c	Sex of the respondent	global question	...
#A7	Sex of the head of the household	global question	...
\$*B2	Overall, how has the health of your family changed over the past year?	global question	...
\$*C2	Overall, how has the amount of food your household produces for home consumption changed in the last year?	global question	...
\$*D2	Overall, how has the amount of money/food your household earns changed in the last year?	global question	...
\$*E11	Overall, has your (or pregnant/breastfeeding women's) diet in the last year...	global question	...

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

!--chapter:end:30760-questions.rmd--

The Closed Question Blocks Table{#xclosed-table}

The closed question table provides 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

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

There are two presets with this table:

- Closed questions by source (basic) gives an overview of how each **individual** respondent answered each closed question, which you can see below.
- Closed questions summary 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.



Table

Factors mentions links statements sources questions closed_questions metrics unfiltered filtered Presets

About this table How to use the tables

Rows Columns Pivot column Summary

question value value Total Unpivot

Fill the cells with the count (number of distinct values) of this field: source_id

12 rows
Click the = signs to filter, click the headers to sort. You can also search 'blanks' or 'notblanks'. Click the pink '...' to edit

question	value	value			
		=	=	=	=
-	+	0	Total		
Overall, how has the health of your family changed over the past year?		19		19	...
Overall, how has the amount of food your household produces for home consumption changed in the last year?	5	14		19	...
Overall, how has the amount of money/food your household earns changed in the last year?	3	16		19	...
Overall, how has the quantity of your household's food changed in the last year?	2	17		19	...
Overall, how has the variety of the different foods your household eats changed in the last year?	4	14	1	19	...
Overall, how has the amount of food your household eats during the hungry season changed in the last year?	3	16		19	...
Overall, do you think the diet of your children in the last year has...		17	2	19	...
Overall, has your (or pregnant/breastfeeding women's) diet in the last year...		15	4	19	...
Overall, how has the amount you can buy as a household changed over the last year?	2	15	2	19	...
Overall how have relationships between men and women in your household changed over the last year?	2	15	2	19	...
Overall, how do you think the way the community works together has changed over the past year?		17	2	19	...
	21	175	13	209	...

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.

More information on how the app recodes your closed question responses can be found [here](#).

!--chapter:end:30770-closed.rmd--

The Metrics Table{#xmetrics}

Summary

The metrics table displays a range of useful measurements about your data. These metrics respond to any filters that have been applied to your data, so consider whether you want to see these metrics `unfiltered` or `filtered`.



Table

filters mentions links statements sources questions closed_questions metrics unfiltered filtered []

[i] About this table [i] How to use the tables [] 🔍

Rows	Columns	Pivot column
name	metric description	

15 rows

Click the = signs to filter, click the headers to sort. You can also search 'blanks' or 'notblanks'. Click the pink '...' to edit

=	=	=
name	metric	description
size	857.00	The number of links in the map.
order	133.00	The number of factors in the map.
links_per_factor	6.44	The number of links per factor.
n_statments	855.00	The number of statements.
n_words	34,975.00	The number of words in all statements.
n_coded_words	24,260.00	The number of words in unique quotes.
n_characters	189,127.00	The number of characters in all statements.
links_per_statement	1.00	The number of links per statement.
percent_coded	31.00	The percentage of statements with codings.
adhesion	0.00	The minimum edge connectivity.
component_count	8.00	The number of unconnected components in the map.
diameter	6.00	The length of the longest geodesic.
reciprocity	0.00	The proportion of mutual connections in the map.
min_cut	0.00	The minimum number of links to remove in order to split the map into two clusters.
mean_dist	2.52	The mean distance between all factor pairs in the map.

What the metrics mean

There are quite a few metrics, some are more self-explanatory than others. The table below expands on the brief description given in the metrics table.

🎓 Metrics with this "mortar-board symbol" are well-known network metrics.

Metric	Explanation
size	The total number of links in the map.
order	🎓 The total number of factors in the map.
links_per_factor	The average number of links per factor. Usually, you want to have plenty of links per factor.
n_statements	The number of statements in the map.
n_words	The total number of words in all the statements.
n_coded_words	The number of words in coded phrases.
n_characters	The number of characters in all statements.
links_per_statement	The number of links per statement.
percent_coded	The percentage of statements which have been coded.
component_count	🎓 The number of unconnected components in the map. Do you have islands in your map? The app first identifies the components -- sets of factors which are all ("weakly", i.e. ignoring the direction of the links) connected to each other by some path; then counts how many separate components there are, i.e. none of the factors in one component are connected to any of the factors in the other component(s).
diameter	🎓 The length of the longest geodesic or path between two factors. For example, the count for the very simple map consisting only of A->B->C->D would be three, as it's three links from A to D. This takes into account the direction of the arrows.
reciprocity	🎓 The proportion of mutual connections in the map. This calculates the proportion of factors in which both factors influence one another. The connection must be between two factors, feedback loops are not counted.
min_cut	🎓 The minimum number of links that you would have to remove to split the map into two clusters.

Metric	Explanation
mean_dist	🎓 The mean number of links between all factors in the map.

[!--chapter:end:30780-metrics.rmd--](#)

Level 3 quiz

Now do the [quiz](#) to see if you've understood about basic analysis.

[!--chapter:end:31000-quiz3.rmd--](#)

(PART*) 4: Advanced coding{-}

[!--chapter:end:40000-PART-advanced-coding.rmd--](#)

Level 4 overview{#xlevel4}

In the pages in this section you will learn how to do advanced coding in Causal Map:

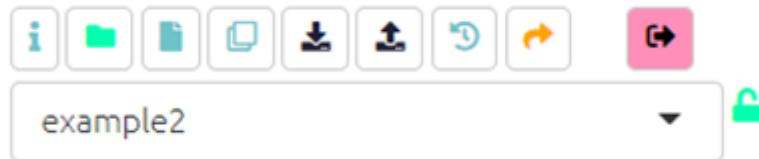
- using [memos](#)
- speeding up your work
- using the Factor Editor
- and lots more.

When you've read the pages, try the [quiz](#) to see if you have understood enough or need to review some things!

[!--chapter:end:40001-level4.rmd--](#)

Managing your file: The top menu{#xtop-menu-advanced}

This section is all about how to manage using the top menu to help you deal with the currently loaded file. Hover over the icons to see what they do.



There is also a dedicated [File Manager](#) to manage all your files.

Sharing files with the Current File Manager{#xpermissions}

Click the green icon to manage your file.



If you rename a file you lose its previous versions and you will not be able to restore them!

file memo

An example file (which has been already coded) for you to view.

view

copy

steve@pogol.net

edit

hannahmishan@gmail.com
steve@pogol.net

Creator (cannot be changed): steve@pogol.net

locked

archived

Save

Sharing and locking files

You can add any number of other Causal Map accounts into the "copy" and/or "view" boxes. There must be at least one editor.

- "Copy" permission means that the other user can make their own copy of this file, which they can then edit.
- "View" permission means that the other user can make only view the file.

You can also use the `locked` switch to lock the file. This means that no-one can change it, not even you, until it is unlocked again.

It is also possible to create larger groups of users. Ask us for help with this:
hello@causalmap.app.

File memo{#xfi}

You can put information about your file in the `file memo` box, and it will be shown below the interactive map to tell other users what this file is about. You can use markdown to format the text, e.g. start a line with # to make level-1 heading. You can even include images using markdown.

The screenshot shows the Causal Map interface with the 'file memo' section highlighted. The 'Name' field contains 'sptestagain'. Below it is a note: 'If you rename a file you lose its previous versions and you will not be able to restore them!'. There are 'Load' and 'Delete...' buttons. The 'file memo' section contains the following markdown:

```
# This is a headline  
  
And this is an ordinary paragraph.  
  
And this is a picture: 
```

On the left, there are sections for 'view', 'copy', and 'edit', each with a text input field. The 'edit' field contains the email 'steve@pogol.net'.

This is in addition to any "newsflash" information about updates which you add using the chat widget.

Making your own copy of a file{#xown-copy}

```
knitr:::include_url("https://player.vimeo.com/video/641927229")
```

Copying a file in Causal Map is easily done. Simply load the file you want to copy using the dropdown menu. Then click the 'Save As' button on the top left of your screen.



You can then append something to the end of the filename to make it yours:

Save a version of this map

You don't need to save, your map is being saved constantly in the background.

But you can save a version under a new name.

Name of new map (will be added to existing name)

Click the 'Save current map' button. This window will then close and your copy of the file will load up. You can now make changes to the file.

Saving a file as currently filtered{#as-currently-filtered}

```
::: {.rmdcaution}
```

Suppose you have applied a filter so that only statements from women are visible. Saving the file as currently filtered means that only these statements remain in the file, and the invisible statements and links are completely deleted from the file. If this is what you want, here's how to do it!

```
:::
```



Or: Save map as currently filtered without changing name

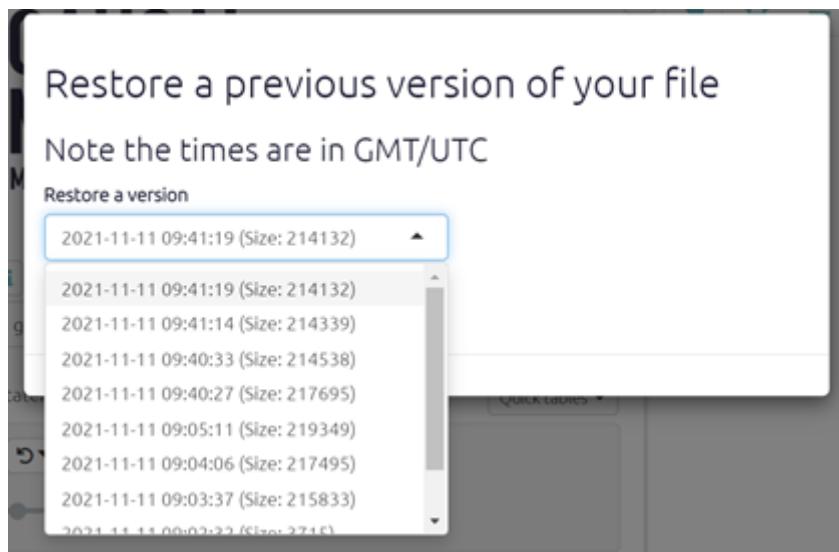
Then press:

Save current map as currently filtered

That's it.

Restoring a previous version of your file{#xversions}

The app is continuously saving your work and so you can restore your file to any previous version, by clicking this icon . This will open the below panel where you can choose which timepoint you wish to revert your map to.



This panel shows a dropdown list of times when you made changes to the mapfile in GMT. Along with the size of your file which can help you identify which timepoint you want to revert to. It can be easy to forget what time you made alterations to your file, so if you're likely to want to restore a previous map it is best to note the time so that you can easily return to it.

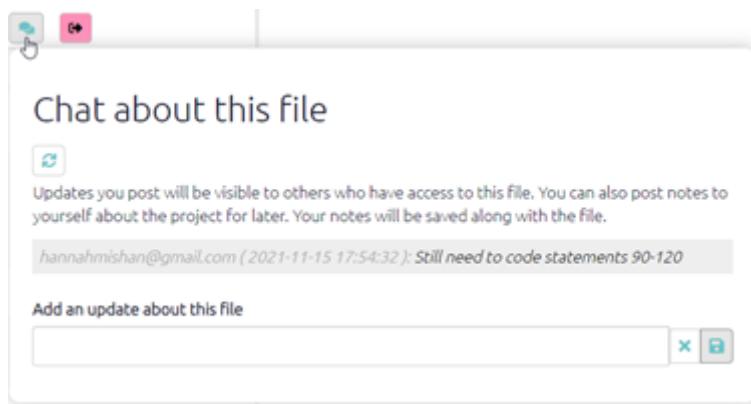
Uploading your data



See [this section](#) for more information.

Chat about this file

By clicking on the icon of two speech bubbles, a panel containing chat about the file will appear. You can see the comments you and other people with access to this file have added. From here you can also add updates on this file, this function is useful when working collaboratively.



Why did I make this version of that file? Is it different from version xzy? Did we already upload those new statements into this version??

This panel can help with these kinds of problems. You and your colleagues -- anyone else with access to the current file -- can type notes and updates about the file. These updates will be listed here sequentially.

Also, these same notes and updates are listed in the main [Updates panel](#), along with automatic updates about when files were opened and edited.

You can just search for type=chat to view recent chat about any files or some particular files:

Updates about your files

Show...

Most recent status for each file All updates for current file All updates for all files



Click a row to load the corresponding file.

[Copy](#) [CSV](#) [Excel](#)

Showing 1 to 10 of 32 entries (filtered from 17,045 total entries)

Show 10 entries

Previous

1

2

3

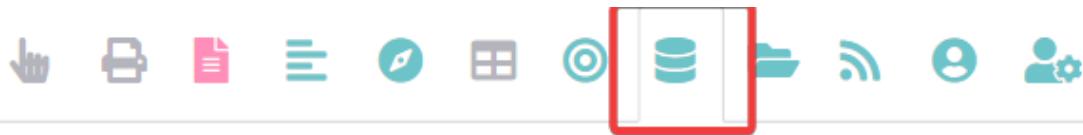
mapfile	user	server	type	content	time_stan
All	All	All	chat	All	All
hannahupload	hannahmishan@gmail.com	CM2test	chat	Still need to code statements 90-120	2021-11-15
odi-plus-minus-standard-tilde	steve@pogol.net	CM2test	chat	this is still usable, I made a kinda transition to tilde notation	2021-10-31

!--chapter:end:41020-top-menu-advanced.rmd--

Data manager {#xdata-manager}



Is your data all uploaded correctly? Does it all make sense? Try the Data Manager tab.



Data manager

Sources in statements table not in sources table:

none

Links which seem to have quotes which are not ac

2, 5, 6, 10, 12, 23, 24, 25, 33, 39, 46, 50, 51, 54, 56, 64, 67, 68, 77, 98, 100, 104, 109, 1
201 223 224 225 227 233 234 235 271 272 273 278 279 287 296 311 313 3

It is good to check this page to ensure your data is in the correct place, especially after you have uploaded your data.

Sources in statements table not in sources table:

You need to check your sources table to see that there is a row for each one of these IDs. You have statements which are supposed to come from these sources but you need to add information about these sources. Fix this by roundtripping: simply fill in the missing `source_ids`.

Links which seem to have quotes which are not actually contained in the corresponding statements:

Quotes associated with each link should always be verbatim sections of the original statements. Sometimes they can lose this association, for example if you accidentally edited a quote to make it different from the original statement. You should check these link IDs. **Fix this by clicking on the link ids**, or by searching for the corresponding IDs like this, and then edit the link by clicking the pink "..." at the end of the row:

Table

links

Rows: link_id

Filter 'link_id' for: 2

Columns: from_label, to_label, quote, source_id, hashtags, simple_bundle, link_id

Pivot column:

1 rows

Click the = signs to filter, click the headers to sort. You can also search 'blanks' or 'notblanks'. Click the pink '...' to edit.

link_id	from_label	to_label	quote	source_id	hashtags	simple_bundle	...
2	~Improved health	Unable/less able to work on farm	[Quantity of food decreased because...] It has decreased because in my family we do get sick, my husband and I likewise for my husbands other wives. We all get sick and when this happened everything is affected and put on hold. [Income decreased because they were all sick and unable to plant.]	MNX-1	~Improved health / Unable/less able to work on farm		

Missing sources in statements table:

The `source_id` field in the statements table is blank for one or more statements. Fix this by roundtripping: simply fill in the missing `source_ids`.

Sources in sources table not in statements table:

You have sources listed in your sources table which have no statements associated with them (there are no statements with the corresponding source IDs).

Statements in links table not in statements table:

The `statement_id` field in the links table is blank for one or more links. Fix this by roundtripping: simply fill in the missing `statement_ids`.

!--chapter:end:41030-data-manager.rmd--

Memos in Causal Map{#xmemos}

Using memos is an important part of qualitative data analysis as qualitative coding usually involves making notes and memos; you can do this in Causal Map too.

In qualitative research, memos are used as short notes about the researchers' thoughts, reflections and ideas during data collection, coding and analysis. So you can use the memos in Causal Map as internal comments to help your research.

In Causal Map, you can create, view and edit memos for just about everything:

- The [whole file](#)
- Statements
- Sources
- Questions
- Factors
- Links

Memos for the **whole file** can be viewed and edited from the [file sharing dialog](#)

Memos for **statements, sources and links** can be viewed and edited from the [memos panel](#).

Memos for **factors and links as well as statements, sources and links** can be viewed from the [corresponding tables](#), and edited using the pink pencil icon at the right of each row:



Memos for **factors and links** can be viewed and edited from the [interactive map](#).

Memos for **links** can be viewed and edited while you are coding or editing a link.

You can use information and/or hashtags in any kind of memo to filter your tables, like this:

Table

The screenshot shows a data visualization interface with a top navigation bar containing tabs: factors, mentions, links, statements, sources, questions, closed_questions, and metrics. A radio button is selected next to 'metrics'. Below the navigation are two informational buttons: 'About this table' and 'How to use the tables', followed by a refresh icon and a download icon.

The main area is a pivot table configuration. It has three sections: 'Rows' (set to 'label'), 'Columns' (set to 'label', 'factor_memo', and 'source_count'), and 'Pivot column' (empty). A 'Filter' section titled 'Filter 'factor_memo' for' contains a search input with 'Health' typed in, and a clear button ('x').

Below the filter is a message: '14 rows Click the = signs to filter, click the headers to sort. You can also search 'blanks' or 'notblanks'. Click the pink ...'.

The data table itself has three columns: 'label', 'factor_memo', and 'source_count'. It contains two visible rows:

=	<i>Health</i>	=
label	factor_memo	source_count
Improved health	Health and hygiene	18

There are edit icons (pencil) and delete icons (trash bin) for each row.

You can use information and/or hashtags in any kind of memo to filter your maps, like this:

Find factors

Keep this panel open?

Include factors how many steps up?

0 10

Include factors how many steps down?

0 10

What field to search in?

factor_memo ▾

Type to find a value or values for this field

Health

Operator

contains ▾

!--chapter:end:41040-memos.rmd--

Collaborating at Causal Map

Overview{#xcollaboration}

One feature we've been asked for most on Causal Map is **live collaborative coding**: if only two people could edit the same file at the same time! Well, we've now made that possible.



Up until now, you and your colleagues have been able to log in to the same or different maps from the same or different accounts^ [depending on your subscription status], including from multiple browser tabs -- so you can have two maps open at once -- but you couldn't both edit the file at the same time. That's what we have enabled. (Or you can even have the same file open yourself in two different browser tabs if you want).

::: {.rmdcaution}

Whether someone else can view or edit your file depends as usual on the [file sharing settings](#).

:::

So for example if you have a large file with a lot of transcripts, you could start coding, say, statements 1–200, and your colleague can start coding statements 201–400. The factors you create will be available to them and vice versa.

This will probably later become a premium feature for subscribers to the Unlimited package only, but right now it is free for everyone on the [Tokyo server](#) only -- let us know (hello@causalmap.app) if you have any questions or comments.

Details

If the file was edited elsewhere since you last (re)loaded it, when making changes you will see a notification that your changes will be merged, but everything should just work, for example:

- adding new codings / links / editing existing links
- deleting or editing links or factors

When your changes are merged, you will have the latest version of everything. Your colleague might not yet have the changes you just made, but the next time *they* make an edit, they will be up to date.

If your file is not up to date with the latest changes, this "refresh" button will be active (usually it is greyed-out).



If you want, you can press the button to get the latest version, but if you are doing your own edits there is no need as you will see the latest version (plus your changes) when you save changes.

We **don't yet support** doing any of these things while the file is open for in different sessions:

- **using the factor editor** (The "Update" button will be greyed if the contents of the editor are not the newest version of the file -- but really, we don't anyway recommend doing any major reorganisation of your factor labels while someone else is working on the file at the same time!)
 - In this case, if you are sure you are ok to continue editing factors with the Factor Editor, for example someone else has left the file open on their computer but gone for lunch, there is a checkbox you can use to override the lock.

Factor editor



- doing advanced file operations like merging or renaming files.

It is also still possible to work on two or more separate files separately and then merge them.

Don't forget, if two or more colleagues are coding the same file at once, perhaps Jo is coding statements 1–100 and Sam is coding statements 101–200, you can get an overview of what (and how much) has been coded in the Coding sub-tab of the Reports tab:

Chart of number of codings per statement.

Mouse over the boxes to see the statement number. White boxes have no codings.



Teams

Your teams: as manager

[Copy](#) [CSV](#) [Excel](#)

Showing 1 to 5 of 5 entries Show 10 entries

Pre

team	member	owner
All	All	All
Causal Map	steve@pogol.net	steve@pogol.net
Causal Map	fiona@bsdr.org	steve@pogol.net
BSDR	fiona@bsdr.org	steve@pogol.net
BSDR	rebekah@bsdr.org	steve@pogol.net
Training 1	jo@bsdr.org	steve@pogol.net

Your teams: as member

Causal Map

- You can create one or more teams and make other people members of that team
- Other people can include you in their teams
- Each team has an owner and a name, and one or more members
- The names of teams you are a member of appear in sha
- There is a special team called `global`: sharing a file with `global` means that any user of Causal Map gets access.

!--chapter:end:41050-collab-teams.rmd--

Uploading your data{#ximport}

This page tells you how to get your data into the right format for uploading to Causal Map. There are different formats to choose from: pick the one which is most convenient for you.

Currently, you can work with 5 different types of formats in the Causal Map App, they are:

1. Text documents: you can upload multiple text files, doc, docx, or pdf files.
2. Standard format Excel: an Excel file with one worksheet called `statements` with one row per statement and optionally other worksheets for `sources` etc.
3. Wide format Excel: an Excel file with one worksheet containing row for each source, one column for each question. e.g. data from Google Forms.

4. Hybrid format Excel: an Excel file with one worksheet containing one row for each answer and includes the statements as one kind of answer but also includes rows for additional data like gender.
5. Links-only Excel: an Excel file with just one worksheet called `links` with just the columns `from_label` and `to_label`. When you want to analyse causal links from elsewhere.

Uploading text documents including Word and PDFs{#ximport2excel}

If your data is in the form of several word processing documents like Word .docx or .odt, or plain text documents or even PDFs, you don't need to bother with Excel files, you can just upload your text files directly.

Preparing your files: If you wish, in your doc, docx and text files, you can separate text into statements by inserting lines containing just -- (that is two hyphen characters together) where you want to separate statements, like this.

--

Please tell me about your health over the last few months.

It's been fine.

--

Ok good, now let's talk about finances.|

Otherwise, doc and docx files will be broken into paragraphs and pdf files will be broken into pages.

You can also use the same trick to add question_id and respondent_id to your upload, like this:

-- question_id: q1 respondent_id: Jo

Please tell me about your health over the last few months.

It's been fine.

-- question_id: q2

Ok good, now let's talk about finances.

That means you can add, at the top of your statement after the --, exactly one or other of these:

- `question_id`: followed by the ID of the question (a single string of characters with no spaces) **and/or**
- exactly the text `respondent_id`: followed by the ID of the respondent (a single string of characters with no spaces)

You don't need to do this for every statement, only where the `question_id` or `respondent_id` changes. So add it for every new question and/or every new respondent.

Wait, you say, what's a respondent? Well we already have `source_id` which is the name of the document. Usually you will have one document for one respondent, so you won't need `respondent_id`. You will usually only want to use `respondent_id` when your respondents change within one document, for example in a focus group transcript.

Uploading like this does not retain information about page numbers in PDFs. If you need to get the page numbers of each statement when uploading PDFs, try [another free utility](#) (Tool 2) which uploads PDFs only but does provide page numbers as an additional column.

This way of uploading your data is only intended for uploading statements for coding. Characters like \$ and # which are used in [hybrid import](#) have no special meaning here.

Problems uploading PDFS?

This upload option may miss text in PDFs which is in boxes and columns. As an alternative you can export a PDF to Excel (one by one) [here](#). The resulting file is also full of pictures etc but does seem to contain all of the text, although in many columns. 80% of the time it identifies suitably large chunks of text, but sometimes you get very small fragments.

We provide another our [free utility](#) (Tool 3) which batch converts any number of these Adobe Excel files into an Excel file which you can then upload to Causal Map.

Uploading a standard format Excel file{#xupload-standard-format}

Causal Map has a standard format for storing data which can be uploaded and downloaded as an Excel (xlsx) file.

That means you can upload your data the first time using this format, and it is also the format you might later use to add more data or to add data about your sources or questions.

The video below shows you how to format and upload a set of statements for coding to the Causal Map App.

```
knitr:::include_url("https://player.vimeo.com/video/637121851")
```

You might find it easier to start from the [Excel template](#).

The app will upload any worksheets in your file called "statements" (required), "sources" (optional) or "questions" (optional), and also, if you need them "factors" and "links"; the app will ignore other worksheets.

Formatting a Standard Format file {#formattingyourfile}

```
knitr:::include_url("https://player.vimeo.com/video/604138709")
```

When uploading data for coding the only data you **need** to upload to the app is an xlsx file with a worksheet called `statements`. Your worksheet must have a header row with the names of the columns; the only required column, which contains the actual responses for coding, must be called `text`.

Each cell in the `text` column contains one "statement" - a chunk of text usually one or two paragraphs long. The optional other cells in the same row give you more information about that

All worksheet names and column names should be lower-case.

	text	source_id	question_id
2	The family welcomed the arrival of mjf-7	a	
3	They have adopted hygienic practices mjf-7	b	
4	The main reason for the changes is mjf-7	c	
5	the respondent earns income through mjf-7	c.1	
6	She is a beneficiary of two social cash transfers mjf-7	c.2	
7	The family eats meal more than an average mjf-7	d	
8	Increased income from mjf-7	e	
9	Money received from the social cash transfer mjf-7	e.1	

◀ ▶
statements
sources
questions
⊕

Pasting data from a Word document into Excel

If you are starting from Word, PDF or other text documents, you may find it easier to [upload them directly](#). But if you want more precise control over exactly what goes into each statement and want detailed information about your sources and questions, this section tells you how to do it using a spreadsheet.

Each paragraph in your Word document will become one statement, so make sure your Word document is already structured so this will work for you. Delete any double paragraph breaks in Word (i.e. delete any empty paragraphs). If you want some short paragraphs to appear together as one longer statement, delete the paragraph breaks between them.

Then

- Copy all the text from your document
- Create a fresh spreadsheet file in Excel (or LibreOffice Calc or similar) and select cell A2
- Paste your text.
- Type the word “text” as column header in cell A1.
- Check the statements are the way you want them. You can ignore any formatting, pictures etc which might also have been pasted in because these will disappear in the next step.

The statements are the texts which appear one by one, for you to read and code. Each statement is one row in your file.

	A
1	text
2	The family welcomed the arrival of a new member in 20xx.
3	She is on her path to self reliance economically
4	The household has become more food secure since becoming a beneficiary of t
5	Important link to Organisation X - Social Cash Transfer: food security, economic
6	Good sanitation practices and a better diet have contributed to an improved he
7	She has more income and more assets now than before.
8	The family is happier now because most of their basic needs are met due to inc
9	The relationship with other members outside the household has on a larger sca

statements +

If you want to upload interviews which are several pages each, you should break each interview into several statements. Ideally a statement consists of between one and five paragraphs – enough to fit on the left-hand side panel when it is displayed there. Usually your text breaks up naturally into sections, for example the interviews might be responses to a number of questions, so your statements might consist of one or a few paragraphs for each question.

We strongly recommend that you don't try to upload statements which are longer than 500 words, as it is harder to code very long statements.

Your statements file must have at least one column with the text of the statements and the header for this column must be “text.” Other columns will be treated as additional data.

Providing additional data like gender

Providing additional data as additional columns in the statements table.

If you want, you can provide additional data like gender **as additional columns in your statements table** when you upload your data. You might want to do that if your data is already structured in this way. However, this means that you have to repeat a lot of data, as you will need to provide, for example, the gender and age etc of each respondent for every single statement they made. It is **usually better to upload additional data about your sources (and questions) using separate tables, as in the next section.**

Just remember that if you do provide additional columns like gender in the statements table, the app understands only text and numerical fields so it will treat dates and currency as ordinary numbers.

Uploading information about sources using a separate source table

Often it isn't convenient to provide all the data in one table.

So in your statements table, additionally (optionally), as well as your `text` column you will usually have a column in the statements table called `source_id` and a corresponding tab called `sources` with at least one column called `source_id` containing each source id, with one row for each source, and additional columns giving more information about your sources.

source_id	#Type	1. age -- A 1. education 1. re
mjf-1	II	36 standard 2
mjf-10	II	37 none
mjf-12	II	35 standard 4
mjf-30	FG	
mjf-4	II	32 standard 2
mjf-5	II	44 standard 2
mjf-7	II	41 standard 5
mjf-9	II	37 standard 4
mjm-4	FG	

statements **sources** questions

The app then merges the new information from the sources table so it is also available in the statements table and elsewhere.

statement_id	text	source_id	age	sex	village
All	All	All	All	All	All
1	The family welcomed the arrival of a new member in 20xx. A baby girl was born to become a newest household member. Overall the family has been healthy during the past year.	mjf-1	36	f	j
2	They have adopted hygienic practices at the household to avoid contamination which may lead to disease outbreak.	mjf-10	37	f	j
3	The main reason for the changes in farming practices during the past growing season has been the adoption of new farming techniques.	mjf-12	35	m	j
4	The respondent earns income through a small scale business she established within the past year.	mjf-30	35	f	j

Uploading information about questions using a separate questions table

Additionally (optionally), you may have a column in the statements table called `question_id`. If you want, this `question_id` can already include the question text, e.g. "q1: tell us about your life". But usually it is more convenient to just use a code like "q1" and also include a corresponding tab in your Excel file called `questions` with at least one column called `question_id` with corresponding values, and an additional column `question_text` with the full text, as well as any other information about your questions, e.g. domain or questionnaire section.

question_id	question_text
a	please tell me about any changes in t
b	please tell me about the health of pe
c	what is the main reason?
c.1	please tell me about any changes in t
c.2	please tell me about any changes in t
d	what is the main reason for any chan
e	please tell me about the other things
e.1	please tell me about the other things

Below the table is a navigation bar with tabs: statements, sources, and questions. The 'questions' tab is highlighted in green.

Uploading a wide-format Excel file{#xupload-google}

This is the kind of data you get if you have been using Google Forms or ODK.

In this format, you only need an Excel file with one worksheet; each row is all the data from one source. Many of the columns will probably be additional data like gender, etc. [Here](#) is an example file

Check your data:

- The app will read the first worksheet and will ignore other worksheets in the Excel file.
- Each row is one source.

Rename your columns:

- Columns beginning with \$ are treated as statements.
- Columns beginning with # are treated as important additional information.

So this means that if you want a column to be treated as a statement, i.e. you want to code the answers in Causal Map, add a \$ at the beginning of the column name.

If you want a column to be treated as important additional data, i.e. you want the information to be easily identified for crosstabulation and other analyses in Causal Map, add a # at the beginning of the column name.

Adding additional cases with wide-format files

You can start coding your data even if you do not yet have all the cases/sources. When you get more data, simply add them to the bottom of the original file and upload again. You don't need to do "roundtripping". All your sources and statements will be overwritten but your coding (links and statements) will be untouched. The questions table will also be untouched. If you want to add information about your questions, for example adding question text as well as question id, you can use roundtripping.

Uploading a hybrid format Excel file{#ximport-quip}

```
knitr:::include_url("https://player.vimeo.com/video/664879474")
```

This section is specifically about uploading from a hybrid-style spreadsheet. This format is sometimes used with QuIP studies.

Here is an example file.

This format has many rows, one for each answer, and includes the statements as one kind of answer; you can upload statements and additional data from the same fieldwork file. The main differences to a normal upload are:

- Rows whose question code does not include a “\$” will be uploaded only as additional data, not as statements.
- A new question ID will be constructed from the question code together with the question text.

Your xlsx file must have just four columns, with column headers in this order:

- `text`, containing the answers
- `source_id`, containing the respondent ID
- `question_id`, containing the question ID

- `question_text`, containing the question text

(Actually the names of the columns do not matter, only the order.)

Your Question IDs may include the following characters:

- A dollar \$ to be treated as a statement.
- A star * to be treated as a closed question.
- A hash # to be treated as important additional data.

Otherwise your question will be treated as unimportant additional data. Also it is worth noting that all the columns of the `sources` table contain metadata about your sources and can for example be used to create filters (e.g. show me a map which only contains statements from people who live alone). Data whose `question_code` contains a # will also be shown in a few additional key place in the app such as the info panel.

You can mix these, so e.g.

`#$* Do you have a house?`

Will be treated as a statement and as closed question (part of a block) and as important additional data.

	A	B	C	D	E	F	G	H	I	J
1	text	source_id	question_id	question_text						
2	HJX	HJX	#A1f	Focus Group Code						
3	Soi Dao	HJX	#A2f	Name of village						
4	Jane Seymour	HJX	A3f	Name of facilitator						
5	20/12/2021	HJX	A4f	Date of meeting						
6	59	HJX	A5cf	Duration of interview (in minutes)						
7	6	HJX	#A6f	Number of participants						
8	We think that Covid i	HJX	\$B1f	Do you feel that Covid-19 is a problem that people in Thailand should care about?						
9	People need a variety	HJX	\$B2f	What do you think people need to live well in this community?						
10	Increased	HJX	*B3f	Overall, how has the amount of money your household earns changed in the last two						
11	Increased	HJX	*B4f	Overall, how has the amount you can buy as a household changed over the last two						
12	Decreased	HJX	*C1f	How has the amount your household saves changed?						

All important additional data, as indicated by the # in the `question_id`, can be found under the info toggle in your left-hand side panel.

```
#A1f_Focus Group Code
#A2f_Name of village
#A6f_Number of participants
```

Uploading closed questions in hybrid format{#hybrid-closed}

Closed questions, as indicated by the * in the question_id, are treated like any other additional data for sources as shown in the picture above. Once imported, they will appear in the special closed questions table (which is really just a special view of the sources table) as well as in the the sources table itself.

question	value = +	value = -	total
All	All	All	All
*B3F_Overall, how has the amount of money your household earns changed in the last two years?	4		4
*B4F_Overall, how has the amount you can buy as a household changed over the last two years?	2	2	4
*C1F_How has the amount your household saves changed?	2	2	4

Plus, the question_code may end with the characters "rank" to mark "ranking organisation" data. [There may be a bug currently such that this data will also be uploaded only if they are not treated as statements, ie the code should not contain a \$].

To upload the hybrid-style spreadsheet click on the upload button on the right-hand-side panel and use the 'hybrid' option.

Appending hybrid data

Suppose you uploaded hybrid data from Excel spreadsheet F, and already did some coding, then you receive more data in spreadsheet G, also in hybrid format.

Find the original spreadsheet F on your computer (don't download from the app!). Paste the data from spreadsheet G below the existing data in spreadsheet F, save it as spreadsheet H, and upload this again. Your existing coding should not be affected; the links and factors you already created will remain.

Caveat: you can't combine roundtripping and appending hybrid data. If you have tweaked or added to your data from file F by roundtripping (downloading the data from the app then uploading again), any changes to statements, sources or questions will be lost if you later try to append data using this hybrid button. So if you need to tweak your data (statements, sources or questions), either:

- wait until you have completed uploading using the hybrid button and only then tweak with roundtripping
- or, manually tweak the old and new data in spreadsheet H.

QuIP recodes for closed questions - live link

Using this [table](#) (which includes spanish) the app recodes your data into + positive, - negative and 0 neutral.

For example:

- better }
• get better }+
• improved
• a bad change }-
• worse
• no change }0
• stayed the same

This data will be displayed in the Tables panel if you select “closed questions.”

Uploading existing causal coding ("links only") {#xlinks-only}

What to do if you have causal information, for example in the form of **edge lists**, from some other source (e.g. AI coding of a larger text) which you want to import?

You can import this kind of file too, using the Standard Format Excel upload button.

Your Excel xlsx file should NOT have a tab for factors. Instead it should have a single tab called `links`, containing two columns called `from_label` and `to_label` which contain the names of your factors.

[Here](#) is an example file.

That's the basic setup. But maybe you also want to provide information about what statement, question or source the link is associated with. In which case, these are your options:

- Optionally, your links tab can also contain other columns as usual, in particular it can contain the column `question_id`.

- If you include a `question_id` column, you should include an additional tab called `questions` containing at least the column `question_id` and other columns providing information about the questions, e.g. full question text, questionnaire section etc, where relevant.
- Optionally, your links tab can also contain a column `statement_id`.
- If you include a `statement_id` column, you can optionally include, as usual, an additional tab called `statements` containing at least the column `statement_id` and a column `text` containing the text of your statements.
 - If you optionally include a `statements` tab, you can optionally also include a `source_id` column and a corresponding tab called `sources` with a corresponding `source_id` column and more columns to provide information like gender, location, or any other information about the source of each statement.
- Optionally, your links tab can also contain a column labelled `quote`, containing the actual quote which underpins each causal claim aka link.
- Usually, this quote will be a verbatim section of the text in a corresponding additional tab called `statements`, see above.

!--chapter:end:41100-import.rmd--

Uploading additional data like age, gender{#import-additional-data}

You often have additional information to upload, as well as the actual statements you want to code. Often this additabout the source, for example about the respondent who provided the information, or about the context of the information e.g. what was the question asked or the name of the interviewer.

Additional data is just any additional fields related to each statement.

You can include them when you do your initial upload, or later, using round-tripping.

- When you upload in Standard Format, you provide additional data fields like gender as `columns` in the `sources` worksheet alongside your `statements` worksheet.
- When you upload in Wide format, you provide additional data fields like gender as additional `columns` in the file you upload.

- When you upload in Hybrid format, you provide additional data fields like gender as additional rows in the file you upload.
- When you upload Text files, the only additional data you can only provide is the ID of question and/or respondent. You will probably want to later add more information about these respondents using round-tripping.

Whatever format you use to upload your data, you can always continue to add data including additional data like age, gender etc using roundtripping. However, remember that any data you upload will overwrite the existing data -- so if you only want to update your sources data, make sure you delete any other worksheet in the file you upload to leave only the `sources` worksheet.

Uploading closed question blocks

This kind of data is common in QuIP studies, which are normally uploaded using the hybrid format. Here are instructions for uploading closed question data in hybrid format.

In Causal Map, closed question data is just another kind of additional data about your sources. So it is possible to use closed question blocks when you import your data in other ways.

As the answer to a closed question like "how often do you go hungry (never/sometimes/often)" is just another field, this means you could ask the app to only show the maps of the women, or the people who often go hungry, or the women who often go hungry.

The point of the closed questions feature, which looks for columns in your sources table whose names contain a *, is first to recode a fixed set of words (like better, worse) into + positive, - negative and 0 neutral, and second to organise and provide the answers to a whole block of questions together in one place, to make analysing the data easier.

This picture shows how a single set of closed responses (like "better", "same" and "worse") are displayed in tables like the one below.

question	value =	value = +	value = 0	value = -	total
	All	All	All	All	All
*amount saved -- F6_Overall, how has the amount you save as a household changed over the last year (net of any debts)?	4	24			28
*assets -- F7_Overall, do you feel the combined total value of all your assets has gone up or down over the period?	4	23	1		28
*farming -- C2_Overall, how have your farming activities changed in the last year?	4	20	1	3	28

This approach only makes sense if you have one larger set of questions with the same set of fixed reply options.

To upload closed question blocks:

Hybrid format: follow [these](#) instructions.

Wide format : simply make sure that the relevant column names contain a *.

Text format: use [round-tripping](#): simply download your file as Excel, add columns to your Sources table whose names contain a *, and upload again.

Standard format: You supply this data simply by adding additional columns to your Sources table whose names contain a *. You can supply this data when you initially upload your data or later, by [round-tripping](#).

[!--chapter:end:41200-uploading-additional-data.rmd--](#)

Uploading your data: special cases{#ximport-other}

Uploading from other software

We are aware of the initiative to make qualitative coding software interoperable, <https://www.qdasoftware.org/products-project-exchange/>. Causal Map has quite a different model and at the moment we don't provide this, but would do if there is a lot of interest.

You can always export all your data at any time using the Download button in the [top menu](#).

If you have existing causal coding in the form of an "edge-list" you can upload it [like this](#).

Exporting for upload into kumu.io

tbc

Exporting for upload into NodeXL

tbc

[!--chapter:end:41250-import-other.rmd--](#)

Appending and fixing data ("roundtripping"){#xroundtripping}

Rather than providing a complicated interface for manipulating and adding data, Causal Map lets you upload data from an Excel file, download it again, make changes or add data and upload it again.

General principles{#import-general-principles}

The 5 core data tables in any Causal Map file are these: factors, links, statements, sources, questions.

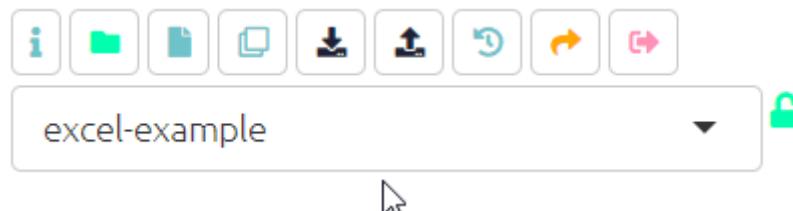
- The downloaded file has these 5 main worksheets (as well as worksheets to show all the data for the current view when you downloaded the file including filtered fields.)
- All 5 standard tables can also have arbitrary additional columns/fields which users can use e.g. to filter, analyse their data.

You can upload your data back into a new or existing Causal Map file.

- The app will read any worksheet with any of the (lowercase) names of the core tables: factors, links, statements, sources, questions and will ignore other worksheets.
- If any of these five worksheets in your Excel file are *empty*, the app will **delete all the data** in the corresponding tables.
- If any of the five worksheets are completely **missing** (e.g. you deleted them), when you upload, the app will just keep the old data in the target file.
- If you want to append data, just download the old file, add the rows you want in any of the tables, and upload again. (Adding additional *links* is a special case, see [xx.](#).)

Downloading a file

To download a file click the download icon in the left panel and your data will be downloaded as an `xlsx`.



What you can expect from your download:

- Factors table: the `factor_id` field is included.
 - If you delete any of the rows, any links pointing to them will be deleted on upload.
 - If you want to add more factors at the bottom of the table, leave the `factor_id` field blank for those rows and the app will add them on upload.
- Links table: the `link_id` field is not included and is not needed.
- Statements table: the `statement_id` field is not included.
 - Don't delete any of the rows, unless you know what you are doing. (You can delete any rows after the currently last coded statement.)
 - You can add more statements at the bottom of the table.
- Statements table on our Test server: the `statement_id` field IS included when you download a file.
 - You can delete rows and reorder rows (by dragging them around in Excel). Don't change the `statement_id`.

- Note: If you delete rows which had links, these links will become orphaned and will be displayed in the maps only when you do not filter by any `statement_id`. This probably isn't what you want -- you might want to delete the corresponding links too. If you choose to add new rows as well as deleting old rows, the `statement_id` of the new rows will be the row number, so make sure this is what you want -- will they then get "gifted" orphaned links?
- Sources table: the `source_id` field is included.
 - You can edit the `source_id`. Keep in mind that the contents of the `source_id` column in this table should match the contents of the `source_id` column in the statements table.
 - If you want to add more rows, you must include a `source_id` for each row.
- Questions table: the `question_id` field is included.
 - You can edit the `question_id`. Keep in mind that the contents of the `question_id` column in this table should match the contents of the `question_id` column in the statements table.
 - If you want to add more rows, you must include a `question_id` for each row.

Uses for roundtripping

You could download a file to:

1. Make any ordinary trivial edits like changing a name, and/or add ordinary columns/fields.
2. Add rows. Note this does **not** add linked entries in linked tables.
3. Delete rows. Note this does **not** delete linked entries in linked tables (except if you delete factors, any links to/from them will be deleted.)

You could upload a file containing only a statements tab:

1. with a column text and optionally `source_id`, `question_id` and maybe other fields.
Note this does not add the source and question ids to the source and question tables.

You could upload a file containing only a links tab:

1. With just the `from_label` and/or `to_label` columns and no factors tab.

- With numerical from and to columns which relate to a factors tab which has a factor_id column and a label column and usually a statement_id column.

What if you have merged cells in your Excel file?

If you have merged cells so e.g. source_id spans many rows, the app will not be able to read this data. It is easy to unmerge your cells without losing your data, just follow the instructions on this [webpage](#).

!--chapter:end:41300-roundtripping.rmd--

Merging files{#xmerging-files}



You want to merge files "foxtrot", "golf" and "hotel" into one. What you are going to do is create a blank file and then use the merge mapfile command to merge in the file foxtrot, and save the map as currently filtered (i.e. including all the contents of foxtrot which arrive via the filter), and then repeat for the others.

1. Create a new file under a new name e.g. "my-merged-file"

2. Type this in the Advanced Editor:

- merge mapfile path=foxtrot

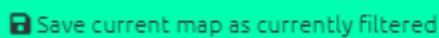
3. Press Save in the Advanced Editor

4. Save the file as currently filtered:

- Press Save As:


Or: Save map as currently filtered without changing name

- Then press:

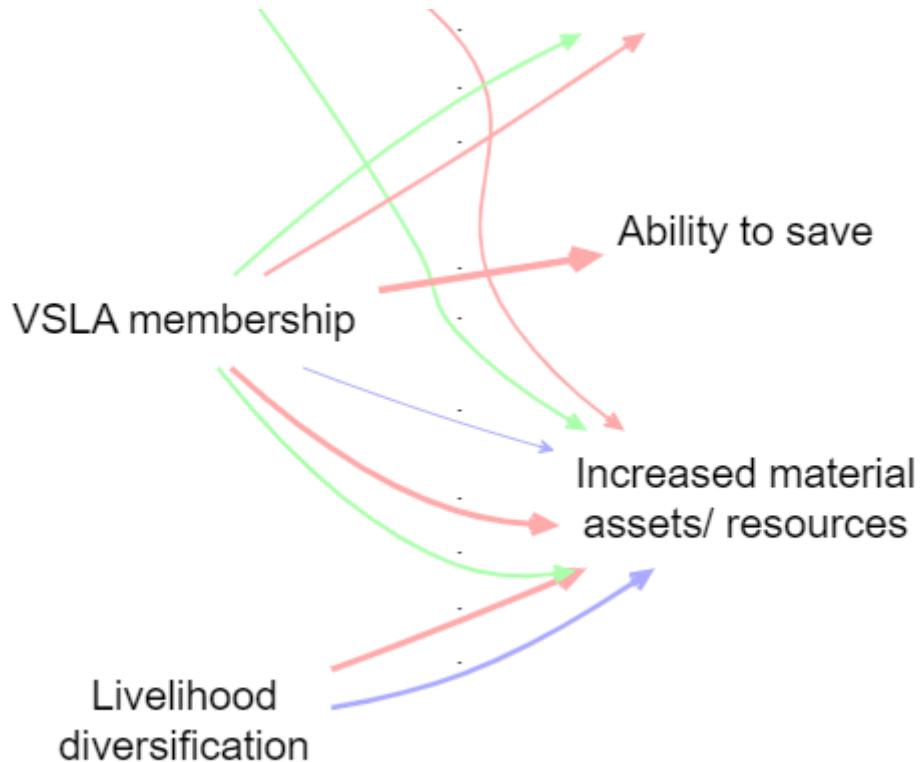
 Save current map as currently filtered

5. You now have a file my-merged-file which has as its contents the file foxtrot.

6. Repeat steps 2–4 exactly as above, but with the command `merge mapfile path=golf`. Your file my-merged-file will now contain both files, and all the tables will have a column `map_id` in the links table, which distinguishes the files, with the value 2 for foxtrot and 3 for golf.

7. Repeat steps 2–4 exactly as above, but with the command `merge mapfile path=hotel`. Your file my-merged-file will now contain both files, and all the tables will have a column `map_id` in the links table which distinguishes the files, with the value 2 for foxtrot, 3 for golf and 4 for hotel.

8. If you want, you can then visualise the different links like this (in this case, the files already had mostly common factors):



The filters here were:

```
bundle links field=map_id  
scale links value=count:link_id  
color links value=unique: map_id lo=#ffaaaa mid=#aaffaa hi=#aaaaff
```

You can verify the number of uploaded links like this:

Links

The screenshot shows a data visualization interface with a table titled "Links". The table has four columns: "from_label", "to_label", "statement_id", and "quote". A filter is applied to the "map_id" column, showing counts for different map IDs. The counts are: map_id = 1 (3330), map_id = 2 (632), map_id = 3 (651), and map_id = 4 (All). The interface includes buttons for "Copy", "CSV", and "Excel", and a "Show 10 entries" dropdown.

map_id	Count
1	3330
2	632
3	651
4	All

⚠️⚠️⚠️ Warning: that filter in the Advanced Editor will merge the data in every time it is used, either when you press Save or when creating a shortlink. So if in doubt, delete the filter as soon as you save the file "as currently filtered".

If you wish, you can [roundtrip](#) the file to add a column called, say, `original_map` in the links table next to the column `map_id` which contains names rather than numbers to make it easier to remember what is what. In any case you can use the `map_id` (or `original_map`) fields to filter and/or format the map however you like.

Alternatively you could first download the constituent files to Excel (roundtripping) and provide a column (with the same name, e.g. "original_file") in at least the statements worksheets in all the constituent files, before merging. You can also do that for the links tab. You can do it for sources, questions and factors too, but the identifiers would get coalesced if there were duplicates (which there might be by design) with the first map winning.

Alternative method

You can also create a blank file and merge in all the constituent files like this:

```
merge mapfile path=foxtrot  
merge mapfile path=golf  
merge mapfile path=hotel
```

remembering to then "save file as currently filtered". However using this method, the `map_id` columns will not be created for you.

--chapter:end:41/10-merging.rmd--

Editing factors and links (advanced)

{#xedit-factor-and-links2}

To see the basic instructions for editing factors and links, click [here](#).

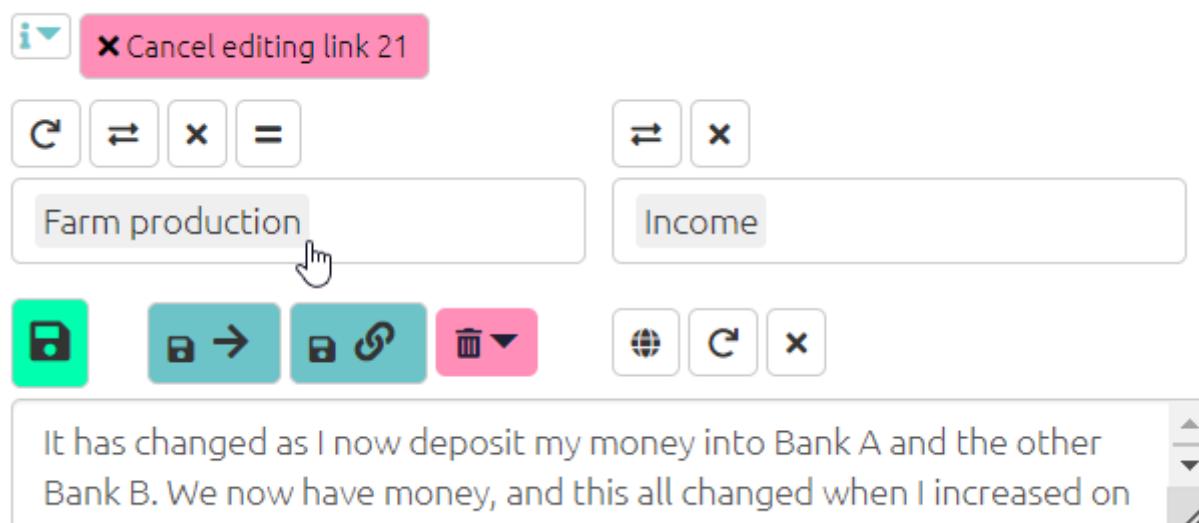
It is likely that you will need to regularly review and edit your links, and there are various ways to do this in the app. You can also delete factors and links in the same way, however you cannot delete factors or links when certain filters are applied such as `combine opposites` and `zoom` as these filters represent multiple factors as one factor label.

Edit factors in the Factors table

To edit factors in the edit factors table, click on the row containing the factor you would like to alter. This will open the edit panel, as shown above. You can use the search function at the top of the factors table to find the factor you want to change.

Edit factors and links in the Links table

In the links table you can edit links and factors by clicking on the row containing the text you want to edit, this will open a panel asking what you want to do with this link. By clicking 'edit interactive' in this panel you will be able to edit the link in the left-hand side panel, where you originally coded the statement.



You can also edit multiple factors using the [factor editor tab](#).

--chapter:end:41500-editing-factors-and-links2.rmd--

Factor labels: consolidating {#xfactor-labels-consolidating}



Consolidating and editing factor labels

After you have created more than about 50 factors you will probably feel the need to start consolidating and renaming them and perhaps merging some.

If you look at the [Mentions table](#) in the app, you will probably find a lot of *pairs* of factors each of which is used only once, as half of this particular pair, and don't appear linked into longer stories. We almost certainly will want to consolidate some of these factors by in some way merging them with others which have similar meaning.

Suppose you have 100 factors which are just mentioned once or twice. It's a lot of work to do all that coding, a lot of work for almost nothing, because those very infrequent factors (we call them “tiddlers”) are unlikely to appear much in any reporting you do. With hierarchical coding, even infrequent factors can contribute to the bigger picture.

Before we even think about consolidating factors, do remember the golden rule that, **when coding a new link, always look at your existing factors and try selecting one of those before creating a new one**. Even if you think an existing factor isn't quite a perfect fit, you can use it and adjust the factor name later. Ideally you will always be able to think of labels for your factors which are at least general enough to cover a few different cases.

It can be helpful to reformulate factor labels so that some common themes come *first*:

- communication difficulties - intercultural
- communication difficulties - lack of internet

This is useful when the factors are listed alphabetically and can be a stepping-stone to formulating factors hierarchically.

Even when doing inductive coding, some people formulate factors with a theory in mind, with some higher-level factors like communication difficulties or resilient outcomes. However we have found that it is often most interesting when you let the theory emerge rather than by starting with a preconceived template in this way. Look at your factors periodically as you code and see if there are any obvious groups.

Also, our recommended way of having your cake and eating it – keeping the detail but seeing the general trends too – is hierarchical coding.

Merging multiple factors into one

In some cases you will simply decide, in the face of too many factors, to merge Engagement with communities and NGO engagement into one, e.g. Engagement.

If you have two factors like this which are very similar but not quite the same, you can sometimes keep some detail by using a composite label which explicitly mentions both i.e. Engagement with communities/NGOs.

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.

!--chapter:end:42410-factor-labels-consol.rmd--

Factor labels: desirable?!{#factor-labels2}

See also [here](#).

If possible, especially when conducting evaluations or research for policy purposes, try to formulate these semi-quantitative factors so that for each one, *more* of it can be broadly thought of as *more desirable*, e.g. Health, Wellbeing; and failing that, as something *undesirable*, e.g. Psychosocial Stress, Mortality. Including factors which primarily code something undesirable as well as factors which primarily code something desirable, is often very natural, but you might sometimes have to think more carefully when making reports and summaries.

Again, it is least useful, but also perfectly acceptable and sometimes necessary, to fall back to using factor labels which are ambiguous as to their desirability, e.g. Moved house, if there is a lack of information or in the case of factors which simply are neutral.

!--chapter:end:42420-factor-labels-desirable.rmd--

Factor labels: more about hashtags{#xhashtags2}

See also the [introduction to using hashtags](#).

Hashtags for bundling factors

It is very convenient to use hashtags when you want to bundle factors. This means you can easily and temporarily collapse together one or more sets of factors sharing a common theme, regardless and independently of any hierarchical structure you are using.

Intervention hashtags

The commissioner of an evaluation will nearly always want to be able to focus on their own interventions, so it is advisable to use a consistent label or category across all factors which are noted as project inputs in order to be easily searchable. These will usually be influence factors with no parents, i.e. factors which have no incoming influence factors.

In the example below, both are similar interventions, but one is clearly attributable to an NGO.

“MSF gave us medicines which healed the children”

MSF gave medicines –> children healed

“Someone gave us medicines which healed the children”

Unknown actor gave medicines –> children healed

You have a range of options to hashtag this difference in the label and enable you to filter for all project-related interventions once the data is coded.

- All such interventions can be labelled with the NGO’s name
- If you are doing nested coding, you can nest such factors within a higher-level factor like Interventions;
- And/or you can use a specific intervention hashtag, e.g. #Intervention. (QuIP has its own hashtags, [I] for Implicit and [E] for Explicit).

You can use any label you want as long as you are sure it is unique, it doesn’t appear accidentally in other factor names, and you use it consistently.

Outcome hashtags

Similarly, if you wish to hashtag up outcomes /consequences as either desired/desirable/valued, or the inverse, you can mark these in a way which makes them easy to identify. You could use symbols like ♥ etc. in the labels to designate this, or use QuIP-style labels [P] for positive and [N] for negative. This can be helpful when undertaking analysis; for example, displaying direct and indirect links from some or all of the intervention factors to some or all of the positive outcome factors or highlighting or counting paths leading to all negative outcome factors, etc.

Again, you can use any label you want as long as you are sure it is unique, it doesn’t appear accidentally in other factor names, and you use it consistently.

Hanging hashtags{#xhanging-hashtags}

A subtle trick is to use hanging hashtags which are additional hierarchical components. So rather than writing, say,

More food #nutrition --> Improved wellbeing

you write

More food; #nutrition --> Improved wellbeing

(note the ;).

If you do this, you can still search for your hashtag `#nutrition` as before.

But if you "zoom out" to the top level of the hierarchy, the hashtag will be neatly hidden:

More food --> Improved wellbeing

!--chapter:end:4.2.4.20-factor-labels-flags2.rmd--

The Factor Editor{#xfactor-editor}

Summary

The Factor Editor is a great way to:

- edit your factor labels in bulk
- combine several factors into one

Use the Split checkbox to edit only the factors you can see in the map as currently filtered, and leave untouched any other instances of the same factors in the rest of the map.

Use the sidebar to:

- see and edit your factor memos
- see and edit every link to and from each factor

Editing factor labels in the Factor Editor

The Factor Editor shows the label of every factor in the map as currently filtered, one row for each factor.

So for example if you are applying no filters and viewing the entire map, the label of every factor will be displayed.

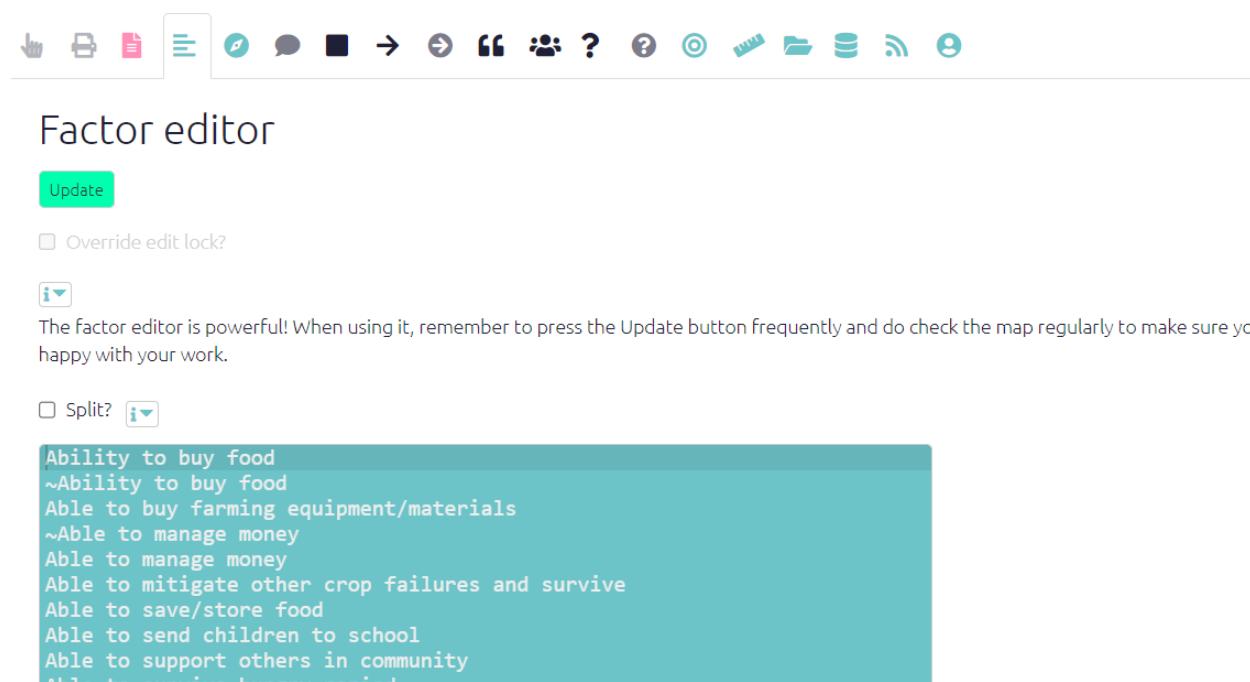
To edit one or more factor labels, click inside the editor and retype the factor labels however you want.

Then hit Update.

::: {.rmdcaution}

You can rewrite parts of labels or entire labels any way you want, but you can't delete or add lines.

:::



::: {.rmdcaution}

Editing factors in the factor editor is **disabled** when you use any of these filters:

- combine opposites
- zoom factors

- remove brackets

With these filters, any changes you made would be ambiguous. So the app makes it impossible to make edits and the save button should be inactive

:::

TIP: press Ctrl F (Cmd F on Mac) to find or Ctrl H (Cmd alt F on Mac) to replace.

TIP: Even if you are not using hierarchical coding, formulate your factor labels so that some common themes come *first*. So rather than Eating fruit (Good nutrition) write, say Good nutrition - eating fruit. The Factor Editor will then help you by keeping similar labels together.

TIP: The Factor Editor is especially useful with hierarchical coding because *related labels are displayed together*.

TIP: Use hashtags like “Problem!” or “#innovation”, unique words which are easy to search for, to make it easy to find particular kinds of factor within the editor and elsewhere.

Merging your factors

You have two or more factors which are more or less the same and you want to combine them?

- pick the best version, tweak it if necessary, and then copy the whole line into your clipboard
- go to one of the other versions, and replace the label (the whole line) with the label you copied, **taking care not to delete or add lines**
- repeat with any other similar lines, so that all the similar labels are now exactly the same.
- press Update!
- ... these factors with the common label will all be combined into one

Split-recoding existing factors: Use the Split checkbox.{#split-recoding}

Usually when you rename your factors, **the names change everywhere** – every time this factor appears anywhere in your file.

But maybe you don't want this!

Suppose you have coded both “I bought a cow”“ and "I bought a sheep" as bought livestock and then you change your mind.

Suppose there is just one statement about buying cows.

You can't just go back to this statement and recode bought livestock as bought a cow, because then the *other* quote (“I bought a sheep”) would also be recoded as bought cow, which isn't what you want.

This is what the Split button is for. If you switch on “Split” and retype bought livestock back to bought cow. and then press Update, only the instance of bought livestock in the current filter (which you can see in the current map) is recoded back to bought cow.

If you switch on “Split”, each factor you rename will be split into a new factor (with the new name which is used only for the links in the current view/filter), *and the original factor is unchanged*.

The Split checkbox applies to whatever is in the current view – so you can use it to split-recode factors just within one statement, but also to split-recode factors in any filter.

So for example maybe you have used more work for both men and women and you want to recode it just for statements from all the men, not the women. Then you can apply a filter to show only statements from all the women, and use Split recode.

Advanced editing tips

- Go to the next mention of a selected word/sentence: **Ctrl K, or Cmd G on Mac**. Add SHIFT to these shortcuts to find the *previous* mention.
- The app will help you consistently type words you have already used with autocomplete. Or just ignore the autocomplete if you want.
- Multicursor: Ctr alt Up/Down: edit multiple lines at once. This is **incredibly useful** if you want to bulk edit many labels at once.
 - Add next occurrence of selected word/sentence to multi-selection: **Ctrl-Alt-Right, or Cmd-Option-Right on Mac**. Use **Ctrl-Alt-Left / Cmd-Option-Left** to add *previous* occurrence.

The factor editor sidebar{#xfactor-editor-sidebar}

The factor editor also has a sidebar, which shows more details for the factor currently under the cursor:

- The label (click on it to:

- edit the factor memo
- focus on the factor
- delete the factor)
- The memo, if there is one (click on it to edit or add a memo)
- Lists of influence and consequence factors, with source and statement ids
 - When hovering over the influence or consequence factors, see the whole original text from the corresponding statement
 - Click on these boxes to edit the corresponding links directly
- Note that the sidebar does show you for each factor in the current filter ALL the links / quotes relating to it, even ones which are not in the current filter. This is so you can make a decision about eg changing factor labels which takes into account all the uses of the factor.

Factor editor

Override edit lock?

 The factor editor is powerful! When using it, remember to press the Update button frequently and do check the map regularly to make sure you are happy with your work.

Split? 

Ability to buy food
 Community groups/learning
 Community works together
 Diet improved; Diversified
 Farm production
 Health behaviour
 Health behaviour; Use better toilets
 Health behaviour; Use mosquito nets
 Health behaviour; Use pits to dispose of rubbish
 Health behaviour; Wash hands; After using toilet
 Improved health
 Improved/new farming techniques; Cultivating in lines
 Income
 Increased knowledge; Farming method/practice
 Increased knowledge; Relationships/gender
 Planted new crop/vegetable varieties
 Shared household decision-making

Farm production 

Memo: Agriculture 

influences:
 Statement: 44
 Source: MNY-1
 Farm production. Quote: We have had some new changes which have led to an improvement in our production. It changed for the better because now we have enough to eat and sell...I recently changed as I cultivated without transferring, sometimes we work in lines.. Question: \$C1 

Statement: 88 

If you use emojis in your factor labels, especially emojis with skin colours, you may find it difficult to find the right place in the editor. Tip: put the emojis at the end of the label if possible.

!--chapter:end:42430-factor-editor.rmd--

Simplifying causal maps with hierarchical coding{#xsimplifying-with-hierarchical}



(See also the [next section](#) for how to implement these ideas in the Causal Map app.)

Summary

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

New intervention; midwife training → Healthy behaviour; hand washing

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

New intervention, in particular the midwife training,

Or we can read it in reverse like this:

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

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

New intervention; midwife training; hand washing instructions

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

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

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

New intervention → Happier health providers

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

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

New intervention → Healthy behaviour

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

Introduction

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

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

This approach is convenient for several reasons:

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

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

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

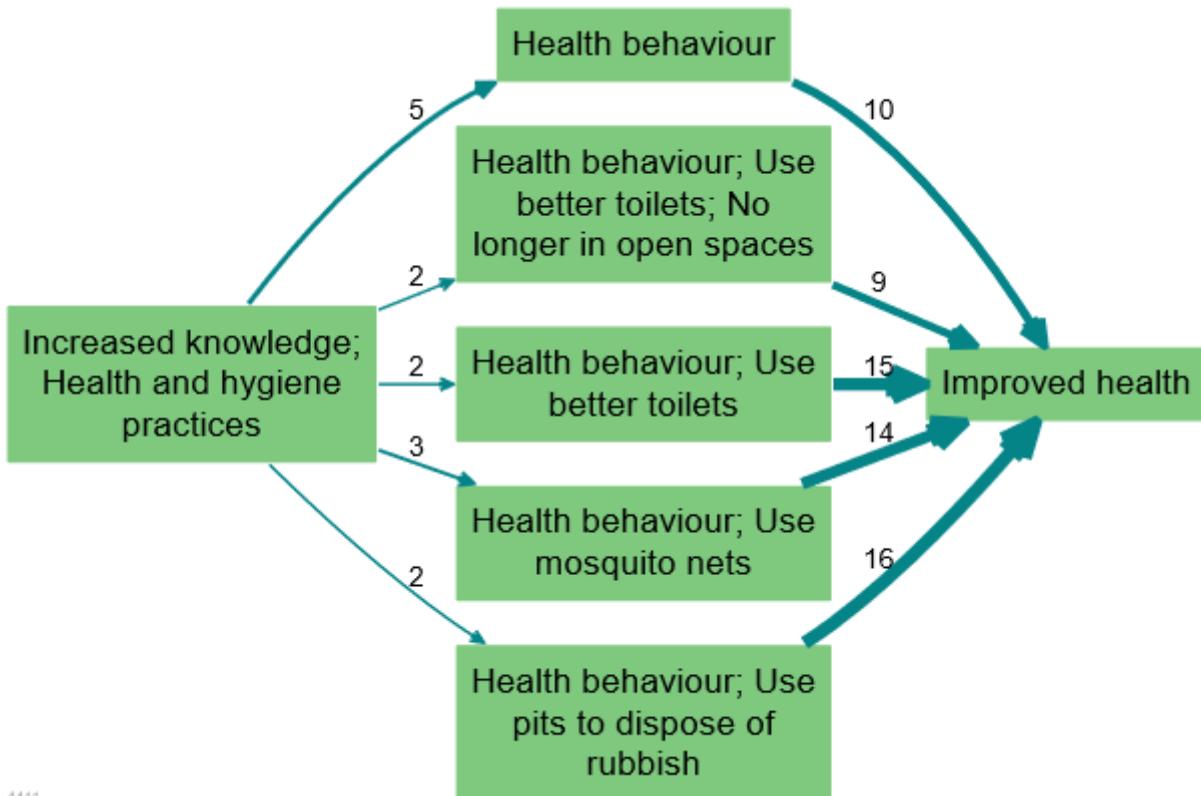
A factor can't belong to two different hierarchies

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

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

If you find yourself wanting to put a factor into more than one hierarchy, try using [quickfields](#) or [hashtags](#) instead.

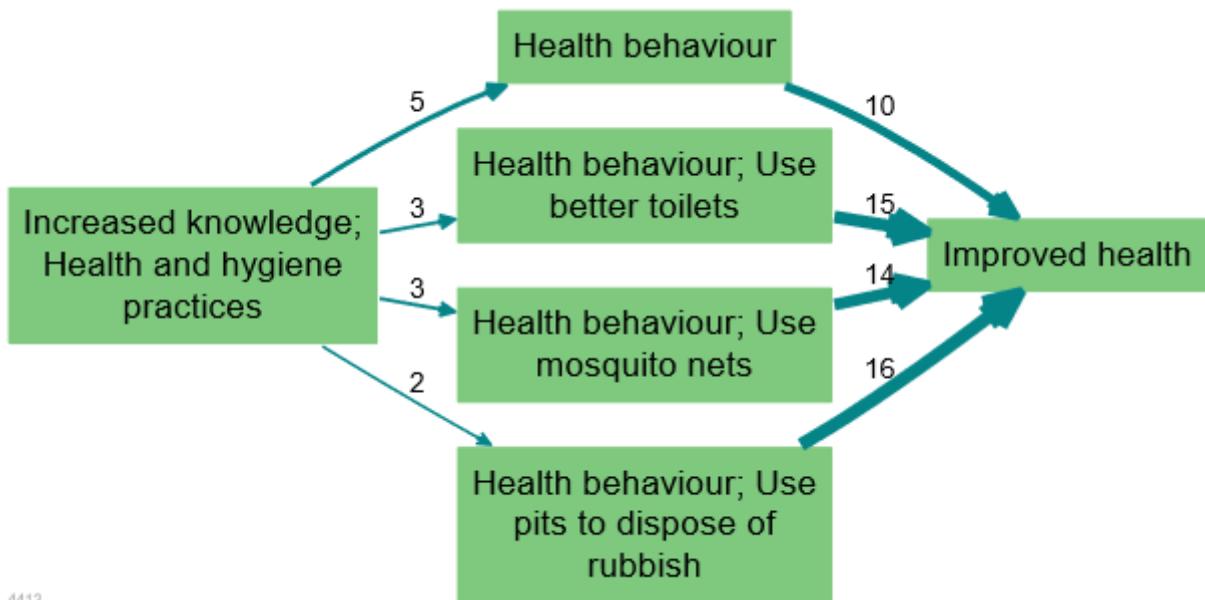
Zooming out



4411

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

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



4412

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



4413

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

Zooming out is like making deductions with the ; separator

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

If we know

New intervention; midwife training → Healthy behaviour; hand washing

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

New intervention → Healthy behaviour; hand washing

and

New intervention; midwife training → Healthy behaviour

and even

New intervention → Healthy behaviour

This actually reflects the dilemma of the analyst often referred to as *granularity*: with how much detail should I code the beginning (or the end) of this causal story?

Expressing a factor as Health behaviour; Hand washing; Before meals shows that this is indeed to be understood as a kind of health behaviour, although of course not the whole of it. By using this approach, the analyst says: if you are just looking for the global picture, I am happy for this factor to be understood as Health behaviour.

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

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

Semi-quantitative formulations work best

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

- Social problems
- Social problems; Unemployment
- Social problems; Addiction
- Psychosocial stressors
- Psychosocial stressors; Fear of job losses
- Psychosocial stressors; Pre-existing mental health issues

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

So we might have:

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

Social problems; Unemployment → Psychosocial stressors

“When people get stressed they often turn to drugs”

Psychosocial stressors → Social problems; Addiction

These could be combined into this story:

Social problems; Unemployment → Psychosocial stressors → Social problems; Addiction

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

Social problems → Psychosocial stressors → Social problems

Higher-level factors are generalisations

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

Health; vaccinations law is passed

Health; mortality rate

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

Vaccinations law is passed #health

Mortality rate #health

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

Don't mix desirability!

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

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

Stakeholder capacity; Lack of skills.

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

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

~Stakeholder capacity; ~Presence of skills.

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

Using higher level factors for "Mixed bags"

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

Politics; increase in populism

Politics; shift to the left

Politics; shift to the right

Politics; more engagement from younger voters.

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

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

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

Don't use a hierarchy when a hashtag will do

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

Soil loss (#Environment)

Eco training courses for NGOs ([Environment])

or even a symbol:

Soil loss 

Eco training courses for NGOs 

We can still search and filter with text tags, even with icons, as long as they are unique.

You can use hashtags in parallel with hierarchies:

Training courses delivered; Eco training for NGOs 

Re-usable factor components as hashtags

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

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

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

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

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

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

!--chapter:end:42432-hierarch.rmd--

Hierarchical factors in Causal Map{#xhierarchical-coding}

(This section is about implementing hierarchical coding in the Causal Map app. See also the [previous section](#) on hierarchical coding in general.)

```
knitr:::include_url("https://player.vimeo.com/video/683770170")
```

Creating labels

Factors can optionally be expressed as part of a hierarchy by using semi-colons.

For readability, it is usual to leave a space after the semi-colon, but this makes no difference to the functionality.

Organising your factors using the Factor Editor

In Causal Map, the process of organising the factors into this kind of hierarchical structure can be conveniently carried out in the [Factor Editor](#) tab which is a simple text editor where you can edit any labels you have created during coding. It lists the currently visible factor labels, sorted alphabetically.

This editor has many features such as global search and replace, and multiple cursors, which make it easy to rapidly edit many factor labels.

Using this panel you can also combine several factors into one and split one factor into several.

[Here](#) are a lot more tips for using the Factor Editor.



Factor editor

Override edit lock?



The factor editor is powerful! When using it, remember to press the Update button frequently and do check the map regularly to make sure you are happy with your work.

Split?

Access to tap
Better community relationships
Health behaviour
Health behaviour; Brush teeth
Health behaviour; Drink clean/treated water
Health behaviour; Increase number of washes per day
Health behaviour; Live in clean environment

Better community relationships

Memo: Relationships

influences:

Statement: 553

Here it is easy to “move” an (incorrectly labelled) factor

Health behaviour; understanding of germ theory

to something like

Health knowledge; understanding of germ theory

simply by retyping it: replace "behaviour" with "knowledge".

Or you might want to "move" your factor somewhere else entirely, relabelling it as

Real-world knowledge; health; understanding of germ theory

... without worrying whether the higher-level factor Real-world knowledge even exists elsewhere in your file.

Additional calculated fields

When you use hierarchies, the app adds some fields to the factors table which help you do searching and filtering with hierarchical ideas.

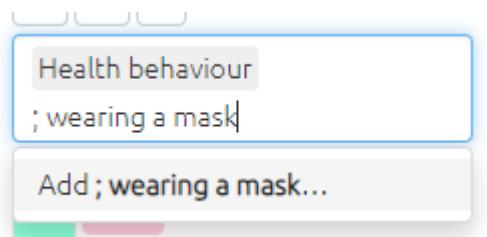
One such field is `top_level_label` which gives just the top level for each factor.

Additional functionality when creating and editing links

Once you have created at least one hierarchical factor, i.e. one with a “;” in its label, the *influence factor* box and *consequence factor* box have some additional functionality to help you.

- Now, when you start to type, the list of existing factors which you can choose from is extended to include existing factor components, even if they have not (yet) been coded as such. This means it is easier to add new detail to existing or implied higher-level factors.

Quick tails: quickly adding new "daughters" of an existing "mother" {#xquick-tails}



Suppose you want to create a new factor Health behaviour; wearing a mask and you know there is an existing higher-level factor Health behaviour but there is not yet any factor Health behaviour; wearing a mask.

- You can select Health behaviour from the list and then just add ;wearing a mask with a leading semi-colon. These two fragments will be combined into the desired factor label. This is quicker and ensures you don't end up with different spellings of the higher-level factors.
- This second component can itself have a semi-colon, so you can do this: health behaviour ;hand washing; before meals.
- Also you can do this multiple times:

health behaviour ;washing hands ;washing pots.

This can be useful when you have to code a list of things which are all examples of some higher-level idea.

This example will produce:

health behaviour; washing hands

and

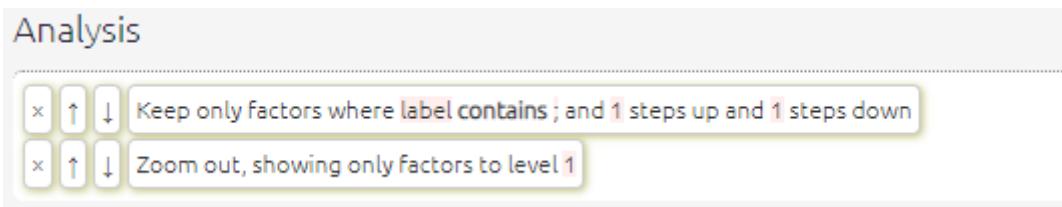
health behaviour; washing.pots.

Additional functionality when searching for and filtering factors

The same principle applies in the Search and Filter Factors box: you can see factors you have already used but also implied higher-level factors (like Health behaviour) and other factor components, beginning with a semi-colon (;).

How can I view just the factors where I have applied hierarchical coding?

Sometimes you only want to use hierarchical coding for a few of your factors. To view just those, you could create a mini map including just the factors which you have hierarchically coded (and one step up or down from them), by searching for ; in the factor labels, but then zooming the factors out to just see the top level.



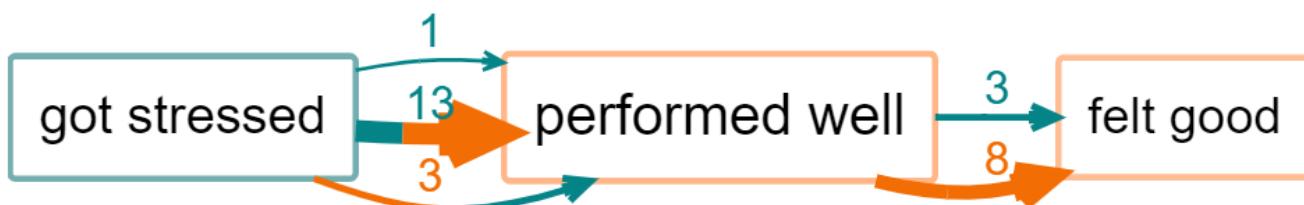
... which is equivalent to

```
find factors field=label value=; operator=contains up=1 down=1  
zoom factors level=1
```

!--chapter:end:42434-hierarch-in-app.rmd--

Coding opposites{#xopposites}

Summary



This map shows that 13 people say they got stressed (the first section of the link is green) and therefore did *not* perform well (second section is *orange*), another person said they also got stressed but performed well (both sections green) and 3 people said they did *not* get stressed (first section of the link is orange) and performed well. The border of the factor "performed well" is mostly orange because most people mentioned *not* performing well.

In the first part of this Guide we have dealt only with undifferentiated links which simply say "C causally influences/influenced E" or more precisely "Source S claims/believes that C causally influences/influenced E".

- Following a "semi-quantitative" convention, Health trainings --> Health - outcomes means that (good / more / present / adequate) health trainings lead to (good / more / present / adequate) health outcomes.
- Without this convention, it means no more than health trainings somehow influence health outcomes. Both "The trainings actually made the outcomes worse" and "The trainings actually made the outcomes better" could both be coded in the same way. We do not recommend coding in this way.

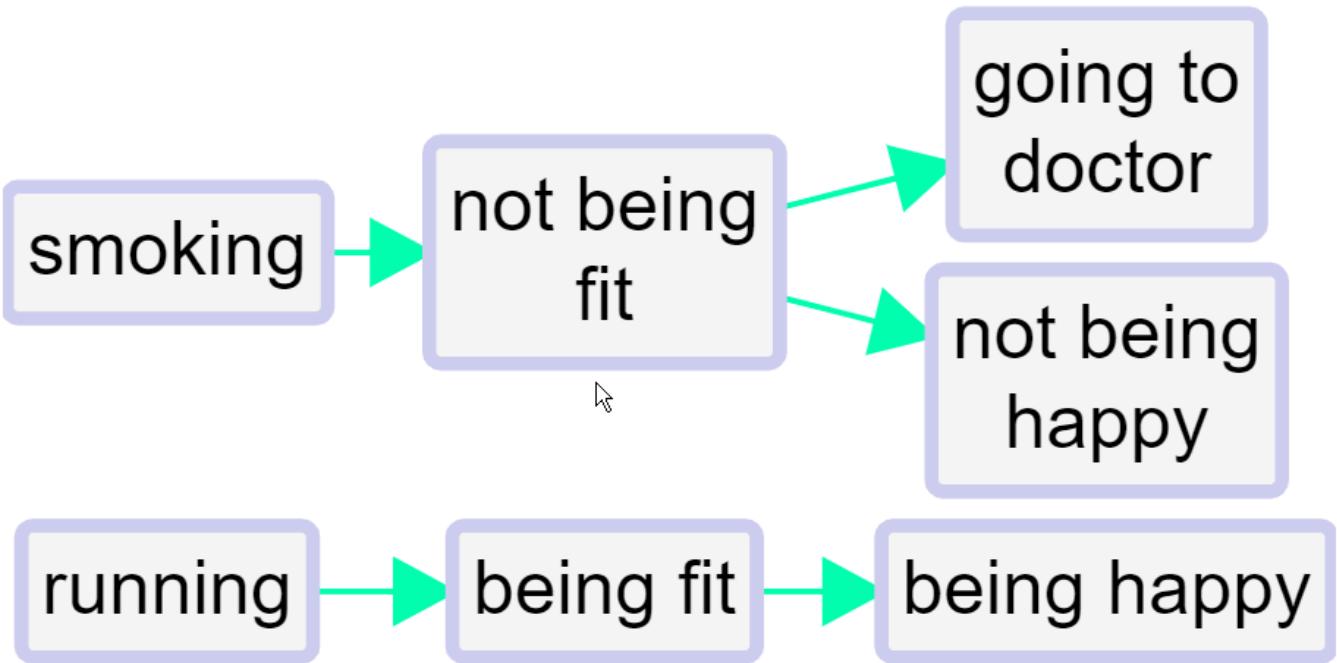
Combining opposites

Combining opposites is a simple and powerful way of dealing with "negative" factors such as poor health which are in some sense the opposite of existing "positive" factors such as good health.

```
knitr:::include_url("https://player.vimeo.com/video/636997200")
```

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

Here is an example of quite barefoot QuIP-style coding in which we have not used strength at all. Nevertheless there are the beginnings of some ideas about (and issues with) polarity: for example, we have `fit` and `not fit`.



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

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

The first step forwards is to follow this convention:

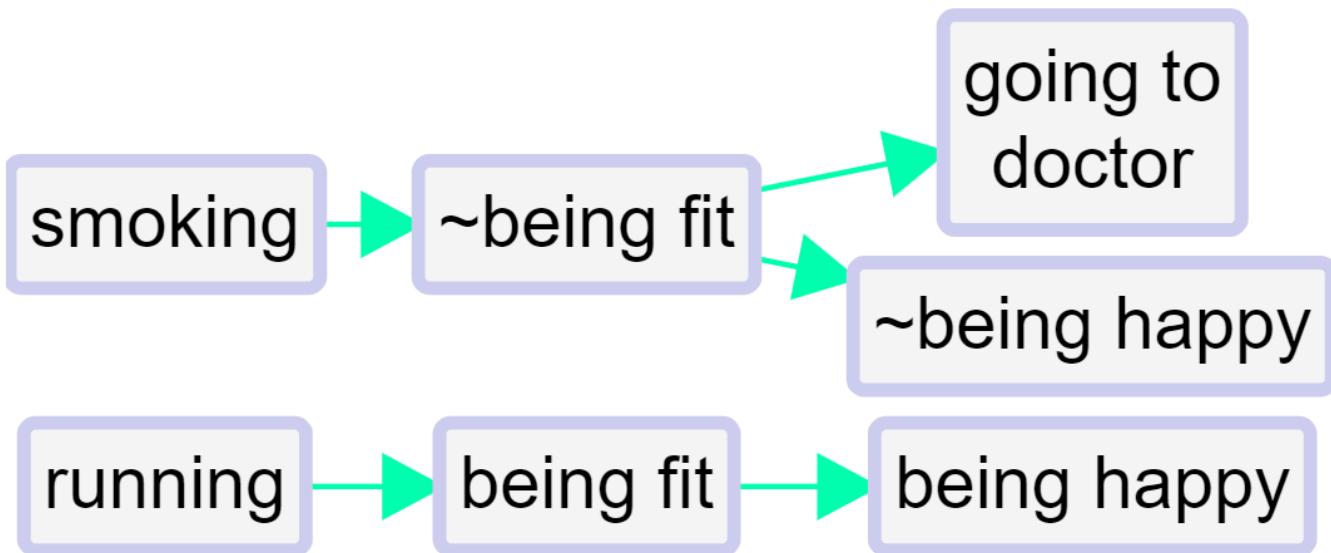
To signal that two factors are *opposites* we formalise the idea we already instinctively used in the above example, where we used the word "not" for one of each pair. Formally, we will code them in the form $\sim Y$ and Y . The \sim may appear at the start of a factor label. This already ensures that when we search for Y we will also find $\sim Y$.

In the factors-editor panel, these two factors will be listed next to each other – the alphabetical listing will ignore the \sim .

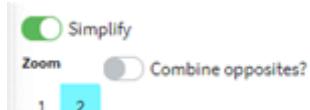
We talk about *opposites* rather than positive/negative or plus/minus because that frees us from any implications about valence or sentiment: smoking is the opposite of not smoking, health is the opposite of not health / ill health / illness.

Where there is some kind of sentiment involved, we do suggest using the ~ sign for the negative member of the pair.

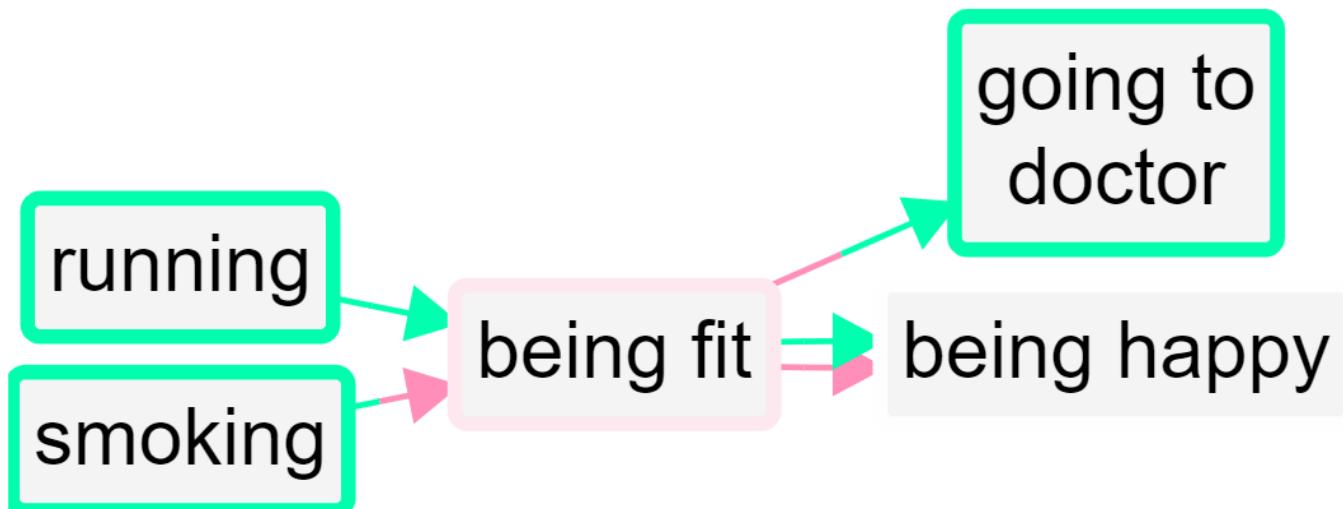
So the same map would look like this, using ~ instead of not.



When you press "combine opposites",



the app combine any pairs of factors which are opposites. It looks at all the factors which begin with a ~ and takes off the ~ where there was one - a process which we call "flipping" the factor. But it only does this if there is already its opposite coded in the file as currently filtered, otherwise there wouldn't be any point i.e. ~running will remain as such is you haven't coded running.



So now there are two factors combined into the "fit" factor and two into the "happy" factor. The "not" factors have their incoming and outgoing links preserved, but when a factor is flipped to match up with its opposite, the part of the link next to that factor is now coloured red. So the lower link from fit to happy is red because the factor at each end of the link has been flipped; the influence factor was originally *not fit* and the consequence factor was originally *not happy*.

So, a link has two polarities: a *from* polarity and a *to* polarity. From a quantitative point of view, the *strength* of the link is their product (if the signs are the same, it's a plus, if they are not, it's a minus). So both arrows from fit to happy have "plus" strength, but they are not the same because one is in a sense a negative example of the other.

This lower link from fit to happy has the same "plus" strength as the one above it but it does not represent the same information. In fact, **no information is lost when you press the "combine opposites" button; you can still always read off the original map from it.**

But there is more: the *border colour* can reflect the overall polarity of the factor: if there were many mentions of ~Fit and only a few mentions of Fit, the border will be nearly red; if there were equal numbers the border will be white, and the more the positive mentions there were, the greener the border will be.

If a factor has a red border, that means that at least mostly, its *opposite* was mentioned. So in this example, happy has a grey border because it was mentioned once and its opposite was mentioned once. Fit has a slightly red border because it has one incoming and one outgoing plusses, and one incoming and two outgoing minuses.

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

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

Two-tone links

In Print view, when combining opposites the links can be two-tone, with a different colour at each end.

Stress → Performance

Original link	Interpretation	Link with combined opposites
Stress → Performance	I felt more stressed, so I performed better	Stress → Performance
~Stress → Performance	I felt less stressed, so I performed better	Stress → Performance
Stress → ~Performance	I felt more stressed, so I performed worse	Stress → Performance
~Stress → ~Performance	I felt less stressed, so I performed worse	Stress → Performance

In Interactive view, two-tone links are not possible. The first and last cases above are the same as in Print view, but in the second and third cases the links are grey.

With larger files you will often want to bundle the links between each pair of factors. When you combine opposites you will want to combine the links into so-called flipped bundles. There can be up to four flipped bundles between each pair of factors just like in the table above.

Filtering

When you combine opposites, the app gives you additional combined fields (in the links and factors tables) which you can use to display and filter your maps in various ways. For example, you can keep only links in which come from or to *positively* formulated factors: `find links field=flipped_bundle value=FALSE | FALSE operator>equals` in your advanced editor or:

The screenshot shows a software interface titled "Find links". At the top right are icons for a magnifying glass, a right arrow, and a hand cursor. Below the title is a green button labeled "Apply filter!". A checkbox labeled "Keep this panel open?" is checked. To the left of the main panel, there are vertical tabs labeled "i", "c", "Sta", and "S". The main panel has a section titled "What field to search in?" with a dropdown menu set to "flipped_bundle". Below it is a text input field containing "FALSE|FALSE". Another section titled "Operator" has a dropdown menu set to "equals". A cursor icon is visible near the "Operator" dropdown.

Opposites coding within a hierarchy{#xopposites-coding}

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

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

~healthy habits; smoking

~being fit

going to doctor

~being happy

healthy habits; running

being fit

being happy

When you press "combine opposites", the app tries to combine any pairs of factors which are opposites. It looks at all the factors which begin with a ~ and *flips each component*, taking off the ~ where there was one, and inserting one where there was not.

healthy habits; running

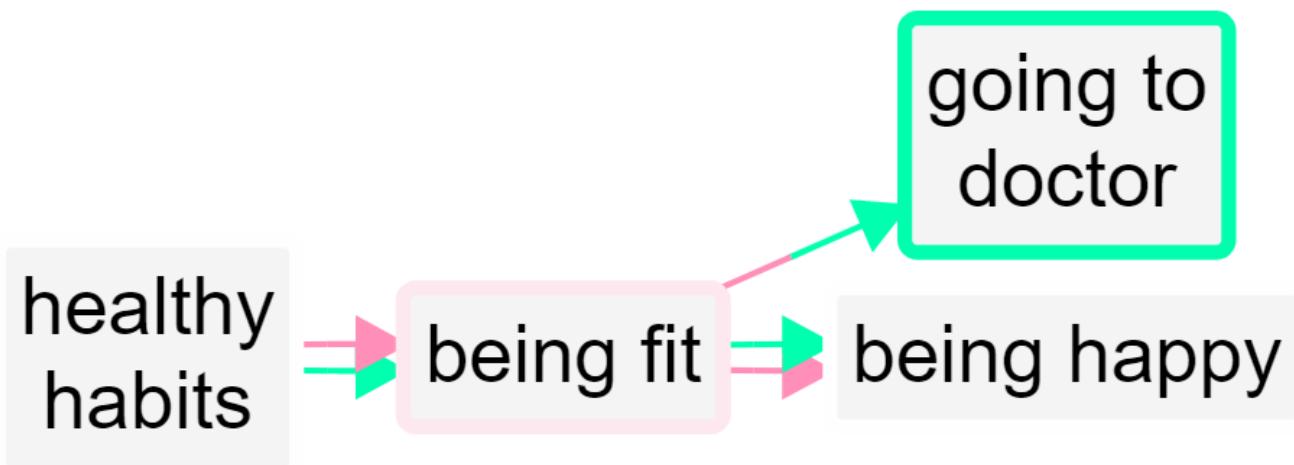
going to doctor

~healthy habits; smoking

being fit

being happy

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



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

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

Opposites coding within components of a hierarchy

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

- ~Healthy habits; ~eating vegetables

is the opposite of

- Healthy habits; eating vegetables

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

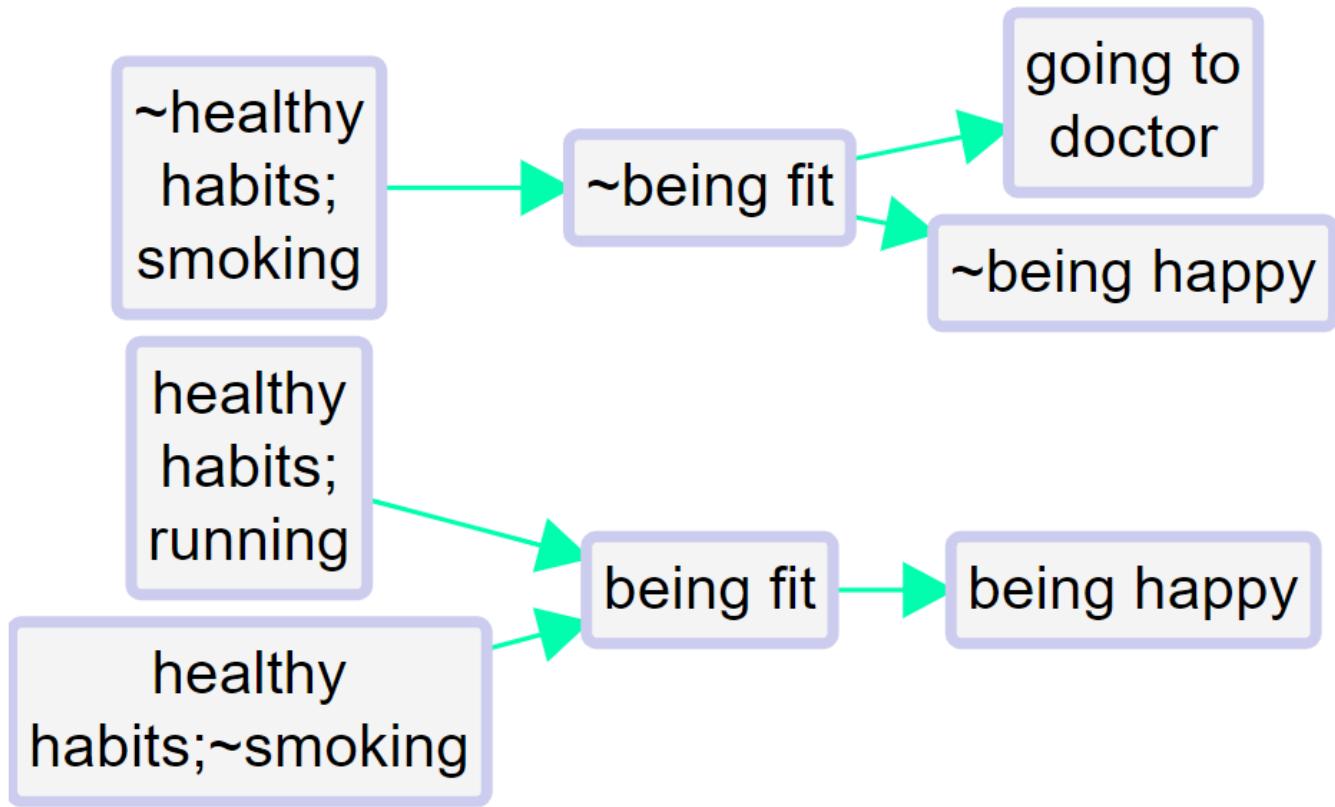
- ~Healthy habits; smoking

is the opposite of

- Healthy habits; ~smoking

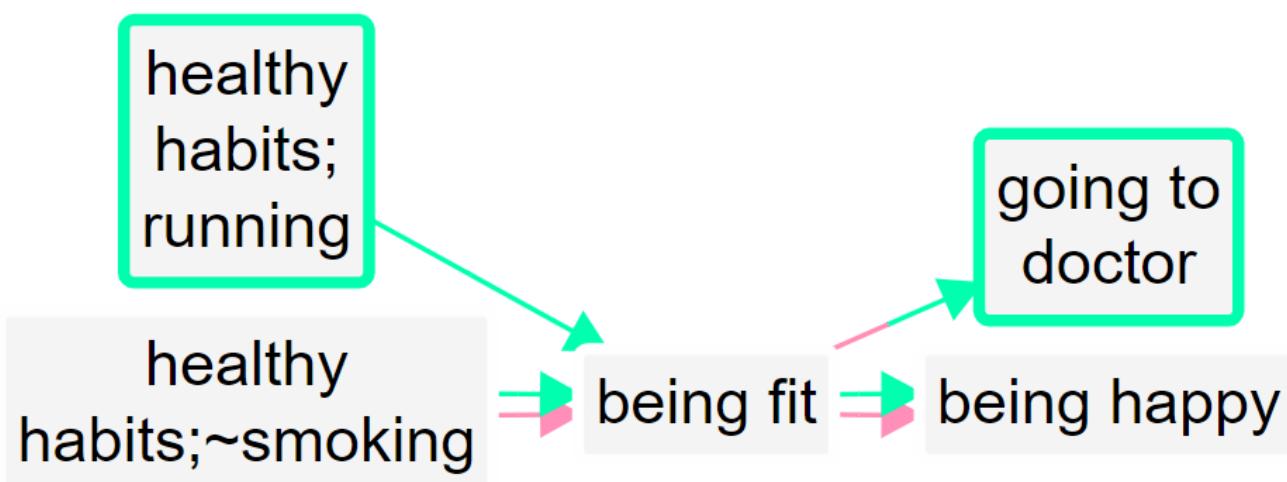
Smoking, which is an example of unhealthy behaviour, is the opposite of not smoking, an example of healthy behaviour.

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



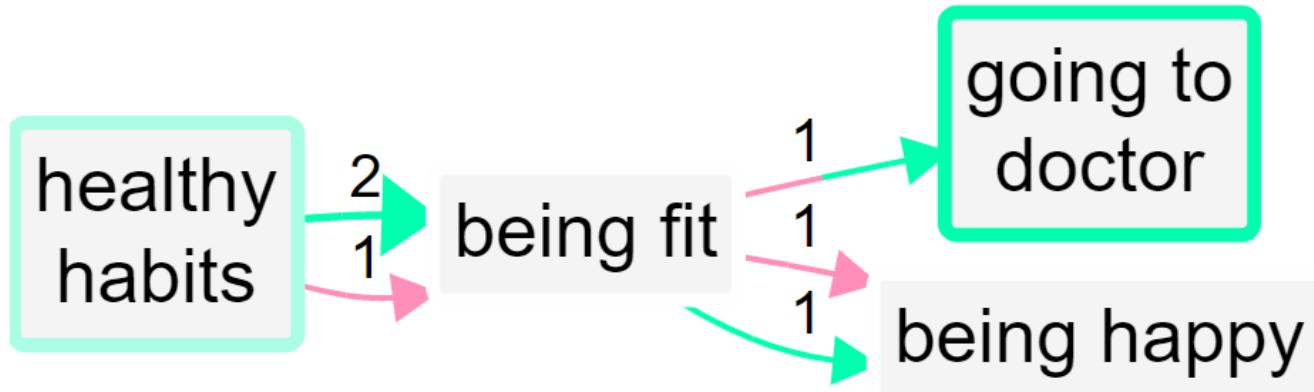
The healthy habit of not smoking leads to being fit.

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

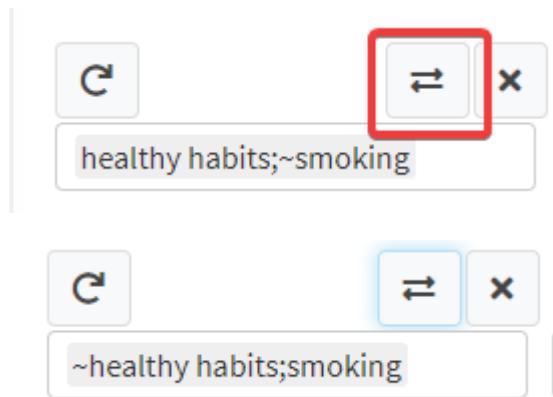


The two arrows at bottom left (one all green, one all red) show that there is one example of this particular healthy habit leading to fitness, and the complementary example in which the opposite of this habit leads to the opposite of fitness.

Zoomed out to the top level:



The app makes it easy to code opposites; while coding, when you have loaded any factor into either of the boxes, pressing the "flip" button above will flip the factor into its opposite.



Unpacking opposites

Occasionally you might decide you'd prefer to view a file where you have used opposites coding but with ordinary-language labels like unfit rather than ~fit – after all, it's easier to read. By all means, create another version of the file and manually recode your labels back again using the Factor Editor. Remember, you'll only want to recode factor labels which *start* with a ~, not labels with a ~ elsewhere. We recommend in this case adding a * at the end of each such factor, so that someone viewing the file that this factor does have a conceptual opposite somewhere.

In this case of course you won't be able to combine opposites any more.

The border colours and how to override them

The factor border colours are calculated so they are more orange as the proportion of contrary mentions of the factor to the total mentions of the factor. So if ~Income was mentioned often *in the map as it was at the point at which this filter is applied* and Income only occasionally, the new combined factor Income will have border which is more orange than blue. This metric depends only on the overall number of times the factors were mentioned, not e.g. on the number of sources mentioning it.

Sometimes these colours can be distracting, so if you don't want them, you can set some other border colour as usual, e.g. by adding a line in the Advanced Editor like this:

```
color borders fixed=white
```

Alternatively in the Dashboard, you should also find an Addon View to add fixed border colours.

!--chapter:end:42435-opposites.rmd--

Importing a codebook of factors{#xcodebook}

Sometimes you might want to work with a pre-determined list of possible factors. This is sometimes called "deductive coding" or "(partially) deductive coding".

Or you might be working on a larger project with colleagues and you want to periodically share your codebook.

To do this you just need to roundtrip your file: start by downloading the file as Excel. Then look at the `factors` tab.

If you already have factors (and links) in your file, you **must not change the existing order of the factors** which are already in the `factors` tab.

If you want to add *new* factors, which do not yet have links, you can add the factor labels you need one above the other in the `label` column of the `factors` tab *after* any existing factors.

You can edit the existing labels and/or add memos about the factors in the `factor_memo` column if you want, and even add arbitrary other columns like, say, `domain`.

- If you add a factor field like "domain", you can only use it in a filter *before* any transforming filters like Zoom.

!--chapter:end:42439-codebook.rmd--

Creating links (more tips)

How the factor dropdowns work

If you are using hierarchical coding, top-level components will appear in the factor dropdown even if you have not used them explicitly so far. For example, if you have used the factors `health; mental` and `health; physical`, the factor `health` will appear even if you never used it explicitly. You can then easily create new nested factors such as `health; spiritual` by selecting `health` and then typing `; spiritual` with a leading `";".` The app will automatically join these for you into the factor `health; spiritual`.

- once you have selected an influence factor, the order of the choices available for the consequence factor will be silently updated, presenting at the top of the list the most frequently mentioned consequences of the influence factor you have chosen.

Creating more than one link at a time

This video shows how to create more than one causal link at a time.

```
knitr:::include_url("https://player.vimeo.com/video/588851236")
```

Chaining Links{#xchaining-links}

```
knitr:::include_url("https://player.vimeo.com/video/588881701")
```

Chaining links is easily done and useful when you have statement where one consequence leads to another, which leads to a third, and so on. To use this function simply click the chain icon below the consequence factor box and the app will automatically fill the influence box with the consequence from your previous link.

“

Since the price of cotton has increased, I have more money and can afford to send my child to school.



Price of cotton increased

Income increased



!--chapter:end:42440-creating-links2.rmd--

Link memos and hashtags

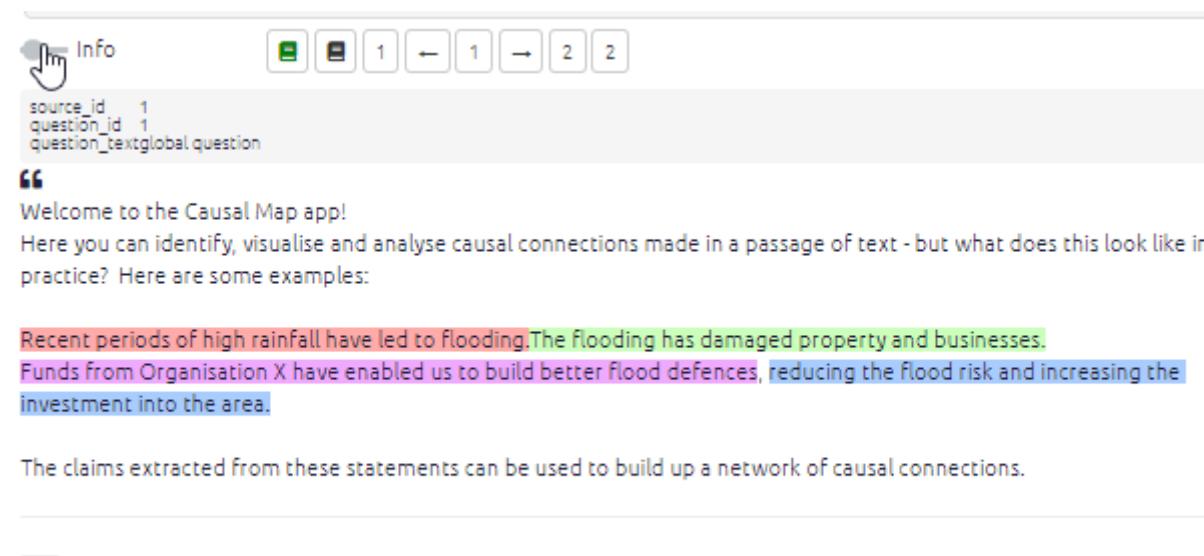


Using memos and hashtags{#xmemosandhashtags}

Memos{#xlink-memos}

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

This toggle opens both additional statement info, i.e. info for any fields which start with #, for example you might have a field #gender in your statements table, and also shows the boxes to edit the memos. (In this example, there is no additional information.)



The screenshot shows the Causal Map app interface. At the top, there is a toolbar with icons for Info, Undo, Redo, and other navigation functions. Below the toolbar, a sidebar displays metadata: source_id: 1, question_id: 1, and question_text: global question. The main area contains a quote: "Welcome to the Causal Map app! Here you can identify, visualise and analyse causal connections made in a passage of text - but what does this look like in practice? Here are some examples: Recent periods of high rainfall have led to flooding. The flooding has damaged property and businesses. Funds from Organisation X have enabled us to build better flood defences, reducing the flood risk and increasing the investment into the area." A note below states: "The claims extracted from these statements can be used to build up a network of causal connections."

Link hashtags{#xhashtags}

There is no need to actually use a hash # at the start of a link hashtag, though you can if you want. Just use any unique single word which is easy to search and filter on, like #nutrition or nutrition# or nutrition--.

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

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 or beginning or ending with specific hashtags (or parts of hashtags), and also for links which do not contain specific hashtags or parts of hashtags.

Conceptually, there are two kinds:

Ordinary link hashtags

You can use any hashtag which does not begin with a ? to record any other information about the link, e.g.:

- respondent doesn't like this connection
- respondent feels good about the outcome
- for you, the analyst, e.g.
 - respondent is answering a different question
 - to tag links you want to come back and review.

Weak hashtags

Weak hashtags are a special kind of hashtag aka link hashtag. They are caveats. If you use weak hashtags, you should make sure that by default your maps do not include any link with a weak hashtag.

They begin with ? and are used to mark any link which you are not sure is always valid across the global context for the whole global map, for example:

- the causal connection is only valid for a specific context, e.g.
 - the respondent says this is only true for their village, not for other villages e.g. ?village_X
 - a link is only projected for the future e.g. ?future
- you are unsure about the claim about the causal connection
 - a link is only a hypothesis e.g. ?hypothetical
 - you as the analyst are not confident in the claim e.g. ?doubtful
 - the source themselves are not sure e.g. ?source seems unsure
 - to add other qualifying information e.g. ?probably_hearsay
- to mark the fact that a connection is weak or non-existent, e.g.
 - Respondent makes a substantive claim that X does not influence Y, e.g. ?zero_influence
 - Respondent makes a substantive claim that X only insignificantly influences Y, e.g. ?weak

Weak hashtags in quickfields.

If you are using quickfields for (some of) your link hashtags, you can usually put the ? at the start of the hashtag as normal:

?context:village-A

?context:village-B

However the "weakness" only applies to one of the values of the quickfield, in which case we recommend the following kind of format, in which the ? comes before the second part of the hashtag:

strength:?none

strength:strong

or

confidence:?low

confidence:high

This makes sense because in your default maps you would only want to exclude one, not both, of each of the pairs in these examples.

--chapter:end:42450-link-flags.rmd--

Ellipses{#xellipses}

Coding your statements in context: referring to evidence from elsewhere

When coding data from an interview it is very important to see the whole interview as your source, and not fall into the trap of only coding the statement in front you, out of context. If your statements are split into separate questions (as is often the case with QuIP), it can be misleading to code just one statement in isolation, without taking account of what the respondent says earlier or later in the interview. For this reason we recommend that all analysts read the whole interview before starting to code, and start to formulate causal stories which fit across the whole interview. In some cases, you may prefer to have the whole interview as one statement, and to code it all in one place. This is absolutely possible, and just depends how the data is uploaded (whatever information is on one row in the .csv file is considered a statement). If, however, the interview is too long, or you would still prefer to code question by question, you need may want to refer to information provided in another statement when coding a link. This is not necessary

but can be useful when sharing the file with a commissioner who may look at quotes associated with links, and not understand the rationale for coding without further context. The same applies for different parts of a paragraphs within one statement; you may want to highlight two separate parts of a longer statement.

If the quote is not exactly verbatim in the statement, the app won't let you save anything. (If this happens, just cancel the quote (press the X) and start again.) The app is really picky about this to ensure your quotes are accurate. However, this rule does not apply for anything between square brackets, a useful feature for this case. If the evidence for a link is mentioned in two different parts of a single statement, like this:

Yes, we got the help from OrgX. Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. And because of that help, now I have a job. qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit. Ut enim ad minima veniam.

You can use square brackets to remove the text you don't want. Highlight the whole piece of text you want to code, and then edit the quote in the quote window, replacing what you don't need with [...]

Yes, we got the help from OrgX. [...] And because of that help, now I have a job.

Be careful not to introduce any new text or even spaces.

You can also use the same technique to add your own comment or quote from a statement elsewhere in the interview:

[this is my note] Neque porro quisquam est, qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit, sed quia non numquam eius modi tempora incidunt.

Ellipses are also helpful to refer to another statement as part of your evidence:

[reference to statement 36: The whole village received help from Org X] Thanks to that help we received, we are now growing our own produce.

!--chapter:end:42470-ellipses.rmd--

Quickfields{#xquickfields}

Summary

Quickfields are a special kind of hashtag for factors and links. They are a refinement of hashtags for when you have a set of categories which is mutually exclusive.

You can add information about one or more dimensions using a format like this:

Crops improved / time:before

Crops improved / time:after

A field (in this case, a field called `time`) will then be created in the factors table which you can then use for example to filter the maps, to search, filter and organise the factors, links and mentions tables, and to format the maps.

Example uses:

- capturing time before/after
- capturing factors which are explicitly or implicitly attributed to a particular project
- capturing the valence of a factor e.g. when a subject explicitly says the liked or did not like the fact that they had more work
- capturing things which apply only, say, to girls or only to boys
- organising ideas within a hierarchy

You can search for and display this information across all your tables and maps, not separately as they would be if you used ordinary hashtags like #before and #after, but together.

Here is more info about analysis with quickfields.

Alternative formulations

These are equivalent, you can

Crops improved time:after

Crops improved /time:after (optionally use a / just to separate the fields visually)

or you can use a hanging hashtag:

Crops improved; time:after (a so-called "hanging hashtag" which is perhaps a bit more readable and moves the quick field to a lower level of the hierarchy).

Quickfields are also ordinary hashtags

Quick field hashtags are still hashtags, so you can still also do all the things you would do with hashtags, for example search for "time:after" as a complete hashtag just as you would if it did not contain an : sign.

find factors label contains time:after

is equivalent to

find factors time=after

Syntax

Quickfields must be composed only of letters and numbers, no punctuation or spaces. There must be no gap on either side of the colon (:).

More than one quick field

You can combine more than one quick field at once:

Crops improved / time:after / valence:2 (the / symbols here are not strictly necessary but improve readability.)

Permanence

Quickfields become part of the factors table, you can even export them, but the values of the fields are always immediately overwritten if you change the labels. Values of fields derived from labels always take precedence, so there is no point roundtripping a quick field as it will be overwritten according to the labels. (You can always roundtrip other fields in the factors table which are not mentioned in the labels. So you can have a field called, say, valence, which is not mentioned in the factor labels in the form, say, / valence:high).

Position in hierarchy

Usually quickfields are put at the end the labels. This makes it easy to find and manipulate the factor labels in the Factor Editor. But quicklinks can appear anywhere in a label and in particular anywhere within a hierarchy:

Anti-discrimination policy / theme:gender; disciplinary rules

Anti-discrimination policy / theme:race; awareness training

This makes particular sense when the lower levels of the hierarchy might differ according to theme. Hierarchy and quickfields are a powerful combination.

Current limitation

When you use filters which merge factors together (zooming and bundling factors) the values of quickfields for any factors which are lost because they are nested into the higher levels are not available.

Quickfields for links{#quickfields-for-links}

Quickfields are now available for links too. You use the same syntax in the "hashtags" box as for quickfields for factors. This might seem like a small tweak but it means you can now systematically include information about the context of each link. Whereas previously you could mark that a particular link refers only to the context of before Covid by marking it with a hashtag like #before-covid, now you can use systematic sets of hashtags like #covid:before and #covid:after, which automatically then adds a column to your links table called covid which can have values like before, after, and anything else you type.

Quickfields for factor labels versus ordinary hashtags for links

It makes most sense to add dimensional information as described above when the distinctions are best understood as part of the causal structure, as part of the hierarchy of factors. For example:

- You want to express two or more different dimensions within your factor system at the same time, e.g. both time and theme.
- There might be qualitative differences in the lower parts of the hierarchy, e.g.
 - Approach /sport:football; dealing with penalties
 - Approach /sport:swimming; different lengths

On the other hand, you might prefer to use link hashtags rather than quickfields if:

- The dimensions you want to record cut right across your entire factor system and don't apply to any particular part of it

- You want to combine the information with information from your sources table. For example, you might have a set of 20 documents, 5 on each of 4 countries, and so in your links table (transferred automatically from the sources table) you already have a field called, say, #country; but sometimes a report on one country, say, China, makes causal claims about a different country, say France. In this case, you might want to add a hashtag #country:France to these links.
 - What should the app do in this case? Should it override the information in the #country column of the links table, which would have been China in this case, rewriting it as France?

--chapter:end:42560-quickfields-coding.rmd--



Plain coding{#xplain coding}

Summary

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

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

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

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

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

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

```
find links field=hashtags value=#plain_coding operator=notcontains
```

There is a shortcut button to apply this filter in the addons section in the dashboard.

Addons (added to existing view)

- ★ Combine opposites ★ Contrary factors get orange text
- ★ Exclude links with only 1 mention ★ Hide plain codings 
- ★ Hide quickfields ★ Links with more than one mention
- ★ Pink borders - fixed colour ★ Pink links - fixed colour ★ Remove isolated
- ★ Smart Zoom, preserve 5 ★ Standard format ★ Standard overview edited
- ★ Standard style ★ Top 10 factors ★ Top 15 links

It would also be possible to format this kind of link in a special way, e.g. by not showing any arrowheads.

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

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

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

What the app does

This implementation of plain coding is unsophisticated and does not involve the underlying Causal Map Functions package at all.

There is just a little button for plain coding.

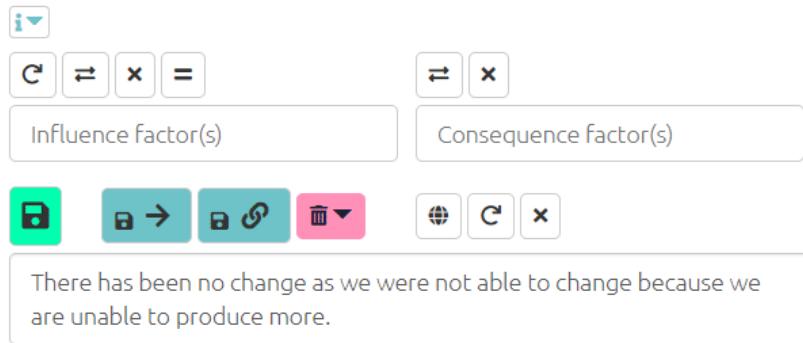


- If you press it, it will toggle the hashtag `#plain_coding` (you could also edit the hashtag manually but why bother).
- When a coding you are going to make, or one which you are editing, has this special hashtag, the consequence factor selector is greyed out and always has the same contents as the influence factor.
- When doing plain coding, if you put more than one factor in the first factor dropdown, each factor will just get one link to itself (rather than links being created between all the factors).

```
source_id MNX-1  
question_id2 E5  
question_id $E5  
question_textglobal question
```

“

There has been no change as we were not able to change because we are unable to produce more.



Future

Potentially it is possible then to use Causal Map to code some text entirely with plain codes.

At the moment, you would have to still have to press the "plain coding" button every time you create a coding, but this could be also made into a default setting.

Should we introduce more functionality around the analysis and reporting of plain codings? Let us know at hello@causalmap.ap

!--chapter:end:43000-plain-coding.rmd--

Document coding score{#xgamification}

Coding a large file can be a daunting task. At the top right of the app, you will see some feedback on your progress in coding your file.



These show you:

- number of (top level) factors. You get a warning if this is more than 100. Ideally you'd have less than 10 or 20 top-level factors, enough so that you can remember most of them and could explain them to someone else without notes. Of course, each top-level factor might have subfactors, but again we'd recommend no more than say 10 lower-level factors within each higher-level factor. So, it doesn't really matter how many factors you have, but the number of top-level factors should be fairly small.
- ratio of links to (top level) factors. You get a warning if this is less than 3. Ideally, each of your top level factors will be mentioned lots of times (including mentions of factors nested within them, like Health behaviour; hand washing) so they will have lots of links. If your factors don't have many links each, consider consolidating them.
- percentage of statements which have been coded. You get a warning if this is less than 60%.

!--chapter:end:48557-gamification.rmd--

Level 4 quiz

Now do the quiz to see if you've understood about advanced coding.

!--chapter:end:48601-quiz4.rmd--

(PART*) 5: Advanced analysis{-}

!--chapter:end:50000-PART-advanced-analysis.rmd--

Level 5 overview{#xlevel5}

To whet your appetite for some advanced analysis, here are some of the variety of maps you can create with Causal Map.

In the pages in this section you will learn how to do advanced analysis in Causal Map:

- using the filter buttons

- using the advanced editor
- and lots more.

When you've read the pages, try the quiz to see if you have understood enough or need to review some things!

!--chapter:end:50001-level5.rmd--

Dashboard (advanced){#xdashboard-a}

Using and sharing a view{#xsharing_view}

Click on:

- A row to load the associated filters into the current file.
- The blue View number to open the view in an other tab.
- Share icon: Copy the link and filter to your clipboard, for pasting into a document or sending to someone. If you send someone a shortlink to a view, when they click on it they will be taken to the same view of the current file (providing you have given them View or Edit rights to the file).
- Edit icon: Edit the view - see below
- Hide icon: Hide this row in future from this table. You may want to use this button to clear unused views from your table. If you already shared this link with someone, it will still work even though it is hidden.

Recommended views versus custom views

Recommended views

When you first create or copy a file, Causal Map will create for you some recommended views. These are marked with a star.

- These recommended views are copies specific for the current file. You can edit them later, or hide them if you don't want them.
- Some recommended views might not be shown in the list if they would not be relevant to the current file -- for example if you do not use hierarchical coding, views which are only relevant to zooming are not shown. Occasionally you might see them appear later as you add features to your file, for example if you add

hierarchical coding.

- Occasionally you might see additional recommended views added to the list as we publish new ones.

Custom views - create your own

You can create your own views for the current file at any time.

Add a shortlink to the current view

Title for your shortlink

Set your map or table up the way you want it then put an (optional) title and press Save.

- A shortlink to this view will also be copied into your clipboard so you can paste it straight into a document or send it to someone else.
- If you put a title, the view will also appear in the list of views for the current file.
- If you don't put a title, it won't be listed, but the shortlink you created will always point to that same view.

Complete views versus "Add-ons"

Complete views replace the existing view and all the filters, included for example analysis filters which change the map itself, for example show only the most frequent links, or zoom out to the top level of a hierarchy. If you are looking at the Advanced Editor, you will see that the entire contents of the editor are replaced.

"Add-on" views are added after any other existing filters and do not replace them. So you can use them for example to add a particular kind of formatting, like making all the links pink, or to add one instruction, like "remove all links mentioned only once".

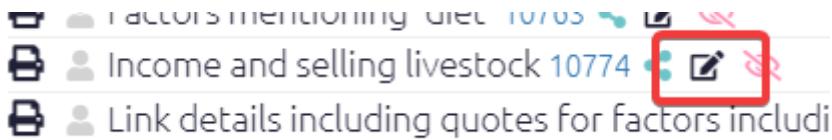
Addons (added to existing view)

- ★ Combine opposites ★ Contrary factors get orange text
- ★ Exclude links with only 1 mention ★ Hide plain codings
- ★ Hide quickfields ★ Links with more than one mention
- ★ Pink borders - fixed colour ★ Pink links - fixed colour ★ Remove isolated
- ★ Smart Zoom, preserve 5 ★ Standard format ★ Standard overview edited
- ★ Standard style  ★ Top 10 factors ★ Top 15 links

Complete views belong with the current file and you can edit them, whereas Add-ons exist separately of individual files (actually, they belong to example-file) and you can't edit them (unless you are an admin).

To edit the content (filter, tab etc) of an existing view: {#xedit_view}

Set up the content the way you want it in the app, so it looks exactly the way it should when someone in the future clicks on the link. Then go to the filter which you want to update and press the edit button:



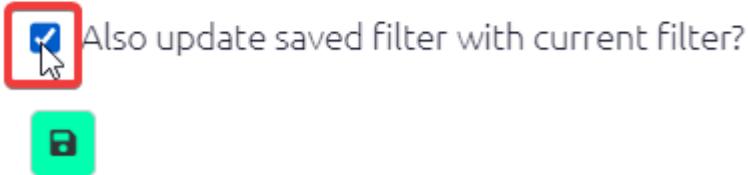
In the modal dialog which opens, make sure you check this box. You can optionally change the title too.

Title

Overview SHA 01

Description

Tip: you can use markdown in your description



If all you want to do is change the title, make sure the checkbox is **not** checked.

Applying a view

💡 Each view has a unique number. They are shown in the dashboard. When you view a map which has been downloaded as an SVG image you will see this unique number in the bottom left-hand corner.

Apply a view

Shortlink nr.



You can paste or type this number into this box at the bottom of the dashboard and press the arrow and the view will be applied. If the view is for a different file to which you have access, the app will first take you to that file.

For admins

Recommended views can only be added / edited by admins, and only from `example-file`.

When you create a recommended view it should be applicable to almost any file, so don't use filters which refer to particular factor labels etc.

Recommended views are then copied to every other file. Editors can then edit, rename etc that recommended filter in the other file. But that won't change the original filter in `example-file`. Editors can hide a recommended filter in any other file except `example-file`, and it won't come back.

Addons work a bit differently from complete views - they can also only be created within `example-file`, but they are not copied to other files. As admin, you can create your own addons just for individual files if you want, by not clicking "recommended" -- maybe useful if you have a lot of steps you want to repeat just within a particular file.

To "delete" an addon, edit it and delete the title.

`!--chapter:end:51000-dashboard-advanced.rmd--`

Filters -- overview

```
knitr:::include_url("https://player.vimeo.com/video/618270513")
```

Applying different filters{#xapplyingfilters}

Filters simplify your graph to help you answer specific questions. You will often apply more than one of these filters in order, e.g. first select only specific factors and then make the most frequently mentioned factors larger.

- **Analysis filters**
 - **Exclusion filters.** These filters simply exclude links and/or factors from the map. So instead of seeing the entire unfiltered causal map, which can be quite bewildering, you see only part of it. For example:
 - show me only information given by women
 - show me only information from the first statement
 - show me only paths leading from My intervention to My important outcomes
 - show me only the most frequently mentioned factors.
 - **Simplification filters.** If you want to combine two similar factors, you will be using a simplification filter. These filters actually involve rerouting links, e.g. if you want to combine the factor Ebola into Infectious diseases the app will need to remove the factor Ebola and reroute all its links so they connect with Infectious diseases instead.
- **Conditional formatting filters.** You can format the links and factors in your map to reflect the data, e.g. you can make factors bigger if they were mentioned more often.
- **Simple formatting filters.** These are simply used to apply overall formatting, e.g. give all the factors a grey border.

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 in the text of the statement. So to apply a find statements filter, you click the button with that name. A dialog will appear, once you have filled in the panel, and you click Apply filter.

Find statements

[Apply filter!](#)

Keep this panel open?

What field to search in?

text

Type to find a value or values for this field

women

Operator

contains

equals

contains



A new button with an explanation appears in the active bar. The corresponding button below it 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:

Find statements

[Apply filter!](#)

Keep this panel open?

What field to search in?

1. Sex

Type to find a value or values for this field

m



Operator

equals

In this case, you are searching not the `text` of the statement but the `sex` of the source who made this statement.

Find factors

Searching and filtering factors

To find factors which contain the word Food in their label, click find factors and type to select factors which you want to find.

Find factors

[Apply filter!](#)

Keep this panel open?

Include factors how many steps up?



Include factors how many steps down?



What field to search in?

label

Type to find a value or values for this field

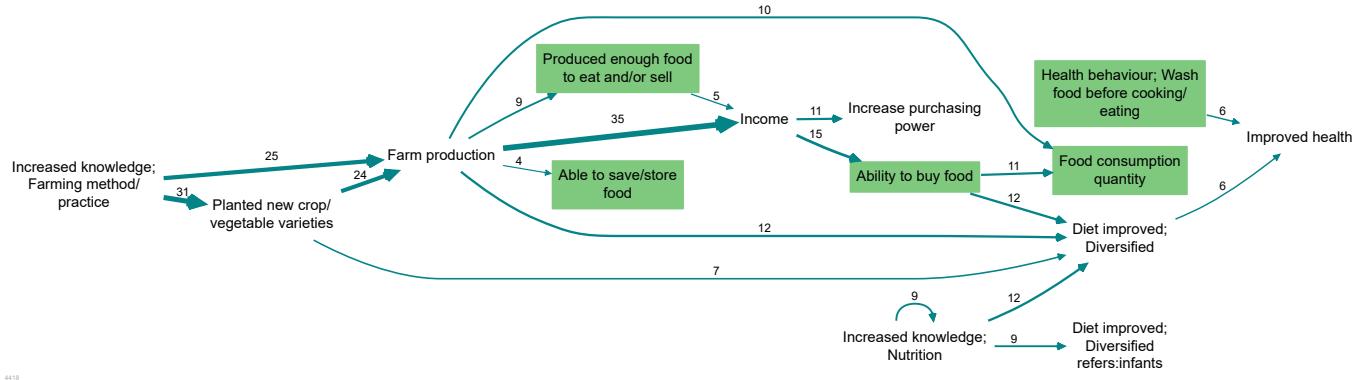
food

Operator

contains

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

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 text 'Flooding'

One step upstream – one step downstream:



4419

One step upstream – zero steps downstream:



This analysis function is particularly useful for searching for important factors such as intervention activities or intended programme outcomes and can be used in conjunction with hashtags.

This also works when you include 0 steps upstream and 0 steps downstream:

Ability to buy food

Find statements, find links {#xfind-statements-links}

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

As each statement is only from one source, filtering for all statements where, say, gender = female is the same as filtering for all links where gender = female.

To use these filters, select a data field from the dropdown menu.

- 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.
- You can select one or more values (e.g. several age groups or several respondents).
- If you are looking for a particular statement, which you know the id for, you can search for `statement_id` to view that statement. Or you can search for all statements with `statement_id` greater than, say, 100.
- **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.

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).

Find links

There are a lot of useful filters specifically for links. For example, rather than using `select_links` to select the most frequent simple bundles of links, you can be more specific by choosing only simple bundles containing more than a certain number of original links. The field you need is called `simple_frequency`, i.e. the number of links in

each simple bundle. So if there are only 74 links in the bundle of links between factor X and factor Y, this bundle will not show up if you use the following filter:

The screenshot shows the 'Advanced Editor' interface with the following configuration:

- Statements: Many
- Show filters
- Analysis
- Keep only links where simple_frequency greater 77

Find hashtags

You can filter for links which contain (or do not contain) a certain hashtag. Note that if a specific link contains more than one hashtag, e.g. doubtful and future you can just that specific link by selecting that specific combination of hashtags from the dropdown, e.g. doubtful, future.

Highlight only

Maybe you want to just highlight the factors or links you find, but not remove the others. You can do this with the advanced editor, adding highlight_only=TRUE:

```
find factors value=risk operator=contains highlight_only=TRUE
```

```
color factors field=found lo=white hi=#7FC97F
```

You can use any of the normal conditional formatting options, e.g. color factors field=found lo=blue hi=red - the hi colour will be used for factors which match the search, and the lo colour for those which do not.

This is a useful way to highlight factors which have the opposites symbol ~:

```
find factors value=~ highlight_only=TRUE
```

```
color borders field=found
```

At the moment you can't use this to colour factor background by one criterion and colour factor borders by another (because they both use the hidden field called found).

Remove brackets

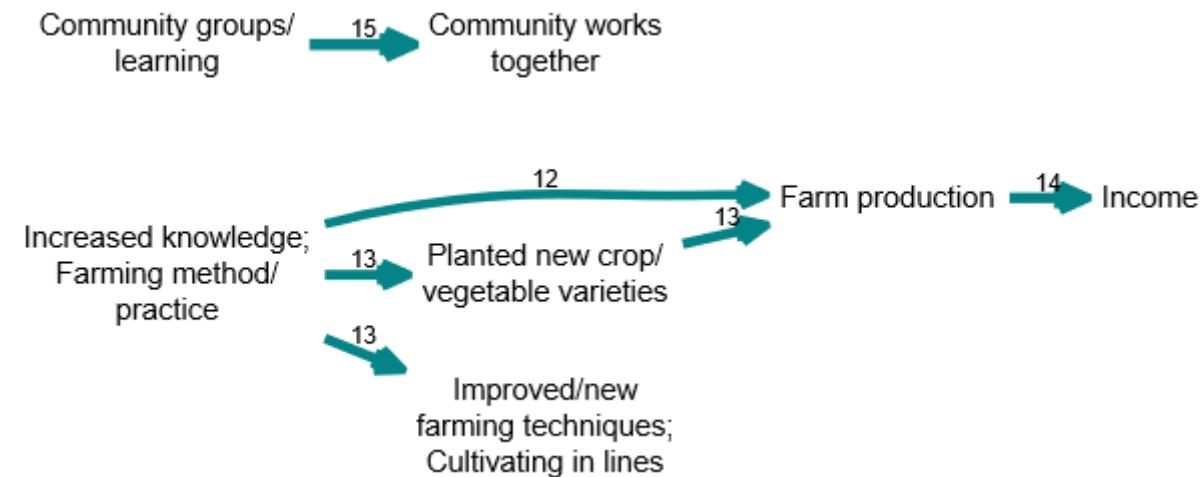
This filter hides any tags or other words written between different kinds of brackets. So instead of this:

Funds from Orgx (E) → Better flood defences built (E) → Flood risk reduced (I) → Increased investment into the area (I)

you see this:

Funds from Orgx → Better flood defences built → Flood risk reduced → Increased investment into the area

Excluding factors, links, statements, sources and questions with notcontains and notequals



Find factors

[Apply filter!](#)



Keep this panel open?

Include factors how many steps up?

0 **1** 10

Include factors how many steps down?

0 **1** 10

What field to search in?

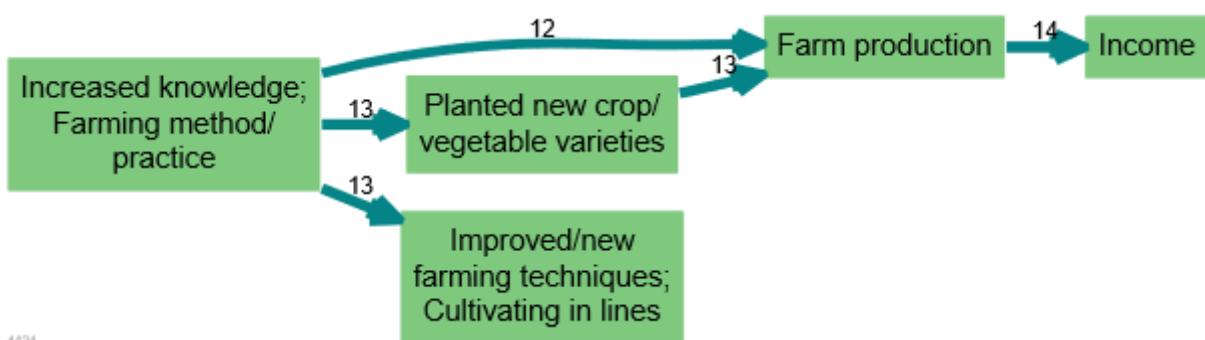
label ▾

Type to find a value or values for this field

Community ▾

Operator

notcontains ▾



4424

Note how, logically enough, following a search for factors which do not contain a certain phrase, all the remaining factors have by default a coloured background.

You can exclude statements, sources etc in the same way. Or you could filter for all links which do not contain the hashtag `doubtful` with this filter:

```
find links field=hashtags value=doubtful operator=notcontains
```

Trace paths

Trace paths

This filter shows all the paths (up to a length which you can specify) from one factor or set of factors to another.

Keep this panel open?

Trace how many steps?

0 10

Remove in-between links?

Begin tracing from which factor(s)?

End at which factors?

What field to use for tracing threads?

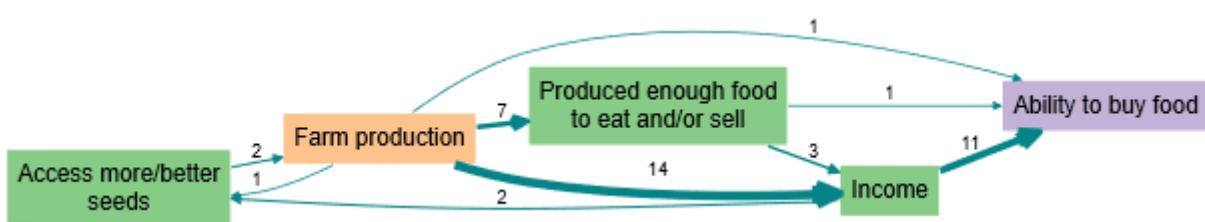


Also trace threads up or down?

none

down

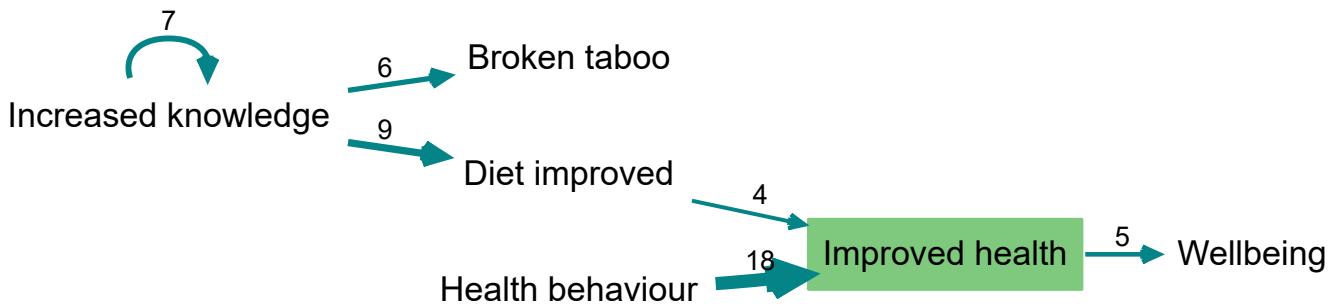
up



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

Bundle factors{#bundlefactors}

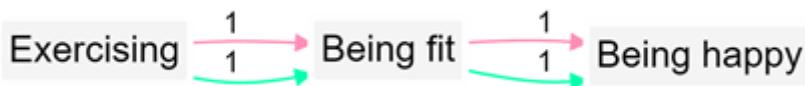
Bundling factors allows us to combine multiple factors into one box. This can be useful if we want to see the influences or consequences of factors with similar attributes. To use this function you just type in the search box what you want the factor label to contain and the app will bundle all the labels containing that text into one. Remember to use the exact spelling and case that is in the factor labels you want to combine. The below screenshot shows all factors including the words 'improved health' bundled into one factor label.



4426

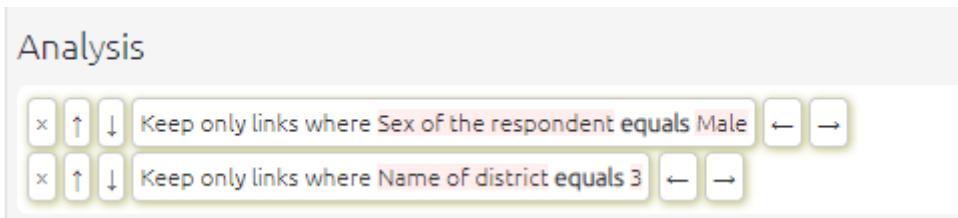
Combine opposites

Combining opposites formats factors so that opposites are displayed in one factor label. Negatives are identified by '~' and must have exactly the same phrasing as their counterpart. In Print view the links are automatically colour coded - red for the negative relationship and green for the positive, so you can clearly see the causal relations.



Filtering group by group

When you use an `equals` filter you are usually filtering for a particular group, for example all men or all statements from a particular district. In this case, "next" and "previous" buttons appear after the full-text description of the filter, as in the figure below. You can use these buttons to click from men to women, or to click through the districts. As in the example, you can use more than one such filter simultaneously.



These "next" and "previous" buttons appear when using `equals` filters with links, factors or statements (except when filtering statements by `statement_id`, as in this special case the next and previous buttons are already visible in the statement navigation panel below).

!--chapter:end:51020-analysis-filters.rmd--

What is the logic behind the Causal Map filters?

Keeping it simple

It's easy to end up with complicated filters. It can get hard to explain to someone else (or even understand yourself) what is going on. Ideally you should have a clear model of what you are doing and be fairly consistent about it, and be able to explain and defend it. There is no difference on this point between qualitative and quantitative analyses.

The order of the filters with calculated fields

We have changed the point in the chain of filters when calculated fields are calculated. Previously, factor frequency meant "the frequency as it was at the start of the chain". Now, it means "the frequency of the factors as they are now at this point in the chain of filters". The same applies to nearly all the calculated fields:

- from_label,
- to_label,
- simple_frequency,
- simple_bundle
- simple_bundle
- to_label,
- from_label,
- betweenness, etc.

The exception is the fields necessary to display source continuity, which is another kettle of fish.

We think this is easier to understand and more intuitive.

Autocorrect{#xautocorrect}

The app does quite a bit of autocorrect for you. It enforces the filter logic. You will sometimes see lines in the Advanced Editor moving around, deleting themselves etc.

An example:

You can't have more than one set print filter, but you can have more than one find factors filter.

Another example:

The app needs to delete preceding label links filters in some cases, but not if preceded by another label links filter if clear_previous=TRUE.

The point is to let you chain labels together if you want say source count and also gender, but to delete extraneous label links filters if you are not chaining the labels.

Where do these islands come from in my map?

It's a really frequent problem that you do either a path tracing or a "find factors containing blah and two steps down etc" and then "select ..." to reduce number of factors or links, and then you get islands of stuff which don't seem to be related.

If you ensure that your select filters come earlier, you never get these islands and the resulting map makes sense as a product of the find or trace filter which you did.

This might seem a bit back to front, because in your working steps you usually apply a Find and then try to simplify it. It is also a bit strange in the sense that the numbers on the Select filter aren't as easy to guess. But it works.

In the same vein, it often makes sense to put any "find links" or "find statements" filters, e.g. looking for a particular village, right at the start.

Causal Map in any case enforces an ordering of the filters:

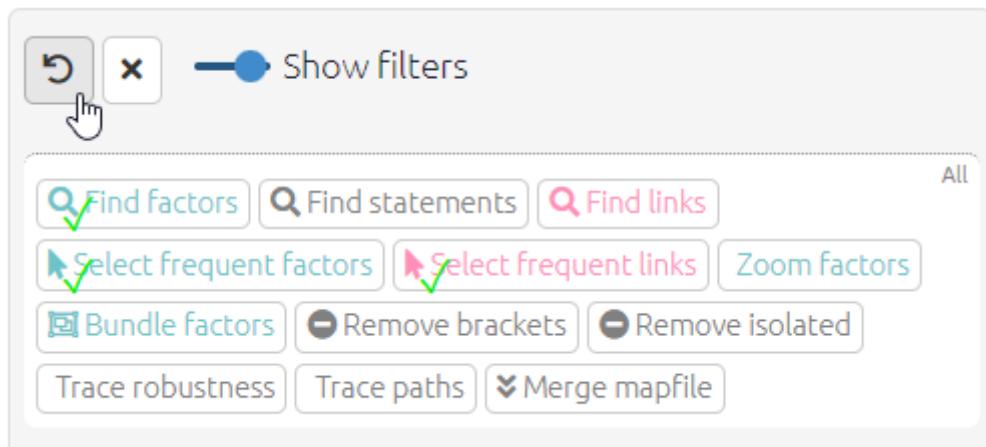
1. Analysis filters (each of these can appear more than once)
2. Conditional formatting filters (each of these can only appear once)
3. Simple filters (each of these can only appear once)

--chapter:end:51030-filters-logic.rmd--

Filters: tips

Restoring previous filter

By pressing this button you will restore the filter you applied previously in the current session.



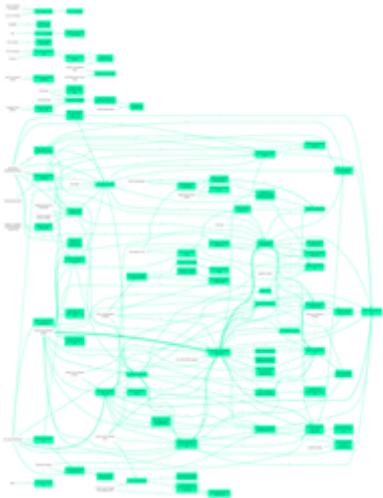
Analysis: comparing maps between particular groups{#comparing}

Filtering by respondent group

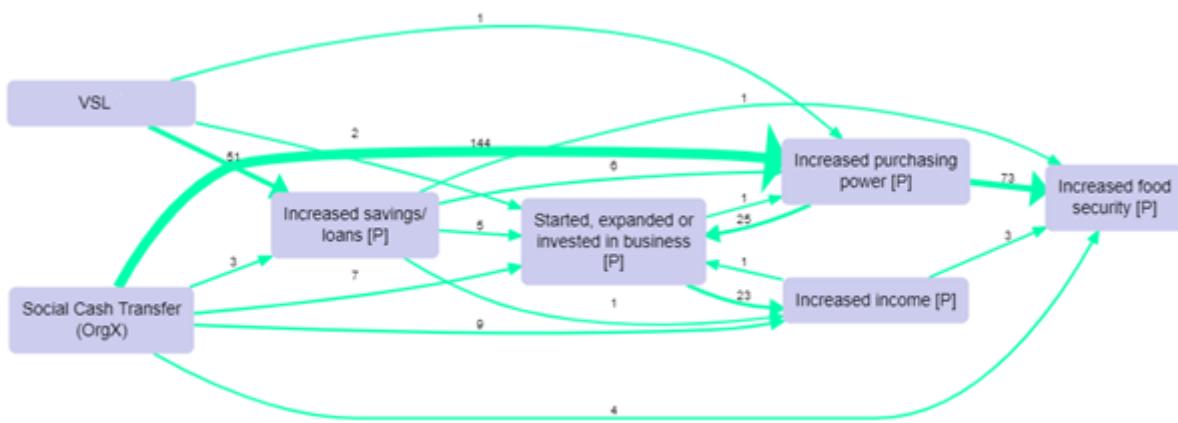
We code a causal map on the basis of text data. That text data can be usefully broken up into statements, usually of a length between a paragraph and a page. Each statement usually has "**additional data**" associated with it, for example the ID or gender of the respondent, the text of a question to 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.

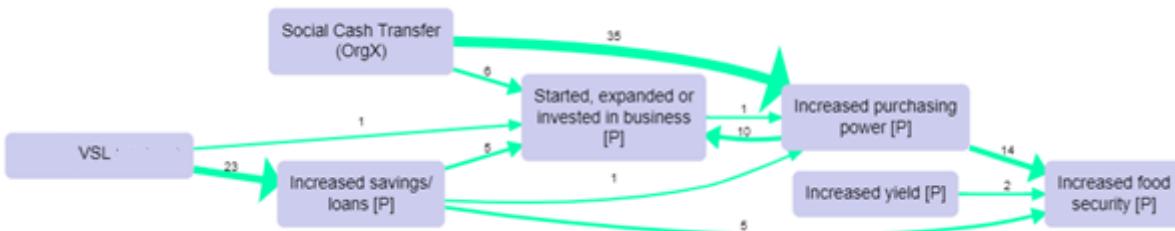
It is easy to filter a causal map by this additional data. This idea goes back at least to (Ford & Hegarty, 1984). For example, here is one map filtered to show all and only the links mentioned by female respondents. We call these the **per-value maps**, e.g. the map consisting of all links mentioned 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:



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



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?

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.

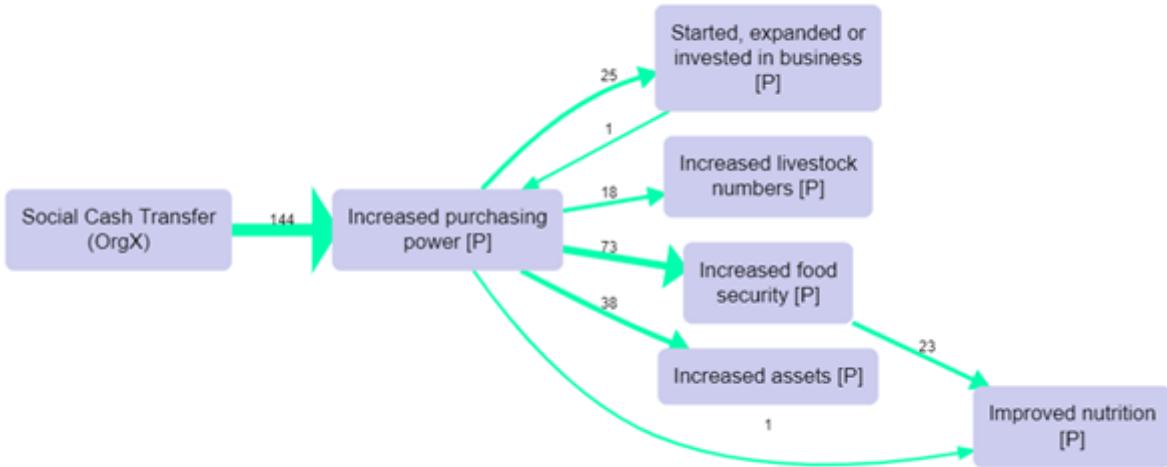
This idea takes a bit of getting used to. A distinctive map for a group filters out all links between pairs of factors which are both not distinctive for that group. That is the same thing as saying that it shows all the factors which are distinctive for that group, plus factors which are linked to those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

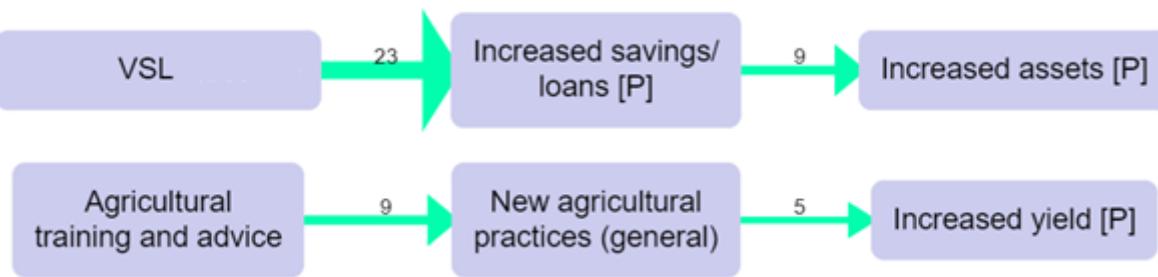
It is also possible, especially with smaller maps, that there are no links which are distinctive for a particular additional data field like gender. This suggests that there are no very interesting differences between the groups. It is also possible that there is a distinctive map for one group, e.g. 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.



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



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.

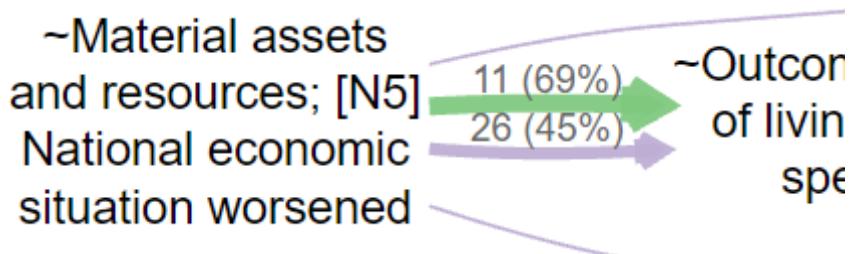
As usual, we should treat the maps just as a gateway into the data, an invitation to read and understand the quotes in more detail and in context.

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.

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.

Percentages{#percent}



It's very useful to bundle many links between every pair of factors into different bundles, one for each gender, say, or one for each village. Then you can compare how many times this particular connection was mentioned by each group (if you count link_id) or how many sources in each group mentioned the connection (if you count

source_id). You can print this number as the label for each link, or you can scale or colour the link according to it.

However, if the base number of sources in each group is different -- for example, if there are many more women than men -- these numbers can be hard to deal with. In this situation, we recommend using percentages of the baseline for each group rather than absolute counts.

There are many caveats to using counts and percentages of mentions of causal links and factors in causal mapping, as they arise (or don't arise) during conversations and are not answers to questions directly asked of respondents.

Percentages in the app

The details of how to do make these calculations are quite open; at the moment, the app uses the relevant baseline within the current map. So if you want to display "source count", or the number of sources who mentioned each link, and bundle the links by gender, the app takes as the baseline the source count for each gender in the current map. So if 7 women and 5 men mentioned a connection between X and Y, and altogether 14 women and 10 men mentioned anything in the whole map, then the percentage will be 50% on each.

If you use percent for labels, the label shows just the percentage.

If you use proportion for labels, the label shows the raw number / the baseline.

If you use percent or proportion for colour or width of the links, the percentage is used rather than the raw number.

So for example you could use a filter like this

```
bundle links group=Sex  
label links value=percent: link_id  
color links value=unique: Sex
```

or you can add both percentage and proportion to the labels

```
bundle links group=Sex  
label links value=percent: link_id clear_previous=TRUE  
label links value=proportion: link_id clear_previous=TRUE  
color links value=unique: Sex
```

Surprise{#xsurprise}

The most interesting filter uses a function which we call `surprise`. Here is an example: [2302](#).



We have bundled the links by name of District (1, 2, 3 or 4) and scaled the links by source count (count of the number of sources mentioning each link): but we have coloured the links by `surprise`. The bright green from poor weather conditions to poor farm production means that this route was particularly often mentioned in district 3 even taking into account the number of times people in district 3 mentioned anything (in the whole filtered map). And the bright purple link for district 1 means the opposite. Whereas for district 4, there are quite a lot of mentions but the pale colour shows us that this is not particularly surprising. Basically, using `surprise` means you can ignore the pale links and just look at the brightly coloured ones.

We recommend adding the initial of the group and the percentage to the link labels.

Sometimes you will see that the baseline is lower than the actual number in your dataset. This is because the percentage printout and the surprise calculation use the number in each group in the current map, not in the whole unfiltered map.

You might get tempted to think that if half the bundles are green for one group, the other bundles for the same group have to be correspondingly purple, but that isn't necessarily the case if you think about it.

!--chapter:end:51100-comparing.rmd--

Quickfields: analysis{#xquickfields-analysis}

Summary

We already saw that Quickfields are a special kind of hashtag for factors and links. They are a refinement of hashtags for when you have a set of categories which is mutually exclusive.

You can add information about one or more dimensions using a format like this:

Crops improved / time:before

Crops improved / time:after

You can search for and display this information across all your tables and maps, not separately as they would be if you used ordinary hashtags like #before and #after, but together.

This feature is automatic and does not need to be switched on with any filter.

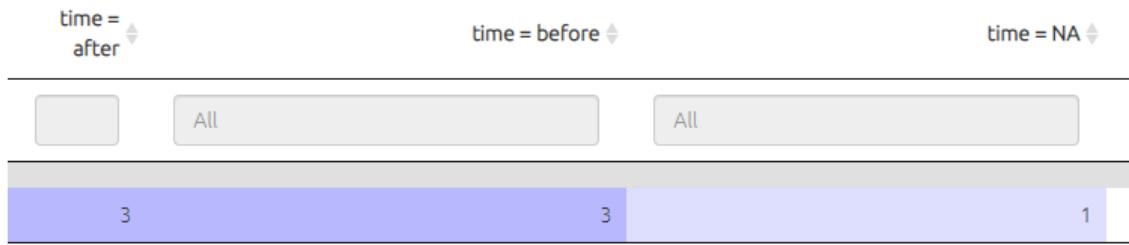
Quickfields can be hidden in labels using the `hide_quickfields` filter. This can provide a kind of zoom.

Uses

- to search, filter, and organise the factors, links and mentions tables

factor_id	label	time
	All	All
2	skills; time:before	before
3	success; time:before	before
4	success; time:after / happiness:2	after
5	received training; time:before /happiness:3	before
6	skills; time:after /newtag:8	after
7	bought house; time:after /happiness:1	after
8	pre-existing things	

- e.g. to summarise all the factors marked as before



- to filter the map to show only factors marked as before.

Find factors

Keep this panel open?

Include factors how many steps up?

0 10

A horizontal slider with a teal track and a grey handle. The handle is positioned at the value 1, which is highlighted with a teal square.

Include factors how many steps down?

0 10

A horizontal slider with a teal track and a grey handle. The handle is positioned at the value 1, which is highlighted with a teal square.

What field to search in?

time ▾

A dropdown menu showing the option 'time'.

Type to find a value or values for this field

after

A text input field containing the value 'after'.

- to filter the map to exclude factors marked as after.
- to colour factor backgrounds, borders and/or text, or to scale the factors, according to the value of the quick field time.

If you set numerical values e.g. `time:0, time:1, time:2` then the quick field `time` will be treated as numerical and so in particular continuous formatting for colours will be used e.g. a fade from red to blue.

Quickfields are also ordinary hashtags

Quick field hashtags are still hashtags, so you can still *also* do all the things you would do with hashtags, for example search for "time:after" as a complete hashtag just as you would if it did not contain an : sign.

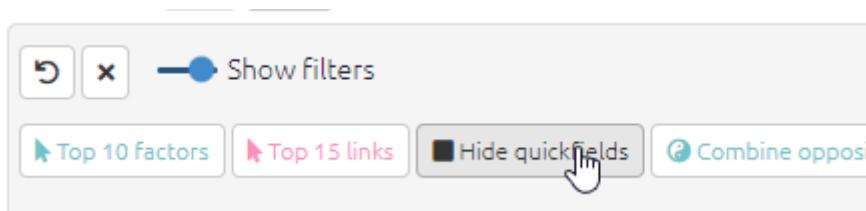
find factors label contains time:after

is equivalent to

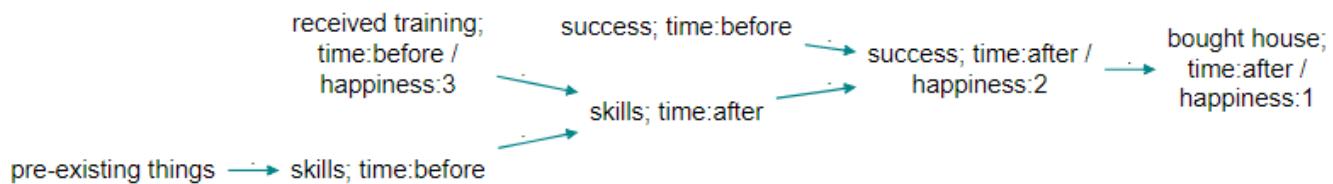
`find factors time=after`

Hiding quickfields{#xhiding-quickfields}

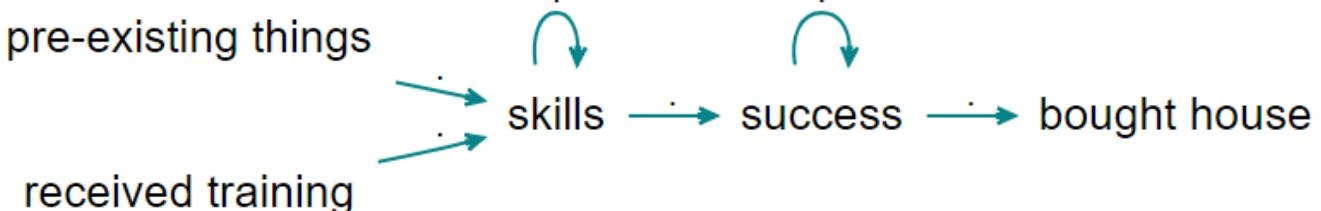
Quickfields can be hidden in labels using the `hide quickfields` filter,



to go from this:



to this:



Slashes and spaces and any leading semicolon will be removed.

Current limitation

When you use filters which merge factors together (zooming and bundling factors) the values of quickfields for any factors which are lost because they are nested into the higher levels are not available.

Quickfields for links

Quickfields are now available for links too.

!--chapter:end:51150-quickfields-analysis.rmd--

Tracing paths {#xtracing-paths}



Summary

How to trace paths{#howtotracepathsandrobustness}

Tracing paths allows you to view full causal pathways and to analyse the relationships between specific causal factors. You will find the `trace_paths` filter in the `analysis` filters section. This will open the `trace_paths` panel which asks you to select how many steps to trace and which factors to trace from and to. Type one or more factors to begin your path (these are the `origin` factors) and which factor or factors you want the path to end at (the `target` factors). If necessary, move the slider to select how many steps down you want the path to go.

Trace paths

This filter shows all the paths (up to a length which you can specify) from one factor or set of factors to another.

Keep this panel open?

Trace how many steps?

0  10

Remove in-between links?

Begin tracing from which factor(s)?

training

End at which factors?

Improved health

What field to use for tracing threads?

source_id

Also trace threads up or down?

- none
- down
- up

If you want you can also leave either (but not both) factor selectors blank, e.g.

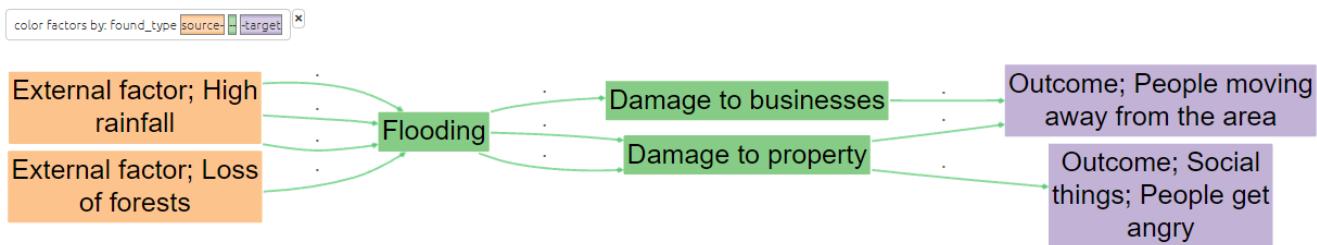
- to search from some specific factor or factors to "anywhere", or
- to search from "anywhere" to some specific factor or factors.

If you want to search from or to the **main** drivers and/or outcomes:

- put "main_drivers" in the first box
- and/or put "main_outcomes" in the second box.

"main_drivers" uses an algorithm which identifies 1-3 factors with the highest driver_score, i.e. and likewise for outcomes.

Once you click apply filter your map will appear, as below. Orange indicates the origin(s) of the path, purple the target factor(s) and green any factors in between.



In this example, there are causal paths from the two external factors to the two outcomes. In real-life examples, these paths may be much harder to follow, so they are best seen and shown by using path tracing, as it simplifies the map and highlights the intervening factors between the two factors.

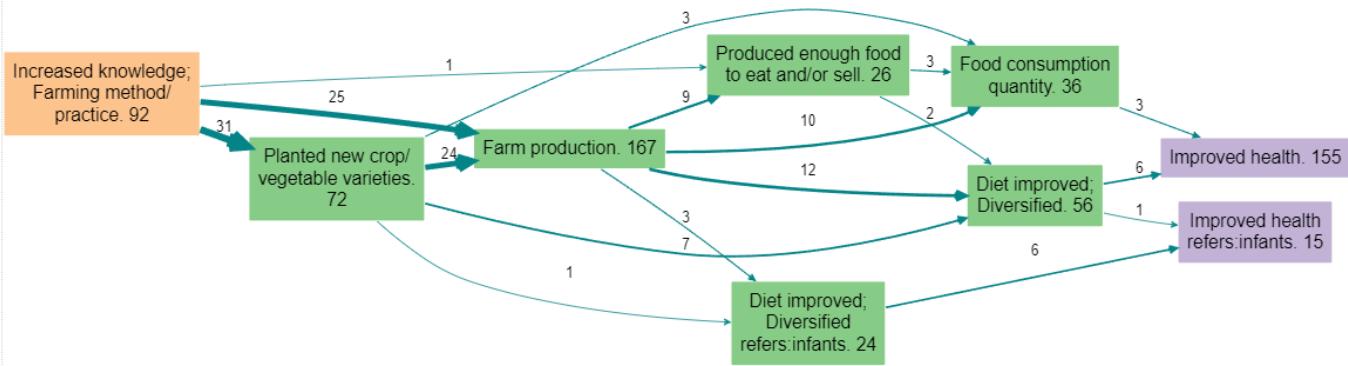
A path length of 1 will only show the one step in the causal chain from/to your chosen factor, i.e. A → B. A path length of 2 will also show the next step in the causal chain (if there is one!), i.e. A → B → C.

Removing unwanted links{#xremoving-unwanted-links}

Suppose you have a map with links B --> D, C --> D and also B --> C. If you run this filter:

```
trace paths length=1 to=D
```

the app will show you also the path from B to C, because both B and C are one link away from D, and there is a path from B to C. But that might have surprised you, as the path from B to C does not get you any closer to D.



So you might prefer this filter:

`trace_paths length=1 to=D remove_links=T` (which is equivalent to clicking this box in the dialog:)

Trace paths

This filter shows all the paths (up to a length which you can specify) from one factor or set

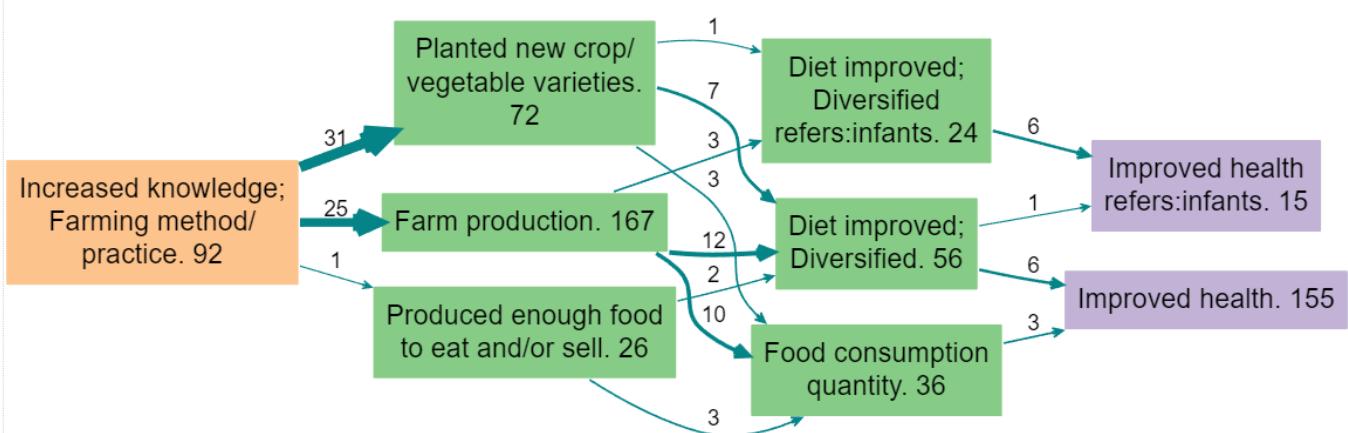
Keep this panel open?

Trace how many steps?

0 1 10

Remove in-between links?

... which will remove any such links.



In other words, `trace paths length=L from=X to=Y` says "show me all the factors which are on paths from X to Y no longer than L, and all the links between such factors" whereas `trace paths length=L from=X to=Y remove_links=F` says "show me all the paths from X to Y no longer than L".

Read more about tracing paths and tracing threads.

<!--### How to calculate robustness

Once you have applied the trace path filter, you can find the robustness calculations in the `robustness` tab. Here is the table for the example above.



The screenshot shows the 'Robustness' tab interface. At the top, there is a toolbar with icons for copy, print, export, and search. Below the toolbar, a message states 'Path Robustness is displayed only for the `trace robustness` filter'. There are buttons for 'Copy', 'CSV', 'Excel', 'PDF', 'Print', and 'Column visibility'. A search bar is also present. The main content is a table titled 'Maximum flow / minimum cut' with three columns: 'All origins', 'External factor; High rainfall', and 'External factor; Loss of forests'. The table has three rows: 'All targets' (values 3, 3, 1), 'Outcome; People moving away from the area' (values 2, 2, 1), and 'Outcome; Social things; People get angry' (values 1, 1, 1). At the bottom, it says 'Showing 1 to 3 of 3 entries' and has navigation buttons for 'Previous', '1', and 'Next'.

	All origins	External factor; High rainfall	External factor; Loss of forests
All targets	3	3	1
Outcome; People moving away from the area	2	2	1
Outcome; Social things; People get angry	1	1	1

In this tab you will find a table displaying the number of links between your `origin` and `target` factors. A larger number means more evidence; more pieces would have to be deleted until we gave up the hypothesis that the Origin factors influence the Target factors at all.

For example, in the top row of the table above you can see that there is more evidence that rainfall influenced both of the targets compared to the influence of loss of forests. In the first column, you can see that there is more evidence of the influence of external factors on people moving away than on people getting angry.

The following section will expand on how to understand these figures.

Understanding robustness{#analysingrobustness}

Robustness identifies how many pieces of evidence (individual mentions of individual links) would have to be deleted so that there is no longer any path from the Origin factors to the Target factors. It is important to understand a large number does not necessarily mean that this path is strong in the sense that the Origin factors have a large effect on the Target factors. It says there is plenty of evidence for the path, whatever its strength.

Subtotals are provided for each Origin factor, when there is more than one, and/or for each Target factor, when there is more than one; it can be useful to compare the amount of evidence for the impact of each Origin factor (and/or Target factor).

These numbers are not very useful on their own. But they are useful for comparisons.

Amount of overlap

If you look at the robustness of the path from interventions B and C to an outcome E, one thing you can do is compare the sum of the score for the two sets of paths and compare it with the overall score. If the overall score is more or less the same as the larger of the score for B and the score for C, you can assume that the (evidence for the) intervention paths overlap - so if the score for B is larger, you could say that C uses many of the same paths. But if the overall score is more like the sum of the two, you can deduce that the interventions work more independently. So for example if you are already funding intervention B, and you have to decide whether to fund C or D, and your causal map says that the score from C to E and from D to E are similar, you'd want to look at the pairs - you'd want to look at B plus C to E and compare it with B plus D to E. B + D might have a much larger score than C + D, because C shares many paths with B, but D doesn't. So you'd want to fund D, as the combination of B and D would be more robust.

Of course, in principle you can see all of this just by looking at the maps and to what extent B, C and D share paths to E. But sometimes it is hard to see and this metric anchors a narrative description in actual numbers.

Tracing paths and calculating robustness by field {#xtracing-paths-by-field}

Tracing paths by field, e.g. "by source"

[this is superseded by thread tracing]

Don't fall into the transitivity trap!

The trouble with the existing maps is that if you ask for paths from A to C, it might include links from sources which never actually mentioned any complete path from A to C. It might even be that there is no complete path from A to C in the sense that no single source mentioned any such path in its entirety.

By default, path tracing shows all the paths from the origin (set of) factors to the target (set of) factors which can be constructed using information from different sources. But sometimes this is not what we want. We only want to see paths which were actually mentioned in their entirety by individual sources. It is easy to construct a single such map from origin to target for a single source. And if we are interested in more than one source, we can combine those maps to construct a shared map for that group of sources.

By setting the field to `source_id` you only show links which are part of a complete path which at least one individual source mentioned in its entirety. The width of the links can be set to a new metric, namely the number of sources whose complete evidence included that path: who mentioned a complete path which ends at the consequence factor of the link. So you could for example set the width of the links to this new metric but for comparison set the link label to a familiar metric (e.g. `count: link_id`): the total number of mentions of the link (regardless of which source they are from, i.e. including sources which never mentioned this link as part of a whole path).

Even these maps can still be misleading: you must not assume that any arbitrary path you can trace through the map was actually mentioned by one particular source. You only know that each of its links was part of some path which was mentioned in its entirety by some source.

Calculating robustness by field, e.g. "by source"

Just because there is a lot of evidence for all the different parts of a path does not mean that any one source, or type of source, actually mentioned all those parts.

For this reason, `trace robustness` also lets you ask that question.

Trace robustness

This filter shows all the paths (up to a length which you can specify) from one factor or set in the Robustness tab.

[Apply filter!](#)

Keep this panel open?

Trace how many steps down?

0

3

Begin tracing from which factor(s)?

Start typing to select a factor, or factors, or part of a factor

End at which factors?

Start typing to select a factor, or factors, or part of a factor

For the robustness calculation, what field to use for counting? Leave blank for none.

Start typing to select a links field

For example you can put `source_id` in the box in order to count how many sources mentioned at least one entire path from the source to the target factor(s).

This is important because it's usually more credible to ask how many people mentioned all the bits of some path than to ask how many pieces of evidence are there altogether.

Or you could put `District` to find out, looking at each district individually, there is evidence (possibly assembled from different sources within that district) for the path.

=>

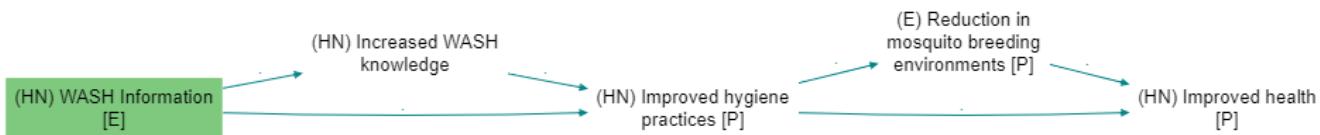
!--chapter:end:51630-tracing-paths.rmd--

Continuity and tracing threads{#xcontinuity}



Summary

When you use the trace threads filter to follow paths of influence across your map, the transitivity trap can make it a challenge to interpret your maps. Some people (or somebody) said that improved hygiene led to reduction in mosquito environments, and some people said reduction in mosquito environments led to improved health, but was this the same or different people?



Tracing threads (aka tracing continuity){#xtracing-threads}

If your map makes it look like Cash Transfer causally influenced Increased Wellbeing further downstream of it, you surely want to know if there are at least some sources who each mentioned the whole of one of the possible paths.

We call this "tracing threads". A thread is a chain of individual "original links" inside the bundles, whereas a path is just a chain of bundles, regardless of which sources mentioned each link.

Trace paths

This filter shows all the paths (up to a length which you can specify) from one factor or set of factors to another.

[Apply filter!](#)

Keep this panel open?

Trace how many steps?

0 10



Remove in-between links?

Begin tracing from which factor(s)?

training

End at which factors?

Improved health

What field to use for tracing threads?

source_id

Also trace threads up or down?

- none
- down
- up

Tracing threads is an option within the tracing.paths dialog. Click to trace threads up or down. Normally you will want to trace by source_id, but sometimes you might want to set, for example, village_id.

"Ah so there are plenty of paths from X to Y, but are there any threads? yes look, here is MXFG-1 who has a thread all the way down".

We can ask "how many threads make it all the way down to Improved Wellbeing"? The threads are named by the source ids, so we can say "Let's look at downstream thread MXFG-1".

When we are tracing threads, we can trace source threads, or village threads, or whatever. But usually we will trace in terms of sources, so if we do not specify which field defines the context, we assume we mean sources.

So a thread is a continuous chain of links with the same ID along a path. Within a bundle of links, each thread is one individual link associated with one ID (or sometimes, more than one link, if the ID mentioned that bundle more than once).

This creates the following calculated fields:

- Links fields
 - has_threads threads
- Factor field
 - threads_surviving
 - threads

You can for example label the links like this `label links value=unique: threads`.

(For compatibility purposes, the deprecated fields `has_downstream_threads`, `has_downstream_threads_surviving` etc. are also available.)

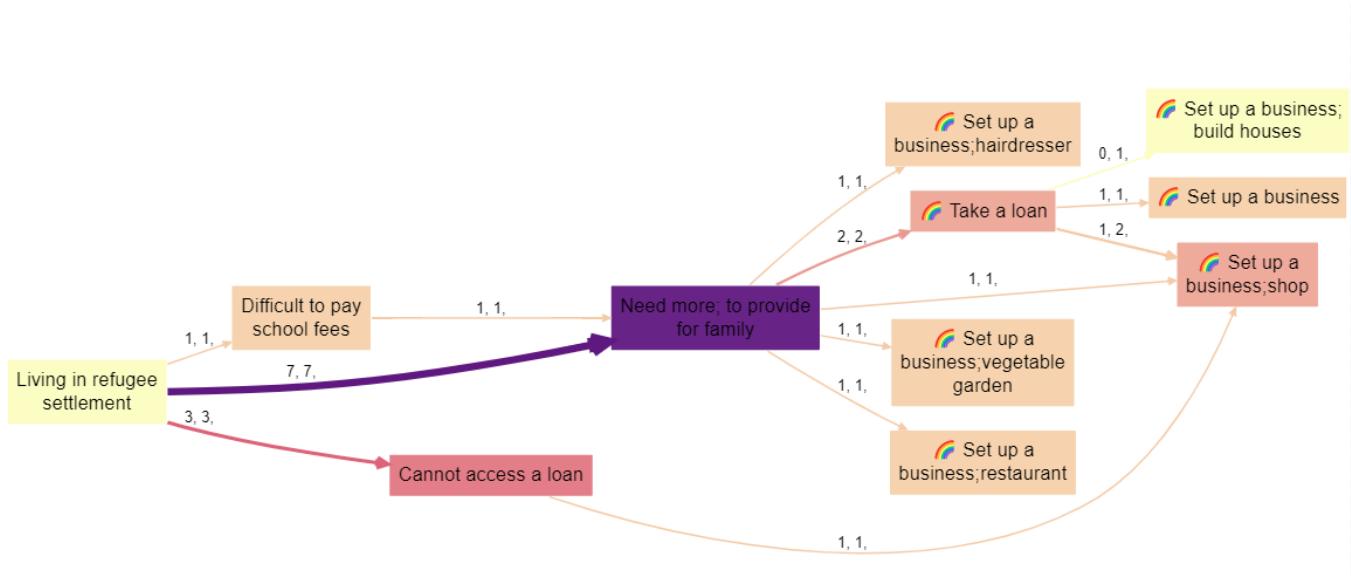
You can for example colour the factors like this `color factors field=threads_surviving`. So if you set a filter to trace `threads downstream` from Cash Transfer C, imagine factor C is coloured completely red, and imagine a factor F immediately downstream of it being coloured with an intensity of red according to the number of sources which mentioned the link from C to F. In fact every factor downstream of C will be coloured with a shade of red in the same way. The central question of whether any source mentioned a whole path from C to, say, "Increased wellbeing" can be answered simply by seeing whether there is any red in the colour of "Increased wellbeing" or whether it is completely white.

... So you can just look at C and watch its influence (in terms of source continuity) seeping downstream and spreading out along all source-complete paths.

It's like ink starting from C and flowing downstream as long as there is source continuity.

Tracing threads upstream asks about causes of effects rather than effects of causes: "Increased wellbeing" is completely red, and the redness of "Cash Transfer C" depends on the number of sources which mention a complete path from C to Increased wellbeing.

Tracing paths anchors the filter to one or more factors, and if you use thread tracing to format an aspect of the map (e.g. to colour the links) then this formatting is only related to this specific set of origin and/or target factors and will change if you trace from and/or to different factors.

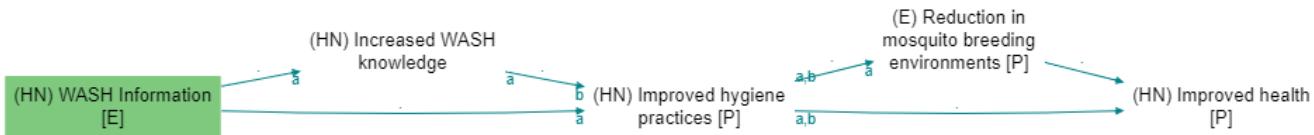


In this example, the width of the links reflects the number of sources mentioning it, but the colour of the links (and factors) reflects how many sources have a continuous story starting from "Living in a refugee settlement" all the way up to that point. This difference is also shown in the pair of numbers on the links: the first number is the source thread count, the second is the source count. So the top-right link was mentioned by one source, but the 0 shows that no single source mentioned a story all the way from "Living in a refugee settlement" which also includes this link.

Advanced diagnostic filter: Mark links for continuity (Print View only)

Showing continuity aka "showing threads" is concerned with adjacent sets of links (or factors). It cannot be used to definitively answer questions about continuity down a longer path (actually, sometimes it does tell us something: if there is zero continuity on a section of a path with no splits, we know there is zero continuity down the longer path; but more generally, if there is some continuity down all the sections of a path, it does not mean that there is continuity down the whole of it).

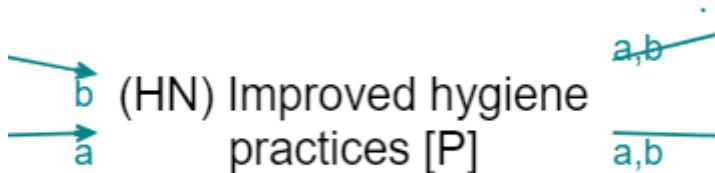
Add the filter `mark links` (there is no button for it, you have to type it) provides the following diagnostics:



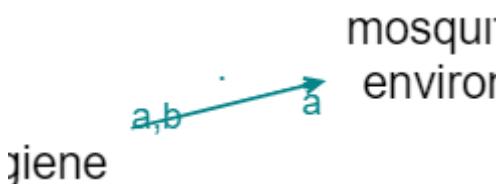
The incoming (bundles of) links to every factor (actually to every factor with outgoing links) are labelled a, b, etc, including when links are bundled e.g. by gender. Then the outgoing links are marked with say a if at least one of the sources who mentioned the outgoing link also mentioned link a.

So we can see that none of the people who said that improved hygiene led to reduction in mosquito environments also said that reduction in mosquito environments led to improved health: there is no label at all on the arrow going out of reduction in mosquito environments. There is no source continuity.

Note that the labels get re-used for each factor, so the as and bs here are related:



but the as here are not:



This also works with all the other fields, e.g. you can type `mark_links field=statement_id` in order to test statement continuity, which is a stricter test of continuity. `source_id` is default so you don't need to type it specially.

Yes, it is a bit difficult to communicate this in a report. But it is important for interpretation. Of course a chain without source continuity isn't an invalid chain per se, it's just something to be aware of.

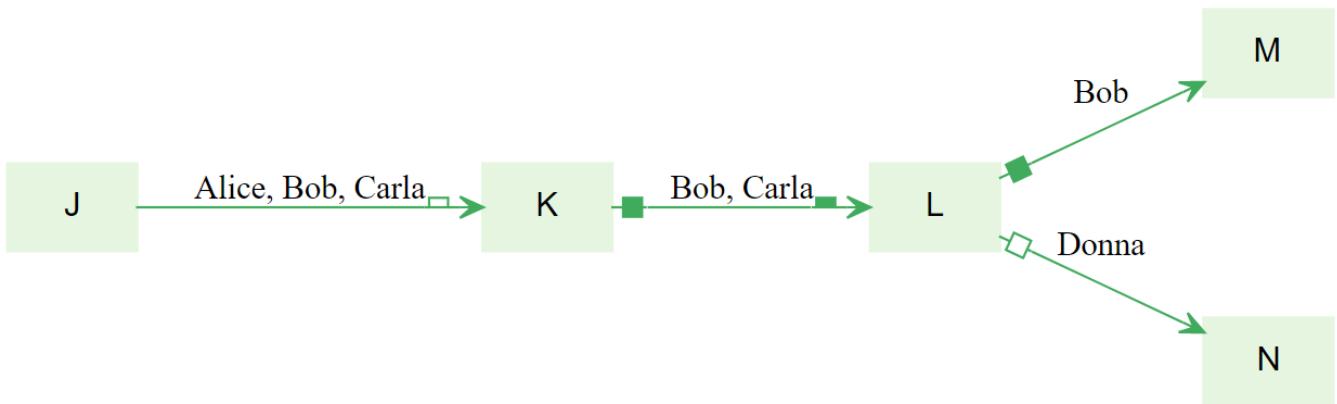
We will probably also add a simpler metric for outgoing links which does not distinguish between the incoming links, something like "Percentage of sources who mentioned a link leaving factor F who mentioned any of the links entering F". This metric could be used to colour or scale the links, or perhaps be printed on the tail of the links.

That's all you'll need for most purposes. Read on for some advanced diagnostics.

Advanced diagnostic filter: Show continuity

Read on only if you are interested in advanced diagnostics!

Summary



Above, the links are labelled with the sources.

The □ open half-box at the end of the first link tells us that at least half but not all of these stories stop here: less than half the sources mentioned any link *out of K*.

The ■ filled box at the start of the second link tells us that all of these stories are continuations: all these sources mentioned some link *into K*.

The — filled half-box at the end of the second link tells us that at least half but not all of these stories continue: Bob mentioned some link *out of L*, but Carla did not.

The □ open box on the link from L to N tells us that this story is not a continuation: Donna did not mention any link *into L*.

There is no UI for this filter yet. You can just type

`show continuity`

in the advanced editor.

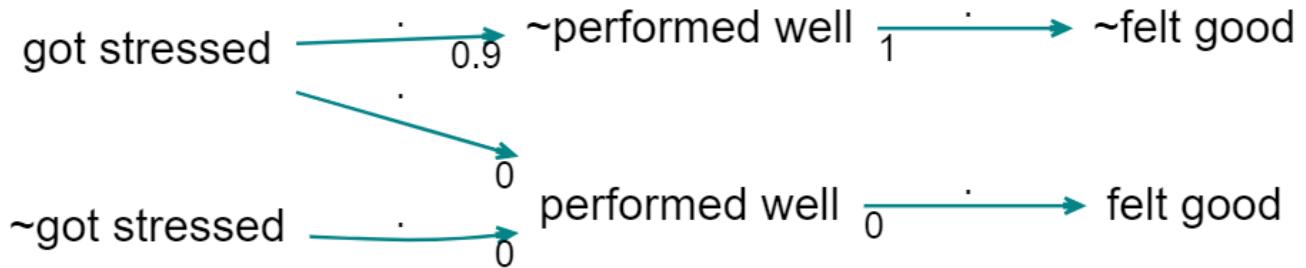
The four kinds of boxes are (possibly aggregated) indicators of continuity, with respect to sources, between stages in a path.

If you want to look at say statement continuity rather than source continuity (the default), type

```
show continuity field=statement_id
```

If you want to see numbers (see examples below) rather than symbols (see examples further below; symbols are the default) then type:

```
show continuity type=label
```



Here, the 0.9 says that 90% of the sources mentioning the link to ~performed well also mention the link from ~performed well. The 1 says that 100% of the sources mentioning the link from ~performed well also mention the link to ~performed well. And the zeros below say that there is no source continuity at all.

What this doesn't tell you is, when there are more than one incoming link, which of them have sources which continue to the outgoing link (that is what the bs and cs are for in `mark_links`). It's just an aggregate.

But what happens with filters which actually transform the map: `zoom`, `bundle factors` and `combine opposites`? Zoom can create its own version of the transitivity trap, if we have:

`eating lemons --> health; no scurvy`

and

`health; fitness --> fast runner`

Health; being physically fit → Ability to run fast

Eating lemons → Health; not getting scurvy

we should be very careful when concluding (when zooming)

eating lemons --> health --> fast runner

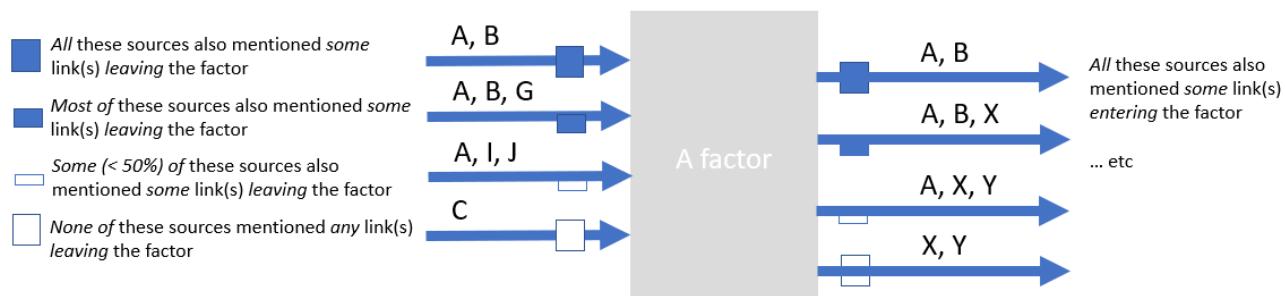
... and indeed, showing continuity highlights this error:

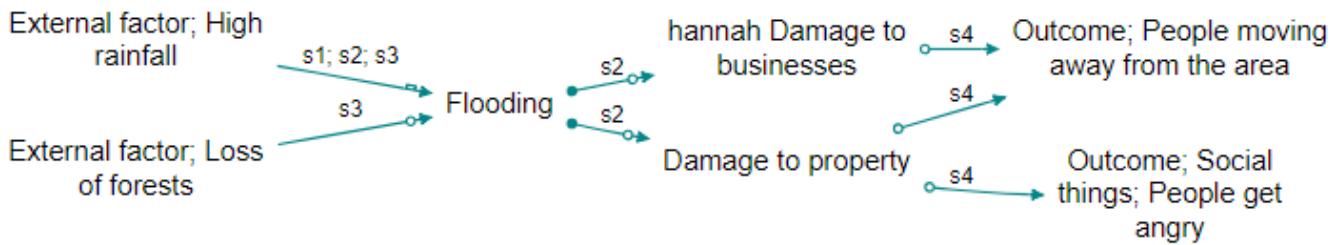
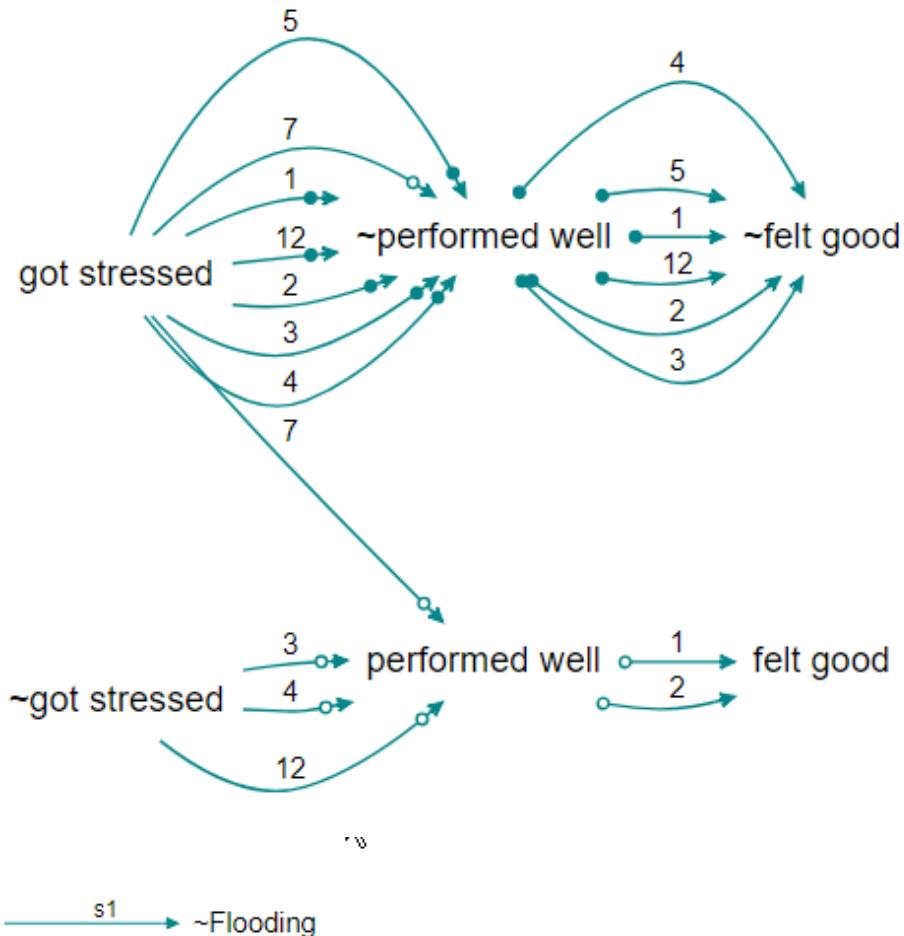


Showing continuity with arrowtypes

Printing actual numbers (from 0 to 1) on the arrows can be very confusing. So the default is to use symbols.

- white box: 0
- half white box: <= 0.5
- half full box: >.5
- full box: 1

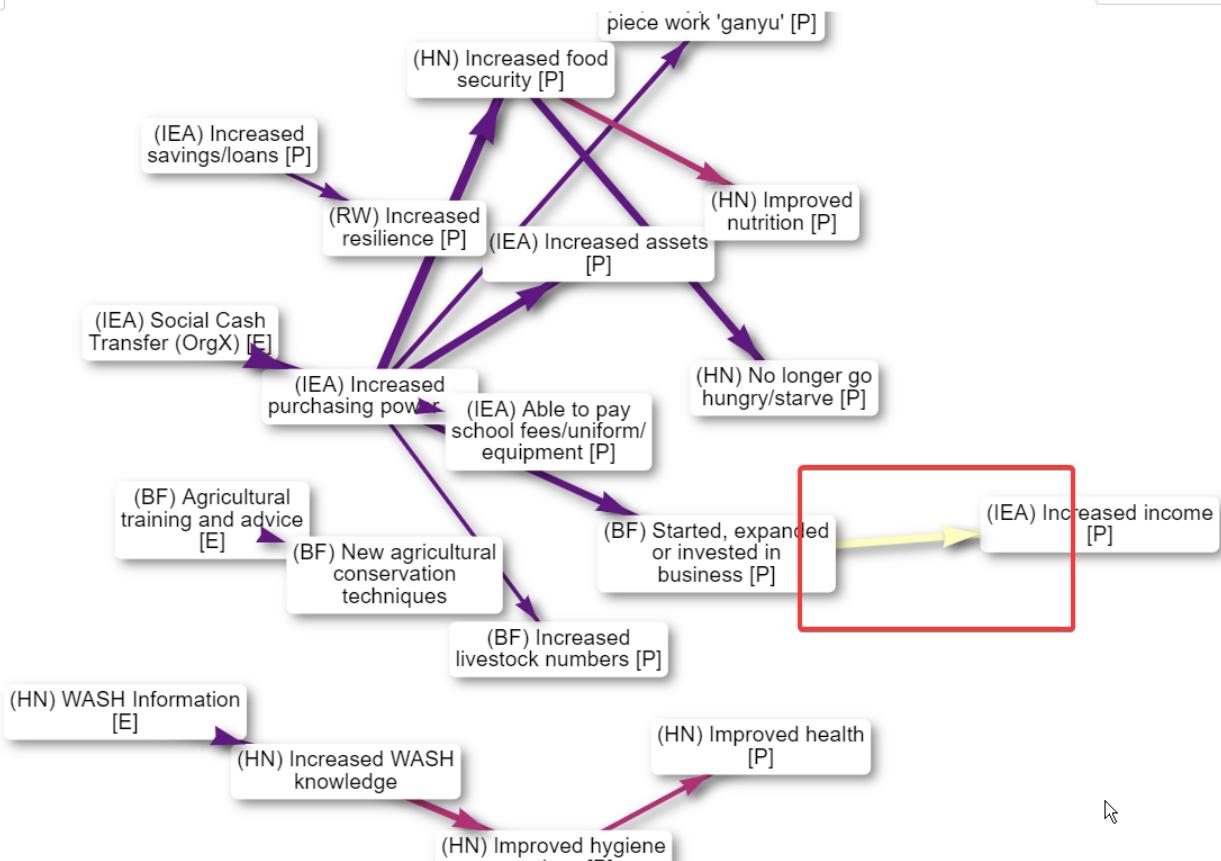




Showing continuity with colours

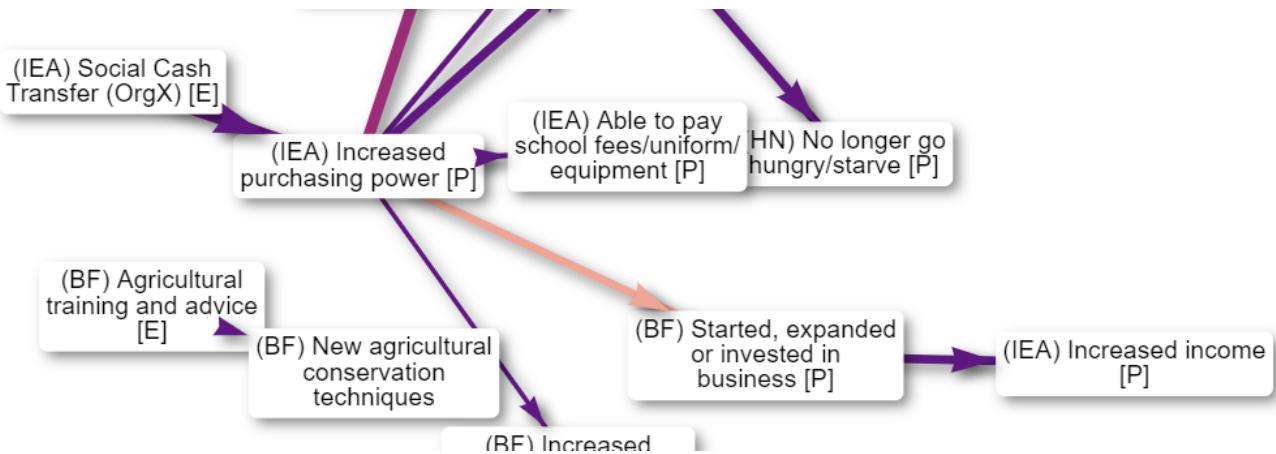
<https://causalmap.shinyapps.io/CM2test/?s=618>

Using arrowheads gives you information about both upstream and downstream flows, but it can be a bit tricky to read. Instead you can use colours to display either downstream (effects of causes) or upstream (causes of effects) continuity.



Here we see that not so many of the people who mentioned the link from business to income mentioned the link from purchasing power to business.

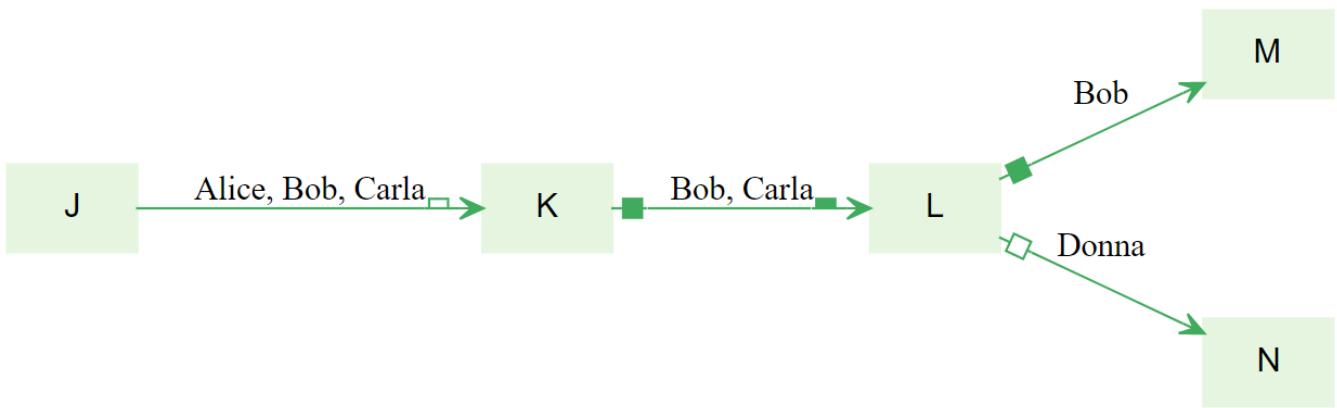
Same, but upstream continuity:



<https://causalmap.shinyapps.io/CM2test/?s=619>

These values are set to 1 at the edges of the map where the metric has no meaning.

Note this is not the same as the non-causal question "how many of the people who mentioned factor C also mentioned factor E".



More about these metrics

Local Continuity factors (simple)	Factors (ego network)
overlap between sources who mentioned links to this factor and sources who mentioned links from this factor	overlap between sources who mentioned links to the causes of this factor and sources who mentioned links from the effects of this factor

And, with links:

Local Continuity	Links
Upstream	overlap between sources who mentioned this link and sources who mentioned links to the cause of this link
Downstream	overlap between sources who mentioned this link and sources who mentioned links from the effects of this link

Each of these metrics can be expressed as a confusion matrix and can be cashed out as different ratios. We can therefore also interpret these metrics in terms of causal necessity and sufficiency. For example, above we can say that K is causally sufficient (with respect to sources) for M because all the sources who mention causes of M (along paths from K to M) also mention effects of K (along paths from K to M).

We need to say "with respect to sources" because all these ideas are generalisable to other fields such as, for example, village or question domain.

Because these metrics (confusion matrices) are defined in terms of source _id (or some other context-relevant link variable) they partly counter the problem with previous versions of these metrics in that they provide a denominator (number of sources) even if this has to be used with some care: as usual, the fact that source S does not mention link L does not mean they wouldn't assent to it, it may just not have appeared in the stochastic interview process.

Many different metrics are possible. These (all?) also have corresponding non-causal counterparts as in QCA, for example:

Local continuity (non-causal)	Factors (ego network)
overlap between sources who mentioned the causes of this factor and sources who mentioned the effects of this factor	

-
These QCA-type metrics (confusion matrices) are inferior to their causal counterparts because they lose the information about what causes what and only use information about co-occurrence.

!--chapter:end:51635-threads-and-continuity.rmd--

Conditional formats{#xconditional-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.

Conditional formatting for links {#xformatting-links}

Conditional formatting gives you more options with links than with factors. It applies to:

- scale links
- colour links
- label links
- dash links



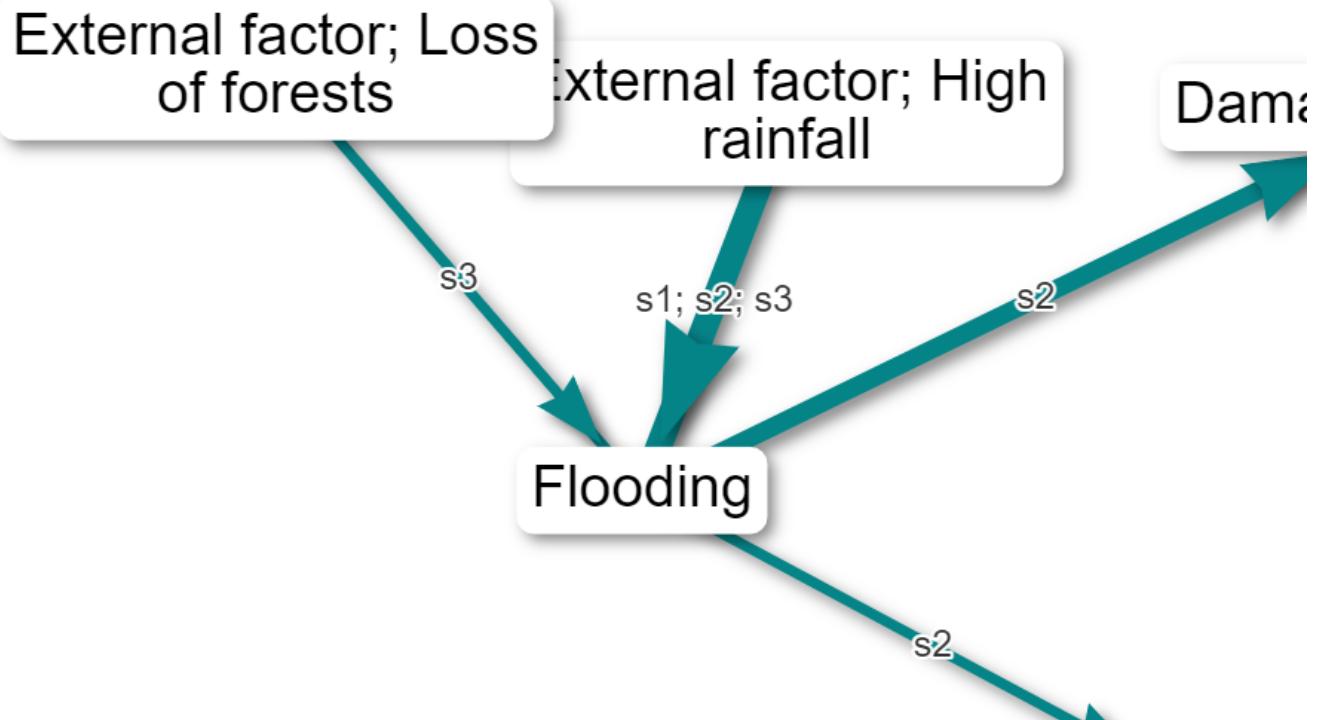
When you click one of the buttons, you are asked how you want to colour, scale or label the links - the choices are the same.

The screenshot shows a 'Color links' configuration panel. At the top, there is a green 'Apply Filter!' button and a checkbox labeled 'Keep this panel open?'. Below this, the title 'Set the color of the links:' is followed by a dropdown menu. The menu contains the following options:
count: link_id
count: source_id
count: statement_id
literal: source_id
unique: source_id

For example, to label the links you can

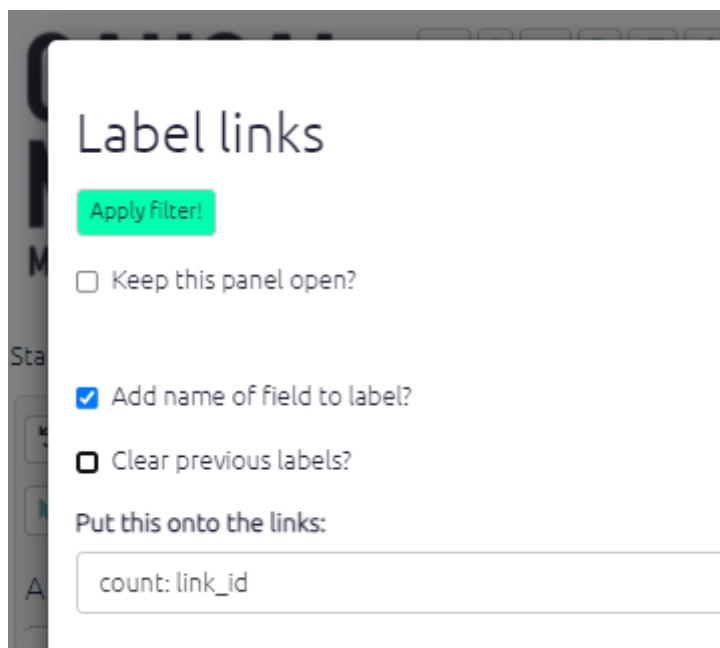
- count the number of links (`count: link_id`)
- count the number of sources who mentioned a link (in a bundle of links with the same influence and confluence factors) (`count: source_id`)
- print the source_id for every link in a bundle of links (`unique: source_id`)
- print literally every source_id, including duplicates, on each link in a bundle of links (`literal: source_id`)

- count the number of villages in which this link was mentioned, assuming you have a field like village_id (count: village_id)



(Here is an example of labelling the links with the IDs of the sources. This can be very useful if you want to know if the same people who mentioned one link mentioned another.)

You can repeat this filter more than once – for example to label bundles of links by gender and by simple frequency: uncheck Clear previous labels?.



You can also opt to add the name of the field to the label, so you could have a label like
Gender: Female, source_id: x12.



The screenshot shows the 'Conditional formats' dialog box with two rules defined:

- bundle links by: simple_bundle
- label links by: count: link_id

Similarly you might want to colour your links with a continuous scale of colours according to:

- the total number of links (count: link_id)
- the total number of sources who mentioned a link (in a bundle of links with the same influence and confluence factors) (count: source_id)

Or you might want to colour your links with a discrete scale of colours according to a non-numerical value:

- show a different colour for every source (unique: source_id)

The app works out for you if your scale is numerical or discrete, so you don't need to worry about that.

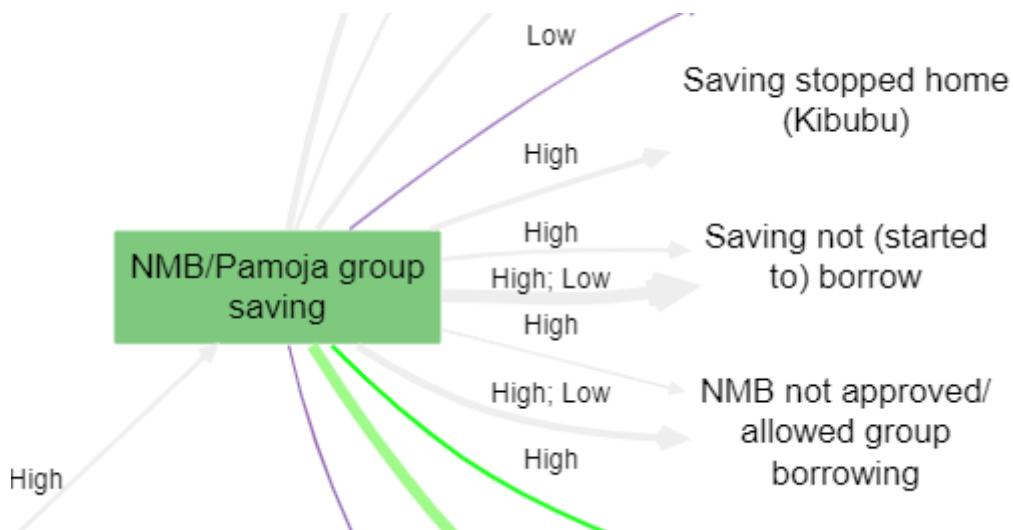
Dashed links

You can mark out some subset of your links as remarkable or doubtful by showing them as dashes instead of solid. Just add a condition the same as you use for finding links, for example:

- links whose hashtag is not empty
- links whose from_factor start with "S"
- links whose source is male

Initials{#xinitials}

Often it is useful to shorten the labels, for example you might have bundled your links by village and want to label the bundles with the names of the villages, but the village names are very long. Or you might have bundled your links by region and want to label the bundles with a list of the unique names of the villages in that region, but this might be quite a lot of text.



In this case you can use the function `initials`.

Label links

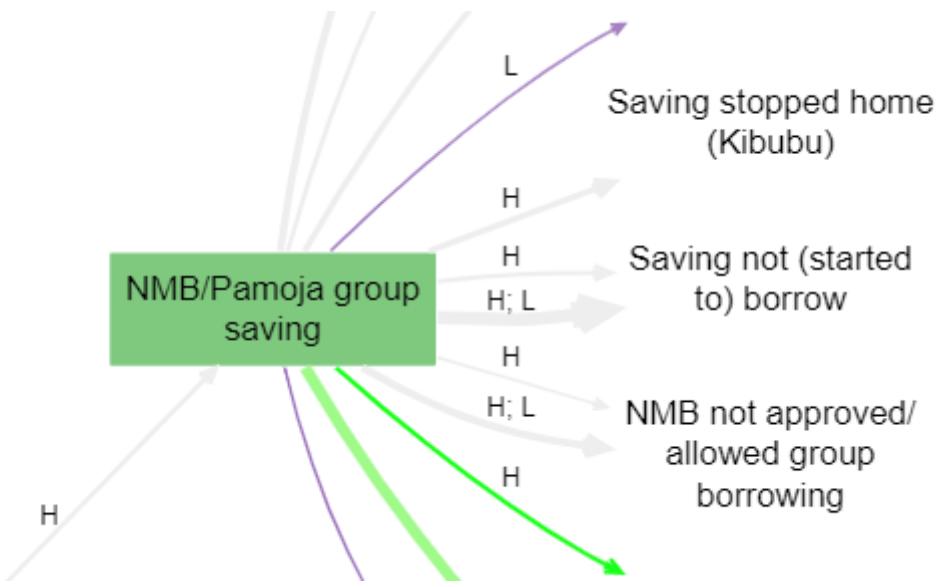
Keep this panel open?

Add name of field to label?

Clear previous labels?

Put this onto the links:

initials: found
initials: #Type

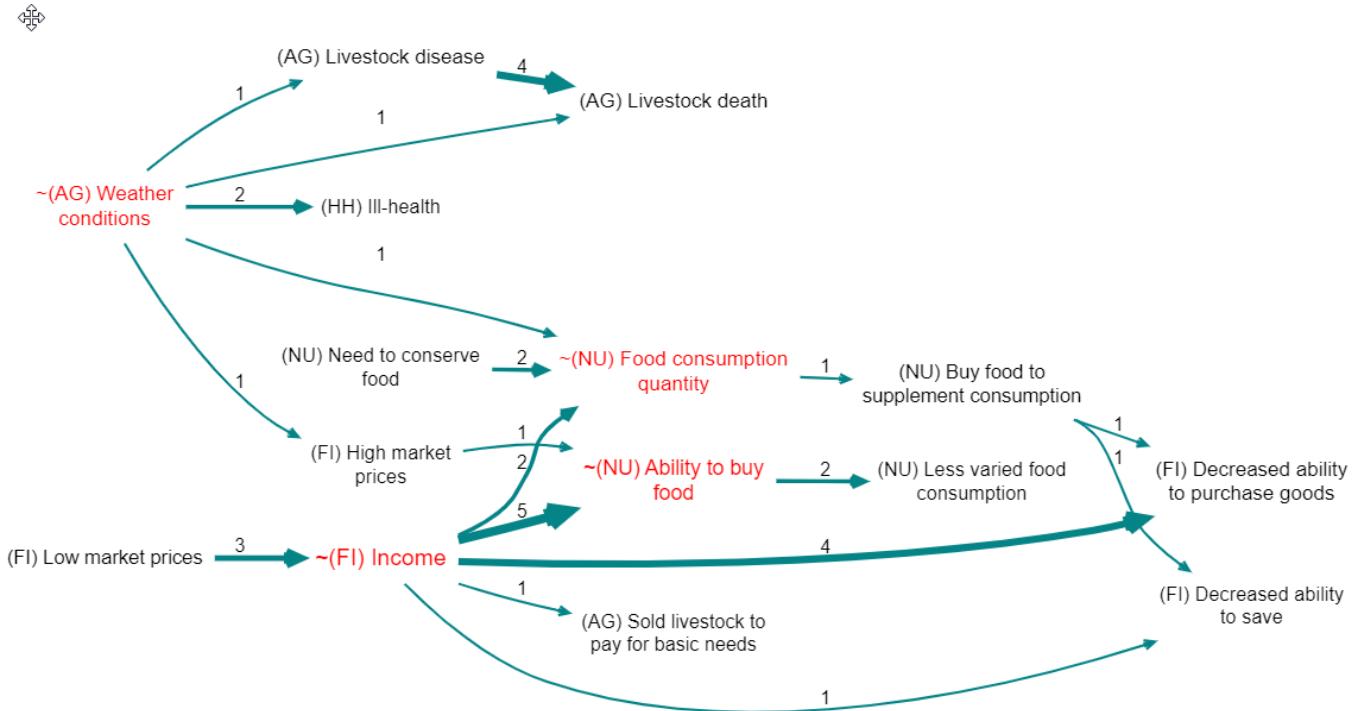


This is especially useful when you do conditional formatting with surprise.

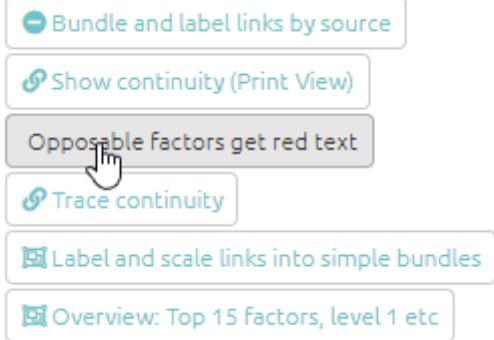
Conditional formatting for factors

Colour text{#xcolour-text}

You can now conditionally colour the text of your factor labels. For example, this can be useful to draw attention to factors which are opposites:



There is a dedicated shortcut button just for this:

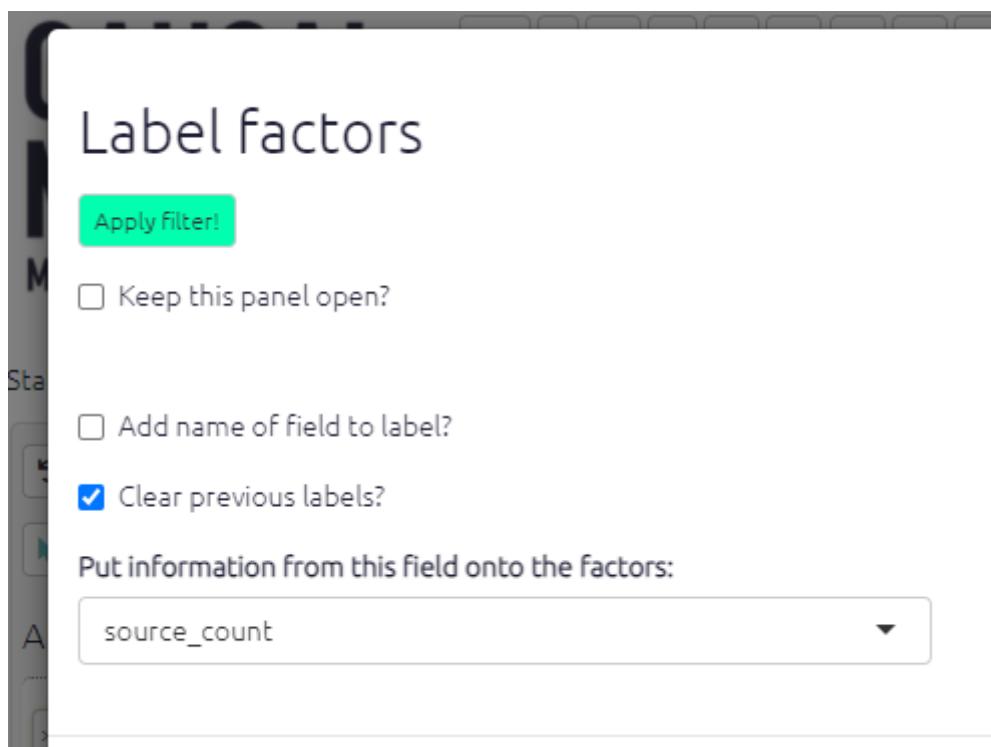


Label factors{#xlabel-factors}

You can label your factors, as well as your links. Labelling your factors is a useful way of adding detail and clarity to your maps. Click on the label factors buttons to open up the filter panel. The app will then ask you to choose which label(s) to add to your factors. Options include frequency, size, and zoom level.

Normally you will want to add this additional information to the existing factor label, but if you want you can check the box to clear previous labels?

You can now also choose whether or not to include the name of the field in your factor labels.



Bundle links{#bundlelinks}

Similarly to bundling factors you can also bundle links. This is really helpful to keep your map neat and easy to read. To start with, you can simply bundle all the links between a pair of factors into the "simple bundle", represented as just one arrow; you will probably want to make the size or colour of this depend on the number of links in the bundle.

Also, you can bundle the links between any pair of factors into more than one bundle, for example one bundle for each gender or district.

If your map is large the app will automatically apply this filter when you are viewing all statements.

When you specify the field in the Advanced Editor, it is enough to only specify only part of the field name. Remember, this name matches any field associated with sources, links or statements.

So #2 will match #2-village-name. We recommend marking your top few fields for analysis like gender etc., or whatever else interests you the most, not just with a leading # (this anyway marks out the field as important) but then with a number, like #2. This makes it easy to create reproducible filters which you can conveniently share across different files.

The Standard Views which are available in the Dashboard also make use of this trick, offering you various views of the file which put these important fields through their paces.

Colours

- 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.

Fixed colours{#xfixed-colours}

Sometimes you might want to set a specific colour for all your factors, or links, or factor borders. This is a bit tricky to set using the buttons, but you can specify a colour easily using the advanced editor:

```
color factors fixed=pink  
color borders fixed=gray  
color links fixed=coral
```

You can use any html colour like red, pink, aliceblue, beige ... and also hex colours like #eee, #00000033 etc.

Changing discrete palettes{#xdiscrete-palettes}

The default palette for discrete fields like gender or village is nice, but maybe you want something different. With continuous fields like source count, you can specify the high, low and mid colours; now you can change the palette for discrete fields too, by adding pal=6 or pal=3 like this:

```
color links field = #area fun=literal pal=6
```

Palette 2 is a good choice if you don't like the paler colours in palette 1 (the default). See also the special sections on conditional formatting with percentages and conditional formatting with surprise.

Also, much of this formatting only makes sense in terms of bundles of links from one factor to another. The app will bundle the links for you in these cases.

Calculated fields{#xcalculated-fields}

The app adds some pre-calculated fields for you and adds them to the tables.

You can use these in various ways to construct filters. For example you can add frequency to your factor labels, or filter the map to show only bundles of links consisting of only one link (find links field=source_frequency value=2 operator=less).

The screenshot shows a software interface titled "Find links". It includes a green "Apply Filter!" button, a checkbox for "Keep this panel open?", and dropdown menus for "What field to search in?" (set to "source_frequency") and "Type to find a value or values for this field" (containing the number "2"). An "Operator" dropdown is set to "less".

Some examples of calculated fields:

Links table

- simple_frequency is just the number of links in each bundle
- source_frequency is the number of different sources in each bundle
- simple_bundle which is just the influence and consequence labels strung together like this: "influence / consequence".
- before_id gives the ids of any links immediately before the current link (i.e. which point to the factor at the *start* of the current link), and likewise for after_id. This information is useful when calculating source continuity.

Factors table

- betweenness: the number of paths going through the factor. Betweenness is a measure of how central a factor is in the map, how many paths go through it. Betweenness is similar to the frequency, but a factor with lots of links on the edge of the map will have a high frequency but won't have a high betweenness.
- betweenness_rank: the rank of the betweenness.
- in_degree: the number of incoming links.
- out_degree: the number of outgoing links.
- role: the number of incoming links minus the number of outgoing links. High values are drivers, low values are outcomes
- frequency: the number of links.

- driver_score: how strongly is this factor a driver?
(driver_score=factors\$out_degree-factors\$in_degree * 2)
- outcome_score: how strongly is this factor an outcome?
- driver_rank: rank of driver_score.
- outcome_rank: rank of outcome_score.
- is_opposable: does the factor label contain a ~.
- zoom_level: number of ; separators in factor label, plus 1.
- top_level_label: the label of the factor's ultimate parent in the hierarchy, if any.
- top_level_frequency: the number of links to and from the top level factor.
- is_in_hierarchy: whether the factor is part of any hierarchy or is just a "singleton" factor.

You can use is_in_hierarchy to focus only on your hierarchical structure, eg you can add a filter

find factors field=is_in_hierarchy value=true operator>equals

to exclude singletons.

When combining opposites

When you combine opposites, the app gives you additional combined fields (in the links and factors tables) which you can use to display and filter your maps in various ways.

For example, you can keep only links in which come from or to positively formulated factors: find links field=flipped_bundle value=FALSE|FALSE operator>equals

-

There are many more.

These fields are calculated at the point in the chain of filters where each filter is applied, so if you, say, filter for bundles with at least 10 links and then add filters which filter out, say, all the older people, you might get bundles with fewer links remaining.

!--chapter:end:51640-conditional-formats.rmd--

Simple formats{#xsimple-formats}

These filters change basic visual elements of your map to make it clearer to interpret.
The data itself is not affected.

```
knitr:::include_url("https://player.vimeo.com/video/651173445")
```

Fixed colours for all links, factors and factor borders

If you want to make all your factors, say, pink, just add the filter `color factors fixed=pink`

See here.

Setting specific colours for links, factors and factor borders

If you want specific factors to have specific colours, for example "Impact on wellbeing" should always be pink, you can do this by roundtripping:



- Create the columns `color.background` and/or `color.border` in the `factors` tab, and type any colour names you want (red, blue, pink, green etc) in this column in the row corresponding to the factors you want coloured. You can also use hexadecimal colours like `#eef55`.

This trick does not work with filters which manipulate the factor structure, if you zoom or combine opposites, as the app wouldn't know which colour you want. In this case, the colours applied in this way are unpredictable; you might want to restore standard colours eg just add the filter `color factors fixed=pink`.

Wrap factor/link labels

Changing the wrap length of your labels can help make your map clearer and easier to read. As with other sliders, you might find it helpful to keep the filter panel open whilst you select your label length. Click on the wrap factor/link labels buttons to open their respective filter panels. Simply adjust the slider to your chosen length for your labels.

Wrap factor labels

Keep this panel open?

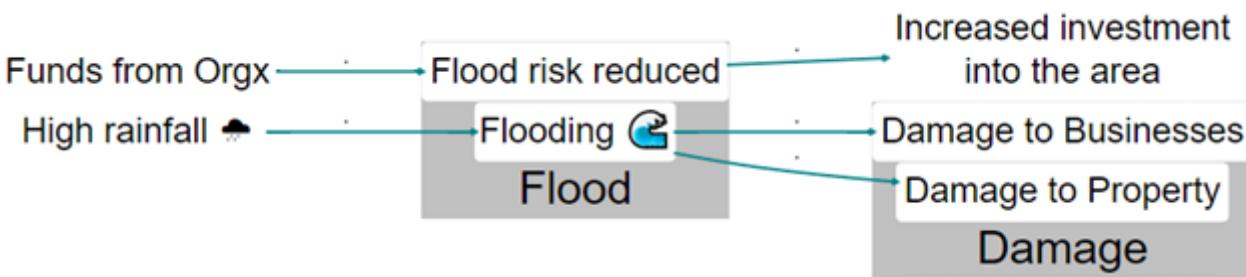
Move the slider to change the line-length of the factor labels

0 22 100

We have now fixed various bugs with wrapping link and factor labels.

Cluster factors{#xcluster-factors}

Sometimes you might want to visually show how factors in the map are similar to one another, without bundling them into one. By clustering similar factor labels, you can highlight similarities and/or differences in the way certain factors interact with other factors in the map without losing any of the details of your factor labels. For example, the screenshot below shows factors containing 'Flood' and 'Damage' clustered together, highlighting the link between flooding and damage.



Click on the cluster factors button to open the filter panel. Simply type in the word/phrase that the factors you want to be clustered contain. Once you click apply filter, all factors containing your chosen text will be positioned close to one another and surrounded by a grey box in the print view of the map.

Cluster factors

In print view, factors with labels containing this text will be positioned close to one another and surrounded by a grey box.

[Apply filter!](#)

Keep this panel open?

Factors with labels containing this text will clustered (case sensitive). You can enter more than one search term in the box and each will be clustered separately.

type the text to search for in factor labels (case sensitive)



Set print format

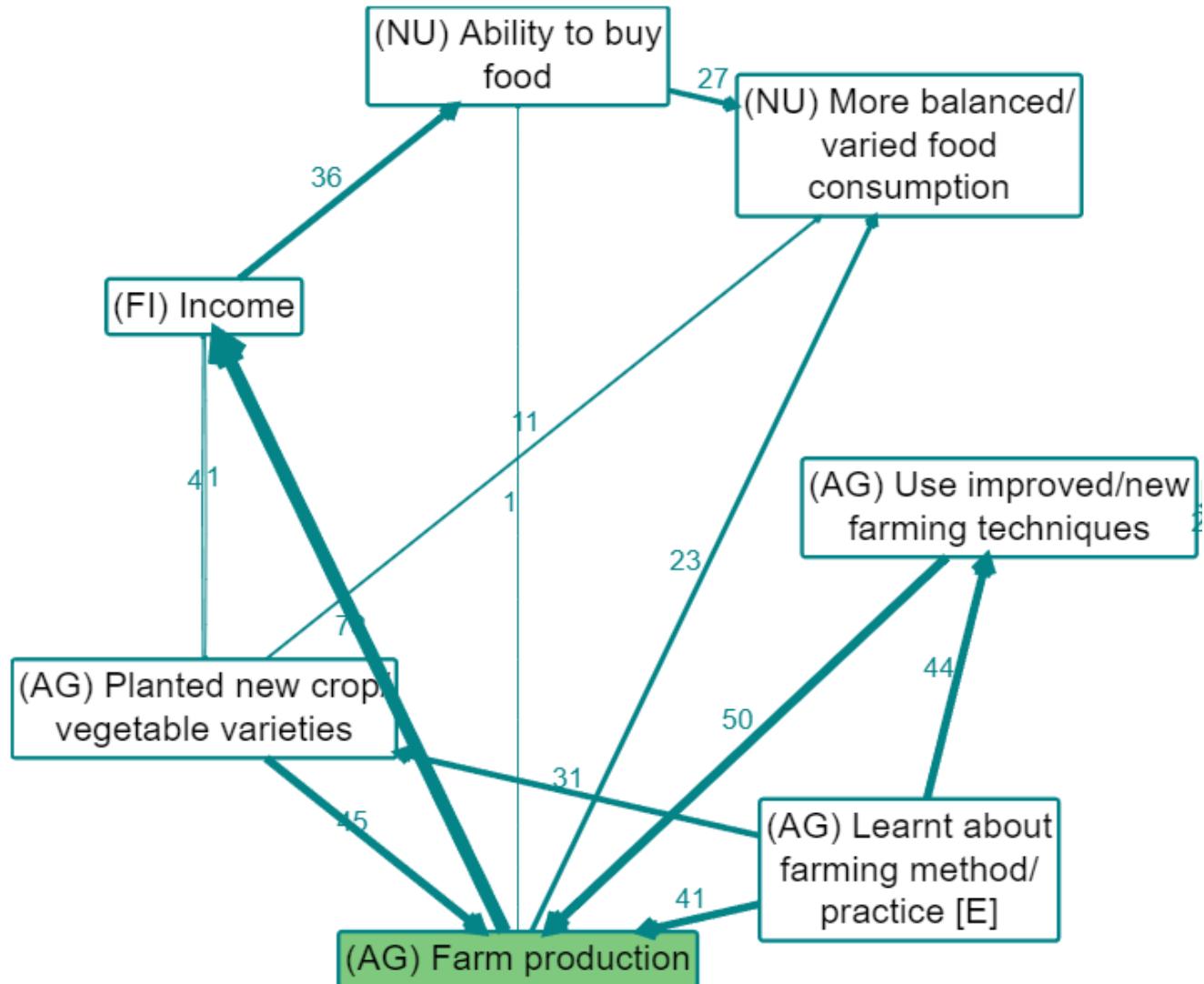
Layout

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

Only the "dot" layout is fully supported. 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.

Beware that the output from other layouts may not look as good as the default "dot" layout, and the tweaks like setting horizontal separation may not work with them.

The screenshot below shows the circo print layout with 1 step up and 1 step down from Farm production.



How do I use this filter?

Click on the set print format button to open up the filter panel. There are various options for map layouts which you can try out. There is also a horizontal separation slider which you can adjust to select how far apart your factor labels are, if the number is higher your links will appear longer and your factor labels will be further apart. This is another filter where it would be helpful to keep the panel open so that you can see how your selections are affecting your map.

Set print format

[Apply filter!](#)

Keep this panel open?

Horizontal separation



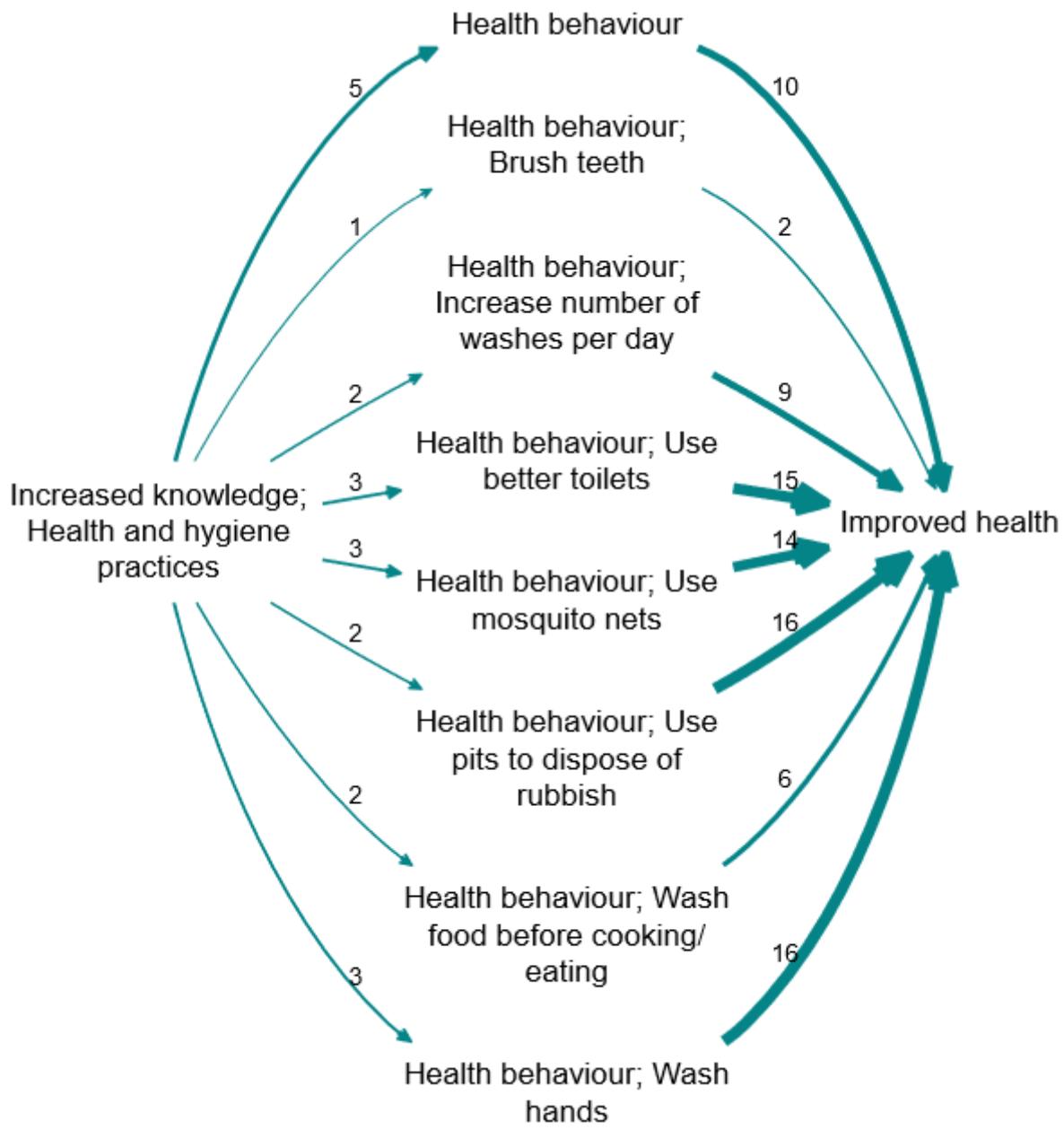
Layout for print map only

- dot
- neato
- circo
- twopi

!--chapter:end:51650-simple-formats.rmd--

Smart zooming{#smart-zoom}

If you are using hierarchical aka nested coding, you can use ordinary zooming to "zoom out" your map to give a simple, more general overview. But sometimes this can mean oversimplifying your map: there are lower-level factors which are so important and popular that it is unhelpful or even misleading to roll them up into a parent factor. Smart zooming solves that problem.



1549

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

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

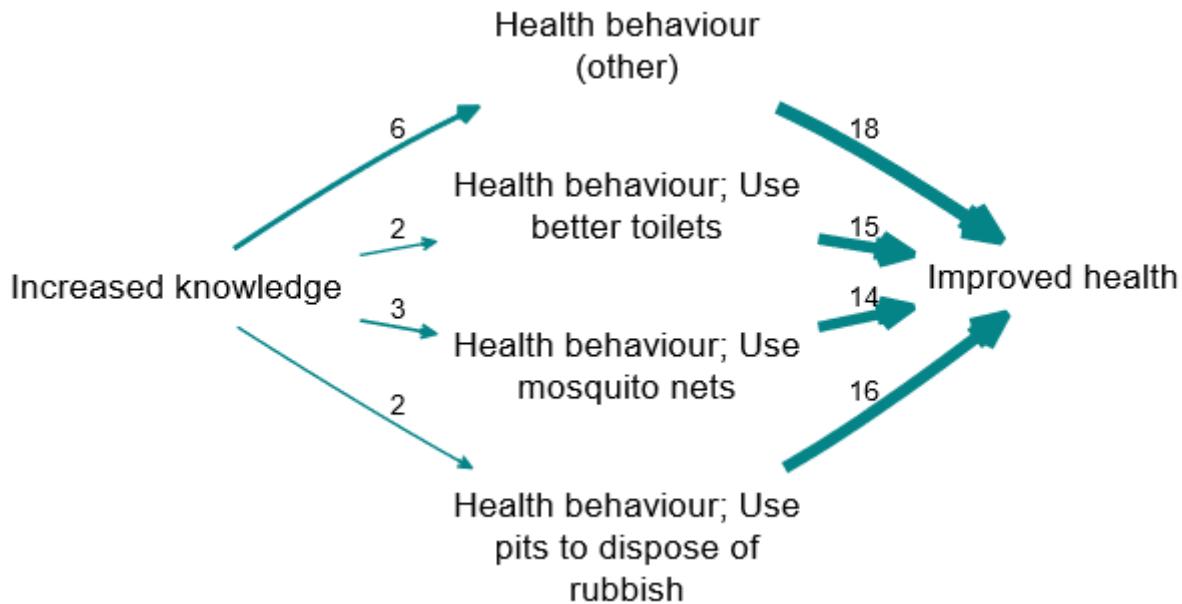
Increased knowledge $\xrightarrow{6}$ Health behaviour $\xrightarrow{18}$ Improved health

A weakness of hierarchical zooming is already visible in the example, we have lost a lot of the important detail. In many cases we may be happy to see all the factors “rolled up” to the top level but in the above example, there are a variety of types of health behaviour discussed and this detail has been lost. So what can we do?

A smarter approach is achievable with this smart zooming, which tells the app: only "roll up" a nested factor into a higher-level factor if it has a frequency (simple frequency, or source count) lower than some cut-off value; otherwise, keep it separate.

This is an improvement over zooming because it keeps the detail where necessary and removes it where it is not. It will roll up infrequent factors into their parents but only if they are infrequent; if you have a very granular factor like "health behaviour; Use better toilets" which actually has a lot of mentions, then it won't get rolled up.

The above example can be rolled up to this:



1552

This is an improvement because it still has a small number of factors, retaining frequently-mentioned factors but wrapping up infrequently-mentioned detail into higher-level concepts.

This isn't available using the buttons yet but you can use the Advanced Editor. The code is:

```
zoom factors level=1 preserve_frequency = 5
```

You can use source_count instead:

```
zoom factors level=1 preserve_frequency = 5 frequency_field = source_count
```

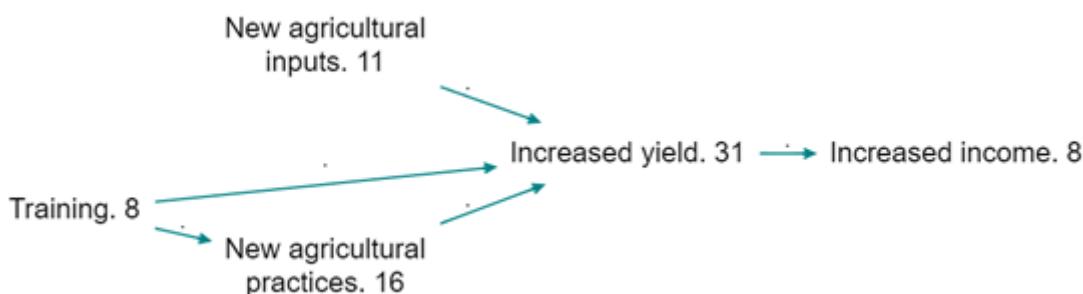
One thing which is not yet clear is whether the resulting pair of factors "Health behaviour" and "Health behaviour; hand_washing" should be distinguished in any way. For example, perhaps it isn't obvious, looking at "Health behaviour", that this higher-level factor excludes in this case its actual daughter, "Health behaviour; Hand washing".

What the app does provide is the additional field in the factors table: frequency_preserved. So if you want, you can highlight factors which are not rolled up, for example like this:

```
color borders field="frequency_preserved" hi=red lo=white
```

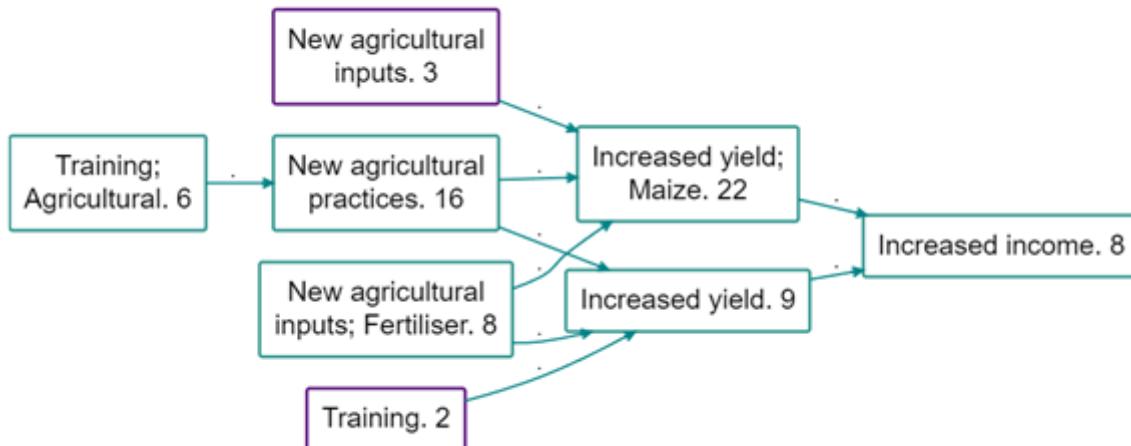
Here is a real-life example.

A top-level map, zoomed out to level 1 (source counts are shown). As you can see this map is quite simple with no additional details.

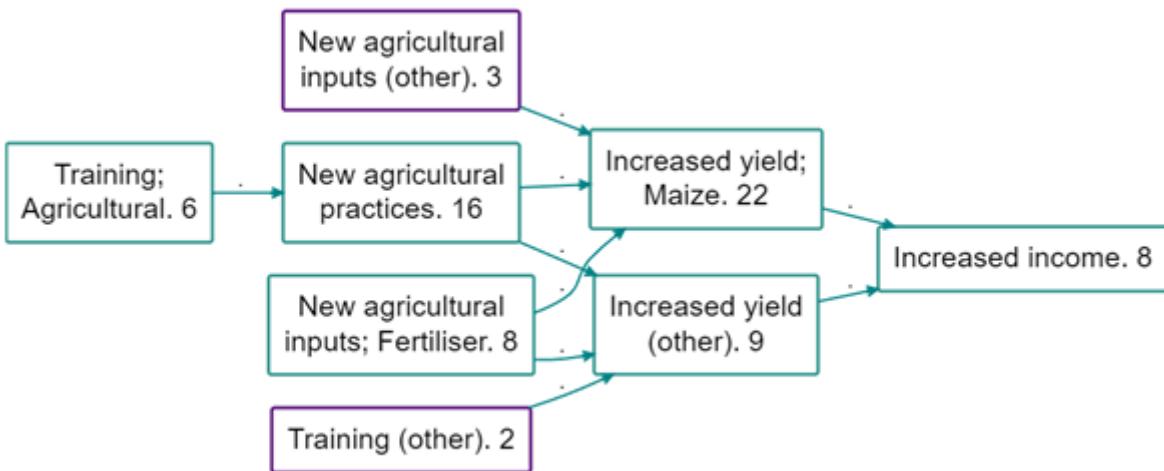


And here is the same map but preserving factors with a source count above 5. This seems like a real improvement as some really concrete and important second-level factors are preserved, including for example that the training mentioned on 6 occasions was agricultural training and that fertiliser was mentioned as a driver of change.

The factors with a purple border have been ‘rolled up’ and only their top-level label is visible. The other factors have been preserved and are displayed in full, as they have a frequency count higher than 5 or because they are a level one label and so cannot be ‘rolled up’.



Probably the best way to do this is to add "(other)", or whatever else you type, on to the ends of the higher-level factor labels which have larger child factors elsewhere in the map. Larger child factors are those which are not rolled up because they are so frequent, like 'New agricultural inputs; Fertiliser' in the below map. To do this add frequency_other= (other) in the advanced editor on the same line as your commands to preserve factors whilst zooming. This will create a map like below:



!--chapter:end:51700-smart-zoom.rmd--

Clustering sources{#cluster-sources}

This new filter groups all the sources in the file (as currently filtered) into clusters of sources which are similar in terms of the maps they tend to produce.

Filtered clusters

You can cluster the sources remaining in a map as currently filtered, at the point at which you apply the filter cluster_sources.

To use it, add a filter like this to the advanced editor:

```
cluster_sources n_clusters=4
```



Here you can see the new field appearing in the Sources table.



You can also ask for cluster sets with 2, 3 and 4 members all at once:

```
cluster sources n_clusters=all
```

These are three attempts to find interesting sets of clusters - one set with only two clusters, one set with three and one set with four.

You should pick the set which you find most interesting and concentrate on that.

The clusters themselves get automatic names a, b, c etc as appropriate. So the set #cluster_set_2 has just two clusters called a and b. They don't mean anything in particular but you should explore each one to perhaps find your own names for them - perhaps "mainly younger farmers" or "mainly older urban residents" or even something more imaginative like "stars" or "stragglers".

The most interesting output will be in the Differences - sample subtab of the Reports tab, where you should see some interesting differences in how members of the different cluster sets responded to different questions -- that is the main point of the clusters.

The Sample subtab of the Reports tab has, after the initial tables for each important field (those beginning with #), now has another longer section with tables of each important field cross-tabulated against the other. If you scroll down to the sections for #cluster_set_2, - #cluster_set_3 etc, you will be able to see if these sets of clusters are easy to interpret in terms of the other fields. Ideally you'll be able to, for example, understand the clusters as "mainly younger farmers" etc.

It is always trivially possible to find at least one cluster set which differs in one link (i.e. in one simple link bundle) -- though whether these differences are significant depends on the numbers involved. It is only interesting when we find common differences between clusters across several different links simultaneously, i.e. the people in cluster A mentioned this and that and this other thing.

In future we will provide output to automatically crosstabulate the clusters with other key fields like gender and district, and to automatically produce maps for each cluster, surprise maps, differences tables, etc., but you can already do this manually.

Here the Sources table has been set up to list the sources according to cluster:



The clustering algorithm

The algorithm is quite simple, it simply treats the links table for each source as a one-dimensional binary vector of length equal to the square of the number of factors. It's a binary vector, so multiple mentions of the same bundle of links by the same source are no different from a single mention.

Bundling links by cluster

You can do anything with these new cluster fields that you can do with any original field like gender, district etc.

Here's an example in which links are bundled and labelled by cluster, and then we also apply a "surprise" filter to colour bundles mentioned by unexpectedly many sources in bright green, and to colour bundles mentioned by surprisingly few sources in bright purple: causalmap.shinyapps.io/tokyo/?s=1177



Finding many solutions at once

You can ask for a specific number of clusters:

```
cluster sources n_clusters=5
```

But how are you going to know if you will get a more interesting set if you ask for five rather than four or three or seven clusters?

This filter with n_clusters=all finds three separate solutions with two, three and four clusters all at once. It's your job to compare and decide which (if any) is the solution which is best and easiest to interpret.

There is also a shortcut button to add this line to your current filters.

!--chapter:end:52150-source_clusters.rmd--

Interpreting your results{#xinterpret}

Like with any other kind of data analysis, in causal mapping you have to be careful to interpret your results correctly.

One thing to think about is, have you coded only explicit causal claims or have you also coded, for example:

- hypothetical links
- dreams / hopes / fears / wishes for the future
- statements that there is NO causal connection between two things, e.g. a quack cure and an illness

Usually we will code this kind of link using dedicated hashtags. You will probably want to exclude links like this from your normal analysis, and treat them separately.

There are quite a few other potential pitfalls! One of them is the transitivity trap, so have a read of that section of the Guide.

You can practice your interpretation skills in this quiz section and the ones after it.

[To be completed]

!--chapter:end:52300-interpret.rmd--

Tables - advanced{#xtables-advanced}

Creating your table

There are three possibilties:

1. A normal table. You simply select the first column you want in the "rows" box and then add any others you need in the "columns" box.
2. Compressed rows. This happens when you select a field for your rows which is "shorter" than the actual number of rows in the table. For example `5043`, in the factors table you can select level `1` label. Because there are fewer of these than there are factors, the table will be "compressed" and we can choose how we want the multiple results for each row to be displayed -- here we have chosen to display all the labels, but we could also have displayed just the number of values.

The screenshot shows the configuration interface for creating a table. At the top, there are three dropdown menus: 'Rows' set to 'level_1_label', 'Columns' set to 'label', and 'Pivot column' set to '---'. Below these is a dropdown 'Fill the cells with' set to 'all values' and a note 'of each column'. A 'Filter' section for 'level_1_label' is shown with a dropdown set to 'Gender'. Below this, a message says '135 rows in total' and 'Click the = signs to filter, click the headers to sort, click the pink "..."' to edit. The main area displays two rows of data. The first row has columns for '=Gender' (set to '=') and 'level_1_label' (set to 'label'). The second row has columns for 'Change in gender relations' (set to '=Change in gender relations; Women/wife more control over finances; Saving relations') and 'label' (set to 'Change in gender relations; Improved understanding, Change in gender relations; Male more involved in childcare, Change in gender relations; Women/wife more control over finances; Spending, Change in gender relations; Women/wife more control over decision-making, Change in gender relations; Women/wife more control over finances; Saving'). Each row has a small '...' button at the end.

Here, we compress the sources table just by the answer to their question on purchasing power `5044`, and display several other columns.

Rows

Overall, how has the amount you can buy as a household changed over the last year?

Columns

source_id Household code
#Name of province #Name of village

Pivot column

Fill the cells with all values of each column

19 rows in total
Click the = signs to Filter, click the headers to sort, click the pink '...' to edit

Overall, how has the amount you can buy as a household changed over the last year?	source_id	Household code	#Name of province	#Name of village		
Decreased	MNX-1, TWX-1	MNX1, TWX1	A, B	1, 2		...
Increased	MNX-2, MNX-3, MNX-4, MNY-1, MNY-2, MNY-3, MNY-5, MSX-1, MSX-2, MSY-1, MSY-2, MSY-3, TWX-2, TWX-3	MNX2, MNX3, MNX4, MNY1, MNY2, MNY3, MNY5, MSX1, MSX2, MSY1, MSY2, MSY3, TWX2, TWX3	A, A, A, A, A, A, A, B, B, B, B, B, B	1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2		...
No change	MNX-5, MSX-4	MNX5, MSX4	A, B	1, 2		...
Decreased pos	MNX-6	MNX6	A	1		...

This is the same thing, but showing just the counts in each cell: 5045.

Rows

Overall, how has the amount you can buy as a household changed over the last year?

Columns

source_id Household code
#Name of province #Name of village

Pivot column

Fill the cells with the count (number of distinct values) of each column

19 rows in total
Click the = signs to Filter, click the headers to sort, click the pink '...' to edit

Overall, how has the amount you can buy as a household changed over the last year?	source_id	Household code	#Name of province	#Name of village		
Decreased pos	1	1	1	1		...
Decreased	2	2	2	2		...
No change	2	2	2	2		...
Increased	14	14	2	2		...

3. **Pivot tables.** It is possible to construct pivot tables by selecting a column to pivot by: 5046. Again, you can fill the cells with all the values of a field you select, like source_id, or all the unique values, or the counts.

Rows

source_id Household code

Pivot column

#Name of province

Unpivot

Fill the cells with

the count (number of distinct values)

of this field:

source_id

19 rows in total

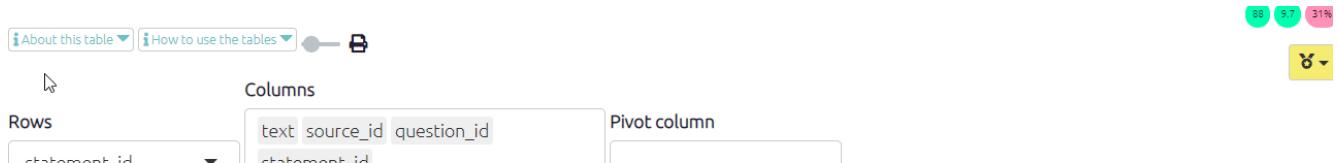
Click the = signs to filter, click the headers to sort, click the pink '...' to edit

#Name of province			
=	=	=	=
#Name of village	A	B	
1	11	NA	...
2	NA	8	...

Pivot tables give us lots of new possibilities, like this display of closed questions: 5047.

Print view{#xprint-view-tables}

Do you want some preformatted quotes for your report?



855 rows
Click the = signs to filter, click the headers to sort. You can also search 'blanks' or 'notblanks'. Click the pink '...' to edit

statement_id	text	source_id	question_id	...
855	The Organisation 1 project has a fruitful end with the works that have been done, they helped our communities a lot.	MSY-3	\$K1	...
854	Important link to Self - Organisation 1	MSY-3	J2d	...
853	Organisation 1 - 0	MSY-3	J1l	...
852	Government - 3	MSY-3	J1k	...
851	Important link to Organisation 1 - Savings	MSY-3	\$J1d	...
850	Important link to Government - Health	MSY-3	\$J1c	...
849	Important link to Organisation 1 - Agriculture	MSY-3	\$J1b	...
848	Important link to Organisation 1 - Monitor households for well being	MSY-3	\$J1a	...

847 more rows ...

Each table can also be shown as simple text for copying and pasting.

If you want printable quotes for your report, there is a table preset for that. Or go manually to the links table, click the  button to display Print View, and choose the columns you want. They will be shown in the Print View in the same order. Then just copy and paste using the "Copy" button.

Formatting your table{#xformatting-your-table}

Below all tables you will see the below buttons which allow you to format the table as you wish. The columns in table with numerical values can be presented as a heatmap. The higher the number, the darker the colour of the cell will be.

127 more rows ...

Rows to show	Size	Heatmap	Style												
<input type="button" value="8"/>	<input type="button" value="50"/>	<input type="button" value="All"/>	<input type="button" value="6"/>	<input type="button" value="8"/>	<input type="button" value="10"/>	<input type="button" value="12"/>	<input type="button" value="14"/>	<input type="button" value="16"/>	<input type="button" value="None"/>	<input type="button" value="Column"/>	<input type="button" value="striped"/>	<input type="button" value="basic"/>	<input type="button" value="bordered"/>	<input type="button" value="hover"/>	<input type="button" value="condensed"/>

Saving your table{#xsaving-your-table}

 Download xlsx

 Clipboard

 Save png

You can copy the data from the tables into your clipboard by clicking Copy, and then you can then paste the data in Word or Excel to create your own tables, graphs, or visualisations.

You can save your table as CSV or Excel. These buttons will export all the data in the current table.

Alternately, you can save the table as a picture or png if you prefer!

Or you can save a link to this table and share it with a colleague just as you with other shortlinks.

!--chapter:end:52610-tables-advanced.rmd--

Top tips on coding{#top-tips}

First of all, there's nothing to worry about, it's fun!

The versioning / backups feature is amazing so you can always go back to any version of your file at any earlier time point.

Also, the Factor Editor makes it easy to rapidly change one or many factors. And you can do it either globally, i.e. changing one factor everywhere in the file, or you can do it for particular statements, if you press the Split checkbox.

Zoom your browser out so everything fits on your screen. Try pressing Ctrl 0 or Cmd 0 to get normal view, then press Ctrl - or Cmd - once or twice if your eyes can stand it.

Don't forget you can get more space while coding by pressing the little resize arrow at the top between left and right panels.

Don't forget you can see additional information about each statement, and make memos on the statement or the source by clicking the info toggle.  Info

Normally you'll want to keep the Filters panel closed because it takes space.

Don't bother coding the same link more than once for the same source, unless they bring up distinctively different evidence each time.

It's okay not to code a statement at all. If there's nothing in it, or if people are just making vague and general statements.

You'll find you're constantly shifting between sometimes creating new factors, and then going back and reviewing them and merging them and organising them in the Factor Editor.

Don't forget you can combine two or more factors into one. Just by pasting the same label over all of them.

Don't forget when you want to search rapidly through already coded links through all of the statements, you can click on the rows in the links table or the statements table. To go back to the relevant statements directly.

Don't forget that the app always knows what question a particular link is in response to, so you don't need to try to sneak the question information somehow into the factors and links, unless you want to.

Occasionally, someone will make a comment about something which is worth coding, even though there isn't actually a causal link. For example, they might make general comments about some outcome without saying what causes it. In this case just make a link like (no influence factor) --> some "|||||outcome or use plain coding. (But if you find you are doing this a lot, you might need to rethink your research design.)

If you are using hierarchical/nested coding (and you probably should) don't forget you can see the whole map zoomed out to the top level: just press the appropriate button in the Filters panel.

!--chapter:end:53400-coding-tips.rmd--

Level 5 quiz

Now do the quiz to see if you've understood about advanced analysis.

!--chapter:end:59900-quiz5.rmd--

(PART*) The rest of the app {-}

!--chapter:end:60000-PART-rest.rmd--

The Updates tab {#xupdates-tab}

In this panel, you can see updates about all the files you have access to.

With the toggle at the top, you can select whether to show:

- The most recent status for each file
- All updates for the currently loaded file
- All updates for all files

You can also click on a row to load the corresponding file.

Updates about your files

Show...

Most recent status for each file All updates for current file All updates for all files

 Copy CSV Excel

Showing 1 to 10 of 94 entries Show 10 entries

Previous 1 2 3 4 5 ... 10

mapfile	user	server	type	content	time_stamp
All	All	All	All	All	All
quip-exercise	steve@pogol.net		mapfile loaded for editing	editing=Yes, ncol_links=30, last_coded_statement=2, nfactors=5, nlinks=13, nstatements=3	2021-11-16 09:10:12
cml2-socio	steve@pogol.net	CausalMap2	mapfile updated	editing=Yes, ncol_links=138, last_coded_statement=590, nfactors=374, nlinks=689, nstatements=2261	2021-11-16 09:10:12

!--chapter:end:60010-updates.rmd--

The File Manager{#xfile-manager}

At the top of the page you will find information about your subscriptions, the files you have created and any teams you manage.

Below, the file manager table shows you all the files you have access to and what level of access you have i.e. copy, view or edit access. You can also see details such as:

- Whether the file is locked- meaning you can view this file but not edit it.
- Whether the file is archived- which is a tag and archived files will remain in your dropdown list.
- Any memos on the file.
- What access other users have to these files.

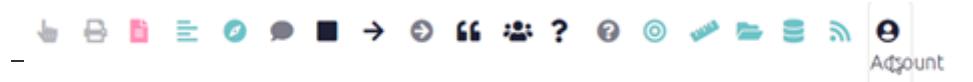
The screenshot shows a user interface for managing files. At the top, there are navigation links: Interactive, Print, Factor editor, Mentions, Factors, Links, Statements, Sources, Questions, Closed qs, Metrics, and Files (which is highlighted with a red border). Below this, the title 'Files for Steve Powell' is displayed, along with a small profile picture and the email 'steve@pogol.net'. A 'Sign out' button is also present. A note states: 'This panel only shows and controls files saved in the Causal Map 2 format.' There are three download buttons: Copy, CSV, and Excel. A search bar and a page navigation bar (Showing 1 to 10 of 86 entries) are located at the top right. The main area contains a table with columns for project, memo, locked, archived, copy, view, edit, can_copy, can_view, can_edit, and access. The table lists two files: 'quip-lorem' and 'example2-'. The 'quip-lorem' file has a checked 'locked' status and a checked 'can_edit' status. The 'example2-' file has a checked 'can_edit' status. Both files have a 'Load' button and a 'Manage' button.

From this page you can either load the map or click manage to open the file manager panel.

!--chapter:end:60020-files.rmd--

Managing your subscription{#xmanage_your_subscription}

To manage, cancel or renew your subscription follow the these steps, go to the Account tab in the app.



The accounts tab will show you:

- Details of any teams you manage
- Any teams you are a member of
- Files you have created
- Your subscription

To change your subscription click the update subscription button at the bottom of the page



From here you can renew, cancel or change your subscription.

--chapter:end:60030-subscription-manage.rmd--

(PART*) Spotlights {-}

--chapter:end:73000-PART-spotlight.rmd--

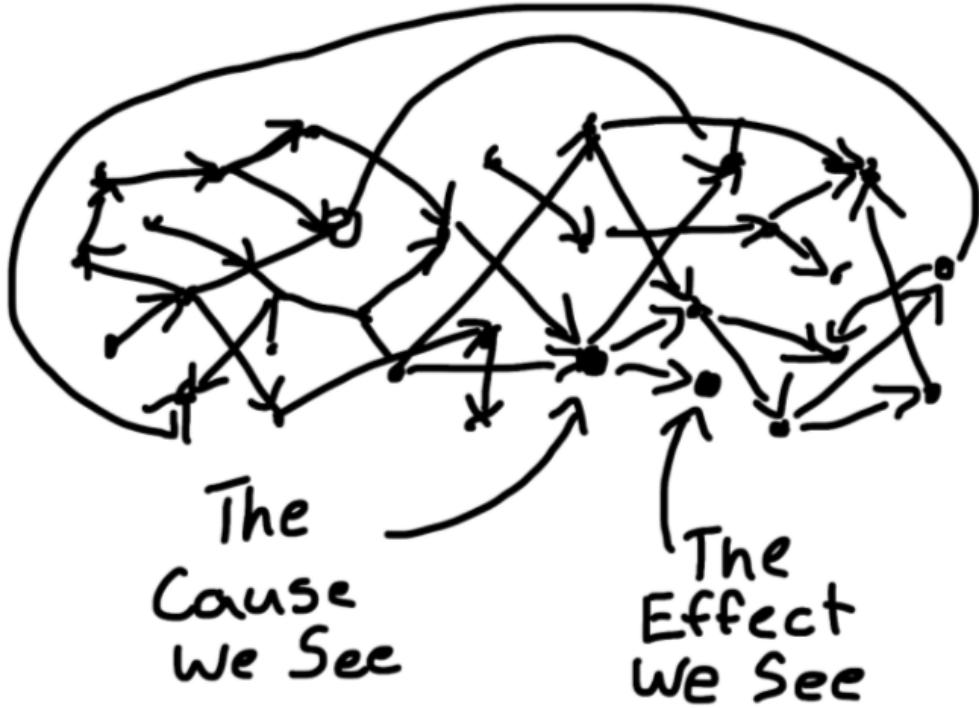
Spotlight{#xlevel6}

This section highlights some thought-provoking issues and quandaries with causal mapping.

--chapter:end:73001-level7.rmd--

bibliography: references.bib

Causal Mapping: Definitions



freshspectrum

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, someone claims that 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.

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 [@pearl2018]. 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 [@axelrod2015], based in turn on Kelly's personal construct theory [@kellyPersonalConstructTheory1955]. 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 [@axelrod2015]. 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:end:73005-defs.rmd--

bibliography: references.bib



Causal mapping for evaluators{#mapping}

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

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.

Features which causal mapping approaches have in common

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.

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.

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.

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.

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.

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.

Advantages of causal mapping

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.

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.

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.

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.

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".

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 return to the story in the original context.

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.

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

!--chapter:end:73008-cm-for-evaluators.rmd--

Causal Mapping and Outcome Harvesting

Causal mapping and Outcome Harvesting can complement each other in a number of ways, combining the particular strengths of OH, such as a clearly documented way to identify and validate specific claims of important changes, with the particular strengths of causal mapping, such as understanding how simple contribution/outcome stories can chain together into a bigger picture.

One way is to use causal mapping after an OH process, perhaps focusing on a subset of the data. [presentation in preparation]

Another way is to use causal mapping when OH has already been used for monitoring, to assemble provisional material and help identify possible outcome statements for further validation.

!--chapter:end:73100-spotlight-OH.rmd--

Context in causal mapping: how to code it

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

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

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

Version 1: context is just a causal factor

In general terms this diagram is ok, because we know the partial function from Spark and Oxygen to Flame:

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

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

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

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

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

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

Version 2: context is a causal factor which causally enables any causal link which has the same colour

In this version, the contextual information is explicitly encoded within each link which is affected by the context.

- It is quite a general solution, because it can cope with a wide variety of cases.

- If there are several links which are all active in the context “Oxygen”, they would all get the same colour.
- If “Oxygen” appears in the map as a causal factor, it could get the same colour; this approach is possible also if the context is never also mentioned as a causal factor, but the colour would have to be explained in another way.
- Visually it is a bit disappointing, because there is no causal arrow coming out of Oxygen, you have to make the visual connection yourself via the colours.
- This approach only works if we think of context as an all-or-nothing thing, which is rarely the case in real life: in fact, in our very example, the contextual role of Oxygen obviously differs according to its strength.

Using context in the Causal Map app

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

- Context: Oxygen
- Context: Science Experiment



At the moment, all hashtags are printed on the Print map. You can also search by and filter for contexts.

!--chapter:end:74500-context.rmd--

Beware the transitivity trap{#xtransitivity-trap}

| *For want of a nail the shoe was lost,*

| *For want of a shoe the horse was lost,*

| *For want of a horse the rider was lost,*

For want of a rider the battle war lost,

For want of a battle the kingdom was lost,

And all for the want of horseshoe nail.

(Thanks to Gary Goertz for remembering this one!)

The transitivity trap



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

from

1. $B \rightarrow C$

and

2. $C \rightarrow E$

we can only conclude

3. $B \rightarrow E$ in the intersection of the contexts of 1 and 2

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

For example:

If

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

and

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

can we deduce

3. $G \rightarrow C \rightarrow H$

and therefore

4. $G \rightarrow S$

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

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

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

If

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

and

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

can we deduce

3. $A \rightarrow C \rightarrow D$

and therefore

4. $A \rightarrow D$

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

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

Not just a problem for causal mapping

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

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

Can we mitigate the trap with careful elicitation protocols?

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

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

Transitivity trap, or identity trap?

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

--chapter:end:74700-transitivity.rmd--

Cases, variables and percentages in causal mapping?{#caveat-numbers}

In causal mapping we tend to in most cases steer away from any kind of "75% said ..." statements. It isn't just that we don't tend to use representative samples.

The point is deeper, namely that most sorts of causal map, the concept of a variable and even the concept of a "case" is a bit vague. In a causal map in general, even one combined from individual inputs, you don't have a clearly defined "case", and correspondingly you don't have clearly defined variables, which makes it more radical and less structured than, say, the world of QCA.

Some causal maps and diagrams are elicited as a group consensus about high-level variables like "economic growth"; others are constructed such that all the causal factors represent variables at the level of the individual respondent such as "improvements in nutrition in my household". Some QuIP maps are however heterogeneous in this respect. The causal chains in individual stories are random walks. Maybe respondent R

mentions a factor like “flooded fields” somewhere in a causal chain but respondent S mentions it nowhere. From this we cannot conclude that flooded fields were not a factor in any causal link for respondent S. We don’t even know if respondent S has fields which might get flooded. If they do, are they specific to this respondent or are they shared across a whole village? Part of a village? In this kind of causal map, our sense of what is a “case” is quite vague, and it is also, correspondingly, difficult to think of the causal factors as variables which are defined for every case. We can only find out whether a causal factor really is a variable which makes sense for each respondent considered as a case (or perhaps for some other set of cases, such as “school which the respondent’s children attend”) by reading the narratives and constructing the cases and variables for ourselves.

This soft approach to cases and variables make causal mapping idea for what are sometimes called “complex” situations or systems. However it makes it more difficult to report some kinds of metrics and draw some kinds of inferences from causal maps.

The notion of cases partly breaks down because while respondents may always answer on the individual level (“My household / my crops / my attendance at training” etc.) often in fact they name causal factors which do not live on this level (like “My village’s water supply / roads”) or even things like “the new law” or potentially non-changes like “Unemployment.” Suddenly, the respondents are not loci of varying information but loci of opinions about things which do not vary at household level – and often they may be something in between (“roads in my half of the village”; “for us pig farmers..”).

More generally, in a causal map drawn up from, say, a literature review, the questions of whether there are any variables or cases is something which only gradually emerges / gets constructed. Sometimes we might find that some areas of a map start to contain cases and variables (perhaps we have data from several similar schools), and perhaps there are different such areas, and they might even overlap (perhaps we have data concerning mothers and data concerning fathers).

!--chapter:end:74800-cases-percent.rmd--

bibliography: references.bib

Different kinds of question chaining, with QuIP and ParEvo as illustrations

We can build a causal map by just asking people, separately or together, for causal statements around a theme.

But there are many advantages in doing this more systematically -- to make sure we get the right information, and to make sure we have data which we can interpret.

discusses two methods – constructing causal links between a pre-determined set of causal factors using pair comparisons, and instructing respondents to draw a free-form causal map.

One way to build up a causal map systematically is "question chaining". For example we can ask respondents about the causes of an event, and then the causes of those causes, and so on.

Four options with Fixed versus Variable question chaining

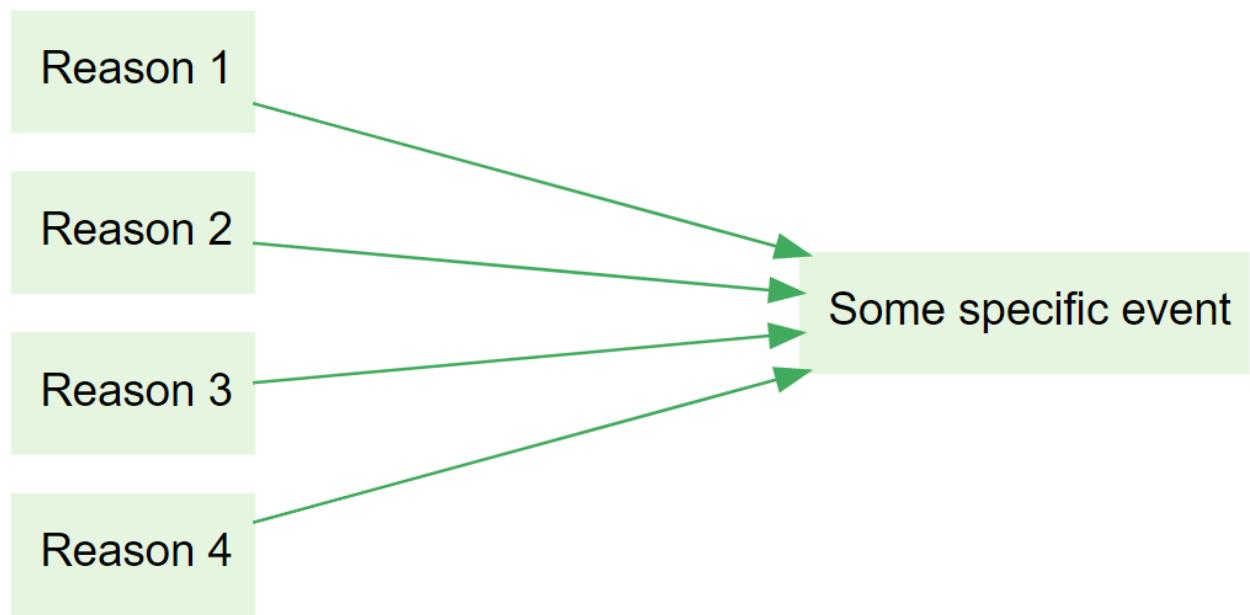
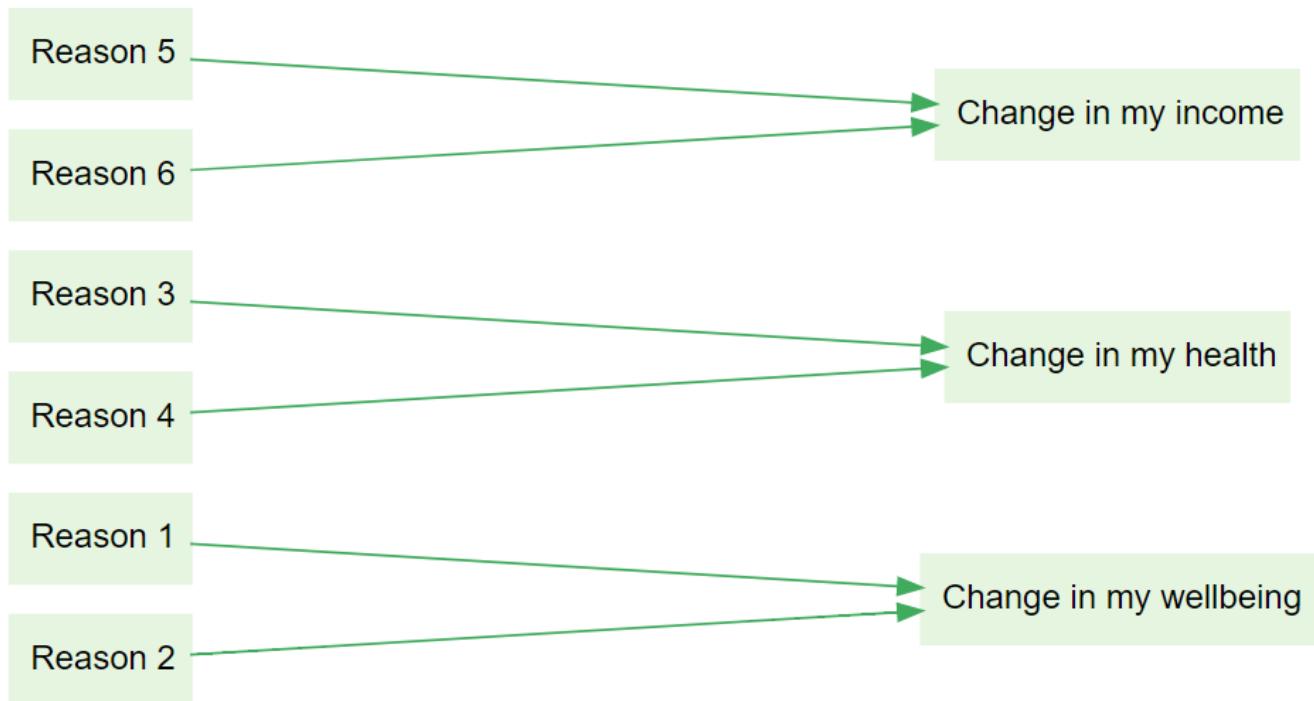
The image shows a set of handwriting practice lines. There are two sets of horizontal lines, each consisting of a solid top line, a dashed midline, and a solid bottom line. In the first set, there are two small vertical arrows pointing downwards between the middle dashed lines. In the second set, there are two small vertical arrows pointing upwards between the middle dashed lines. The lines are spaced evenly across the page.

| Fixed | Please give causes/reasons for event or situation X; then give causes/reasons for those causes/reasons, etc. (example: -- what were the causes of / reasons for Brexit, etc.) | Please give consequences of event or situation X; then give consequences of those consequences, etc. (example: **ParEvo** -- read this sentence describing the situation in the USA now, then suggest consequences of it, etc.) |

| Variable | Please name one or more X's, then proceed as above (example: QuIP -- name one or more changes in your life/wellbeing, and one by one give causes/reasons for those, etc.) | Please name one or more X's, then proceed as above (example: name one or more features of the current situation, and one by one give consequences of those, etc.) |

Causal mapping in the future is often used when constructing scenarios.

The distinction between Fixed and Variable options is subtle but important. In both, the respondent may offer multiple causal claims in the first step; the difference is that in Fixed questioning, these claims all involve event X, whereas in Variable questioning, the initial claims are not causally linked. This means that the causal maps look different: Variable maps can have more than one focal event.



In each case, the precise wording makes a difference as to whether respondents are able to answer the questions in the way we intended, and to how we are able to interpret and make deductions with the resulting maps.

When forward-chaining, we can, for example, ask for

- consequences,
- likely consequences,
- the most important possible consequences in your opinion, etc.

When backward-chaining, we can, for example, ask for

- reasons,
- causes,
- the most important contributory factors in your opinion, etc.

Single versus multiple

The rows in the table above actually make a double distinction: Fixed chaining usually begins from just **single** events, and open chaining usually elicits multiple. We could also cover multiple closed events by asking about each one in sequence, but this does not seem to happen often in practice. And while Variable chaining usually begins **multiple** chains "out of the box", we can insist on just one if we want to.

Future in past, past in future

We can also ask for (hypothetical) causal factors leading up to a future or hypothetical event. This is often used when constructing a theory of change: "What Y needs to happen for X to happen? What needs to happen for Y to happen?"

Again, the wording of our prompts is crucial: perhaps we are looking specifically for necessary conditions, or sufficient conditions?

Similarly we can ask for (counterfactual) causal factors following from a past or hypothetical past event. "What Y would have happened if X had happened? What would have happened if Y had happened?"

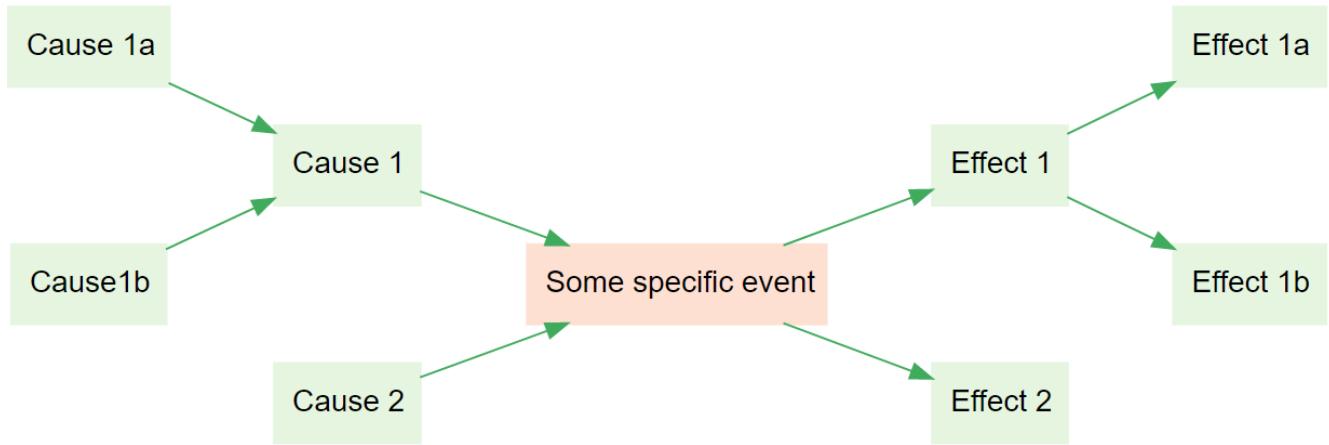
Closed vs open questions

The examples above all assume that respondents are invited to produce their own factors in each step. It is also possible to allow them only to select from a pre-determined list. [[@hodgkinson2004](#)]

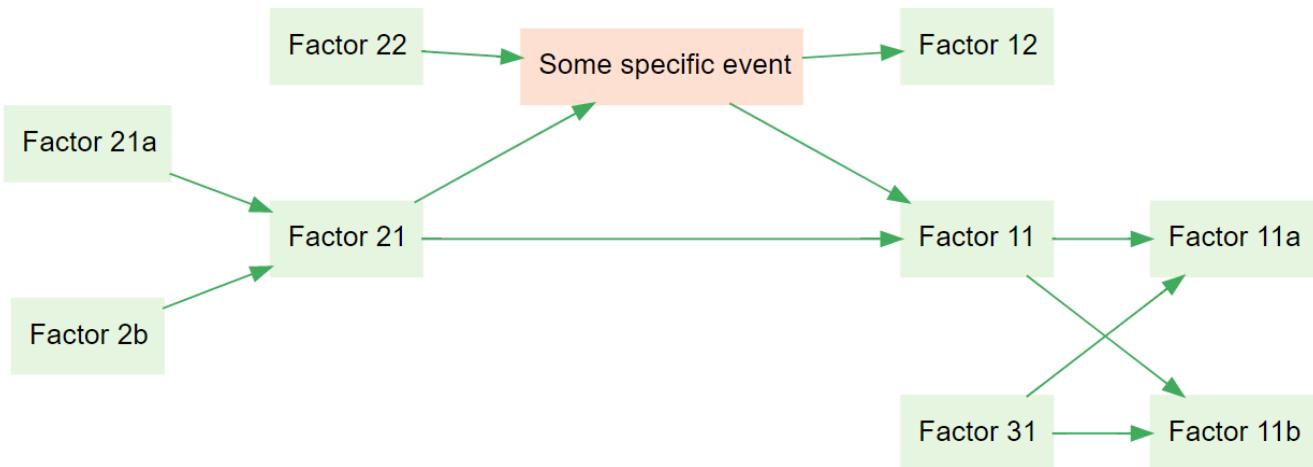
Going both ways: bi-directional chaining

Finally, we can combine backwards and forwards chaining into bi-directional chaining, of which there is a weak and a strong variant.

In the weak variant, we ask separately for causes of causes, and for effects of effects. The resulting causal map has a very specific structure: two hierarchies, joined only at the focus event or events.



In the strong variant, we can ask also for effects of causes of causes, and for causes of effects. This can result in any arbitrary structure.



Evaluation

It's possible to ask respondents to evaluate the factors at any point, for example to ask them if they see something as beneficial or positive. However, it's most usual to ask only for evaluations of the temporarily latest factor(s) – so, when back-chaining, to ask for an evaluation of the focus event(s), as in QuIP, and when forward-chaining to ask for an evaluation of the terminal events, as in ParEvo. This is parsimonious as one can, in each case, arguably derive an evaluation of the preceding factors by following the causal links in reverse direction. However it's maybe also not enough, as Simone Ginsburg points out and as discussed here.

!--chapter:end:75000-question-chaining-.rmd--

bibliography: references.bib

How to get to the shadow side: a problem with back-chaining

"The conscious mind knows nothing beyond the opposites and, as a result, has no knowledge of the thing that unites them. For this reason all uniting symbols have a redemptive significance." C.G. Jung, Archetypes & the Collective Unconscious, para. 285

This kind of question,

... and what were the causal influences on that? ...

... and what led to / contributed to that happening? ...

... what were the main reasons for that ... ?

repeated successively, is part of what we can call "causal back-chaining".

It is a core part of the QuIP data-gathering approach, and now it is also a key part of how the StorySurvey "question bot" works.

QuIP researchers are trained to do more than simply repeat this question and then stop. Even the StorySurvey question bot is more subtle than that. But at its core, this question, with light adaptations, gives us a single iterative step which can help construct an entire causal map.

This causal back-chaining algorithm is not perfect. It can't be: Causality is a vague concept with overlapping demands on it. It is homo sapiens' way of storing knowledge about which buttons to press but it is also part of our systems of judging, rewarding and punishing. Any algorithm we come up with for eliciting a causal map from a respondent will have flaws. We just have to do the best we can.

One key problem with the algorithm, which we address here, is that it primes us to think of only things which were in a sense positive, which "pushed the consequence factor in the same direction". There might be good reasons for this in terms of legal and moral thinking, but if we want to find out more about "the complex web of causal factors affecting this outcome", it's simply a strong and unwanted bias which we need to remove.

In a minute, we'll see an example of where the back-chaining algorithm falls down.
First, here's a closely related example where it seems to work fine.

Concert example

Jo wants to save up money for a concert. She succeeds. What was the cause of her success?

- She saved the money from Uncle Nick
- She saved nearly all her pocket money
- **She gave up buying sweets**

These are three factors which contribute to a (possibly implicit) numerical intermediate factor (amount of money saved) which is then dichotomised into "saved enough / didn't save enough".

The third one, giving up buying sweets, seems like a natural answer even though you could say it's a counterfactual: nothing actually happens (she simply doesn't spend money on sweets) but if she hadn't given them up, she wouldn't have bought the ticket.

Here, back-chaining is successful in implicitly identifying a negative factor buying sweets, which makes the outcome less likely. But it only succeeds in identifying it because it is contained within its opposite, giving up buying sweets, which "pushes the consequence factor" in the same direction as the others.

Football example

Liverpool and Everton are 0-0 until almost the final whistle, and then Everton get a goal in the last minute. What were the causes of Everton's success? The pundits say:

- Everton were more clinical in the final third
- Closing down the Liverpool strikers
- ... etc

Whereas, reasonably enough, if Liverpool had got the winning goal, the explanations would probably have been different, for example:

- Liverpool made more chances, and something had to go in
- Greater width
- ... etc

There's nothing wrong with this. But we have to be aware that the list of explanations given by the pundits is liable to flip at the last moment, depending on who gets that final goal. If we want to actually understand the causal dynamics of the game which was played, we really need the contents of both lists, the one where Liverpool win in the end, and the one where Everton win in the end, and we'd also like to know how these factors combine.

If we want to understand the game well enough to, say, advise a manager on a future encounter between the same clubs, it would be absurd to just pick one of the pundit lists we just mentioned, and throw the other list away.

One solution: pairs of initial questions

If we say that QuIP and StorySurvey can help understand the web of causal factors surrounding an intervention and/or an outcome, we have to do more than simple back-chaining, because we have to get at those negative factors too.

One way we can do this is, say, ask a pair of initial questions:

- What were the best moments of Liverpool's play, and what factors caused them, and what influenced those factors ...
- What were the worst moments of Liverpool's play, and what factors caused them, and what influenced those factors ...

we can even add:

- What were the best moments of Everton's play, and what factors caused them, and what influenced those factors ...
- What were the worst moments of Everton's play, and what factors caused them, and what influenced those factors ...

This is a useful approach (and it's easy to do in StorySurvey). It's an interesting challenge to then combine, where appropriate, factors which are common to the different sets of answers, including factors in one set which are contrary to factors in another set (e.g. "Liverpool seizing chances" and "Liverpool not seizing chances"). In particular, in StorySurvey, when respondents are asking question Q, we sometimes also suggest to them existing factors which have emerged within answers to Question Q; should we also suggest to them factors which other respondents already mentioned when answering a different initial question?

Solving the deeper problem

Using pairs of initial questions can really help. But it only solves our problem for the first step, the initial question. For each subsequent step, we get the same problem: asking simply "what were the influences on that ... " is biased to pick up the positive contributions but not the negative contributions.

There are different kinds of "negative contribution", which overlap.

- Factors which are a drain on a consequence factor (eg when we are thinking of the consequence factor as a linear sum of different influences, like a bank account or a flood level) and pull it "downwards".
 - For example, "lending money to little sister" in the Concert example above
- Factors which were a danger to a consequence factor but never actually had a decisive effect on it
 - For example, "This season, Liverpool are good at getting goals in the last five minutes" in the Football example above
- ...

So what can we do?

Can we add a "shadow side" to the back-chaining question armoury? It's possible to do in individual cases, but is there a general question form which can help us?

Counterfactual worlds

How about this:

You mentioned X. If X had not happened, what would have been the reasons for it not happening?

The trouble with this one is that it opens up a counterfactual world and if we back-chain from hypothetical events we might get even more hypothetical events ... We don't want to know about an alternative universe, we want to know about this one.

Best solution?

So my final suggestion is this, our best attempt at an additional question for causal back-chaining which can also catch "opposite influences" - drains and dangers.

You mentioned X. Were there any "opposite influences" on X: factors which tried to make X not happen, or tried to make the opposite of X happen, or which were a drain on X, or might have blocked X?

It's a bit strange, but could it work? Does it work with these typical factors mentioned in a StorySurvey on the response to Covid-19?

Pfizer Vaccine

Money

Rapid vaccine rollout

misinformation

political motivations

medical research

development of vaccines

countries not sharing data

countries slow to react

secrecy surrounding virus

quick spread

human deaths

countries working together

human nature

united people

Early lockdowns

The cdc eventually said to wear them.

We have to be careful not to label these drains, blocks and dangers (words which already have a negative sentiment) as "negative factors" (as I did above). This would be very confusing with factors which already have a positive rather than a negative sentiment, like "human deaths". You wouldn't naturally think of a factor which blocked or reduced human deaths as negative. That's why I suggest "opposite influences".

The important thing is that back-chaining with shadow questions should only surface factors which are/were actually present, and should not open up a parallel universe.

Encoding the results

When we back-chain using this "opposites back-chaining" question:

You mentioned "Lockdown happened": Were there any "opposite influences" on that: factors which tried to make it not happen, or tried to make the opposite of it happen, or which were a drain on it, or might have blocked it?

and someone says "pressure from sceptics" we should could encode this as "Pressure from sceptics --> ~Lockdown happened" using our "Opposites symbol, ~". After that, we can if we wish continue to back-chain from "Pressure from sceptics" in the normal way.

But it's not quite clear that the Opposites Symbol is the right way to go. Because up to now, with "barefoot" coding, from any causal claim such as

Source 1: Government intervened --> Lockdown happened

we can infer not only the causal link but also the propositions at each end of the arrow:

Source 1: Government intervened

and

Source 1: Lockdown happened

whereas here, it isn't actually the case that our source claims that Early Lockdowns didn't happen.

Conclusion

Can we add this kind of shadow question (sometimes) into the StorySurvey question bot algorithm? What about QuIP?

Finally ...

I'm not arguing here for a variable-based approach as in quantitative causal analysis or systems dynamics. We don't need to think of the factors we identify as quantities which are eternally present within the context and simply vary over time. We are of course still thinking in terms of propositions rather than variables, which means in particular we can embrace narratives, e.g. we can include factors which only emerge half-way through the game (Mo Salah was sent off!) and which spawn a lot of other factors in their wake.

Also, I've used the word "blocker", which is sometimes used for a third, moderating factor F which changes the influence of a first factor C on a second factor E. But here, "blocker" simply means a factor C which endangers or drains a factor E, makes it less likely. Moderating factors are still not specifically captured by our algorithm, even when extended as suggest

!--chapter:end:75000-question-chaining-problem.rmd--

 **Strength: Adding additional information like strength of a link in a causal map**



Let's keep it simple.

Causal mapping involves modelling people's causal claims and beliefs as links or arrows between factors. In Part 1 we argued that a causal mapping approach is represented first and foremost by sets of rules which say how causal information is encoded by specific features of the elements of the map. In this second Part, we will present some new rules for encoding causal claims within causal maps. We will see that there is no one perfect set of features we can use for encoding all or even a satisfactorily broad selection of all possible causal claims. Even with some quite basic possible features like encoding the "strength" of a causal link, it is not hard to dream up a range of similar but different sets of rules for encoding and manipulating causal maps using this feature. To encode strength of a link, should we use numbers from 0 to 1, or maybe from -1 to +1? If several links point to a factor, should we simply add up the strengths, or use some function which keeps the value of the factor somewhere lower than 1? And so on.

These kinds of problems are common to all attempts to model phenomena. But the problem is particularly acute with causal mapping, especially in causal evidence mapping, where we try to encode many individual claims or beliefs. For example, when viewing actual causal claims we sometimes (but not very often, unless the sources are

academics) come across claims about how a set of two or more factors interact to influence another in a way which cannot be captured just by sets of individual links: one of many phenomena sometimes called “non-linear”. For example, "you need fuel AND oxygen AND a spark to get a fire". The trouble is, are we going to activate a whole new set of conventions and rules just to cope with this minority of causal claims? Probably not.

In practice, we suggest that any one causal map should select an only *one* approach (only *one* set of rules for encoding causal ideas with specific features of the map's elements) and that this approach should be no more complicated than is necessary to encode most of the causal claims. Mixing approaches within a single map is endlessly tiresome and complicated. It is not a good idea to introduce an additional feature (like "C is necessary but not sufficient for E") just to cope with a minority of causal claims in the map.

!--chapter:end:75001-strength.rmd--

The last minute principle

A principle we try to follow at Causal Map:

"leave the evaluation (the evaluative judgements) to the evaluator":

We see the job of the software/algorithms as preparing everything data up to the point of comparison. For example, to help prepare information like this:

"you'd have to destroy 37 pieces of evidence in order to destroy any argument from intervention X to outcome Z, but in the case of intervention Y, there are only 8 pieces which you'd have to destroy, and the chains in the first case are shorter and involve more respondents. Let's now discuss what this means"

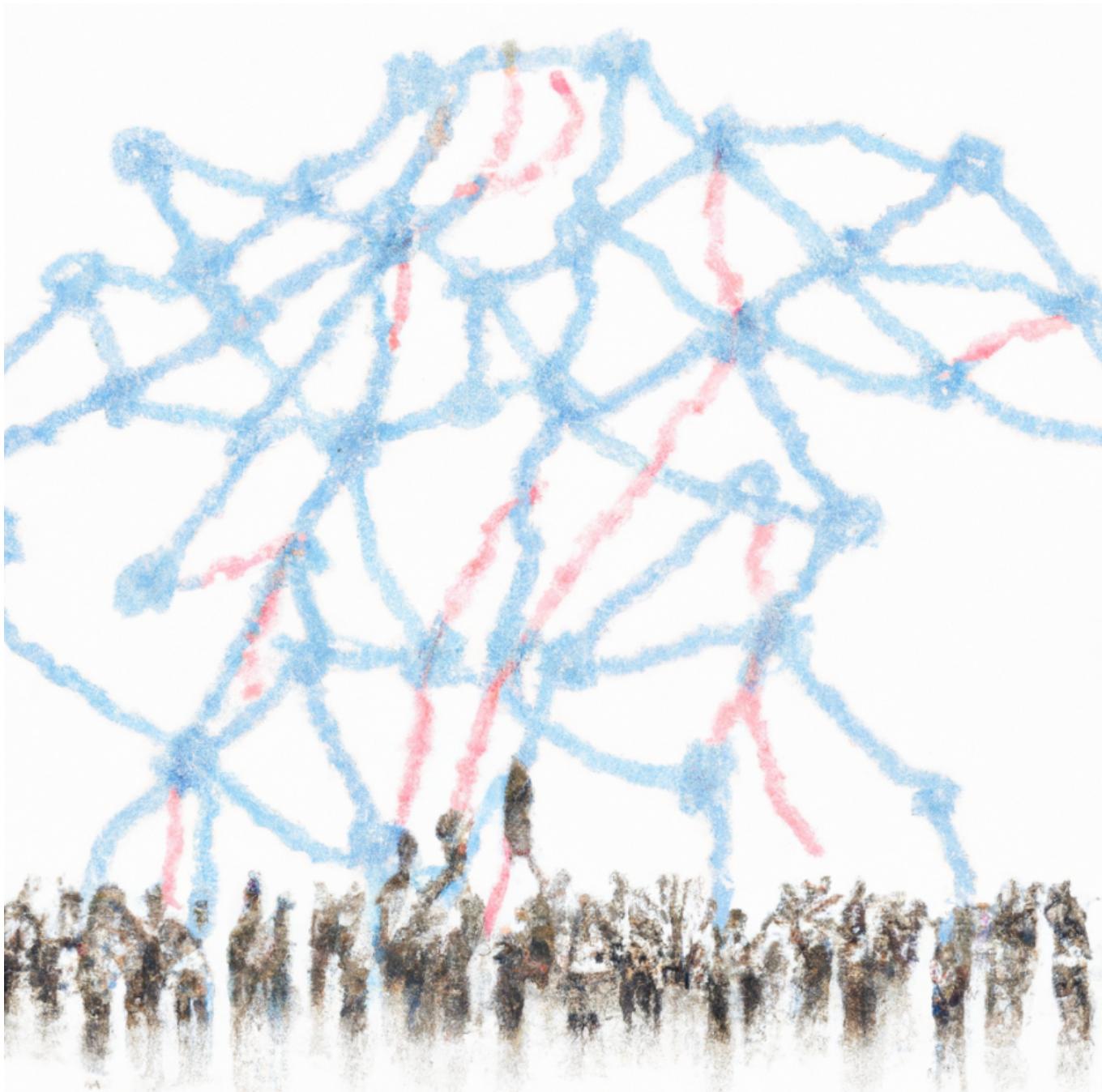
... rather than this:

"the software says Intervention X had a bigger effect than Intervention Y".

!--chapter:end:75300-last-minute.rmd--



Causal maps, systems maps, what's the difference? Does it matter?



What is the difference between a causal map and a systems map?

Causal mapping is related to systems mapping, but we are specifically interested in causal connections, and we specifically map what individual sources think/believe/claim.

The focus on causal connections means that causal maps are useful to "cut to the chase" and can be helpful in addressing questions like

- "what are the most important drivers of X - according to the stakeholders"
- "what are some of the unintended consequences of Y- according to the stakeholders"

The focus on individual sources means that we can aggregate and simplify information from quite large numbers of sources and can address questions like

- do people agree about this?
- what are the subgroups?
- do women see this differently from men?

Approaches called "systems mapping" often use causal links, and may or may not mix causal and non-causal links within a map. However, systems mapping approaches do not usually have this cognitive aspect, each link just encodes "X influences Y" whereas a link in a causal map encodes "X influences Y according to source S."

Are systems maps about systems?

-
But also, many approaches have a way of mapping what-causes-what in which the resulting map is called a "systems map" without reflecting very much on the "system" part of the name.

How do we know it is actually a system? Do you sometimes look at the map and say "nice map, unfortunately it isn't a system, so we will discard it, or at least not call it a systems map"?

Or, perhaps the resulting map is always a systems map by definition, because any set of (interconnected) causal drivers of something is a system? Surely not?

Or, are there exclusion criteria which are applied during the process of constructing the map, according to which someone (the moderator?) says "sorry, X is indeed a driver of Y but we can't include it because it isn't part of the system"? If so, who decides this and how?

"Causal map" seems to us a conceptually more modest name: it's just a map of causal connections.

However, as Bob Williams pointed out in an email exchange on this question:

The systems field was born in the practical task of solving real world problems, and many of the methods and methodologies are deeply investigative, reflective and solution oriented.

When we do causal mapping, what toolkits and ideas are we relying on which causal mapping brings to the table? Do we add anything new apart from e.g. using metrics like Centrality from network science to identify important factors? In particular, causal mapping (arguably) started from the attempt to map the internal causal models of decision-makers, and has grappled with problems like "are the rules for modelling the strength of a causal connection in the real world related to the rules for modelling how certain we are about it"? How do we model the evidence that a decision-maker believes X does not cause Y? When can we use evidence that stakeholders spontaneously mentioned Intervention X as an upstream cause of Outcome Y as evidence that X in fact meaningfully influenced Y? Some of these issues are already mentioned elsewhere in this Guide.

So if we've bothered to get people together to do some causal mapping to tackle some recurring problem, is there already a hint of something system-like which is e.g. stable and pushing back at us?

As a phrase, "causal map" is nice and clear: a map of the causes and effects of things. If a map consists entirely of causal links then it's certainly a causal map. It might be a systems map too e.g. if someone has done the work of highlighting/curating the map to focus on systems phenomena.

But "causal mapping" (compared to systems mapping) as an activity sounds a bit flavourless. Why would you want to do it, when would you stop? Where do we draw the boundaries, how do we exclude factors that might not be salient?

What useful tools/ideas do we offer to help move from (causal) modelling to solving a problem?



Coding time in causal mapping

Many of the questions we want to answer when doing causal mapping have to do with time. Of course, there is always some kind of flow of time through our causal maps, because effects come after causes. But the actual “timings” – specific time points or periods associated with particular things – can take almost any order. In causal mapping, we don’t have any fixed set of options in order to allow the analyst to capture causal stories free form, without any preconceptions. In quantitative research, time appears in research designs in quite straightforward ways, usually as a continuous variable (which may be left continuous or transformed into a discrete variable with values like “before” and “after”); research conclusions can be, for example, “the dependent variables decreased over time”. Often it will miss stories which are not linear, for example when things get better, then worse, then better again. When designing research to analyse with causal mapping, we may also have a research design which already anticipates temporal questions, for example we want to compare people’s causal maps before and after some intervention. However, the really exciting (and for the analyst, really demanding) potential of causal mapping is to identify temporal aspects to stories which only emerge when listening to the respondents.

Temporal features which can emerge without explicit coding of time

It is useful to note some common patterns and features which can emerge even before we have any system to explicitly code timings:

- Some factors might have a particular time point or period, e.g. “the Launch of the London Olympics”. Or they might be things which repeat, like “Start of the school year”. We often see causal links between repeating or generic events, like “Good start to school year → Good progress in first semester”. These are about generic or circular time. Some maps may include factors with fixed timepoints as well as ones with generic timepoints.
- Some factors might be about background conditions and are not events at all, e.g. “unemployment”. Or they might be continuous variables like “level of aggression” or “client satisfaction”.

- When all factors are in a *single chain*, then there is a very simple time progression through the whole map, and we know for any two factors that one is before or after (or simultaneous with) the other. However, this is quite rare.
- There is a big distinction between maps which do or do not contain any loops. We call maps without loops *acyclic*. Maps with loops can be used to capture things like feedback, creating vicious and virtuous circles.

A very common case is when respondents report an intervention or some other event as a cause, a factor which influences other factors. Those consequence factors might only be mentioned as a consequence of the event, or they might also be mentioned (but perhaps less often or with less emphasis) before the event.

For example:

“Before the training, my skills were ok, and they improved with self-study, but after the training they were much better”

This could be coded in various different ways, but perhaps most simply like this:

Self-study → skills

Training → skills

All causal coding automatically contains temporal information, because the consequences (at least in some sense) happen after the influences. In many cases, this is all we need. For example, looking at the downstream consequences of “Training” may be enough to qualitatively answer questions about “What was the effect of the training”: we can look at the map downstream of it, especially looking at effects on factors we are particularly interested in (usually some kind of outcome like “student progress”), and compare this with the rest of the map; and we can look upstream of important factors to compare how much of any positive influence on them can be traced back to the intervention as opposed to other factors.

When we think about influences and consequences in this way we can often ask and answer questions about time and effect without ever mentioning time or temporal aspects explicitly.

The rest of this paper deals with other cases when this is not enough.

Temporal information attached to additional data

The statements we process when doing causal mapping usually have additional data attached to them, for example respondent ID and gender, question number, etc. In some cases, the research may have been planned (and the data explicitly gathered) with temporal questions in mind, so that we have *temporal* additional data. For example, we might have gathered data at separate time points, before and after an intervention and then in a longer-term follow-up. In this case, for every statement, we know from its additional data which of these three timepoints it belongs to. This way, it is possible (if desired) to use the same set of factor labels for all of the timepoints but quickly and automatically identify just those causal links which are associated with just one time point (because they come from statements which were actually made at that time point); or we can compare maps between timepoints, etc.

Alternatively, we may have asked respondents *questions* which concern time. For example, we might have some questions which are explicitly about:

- one single timepoint (like the launch of something)
- a specific period (like the time after a launch)
- generic or circular time (like “start of the school year” or “after a child is born”).

If we have a number of questions which share some of these aspects (for example several questions about “start of the school year” and several about “rest of the school year”), we can add an additional data field to capture this information (for example a field called “school year time” with two possible values “start” and “rest”). This additional data for each statement can be referred to by the analyst when coding, and can be used to filter the causal map as described above. It could also be indicated in the map itself, e.g. by using different colours for links associated with claims from different time points.

We may have planned our research with temporal questions in mind, and so our additional data already has neat fields to capture temporal data. In other cases we may decide during the analysis that we need to make use of temporal information which is implicit in the additional data – for example, we may be able to work out from the interview number when an interview was held, and so we might want to go back to the additional data and add something like “interview week” as an explicit additional data field.

Recording temporal information in causal links

The analyst may decide that the temporal aspects of the *links* in the emerging story are particularly important. Sometimes it is enough to merely mention them in each causal link, for example to attach a hashtag like “before intervention” to a link – perhaps without recording any temporal information in the factor labels. Sometimes a more systematic approach is required, for example to establish two hashtags “before intervention” and “after intervention”. (It might even be required that the hashtag always be applied to every coding, though this is not necessary). So, for example, the statement:

“Oh, before the intervention, as soon as we had a discussion on a difficult theme, we’d get into an argument, but now we don’t, we actually make decisions instead!”

might be coded as follows. Here, the text in brackets after the arrow is understood as being attached to the link, not to the factor labels:

Family discussion on difficult theme →(#before intervention) Family gets into argument

Family discussion on difficult theme →(#after intervention) Family makes decision

Text like this can actually be displayed on the links in the resulting causal maps if desired.

Note that the factor labels are written in a way that they make sense for either application; they don’t also need to include temporal information.

If we are also coding other aspects of the causal links like strength or direction, we might even code like this, where the very same pair of factors is involved in each case, and the difference is stored only in the links:

Family discussion on difficult theme →(#before intervention; high probability) Family gets into argument

Family discussion on difficult theme →(#after intervention; low probability) Family gets into argument

This is just an example; temporal information might not be coded simply as the names of discrete phases like “before” and “after”; for example, hashtags like “rapid” or “unchanging” or “repeated” are also possible.

Recording temporal information in factor labels

Likewise, the analyst may decide that the temporal aspects of the factors in the emerging story are particularly important.

Coding temporal aspects within links may be more suitable when there is an overall causal structure which does not itself change much over time, whereas coding temporal aspects within factor labels may be more suitable when respondents are thinking in terms of different (although maybe similar) sets of causal factors time which differ in time.

How to code temporal information in factor labels?

Sometimes it is enough to merely mention them in a factor label, for example “improved concentration after intervention” in an ad-hoc way. Sometimes a more systematic approach is required. There are several possibilities.

Use a hashtag

It might be enough to simply use a consistent set of hashtags, appearing anywhere in the factor label, for example:

“Oh, before the intervention, as soon as we had a discussion on a difficult theme, we’d get into an argument, but now we don’t, we actually make decisions instead!”

... might be coded like this, where the hashtags are understood as being attached to the link, not the factor labels:

Family discussion on difficult theme #before-intervention → Family gets into argument #before-intervention

Family discussion on difficult theme #after-intervention → Family makes decision #after-intervention

This way, as the hashtags are unique and searchable, it would be possible to show a map consisting only of #after-intervention factors.

Sometimes a more systematic approach is required, for example to establish two hashtags “#before-intervention” and “#after-intervention” which are used systematically.

You can also use hanging hashtags like this:

Family discussion on difficult theme; #before-intervention → Family gets into argument #before-intervention

Family discussion on difficult theme; #after-intervention → Family makes decision #after-intervention

Use components in hierarchical coding

This approach allows us to filter just for factors with or without one or more specific hashtags. However, what if we want to zoom out to a causal map which does not make this distinction, because we want to see family processes in general, regardless of timepoint? We would like to be able to make a higher-level factor “Family discussion ... “ which generalises from the specific timepoints. We can do that by making the timepoints into components in hierarchical coding style:

Family discussion on difficult theme; before-intervention → Family gets into argument; before-intervention

Family discussion on difficult theme; after-intervention → Family makes decision; after-intervention

This way, when we zoom out to view only first-level factors, all the causal links to and from the specific factors for the two timepoints will be rerouted to the top-level factor, to give a time-independent overview. Effectively we create “virtual” factors like Family discussion on difficult theme and so on

Use higher-level temporal components in hierarchical coding

Another approach is to do this the other way round, with factors like these.

Family stressors before intervention; discussions on difficult themes

Family stressors before intervention; bedtime

Family behaviours before intervention; making decisions

Family behaviours before intervention; getting into arguments

Family stressors after intervention; discussions on difficult themes

Family stressors after intervention; bedtime

Family behaviours after intervention; making decisions

Family behaviours after intervention; getting into arguments

Here, the temporal aspect is moved to the top of the hierarchy, so we could zoom out to create “virtual” factors like Family stressors before intervention and so on.

These components can also be mixed in with any other components used during coding, for example components for different regions, contexts or genders.

All of these suggestions can in principle be mixed in with one another, but the analyst is advised to be very careful not to make things too complicated.

The temporal aspects mentioned here are just examples; temporal information might not be coded simply as the names of discrete phases like “before” and “after”; for example, concepts like “rapid” or “unchanging” or “repeated” are also possible.

But what if factor labels themselves express changes?

We have already recommended using heterogeneous, "in-vivo" factor labels and to resist the temptation to formulate all factor labels in a strictly similar way: when in doubt, 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. In particular, this means that sometimes you may find yourself using some labels which express a difference, even a difference over time, like "improvement in X" and some which do not. It can get confusing using these kinds of labels at the same time as some of the temporal approaches mentioned here. But, we argue, that is the way we humans think about how things work, as a heterogenous pile of fragments of connected causal information. This is always a balancing act, but we encourage you to side with your respondents and try to capture what they are saying rather than impose a schema.

Lateness{#xlateness}

Evaluations often report timeliness and/or lateness. It can be difficult to code a sentence like:

Thanks to UNICEF's intervention the material was delivered but with some considerable delay

Three ideas how to code this:

- use a hashtag like #late on the causal link

- use a hashtag like #late in the factor label:
 - UNICEF's intervention --> Material delivered #late
- use a hanging hashtag :
 - UNICEF's intervention --> Material delivered; #late

Your choice of what to use might depend on what you think is here the primary story: is it the fact that the material was delivered at all, and the lateness is just a qualifier? Or is the lateness the main story?

Analysing and displaying the results of temporal coding

All the usual tools are available to us to analyse and display these results. For example, we can:

- make tables of frequencies of factors or links broken down by hashtag
- show submaps for
 - o one or more values of a temporally relevant additional data field
 - o and/or one or more temporally relevant link hashtags
 - o and/or one or more temporally relevant factor label hashtags/hashtags
- and/or we can zoom in or out if we have used hierarchical coding with one or more temporally relevant components.

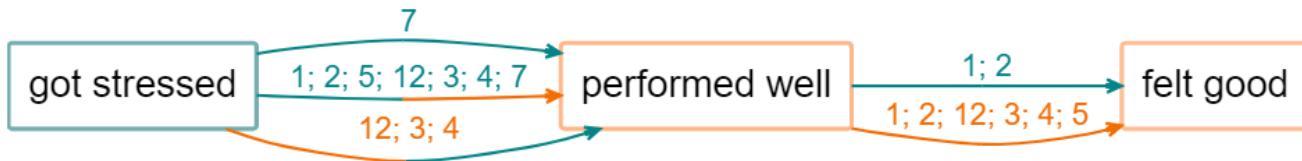
In particular, if a small number (2, 3, 4) temporal categories (like, “before, during, after” or “year 1, year 2, year 3” have been used in additional data fields, and/or when coding links and/or when coding factors, it is possible to arrange the above kinds of output in a left-to-right, storytelling fashion. For example, three simple, top-level maps for three time periods can be displayed in a row. Or, the time periods can form the columns of a data table showing for example how different factors were mentioned with different frequency at different times. In the simplest case, as mentioned on page 2 of this paper, the map itself forms a left-to-right causal story, which can be read, with some caveats, according to a schema like “this happened, which made this happen, which made this happen”.

!--chapter:end:76500-time.rmd--

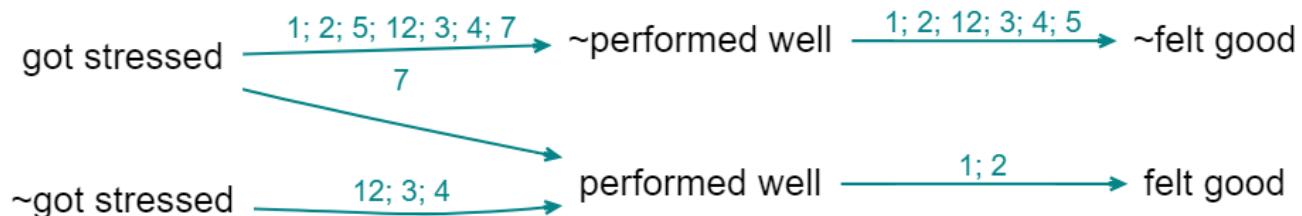
More about coding opposites

Reversible?

Combining opposites is not quite reversible: there is more information in the uncombined map than in the combined map.



Above (the links are labelled with source ids), you don't know whether source 12 mentioned the whole route from "got stressed" to "felt good", whereas below you can see that this is *not* the case.



So, sometimes you might want to check the uncombined map.

Combining opposites is a (mild) kind of simplification and every simplification is a potential transitivity trap.

Combining opposites and strengths

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

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

Graded factors: differences of degree

With coding of opposites we decided to not do any parametric coding e.g. by applying some quantity to the links but instead to capture the idea in the factor labels and, where necessary, do some tricks to make the app itself to combine opposite factors. We can use the same idea to capture gradation in a barebones way, given that we rarely have any kind of very comparable or numerical information in causal mapping.

Suppose that usually training leads to increased skills and then we have one piece of information that the training did lead to increased skills but not by very much. Then following the way we introduced the ~ sign, we could use a qualifying sign like, say, @ in front of the factor label: training held → @skills increased. However, we recommend **coding gradations by using hierarchical coding:**

training held --> skills increased; small.

After all, a small skills increase is still a kind of skills increase.

Philosophical discursion: coding with propositions

On the one hand, with the app we want people to be able to code using their own philosophy and we just provide a broad range of compatible tools. But it's also good to have a recommended style and today we are talking about that.

The most important task is also to help depict the causal stories that people are actually telling us in a way which is intuitively readable in the overall map and in filtered maps and reports.

Nearly all causal mapping uses, in some sense, variables. And the thinking behind causal map as an app comes from this tradition. Although we softened the idea to talk about semi quantitative variables, which are just in some sense, "more" or "less", but cannot usually be numerically measured. Nevertheless, when you have variables, then the links between them are in some sense functions, telling you how the level of the variable at the far end of the arrow is controlled by the variable at the start of it. For example, increasing or decreasing. And, indeed, with the app as it is now, we can do that by coding as strength between minus one and plus one on the arrow. And this is a perfectly workable way to do coding. And to some extent, the app also uses that information for synthesis and visualisation.

You can also call this "parametric" coding: a parameter is some kind of number on each arrow, which tells you how strong it is.

We also have the ability to put something like qualitative parameters or hashtags on the arrows. These are simply ways to enable us to filter certain arrows or sets of arrows into or out of a map. But they're definitely not numerical, so you can't do any kind of calculation with them.

However, we've moved to a still more primitive or barebones way of understanding what we already do in most QuIP coding

which we think better reflects what people actually say and is not parametric at all, i.e. the boxes are not any kind of variable not even semi quantitative variables, but propositions in the sense of declarative statements like "sold cow" or "income increased". These are simply short sentences, which can be true or false. So our usual causal claim, will be between

two propositions or sentences. This way, we don't need to make any kind of distinction between factors which

code true/false events like "the law is passed" and factors which are somehow semi quantitative like "amount of income", because we collapse both into simple propositions. So for example, "new tax law passed --> had less income", both of which are true/false propositions. This approach would seem to be just stupidly easy. And it carries a lot less logical baggage with it. So we don't need to worry about the type or nature of the factors or the type or nature of the links between them. Because the links are simply uninterpreted. And just mean, causally influenced, nothing else. This way, a causal map is nothing but structural and encodes only the information about what influences what in some way or other. (It may also be closer to the way we think in terms of the theory of cognitive schemata but that is another story...).

Probably 90% of claims within a set of relatively wide ranging QuIP interviews can be coded, and in fact are coded, this propositional way, without any further ado, no need to rethink "got more cash" as "level of cash".

The problem of course, comes when we get factors which in this way are coded differently, but somehow seems to relate to the same thing. A classic example would be "income increased" and "income decreased"

(There is a related issue about the extent that these are referring to the same thing, for example, are we talking about one household here or several households or a whole country or whatever, and if each respondent is talking about their own household, can we merge this information, etc)

With a strength-based variable-based way of coding, we can encode them with a single variable and differing strengths.

Whereas with propositional coding, it might somehow be a common variable but it's coded with two different labels. And we had no formal way of connecting increased income with decreased income.

(We can imagine almost a social history of measurement and the social construction of variables in the sense of Wittgenstein's language games at the beginning of the Philosophical Investigations, where people use propositions like "income decreased" and "income increased" without worrying about it, and then when it comes to maybe making financial plans for the future to realise that we have to have some kind of arithmetic which enables us to combine these ideas in a more formal way.)

The point is that in causal mapping of the kind that we do, these occasions are relatively rare. And most of the time we can make do with the propositional approach which is just simpler and faster and more natural in most cases.

There could be much more focused QuIP protocols designed by academics to tease out whether people think X is necessary for Y or only sufficient, etc, but we aren't addressing that here.

So in short, with propositional coding, Decreased Income is coded as ~Increased Income and you can have your cake and eat it: sometimes view this as a separate factor, but when you want you can flip it and combine it with Increased Income, and the app will remember what you flipped when doing summaries etc, e.g. colouring the links differently.

Of course, it isn't just oppositely named factors like income increased in income decreased: a big challenge in causal coding is also differences of degree. And also the claim of zero influence.

Zero influences.

We still have no way to encode a zero influence.

--chapter:end:76600-opposites-extra.rmd--



Puzzles in causal mapping

Distinctive Groups

- How to find if there are important supergroups of similar groups, e.g. can we put villages A and B together into one supergroup with similar submaps and villages C and D into another?
- Can we use the same idea to locate arbitrary subsets of statements (regardless of the additional data) which produce sets of maps which differ the most?
- What is a suitable cut-off for a significance test of differences between the groups, given the size of the map?
- What does it mean if a small, or a large, proportion of factors are distinctive for a group?
- What does it mean if we get many more distinctive factors for one group than for another even when they are of similar size?
- How to interpret results for groups of very dissimilar size?

Quantity of Evidence

- What is the relationship between plus and minus QE on the one hand and the total QE on the other?

Joining rules

Combination rules showing how to build up larger maps. The rules each have two parts saying (syntax): mini-maps joined in such-and-such a way also count as causal maps and (semantics): what does the composite map mean in terms of the meanings of its elements? Combination rules are not just for actually combining information, say from different sources; more fundamentally, they enable us to interpret any arbitrarily large map by using the mini-map rule and the combination rules.

The different ways of combining elementary maps into composite maps embody some interesting social science questions.

- joining two maps on a **shared item** (what counts as the same thing? is "hunger" according to one person's report the same as "starvation" in another?)
- joining two maps on a **shared consequence item** (this is surprisingly difficult in the general case which includes problems of multi-causation or overdetermination). Many approaches make the assumptions that different influences can in some sense just be added together, so that the influence of X on Z is always independent of the influence of Y on Z. But other approaches (like perhaps QCA) specialise in combining influences which are not independent of one another

- joining two maps on a **shared influence item** (usually unproblematic in approaches like QuIP and DAGs, but in Systems Dynamics, where the boxes are containers and the arrows represent flows, this is a big issue).
- merging **co-terminal arrows** from different maps into one (what does it mean when we have overlapping information from different sources about the same causal influence? how do we combine them if they agree or if they disagree?)
- whether or not to allow joining two mini-maps in such a way that a **loop is created**, and if so, what does this mean? Fuzzy Cognitive Maps specialise in these kinds of loop.

Causal map approaches have their own versions of each of these rules. Plus, individual approaches will have their own rules about how to combine any additional information which they allow for (e.g. level of trust).

Does believing that X influences Y mean you have to believe that not-X causes not-Y? Causal mapping and counterfactuals.

We can only say C causally influenced E if it made some difference to E, which means that *believing causation might imply believing a counterfactual*: if the Cause had not happened, the Event would not have happened[16]. It's important to understand that this only makes sense in the narrowest way. It might be true that these protests were one of the direct causes of that revolution, in the narrow sense that if they hadn't happened in just the way they did, but everything else had stayed the same, then the revolution would not have happened, at least not in the way it did happen; but it certainly doesn't mean that protests cause revolutions or that there are no other causes of revolutions. This also doesn't mean that the truth of this counterfactual is part of a definition of causation. It just means you can't say C caused E if it didn't make any difference to E.

We also stress that *evidence about causation very often does not come from evidence about a counterfactual. Reasoning about causes, and gathering evidence about causes, are only possible as part of a whole lot of interlocking knowledge we already have about how things work.*

What if someone says, it's the white chocolate which makes this cake so delicious, and goes on to say, it's the dark chocolate which makes this (otherwise identical) cake so delicious. If white and dark chocolate are mutually exclusive alternatives, this doesn't make causal sense?

Homogeneity of paths

Can we think about transitivity in terms of a metric?

Scenario 1) 20 people might have said that A links to B and another different 20 might have said that B leads to C with no overlap between the groups of people.

Scenario 2) the same 20 people say that A links to B and that B leads to C.

Conventional ways to combine the information from the different sources would produce the same diagram in both scenarios. But we want the user to see that there is some kind of weakness in scenario 1. Can't we just show this on the arrows somehow?

In the case of isolated paths with no forks in them, which of course are very rare, this wouldn't present a big problem. Assume that we are showing, as we do, now the total number of mentions for each section represented by for example the width of the arrow.

We can construct a measure which we could call "homogeneity" for each section. So if in a long path, each of the sections were mentioned by more or less the same people the stretch would have high homogeneity and you could for example show it unbroken. In contrast, for an arrow in which separate sections were mentioned by different groups of sources, it would have low homogeneity and you could show that for example with a very broken arrow with lots of gaps in it. Or by making it almost transparent.

However I don't think this works in the much more common case when paths are of course constantly diverging and rejoining.

For example, suppose you have an arrow from B to C and then an arrow from C to X and another from C to Y; suppose there was high homogeneity from B to C to X in the sense that both of these sections were mentioned by many sources but the section from C to Y was mentioned by a different bunch of sources. How would you mark this?

The problem is that homogeneity as I've described it as a metric of entire paths and so you can't really show it in individual sections when there are forks, i.e. when one section can have different homogeneities because it is part of different paths.

(non-) solution 1)

What one could do is see whether there is any clustering within sources rather than within variables. So you might find there is a bunch of people who tend to mention many arrows the same and another bunch of people who mention a different set of arrows. It would certainly be possible to automatically or manually create subgroups of respondents. Then there are ways to show how different sections of paths were mentioned by these different subgroups, but it doesn't really answer the question.

Solution 2)

You can certainly report the homogeneity of each individual path in the Report tab of the app, but that would be quite long-winded.

Solution 3)

You could show the same information interactively. For example when you click on B in the example above, you would see the width of the section from C to Y shrink but the section from B to Y would stay the same. However I am not a big fan of information which you can only discover by twiddling.

Solution 4)

You might want to summarise the most important results in Solution 3 e.g. in a legend on the diagram, where you could (automatically) mention individual paths with particularly high or low homogeneity. You could probably develop a metric for the overall homogeneity of a whole diagram, and you might want to mention if a whole diagram had unusually low homogeneity.

Solution 5)

Allow the viewer to "play" the different sources one by one, if there aren't too many of them.

--chapter:end:77000-puzzles.rmd--

Is one direction enough?

Alfred Barr's famous diagram poses questions for those trying to plot causation, primarily whether one-directional arrows over-simplify relationships and seem to reduce stakeholders agency?

Barr elegantly organises the history of cubism and abstract art styles, artists and places are mapped and flow down through the years. Colours are used to clearly separate internal and external influences, whilst the size of the font reflects their historic relevance.

![TufteOG](__assets/TufteOG-16335945994571.jpg)

As Tufte (2006) highlights in his book "Beautiful Evidence" Barr made many interesting and, in some cases, controversial choices to fashion this map - such as the lack of arrow between Bauhaus and geometrical abstract art. Those used to causal mapping will be familiar with the multitude of decisions surrounding the creation of such a map. Barr's choice to have single direction arrows is a decision that has attracted attention and we find very intriguing.

First, this map implies that the relationships are all one-way. Expressionism influenced Dadaism but Dadaism did not influence Expressionism. Whereas, art movements are not confined to a set period of time; reality is more complex than that and movements are often interrelated and constantly evolving. Perhaps the use of unidirectional arrows here makes art history less interesting than it really is?

Second, and more fundamentally, the single direction of the arrow can imply an active 'giver' and a passive 'receiver', which can be both problematic and inaccurate as Tufte points out, referring to Baxandall's "Patterns of Intention" (1985). In the case of Barr's map, this means suggesting that artists simply passively receive influence rather than actively learn from other schools of thought. Should we read an arrow as "Purism influenced Modern Architecture" or as "**modern architects extracted inspiration from Purism**"? This second objection cannot be solved by providing additional arrows in the reverse direction; it concerns the active versus passive interpretation of unidirectional arrows.

In international development contexts, theories of change often depict beneficiaries of programmes as passive agents. With unidirectional arrows, organisations often imply that beneficiaries receive inputs such as new skills whereas, as we all know, the success of this kind of activity depends equally on people attending and engaging with training. By acknowledging people's agency we can ensure a better understanding of programmes and respect for all stakeholders involved.

How do your maps frame the relationships between stakeholders?

(The causal diagram for this post is:

Us —>email from Gary Goertz —>Tufte's book —> Baxandall's argument,
or perhaps:

Us <— email from Gary Goertz <— Tufte's book <— Baxandall's argument.)

--chapter:end:77500-directional-arrows.rmd--

(PART*) Technical reference {-}

--chapter:end:80000-PART-tech-reference.rmd--

Technical reference{#xlevel7}

This section contains technical reference for the Causal Map app.

--chapter:end:80001-level8.rmd--

Conventions{#xconventions}

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

--chapter:end:80012-conventions.rmd--

Screenshotting your maps

When creating a report using the Casual Map app you may find it useful to screenshot your map or other parts of the app.

To screenshot using Windows tap PrtSc once to copy an image of your entire screen to the clipboard. You can then hit Ctrl and v to paste it into your document. To screenshot a snippet of your screen press Shift, Windows Key and s to open a small toolbar at the top of the screen. You may then choose between capturing entire screen, a window or a selected portion.



To screenshot on a Mac press the Command, Shift and 3 together will take a screenshot of your entire screen. Alternately, hold Command, Shift and 5, which will open the screenshot tool menu. You may then choose to capture the entire screen, a window or a selected portion.



!--chapter:end:80017-screenshots.rmd--

FAQs

Logging in

- ▶ Can't log in, getting an error about "failed to execute 'setitem' on 'storage'"

Immediately after you get the error, open the Chrome console by pressing ctrl shift J in Chrome,

then at the blinking cursor, paste this:

```
localStorage.clear();
```

and refresh.

If that doesn't work, please contact support at support@causalmap.app

Setting Up

► Is there a security audit for Causal Map?

Please see our T&Cs at <https://causalmap.app/terms-and-conditions-for-the-causal-map-app/>

► How do I delete my account?

If you have created any files which you want to delete, you can delete them using the green file sharing / file management button at the top left of the app.

If you have a recurring subscription, you can cancel it from the Account tab in the app.

► How do I import and export maps from Causal Map?

You can download your map from Causal Map as an Excel file, tweak it and then upload it Causal Map.

► I have done something wrong, how do I know which version of my mapfile to restore?

The Versions dropdown shows you the filesize, so you can see where the file got larger or smaller. Also, have a look in the Updates tab. You can't actually restore the file from here but you can maybe guess which one you want by looking at the number of factors, the last coded statement etc.

Updates about your files

Updates about your files					
Show... <input type="radio"/> Most recent status for each file <input checked="" type="radio"/> All updates for current file <input type="radio"/> All updates for all files					
mapfile	user	server	type	content	time_stamp
All	All	All	All	All	All
cmi2-socio	steve@pogol.net	CM2test	mapfile loaded for editing	editing=Yes, ncol_links=138, last_coded_statement=1037, nfactors=601, nlinks=1331, nstatements=2261	2021-11-19 10:54:53
cmi2-socio	irena.rozajac@live.com	CM2test	session ended	editing=Yes, ncol_links=138, last_coded_statement=1037, nfactors=601, nlinks=1331, nstatements=2261	2021-11-19 10:48:48
cmi2-socio	irena.rozajac@live.com	CM2test	mapfile updated	editing=Yes, ncol_links=138, last_coded_statement=1037, nfactors=601, nlinks=1331, nstatements=2261	2021-11-19 10:30:48
cmi2-socio	irena.rozajac@live.com	CM2test	mapfile updated	editing=Yes, ncol_links=138, last_coded_statement=1037, nfactors=601, nlinks=1330, nstatements=2261	2021-11-19 10:30:23
cmi2-socio	irena.rozajac@live.com	CM2test	mapfile updated	editing=Yes, ncol_links=138, last_coded_statement=1036, nfactors=601, nlinks=1329, nstatements=2261	2021-11-19 10:29:29

- I made changes to file access with the Files panel to allow my colleague to see my file, but she still can't see it.

Tell her to refresh her browser and your changes should appear.

- Can I upload data with special characters like emojis, or characters from non-English languages?

Yes, the app will upload them.

Uploading and tweaking data

- When I upload new or updated data into the app, does it matter what the Excel / spreadsheet file is called?

No. The only thing that matters is what file you are uploading into at Causal Map: the currently active file.

Coding

- When I highlight a single word or very short phrase, why is the same word highlighted but in a different part of the statement?

If you highlight a very short piece of text which happens to appear (exactly the same) more than once in the statement, it is always the first occurrence of this piece of text which is actually highlighted.

- I have some factors with no links at all, how can I get rid of them?

Sometimes you might want to keep these factors because they are part of a fixed codebook. But if you don't want them, go to the Data tab and click the link to delete them.

- I have over 500 factors and when I try to type a new factor label, the app gets slow.

Causal Map is not really designed for maps with more than about 500 different factors. The app will work, but yes it will get gradually slower.

- The quote doesn't match up with the statement number?

If you are looking at statement 99, looking also at factor editor, and click a quote to edit that link, which is actually say link 66 from statement 11, then the text from statement 11 will load up and link 66 will load up but the statement navigator above still says 99. That isn't wrong, it is just confusing! The filter you originally set to look at statement 99 hasn't changed.

► When should I use a factor hashtag, and when should I use a link hashtag?

Use factor hashtags, including quickfields, to add extra information about the causal factors, especially information which cuts across any existing factor hierarchy you are using, for example themes like environment. Use link hashtags specifically if it relates to the claim rather than the factors e.g. respondent is suspicious / uncertain / answering a different question / respondent doesn't like this connection, etc.

► Can I autofill influence factor label if its the same as the last consequence factor?

Yes, that is called chaining links! You can chain links by clicking the chain icon, after you have saved your previous link. If you would like more guidance check out this video.

► I'm all zoomed out, how do I get a normal zoom level in my browser?

If you have Windows press ctrl and o. If you have a Mac press option, command and o.

► Can I create more than one link at a time?

Yes. You can create multiple links at a time just by putting more than one factor in either or both boxes. If you want more information check out this short video.

► How do I link two existing chains?

To link a chain a --> b --> c to a chain d --> e, highlight the part of the statement which provides the evidence and then code a link from c to d using the dropdowns as usual.

► If a respondent makes the same causal claim twice, should I code it twice?

If they are just repeating the same point, don't bother. But if they are bringing different evidence for the same link, code it twice.

► Should I put spaces after `;` in hierarchical coding? What about after `~` when coding opposites?

The convention is this: blah;blah

blah;~blah

etc, i.e. we put a space after the ; but no space after the ~

► Can I trace paths from one factor to "everything"?

Yes! Just leave the target box empty.

► What do the functions `unique` and `literal` mean when conditionally formatting links?

You use unique and literal when you don't want to do any counting, you want to print or otherwise use the actual values. So if you have a bundle with sources a, a, b and c, and you want to print them as labels, then unique: source_id will print a; b; c whereas literal: source_id will print literally a; a; b; c

Known Bugs

► When I highlight text which overlaps with a previously highlighted section no new highlight appears.

Unfortunately there is no easy fix for this bug but it is just cosmetic. You can check you have highlighted the right section by checking the text box below the factor label boxes contains the text you want to highlight. Once you have saved the link the highlight will appear on the statement.

► Depending on the size of the screen, when I go to press one of the tabs like the one with the question icon, the icon then changes position to go to the next level - which I then can't reach. Or the tabs flicker and slide around.

Yes that's annoying. Just zoom out your browser e.g. by pressing ctrl -

► I'm getting an "Ajax error" when viewing tables.

Have a look to see if the URL of the app has a lot of parts like this:

https://causalmap.shinyapps.io/CM2test/w_741fed8d/w_q4eed8d/_w_346fed8d/. If it does, just visit the bare URL: (https://causalmap.shinyapps.io/CausalMap2 or https://causalmap.shinyapps.io/CM2test) and start again from there.

► My link label is appearing slightly away from the link

Unfortunately when a map has a lot of links link labels may appear slightly removed from the link. Currently there is no way to change this with formatting, only by simplifying your map.

QuIP questions

► How do I upload my QuIP project into Causal Map?

Follow the instructions for hybrid files. You must create an excel file with one worksheet and four columns. The column headers must be in this order; text, source_id, question_id and question_text. Then upload this to an empty file.

Downloads

► Where can I find example files to use as templates?

They can be found and downloaded here.

!--chapter:end:80021-faq.rmd--



Glossary{#glossary}

Generic terminology for causal mapping

Term	Definition
Active filter bar	This is the first white bar in the filters panel, if your filter is there it is active.
All filter bar	This is the second white bar in the filters panel, if your filter is there it is not active, or it is active but can also be used again.
Bundle	You can bundle links or factors. Bundling factors merges factors containing a chosen word or phrase. You can also bundle links which are between the same influence and consequence factors, which are usually mentioned in different statements by different sources. A thicker line or a label can be used to show the number of links in the bundle.

Term	Definition
Causal links	Causal links are the links between the consequence and influence factors. Each link represents a causal claim.
Combining opposites	When we combine any pairs of factors which are opposites. This means merging two factors which are identical apart from one starts with a ' \sim ', to indicate a negative. On print view the link will show green and pink to indicate that both the positive and negative of the factor is true.
Contrary factor	With a pair of opposite factors like <code>Progress</code> and <code>\simProgress</code> , we call <code>Progress</code> the <i>ordinary</i> factor and we call <code>\simProgress</code> the <i>contrary</i> factor. (We prefer this terminology to, say, "negative factor" as it has fewer unwanted connotations: contrary factors don't have to be bad.)
Consequence factor	A causal factor at the end of a link, that has been influenced by another factor.
Field	Additional data associated with a statement such as respondent ID number, respondent age or question number.
hashtags aka themes	A recurring subject found within the factor labels i.e. farming or health. Often we use symbols like # to make sure that the hashtags are uniquely searchable.
Hierarchical coding	Nested factors (see below) put factor labels in a hierarchy from more general to specific. See more here.
Hybrid format	(Mostly used in QuIP/BSDR coding.) When uploading data into Causal Map, a data table in hybrid format has some rows which are statements and some which are metadata.
Influence factor	A causal factor at the beginning of the link, that affects another factor.
Mapfile, File	When working in Causal Map, all your original data (the statements, source and question data) as well as the coding you have done (factors and links) are stored together in one file which we call a mapfile , or just file .
Memos	You can make short notes or memos on each statement, as well as about each causal link, each factor, each source, each question and the project globally.

Term	Definition
Nested factors	Factor labels can be nested meaning there are higher level factors i.e <u>New intervention</u> ; <u>midwife training</u> ; <u>hand washing instructions</u> . See <u>more here</u> .
Original links	Often, there is more than one link between any pair of factors. We can call these the original links especially when we bundle them together for display purposes so they visually look like one. In this case we can say that the bundled link on the map actually consists of several original links.
Shortlink	A weblink (URL) to a causal map which has been shortened, usually to make sharing it easier. You can easily make <u>shortlinks</u> using the dashboard.
Statement	Refers to each individual piece of text to be coded in Causal Map, usually this is a sentence, a paragraph or a few paragraphs. Each statement has an ID; a number like 1, 2, 3 etc.
Source	Sources are where your statements come from, this could be respondents you have interviewed or documents you have collected.
Weblink	A complete URL like <u>https://causalmap.shinyapps.io/CausalMap2/</u>

!--chapter:end:80030-glossary.rmd--

Filter reference

A full guide to the filters can be found here: github.

!--chapter:end:80050-function-ref.rmd--

How Causal Map is hosted

The **Production version** of CM2 is at <https://causalmap.shinyapps.io/CausalMap2/>

There is also a **development version** at <https://causalmap.shinyapps.io/CM2test/> This version receives updates much more frequently. You may sometimes get a warning to say that the server is about to restart.

At the moment, both versions are hosted at shinyapps.io.

<!-- Video embedding:

```
knitr:::include_url("https://player.vimeo.com/video/618270513")
```

-->

!--chapter:end:80100-hosting.rmd--

Reports tab {#xreports}

Sometimes all the filters which are available in Causal Map can be bewildering and perhaps you just want to get your report finished already. The main purpose of this new tab is to provide an overview of a mapfile which makes sense and is immediately useful without any complicated searching or setting of filters. It provides for example sample description tables and some of the most useful tables which you can just copy and paste as needed. It's work in progress, so let us know what you want.

Report

[Overview](#) [Sample](#) [Hierarchy](#) [Rankings](#) [Opposites](#) [Differences - sample](#) [Differences - questions](#) [Hashtags](#)

Level 1 map

We do not plan to provide much customisation of these outputs. Possibly the different outputs here could also provide links to the existing, equivalent filters in the rest of the app, so the user can tweak them there instead.

Export to Word and PDF files might be possible in the future.

Here we describe the different sub-tabs. It's work in progress...

Sample

List total numbers and %s for all key groups, i.e. columns in the sources table which begin with either # or x.

Hierarchy

If the mapfile includes hierarchical coding, show maps zoomed out to level 1 and level 2.

Opposites

If the mapfile includes opposites coding, show a maps in which opposites are combined.

If the mapfile includes hierarchical coding, show a map in which are opposites are combined, then zoomed out to level 1.

Rankings

If the sources table (or statements table) includes any columns which end with the word "rank" or "rank#", then a QuIP-like table of ranked organisations is displayed in one of the sub-tabs.

Differences - questions

Shows which links were preferentially mentioned in answer to different questions.

Differences - sample

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

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

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

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

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

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

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

At the moment, the tests for this are chi-squared tests which would not give special treatment groups which are actually ordinal e.g. low income, medium income, high income: the chi-squared test is weaker than it should be.

Differences - sample

Each row .

Differences - question

For the unfiltered map, level 1 map and combined opposites maps, every link is displayed which is answered significantly differently between qu

Hashtags

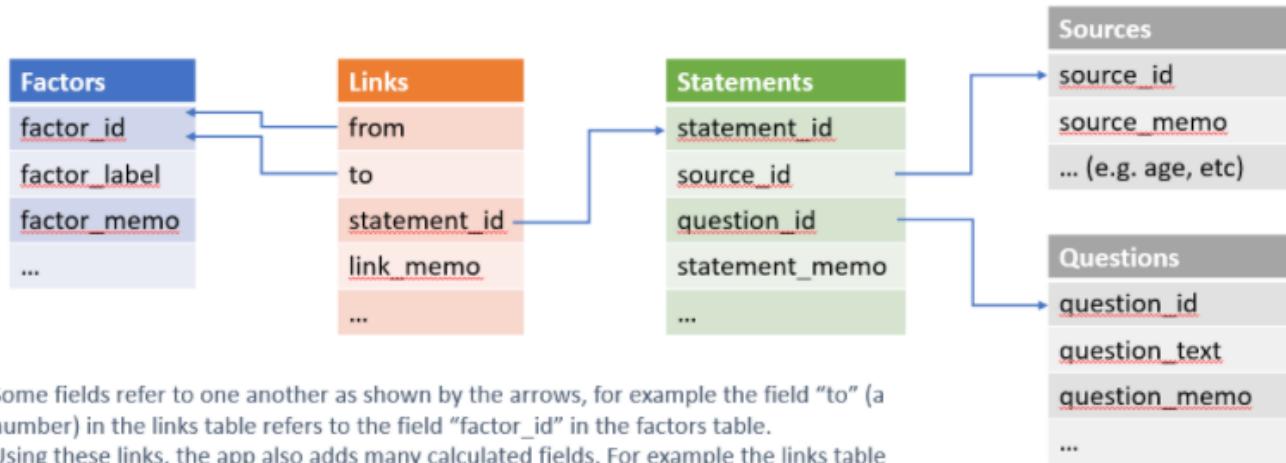
Custom

This section shows maps from the most recent 10 filters for this map which have been saved as shortlinks.

!--chapter:end:80200-reports.rmd--

Core tables in Causal Map{#core-tables}

How the tables in Causal Map are joined together



Some fields refer to one another as shown by the arrows, for example the field "to" (a number) in the links table refers to the field "factor_id" in the factors table.

Using these links, the app also adds many calculated fields. For example the links table also gets the field "from_label" which is the label of the factor to which the "from" field points. And the links table has many fields from the sources and questions tables, e.g. age of respondent, which it gets via the statements table.

Finally, the app also adds some calculated metrics like frequency and betweenness.

Users may also import their own fields, which can also be used in analyses and

visualisation, for example the factors table might have a field called "domain".

This schema is followed by Causal Map 2 and to a lesser extent by Causal Map 1.



Each table has the following types of field:

- ID field
- Standard fields e.g. statement_memo. Users enter data into these fields.
- Custom fields which users can add, as long as their names are not reserved field names
- System fields like frequency which are added by the system. These are not included if you download a file and are ignored when uploading a file.

These are the standard fields. You can add other custom fields too if you wish, and they will be available to the app.

The main five tables and the most important fields / columns

Factors

- label. Has to be unique – you can't have two factors with the same label
- factor_id. This is just a number like 1, 2 etc. This number doesn't change even if the label does.

Links

- from. Contains numbers which match the factor_ids of the factors.
- to. Contains numbers which match the factor_ids of the factors.
- statement_id. Contains numbers which match the statement_ids of the statements.
- link_memo. For making notes about a link as you code.
- hashtags. Re-usable tags which are usually common across several links, for example #hypothetical or #certain or #check.

Statements

- statement_id. A running number from 1 upwards.
- source_id. A code which matches the field source_id in the sources table.
- question_id. A code which matches the field question_id in the questions table.
- statement_memo. For making notes about a statement as you code.

Sources

- source_id. A code which matches the field source_id in the statements table.
- source_memo. For making notes about a source as you code.
- Any other fields which you have uploaded.

Questions

- question_id. A code which matches the field question_id in the statements table.
- question_memo. For making notes about a question as you code.

- Any other fields which you have uploaded.

Special case: Closed question blocks

This QuIP-oriented special feature is only provided for one single block of questions which each have the same set of answers, or could be recoded to have the same set of answers, e.g. twenty questions which are all answered or could be recoded as "always/sometimes/rarely/never" or whatever.

--chapter:end:80300-core-tables.rmd--

Medals

Your medals!



This section gives a whole load of medals to reflect all the work you've been doing and encourage you to try out new things.

##

--chapter:end:80400-medals.rmd--

Using the Advanced Editor to create and edit filters{#advanced-editor}



Show filters

Advanced editor

```
select links top=15
bundle factors value=
color factors field=frequency lo=#FCFDBF hi=#5F187F mid
  =#D3436E
```

[Clear all](#) | [Clear line under cursor](#)

The commands are applied one by one in sequence, in a pipeline of commands, such that after each command, such as each command starts with the map defined by the previous line and produces a new one.

The order of the lines may matter! You can move lines up and down by pressing [Alt+UpArrow](#) or [Alt+DownArrow](#).

This text can also be freely edited and typed independently of using the buttons. You don't need to fear doing anything wrong because there's nothing to break and commands which don't fit the intended syntax patterns will just be ignored.

When you add a filter using the buttons, the filter replaces any similar filters already in the text window. Occasionally, you might want two different versions of the same filter. For example, you might want to filter using `find statements field=text value=women operator=contains` and also `find statements field=#SourceID value=mkf-8 operator=contains`. In this case, uncheck [Replace existing similar filters?](#) before applying the filter.

The operators work as you'd expect, so you can for example have a filter like `find statements field=text value=women OR girls operator=notcontains` to find statements whose text does not contain either 'women' or 'girls'.

To get an AND rather than an OR, just put one filter after another. To find statements containing women AND girls:

```
find statements field=text value=women operator=contains
```

```
find statements field=text value=girls operator=contains
```

Read about them all at: [CausalMapFunctions](#).

Keyboard shortcuts {#xkeyboard}

General

There are only a couple of shortcuts so far:

- Esc should get you out of any modal dialogue
- Alt+q opens the Gallery

Factor Editor and Advanced Editor

For more information about working with the Ace editor,
visit <https://ace.c9.io/#nav=about>.

Ace provides the following keyboard shortcuts to streamline the coding process.

Line operations

Action	Windows/Linux	Mac
Remove line	Ctrl-D	Command-D
Copy lines down	Alt-Shift-Down	Command-Option-Down
Copy lines up	Alt-Shift-Up	Command-Option-Up
Move lines down	Alt-Down	Option-Down
Move lines up	Alt-Up	Option-Up
Remove to line end	Alt-Delete	Ctrl-K
Remove to linestart	Alt-Backspace	Command-Backspace

Action	Windows/Linux	Mac
Remove word left	Ctrl-Backspace	Option-Backspace, Ctrl-Option-Backspace
Remove word right	Ctrl-Delete	Option-Delete
Split line	---	Ctrl-O

Selection

Action	Windows/Linux	Mac
Select all	Ctrl-A	Command-A
Select left	Shift-Left	Shift-Left
Select right	Shift-Right	Shift-Right
Select word left	Ctrl-Shift-Left	Option-Shift-Left
Select word right	Ctrl-Shift-Right	Option-Shift-Right
Select line start	Shift-Home	Shift-Home
Select line end	Shift-End	Shift-End
Select to line end	Alt-Shift-Right	Command-Shift-Right
Select to line start	Alt-Shift-Left	Command-Shift-Left
Select up	Shift-Up	Shift-Up
Select down	Shift-Down	Shift-Down
Select page up	Shift-PageUp	Shift-PageUp
Select page down	Shift-PageDown	Shift-PageDown
Select to start	Ctrl-Shift-Home	Command-Shift-Up
Select to end	Ctrl-Shift-End	Command-Shift-Down

Action	Windows/Linux	Mac
Duplicate selection	Ctrl-Shift-D	Command-Shift-D
Select to matching bracket	Ctrl-Shift-P	---

Multicursor

Action	Windows/Linux	Mac
Add multi-cursor above	Ctrl-Alt-Up	Ctrl-Option-Up
Add multi-cursor below	Ctrl-Alt-Down	Ctrl-Option-Down
Add next occurrence to multi-selection	Ctrl-Alt-Right	Ctrl-Option-Right
Add previous occurrence to multi-selection	Ctrl-Alt-Left	Ctrl-Option-Left
Move multicursor from current line to the line above	Ctrl-Alt-Shift-Up	Ctrl-Option-Shift-Up
Move multicursor from current line to the line below	Ctrl-Alt-Shift-Down	Ctrl-Option-Shift-Down
Remove current occurrence from multi-selection and move to next	Ctrl-Alt-Shift-Right	Ctrl-Option-Shift-Right
Remove current occurrence from multi-selection and move to previous	Ctrl-Alt-Shift-Left	Ctrl-Option-Shift-Left
Select all from multi-selection	Ctrl-Shift-L	Ctrl-Shift-L

Go to

Action	Windows/Linux	Mac
Go to left	Left	Left, Ctrl-B

Action	Windows/Linux	Mac
Go to right	Right	Right, Ctrl-F
Go to word left	Ctrl-Left	Option-Left
Go to word right	Ctrl-Right	Option-Right
Go line up	Up	Up, Ctrl-P
Go line down	Down	Down, Ctrl-N
Go to line start	Alt-Left, Home	Command-Left, Home, Ctrl-A
Go to line end	Alt-Right, End	Command-Right, End, Ctrl-E
Go to page up	PageUp	Option-PageUp
Go to page down	PageDown	Option-PageDown, Ctrl-V
Go to start	Ctrl-Home	Command-Home, Command-Up
Go to end	Ctrl-End	Command-End, Command-Down
Go to line	Ctrl-L	Command-L
Scroll line down	Ctrl-Down	Command-Down
Scroll line up	Ctrl-Up	---
Go to matching bracket	Ctrl-P	---
Scroll page down	---	Option-PageDown
Scroll page up	---	Option-PageUp

Find/Replace

Action	Windows/Linux	Mac
Find	Ctrl-F	Command-F

Action	Windows/Linux	Mac
Replace	Ctrl-H	Command-Option-F
Find next	Ctrl-K	Command-G
Find previous	Ctrl-Shift-K	Command-Shift-G

Folding

Action	Windows/Linux	Mac
Fold selection	Alt-L, Ctrl-F1	Command-Option-L, Command-F1
Unfold	Alt-Shift-L, Ctrl-Shift-F1	Command-Option-Shift-L, Command-Shift-F1
Fold all	Alt-o	Command-Option-o
Unfold all	Alt-Shift-o	Command-Option-Shift-o

Other

Action	Windows/Linux	Mac
Indent	Tab	Tab
Outdent	Shift-Tab	Shift-Tab
Undo	Ctrl-Z	Command-Z
Redo	Ctrl-Shift-Z, Ctrl-Y	Command-Shift-Z, Command-Y
Show the settings menu	Ctrl-,	Command-,
Toggle comment	Ctrl-/	Command-/
Transpose letters	Ctrl-T	Ctrl-T
Enter full screen	Ctrl-Enter	Command-Enter

Action	Windows/Linux	Mac
Change to lower case	Ctrl-Shift-U	Ctrl-Shift-U
Change to upper case	Ctrl-U	Ctrl-U
Overwrite	Insert	Insert
Macros replay	Ctrl-Shift-E	Command-Shift-E
Macros recording	Ctrl-Alt-E	---
Delete	Delete	---
Center selection	---	Ctrl-L

© Copyright Trivantis 2019

Thanks to Trivantis for the list.

!--chapter:end:82000-keyboard-ace.rmd--

Causal Map functions

Causal Map is powered by our free / opensource R package CausalMapFunctions, where you can also find out about the commands to use in this Viewer.

See the Vignettes at [github](#).

!--chapter:end:83000-cm-functions.rmd--

Video list{#xvideo-list}

These videos are all included at different points in the Guide. If you love videos, here they are again in one place. Feel free to binge. Tip: if you play a video, you can press 'Picture in picture' and then switch to the app, and you can follow (and even resize) the video.

```
knitr:::include_url("https://player.vimeo.com/video/689223728")
```

```
knitr:::include_url("https://www.youtube.com/embed/S2acbJJ9t7E")
```

```
knitr:::include_url("https://www.youtube.com/embed/6DBzadDMXew")
```

```
knitr:::include_url("https://player.vimeo.com/video/674369121")
```

```
knitr:::include_url("https://player.vimeo.com/video/674383670")
```

```
knitr:::include_url("https://player.vimeo.com/video/641927229")
```

```
knitr:::include_url("https://player.vimeo.com/video/604099226")
```

```
knitr:::include_url("https://player.vimeo.com/video/580212681")
```

```
knitr:::include_url("https://player.vimeo.com/video/671894620")
```

```
knitr:::include_url("https://player.vimeo.com/video/637121851")
```

```
knitr:::include_url("https://player.vimeo.com/video/604138709")
```

```
knitr:::include_url("https://player.vimeo.com/video/664879474")
```

```
knitr:::include_url("https://player.vimeo.com/video/683770170")
```

```
knitr:::include_url("https://player.vimeo.com/video/636997200")
```

```
knitr:::include_url("https://player.vimeo.com/video/588851236")
```

```
knitr:::include_url("https://player.vimeo.com/video/588881701")
```

```
knitr:::include_url("https://player.vimeo.com/video/618270513")
```

```
knitr:::include_url("https://player.vimeo.com/video/651173445")
```

```
knitr:::include_url("https://player.vimeo.com/video/599620015")
```

--chapter:end:83500-video_list.rmd--

Causal mapping with Excel{#xexcel}

Now we are going to back to the basics of causal mapping. You don't need Causal Map App, or any other dedicated application, to do causal mapping. You can use Excel or any spreadsheet program.

Causal maps store information about links of causal influence between two or more factors alongside the evidence for each link. We call the chunks of text, for example from interview transcripts, which potentially contain claims about causal connections, "statements".

Here is an example, it contains example- statements from a newspaper article and an interview concerning flooding and coastal erosion. These links that you see here, otherwise known as an edge-list, the starting point for any causal map. You can clearly see the factors in the first column affect the factors in the second. For example, Column A contains influence factors such as 'high rainfall' which influence the consequence factor in the next column, like 'flooding'. It is these links that are then displayed in a causal map.

You will notice, in this example, that the columns in the links tab are dropdown lists when you click on the triangles in each cell. These dropdown lists are fed from the labels listed in the factors tab. So you can change the options available by editing or adding labels in the factors tab.

Optionally, on the statements tab the statement_id column works in a similar way and is fed by the sources tab. The question_id is fed by the questions tab. So the sheets are programmed to link together.

You can download the example and play around with it or use it as a template for your own map.

Also once in excel format causal map becomes compatible with other data visualisation tools like Gephi, and if you want to know more about that check out the video below video on using the causal map app alongside other data visualisation tools.

```
knitr:::include_url("https://player.vimeo.com/video/599620015")
```

You can download this file [here](#).

```
--chapter:end:83600-excel-coding.rmd--
```

Using StorySurvey: for survey editor

This documentation is work in progress.

You have to log in with one of the accounts which is hardcoded as admin.

The SQL table `ss2embeddings` has columns for `survey("project")`, `original phrase ("old")` and then hundreds of columns containing the embeddings for each original phrase.

The embedding for each new phrase is added to the table `ss2embeddings` when the respondent enters it.

This table can be (re-)constructed with the function `generate_embeddings()` (there is a button in the Cluster panel). This should not be necessary as embeddings are added as above.

The SQL table `ss2recodes` has xxxxxxx columns for `survey("project")`, `original phrase ("old")` and then hundreds of columns containing the embeddings for each original phrase.

Explorer

Here you can click to explore the factor structure

Map

A view of the complete map with all links with frequency of 2 or above

Project recodes

All the current recodes on the current project, including factors which have been hidden.

Project settings

TBC

Project answers

TBC

Project questions

If you change the text of the questions and there are already answers recorded, these answers will become detached from the original questions.

Each question can also have a short version which is used instead of the long question in the maps.

Project recodes ace: The recodes advanced editor

The tab you need is Project recodes ace which is the first tab.

Basic idea

The idea is that you recode daughters like Y into mothers like X by adding X; in front of Y.

This is very similar to the idea from Causal Map that increased income; cash from sale of pig is an example of increased income, but with StorySurvey the more detailed part of the label is already created by the respondent, and the analyst's job (or the NLP's job) is only to classify cash from sale of pig as an example of increased income.

Also, in Causal Map, the different lower-level labels should be meaningfully different variations of a central idea, whereas in StorySurvey that is only half the job, the other half is simply about merging different formulations of exactly the same idea.

Unlike in Causal Map, in StorySurvey the purpose of the recodes is not mainly to be able to sometimes display outputs with all the different variations. In StorySurvey, the daughters in a recode set are more like the different quotes in Causal Map -- they serve to give colour to, and illustrate, each individual factor.

!--chapter:end:84000-storysurvey.rmd--

Features of Causal Map{#x}

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
- 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

Additional new features in Causal Map 2

- Anyone can use it, and it is free for small projects
- Conditional formatting makes your findings more visible
- Bundle factors together to make your map neater and easier to read
- Bundle links by attributes such as stakeholder, district, question and many more
- Infinite versioning means you can restore any change made to a file from any timepoint. This isn't just backups, it is every single change
- It is easier to upload and manage your data
- You can upload a standard-format Excel file into CM2, meaning you can download a file, edit it, and upload it again (this is called "round-tripping"). You can also include any additional fields or columns you want and they will be uploaded and available for searching, conditional formatting etc
- Build a map by uploading any combination of tables or just some of them - factors, links, sources, questions and statements, in any order, at any time
- All outputs transparently recreatable from the filters
- Share or save links to maps and filters
- Clickable legends and ordinary-language explanations of filters
- Merging maps means you can compare multiple 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
- You can restore, lock, delete, archive and share any of your files
- Log in to the same or different maps from the same or different accounts, including from multiple browser tabs
- Collaborate on a file asynchronously
- View, filter and sort all the tables - factors, links, sources, questions and statements - making up a map. These tables provide all the existing tables functionality plus more.
- Delete factors and edit or 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
 - New colour scales are colour-blind friendly and make sense when printed in black and white
 - 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.
-

!--chapter:end:85000-features.rmd--

Recent changes{#changelog}

Version 0.85, 1/12/2022

Tables

The biggest change is the full release of the new tables functionality. Now all your tables (factors, links, mentions, sources, statements, questions and metrics) use the new features. There are lots of tables you can create, from a table simply showing the total links to and from each factor to a summary of each respondent's response to a closed question. The possibilities are large – and increase with the amount of additional data your project has.

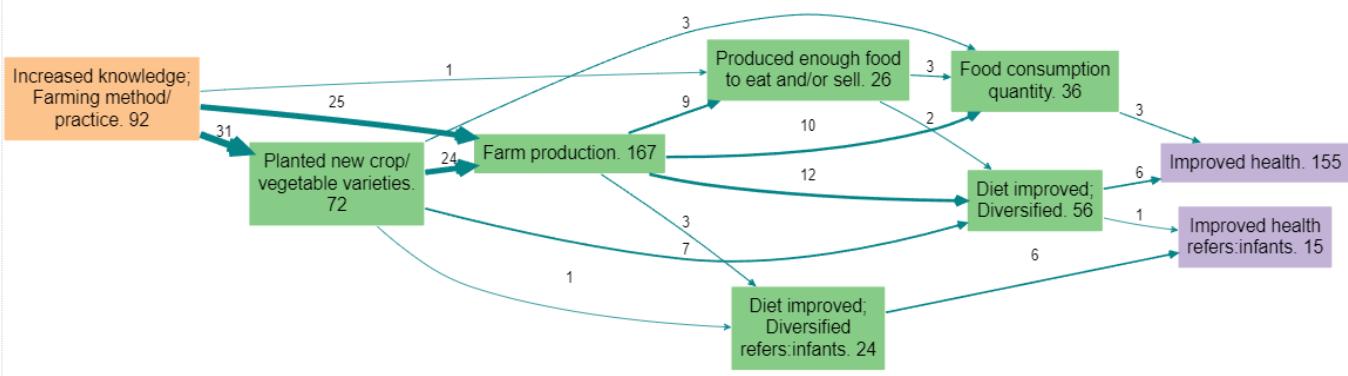
- Complete set of table presets in the Dashboard
- Filter, search and sort
- Pivot tables
- Lots more

Removing unwanted links

Suppose you have a map with links B --> D, C --> D and also B --> C. If you run this filter:

trace paths length=1 to=D

the app will show you also the path from B to C, because both B and C are one link away from D, and there is a path from B to C. But that might have surprised you, as the path from B to C does not get you any closer to D.



So you might prefer this filter:

`trace_paths length=1 to=D remove_links=T` (which is equivalent to clicking this box in the dialog:)

Trace paths

This filter shows all the paths (up to a length which you can specify) from one factor or set

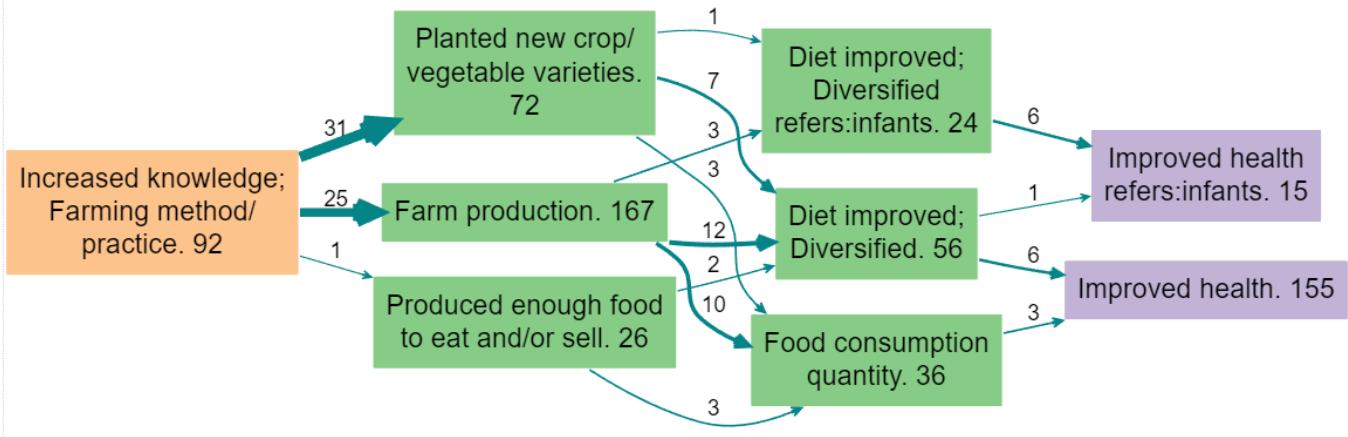
Keep this panel open?

Trace how many steps?

0 1 10

Remove in-between links?

... which will remove any such links.



In other words, `trace paths length=L from=X to=Y` says "show me all the factors which are on paths from X to Y no longer than L, and all the links between such factors" whereas `trace paths length=L from=X to=Y remove_links=F` says "show me all the paths from X to Y no longer than L".

[Read more here.](#)

Quickfields for links

Quickfields are now available for links too. You use the same syntax in the "hashtags" box as for quickfields for factors. This might seem like a small tweak but it means you can now systematically include information about the context of each link. Whereas previously you could mark that a particular link refers only to the context of before Covid by marking it with a hashtag like `#before-covid`, now you can use systematic sets of hashtags like `#covid:before` and `#covid:after`, which automatically then adds a column to your links table called `covid` which can have values like `before`, `after`, and `anything`.

Version 0.81, 9/8/2022

Importing a simple, wide-format file

There are many ways to upload data into Causal Map. We have just introduced perhaps the simplest way to upload your data: wide format. This is the kind of data you might often get if you have been using Google Forms.

In this format, you only need an Excel file with one tab; each row is all the data from one source. So, many of the columns will probably be additional data like gender, etc.

Adding additional cases with wide-format files

You can start coding your data even if you do not yet have all the cases/sources. When you get more data, simply add them to the bottom of the original file and upload again. You don't need to do "roundtripping". All your sources and statements will be overwritten but your coding (links and statements) will be untouched. The questions table will also be untouched. If you want to add information about your questions, for example adding question text as well as question id, you can use roundtripping.

Bundling links

When you specify a field to bundle by, it is enough to only specify only the first part of the field name. Remember, this name matches any field associated with sources, links or statements.

So #2 will match #2-village-name. We recommend marking your top few fields for analysis like gender etc., or whatever else interests you the most, not just with a leading # (this anyway marks out the field as important) but then with a number, like #2. This makes it easy to create reproducible filters which you can conveniently share across different files.

Improved percentage labels when calculating surprise

If you use percentages for labels, the label now shows the raw number, the baseline, and the percentage.

Version 0.80, 10/6/2022

Smart zoom

If you are using hierarchical aka nested coding, you can use ordinary zooming to "zoom out" your map to give a simple, more general overview. But sometimes this can mean oversimplifying your map: there are lower-level factors which are so important and popular that it is unhelpful or even misleading to roll them up into a parent factor. Smart zooming solves that problem.

Collaboration

One feature we've been asked for most on Causal Map is live collaborative coding: if only two people could edit the same file at the same time! Well, we've now made that possible.

The Dashboard

The Dashboard replaces the Gallery, which many people found hard to use. It still packs plenty of functionality! Read to find out more.

The screenshot shows the Causal Map Dashboard. At the top, there's a yellow header bar with a gear icon and a dropdown menu. Below it, a section titled "Your medals!" displays various achievement icons and counts: 1360, 82, 5, 53, 3, 10, 3279, and 35. There are buttons for "Load example", "Load uncoded example", and "Help".

The main area is titled "Explore views of this file" and lists a series of items, each with a small icon, a title, and a count:

- Untitled 1170
- A3 comm 1048
- Changes to cocoa yield 1096
- Deforestation 1101
- Top level 1095
- Up to A3 top 18 links 1049
- 10 Overview 1226
- 12 Overview + smart zoom 1227
- 20 - Factors by Source Count 1236
- 26 - Search factors and quotes 1239
- 28 - Quotes 1240
- 30 - Factors - source counts 1241
- 32 - Factors - levels 1242
- 34 - Factors - codebook 1243

On the right side, there's a vertical sidebar with a "Report a Bug" button. At the bottom, there are sections for adding a shortlink and jumping to a specific shortlink.

Reports Tab

Sometimes all the filters which are available in Causal Map can be bewildering and perhaps you just want to get your report finished already. The main purpose of this new tab is to provide an overview of a mapfile which makes sense and is immediately useful without any complicated searching or setting of filters. It provides for example sample description tables and some of the most useful tables which you can just copy and paste as needed. It's work in progress, so let us know what you want.

Document coding score!

Coding a large file can be a daunting task. At the top right of the app, you will see some feedback on your progress in coding your file.



These show you:

- number of (top level) factors. You get a warning if this is more than 100. Ideally you'd have less than 10 or 20 top-level factors, enough so that you can remember most of them and could explain them to someone else without notes. Of course, each top-level factor might have subfactors, but again we'd recommend no more than say 10 lower-level factors within each higher-level factor. So, it doesn't really matter how many factors you have, but the number of top-level factors should be fairly small.
- ratio of links to (top level) factors. You get a warning if this is less than 3. Ideally, each of your top level factors will be mentioned lots of times (including mentions of factors nested within them, like Health behaviour; hand washing) so they will have lots of links. If your factors don't have many links each, consider consolidating them.
- percentage of statements which have been coded. You get a warning if this is less than 60%.

Copying your map as a vector image

You can now export your map as SVG using the export SVG button at bottom right.

When you export an SVG image, a shortlink is automatically created for you. The ID of the shortlink is printed in grey at the bottom left of the image, and the link is also copied into your clipboard so you can paste it straight into a report if you want. Anyone who clicks on the shortlink will be taken to the app with the same image with the same filters.

Clustering sources

This new filter groups all the sources in the file (as currently filtered) into clusters of sources which are similar in terms of the maps they tend to produce.

To use it, add a filter like this to the advanced editor:

```
cluster sources n_clusters=all
```

...

So the maps for each cluster should be maximally different from one another. This filter adds corresponding fields to the sources table (and other tables like the links table), which you can use to filter and format your maps.

Quickfields

Quickfields are a special kind of hashtag for factor labels. They are a refinement of hashtags for when you have a set of categories which is mutually exclusive, like before/after which are not attributes of your sources and which you don't want to include in your hierarchy of factors.

You can add information about one or more such dimensions using a format like this:

Crops improved / time:before

Crops improved / time:after

A field (in this case, a field called `time`) will then be created in the factors table which you can then use for example to filter the maps, to search, filter and organise the factors, links and mentions tables, and to format the maps.

Example uses:

- capturing time before/after

- capturing factors which are explicitly or implicitly attributed to a particular project
- capturing the valence of a factor e.g. when a subject explicitly says they liked or did not like the fact that they had more work
- capturing things which apply only, say, to girls or only to boys
- organising ideas within a hierarchy

Plain coding

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

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

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

New palettes for conditional formatting

The default palette for discrete fields like gender or village is nice, but maybe you want something different. With continuous fields like source count, you can specify the high, low and mid colours; now you can change the palette for discrete fields too, by adding `pal=6` or `pal=3` like this:

```
color links field = #area fun=literal pal=6
```

Palette 2 is a good choice if you don't like the paler colours in palette 1 (the default).

See also the special sections on conditional formatting with percentages and conditional formatting with surprise.

Quick tails

 [image-20211005153410088](#)

Suppose you want to create a new factor “Health behaviour; wearing a mask” and you know there is an existing higher-level factor “Health behaviour”, you can select “Health behaviour” from the list and then just add “; wearing a mask” with a leading semi-colon. These two fragments will be combined into a new factor label. This is quicker and ensures you don’t end up with different spellings of the higher-level factors.

Search for main drivers and/or outcomes

If you want to search from or to the **main** drivers and/or outcomes:

- put "main_drivers" in the first box
- and/or put "main_outcomes" in the second box.

Filtering print view of tables

Have you ever tried pressing the toggle for Print View in a table? It provides you with a text version of text in your table. This can be really handy for example for getting a list of quotes related to your current map. Now, you use the additional table filters in combination with Print View to filter the information in your (text) table. Don't forget you can also include and exclude the fields which appear in Print View, and change their order, using the Show Columns box. Only the first 100 rows are shown.

Version 0.78

The Gallery

The Gallery loads as the app loads and gives you the opportunity to jump to another file and/or filter before loading the file which was already loading.

You can use it as a dashboard from which you can check out and revisit your previous work and get ideas from others.

You can also reach this dialog by pressing the  icon in the top menu, or with the key combination Alt+q.

The Gallery contains the history of all the filters you used in the current session as well as filters from shortlinks saved by you or other users. Search to find the filters you need. For example, if you want to get some ideas for using the `combine opposites` filter, start to type "combine opposites" in the Search box on the left or (if you want to be more precise) in the search box above the filter column. This will show you examples from your own history or from when others have saved a filter in different projects.

Then you can either load just that filter into your current file or go to see how that filter works in the file where it was used.

The screenshot shows the Causal Map application interface. At the top, there's a toolbar with various icons for file operations like Open, Save, Print, and Share. Below the toolbar, the title "CAUSAL MAP" and the subtitle "MAKING CONNECTIONS" are displayed. A dropdown menu shows "example2". On the left, there's a sidebar with "Statements: One" and "Many" buttons, a "Quick tables" dropdown, and "Show filters" and "Advanced editor" buttons. Below this is an "Info" section with a table showing "source_id": 1, "question_id": 1, and "question_text": "global question". A large text area contains a quote from the app: "Welcome to the Causal Map app! Here you can identify, visualise and analyse causal connections made in a passage of text - but what does this look like in practice? Here are some examples: Recent periods of high rainfall have led to flooding. The flooding has damaged property and businesses. Funds from Organisation X have enabled us to build better flood defences, reducing the flood risk and increasing the investment into the area." Below the quote, it says: "The claims extracted from these statements can be used to build up a network of causal connections." At the bottom left, there are buttons for "Influence factor(s)" and "Consequence factor(s)". To the right, there's a "quote" input field with a list of icons above it. On the right side of the interface, there's a "Updates about your files" section. It shows a table with columns "mapfile" and "user". The table lists several entries, all associated with "example2" and "steve@pogol.net":

mapfile	user
All	All
example2	learning@dorcus.net
example2	steve@pogol.net

Continuity

Continuity: several new features to help you visualise the effects of causes and avoid the transitivity trap. The transitivity trap can make it a challenge to interpret your maps. Some people (or somebody) said that improved hygiene led to reduction in mosquito environments, and some people said reduction in mosquito environments led to improved health, but was this the same or different people? Our favourite filter is tracing threads.

Quickfields

Quickfields are a special kind of hashtag for factor labels. They are a refinement of hashtags for when you have a set of categories which is mutually exclusive. You can add information about one or more dimensions using a format like this: Crops improved / time:before or Crops improved / time:after. A field (in this case, a field called time) will then be created in the factors table. Useful for coding a wide range of concepts which cut across your existing factor coding scheme.

Other features

The Factor Editor Sidebar has added functionality

- view and edit factor memos
- when hovering over influence or consequence factors, see entire statement text, and click to edit the links

Print view for tables: If you want printable quotes for your report, there is now a table preset for that. Or go manually to the links table, click the  button to display Print View, and choose the columns you want. They will be shown in the Print View in the same order. Then just copy and paste using the "Copy" button.

Colour text: You can now conditionally colour the text of your factor labels. For example, this can be useful to draw attention to factors which are opposites.

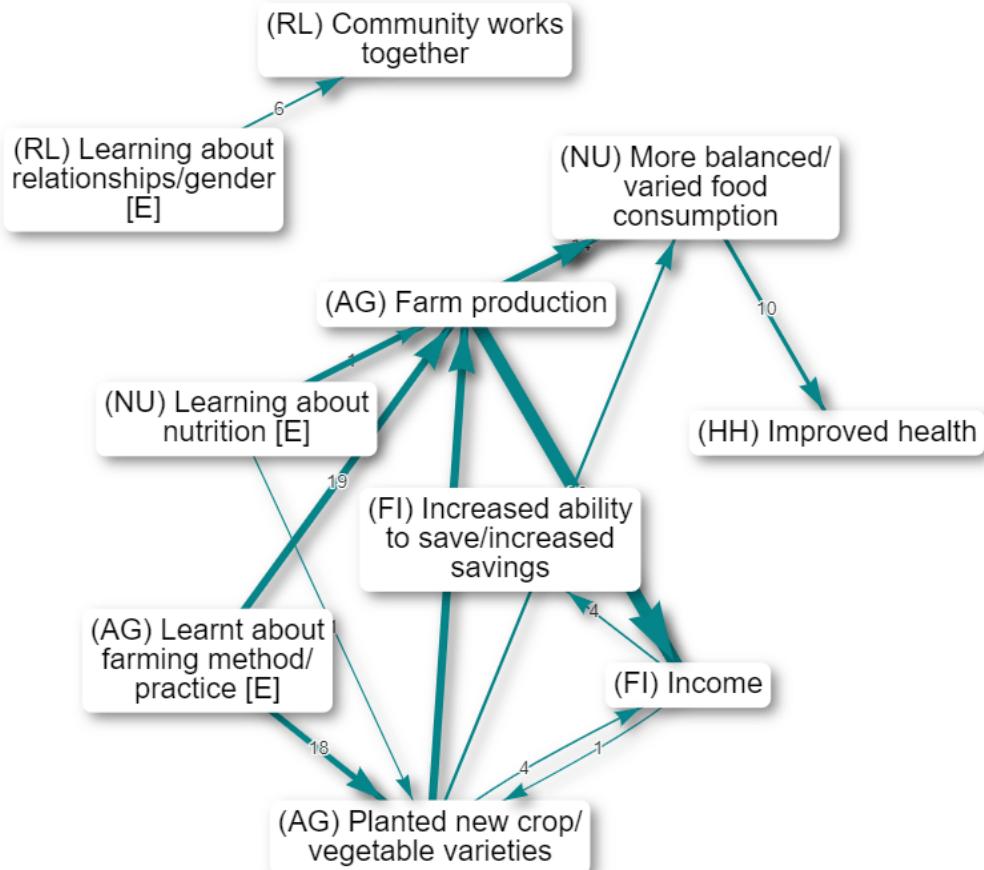
Factor labels: You can now also choose whether or not to include the name of the field in your factor labels.

source_frequency is now available for links: the number of different sources in each bundle

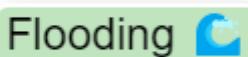
Improved application of the order of the filters with calculated fields: We have changed the point in the chain of filters when calculated fields are calculated. Previously, factor frequency meant "the frequency as it was at the start of the chain". Now, it means "the frequency of the factors as they are now at this point in the chain of filters". The same applies to nearly all the calculated fields

Initials: Shorten link labels in maps with the new initials function.

Click on factors in interactive view to focus on them, trace paths from/to them, edit their memo, or delete them.



Finding factors also works when you include 0 steps upstream and 0 steps downstream:



Reorder columns: You can reorder the columns just by dragging the column headers.

Version 0.75, 27/12/2021

Print view: exporting and zooming

You can now export your map as SVG using the export SVG button at bottom right.

You can now change the size of the Print map by using the + and - (and Reset) buttons.

You can also drag it around.

Label links

You can have multiple labels on your links (as you always could with factors), for example, count: link_id as well as unique: source_id: set clear_previous = F if you want to append a label to any existing label rather than replacing it. The app will correspondingly delete filters which are not needed.

Highlight factors without removing other factors

Maybe you want to just highlight the factors or links you find, but not remove the others. You can do this with the advanced editor, adding highlight_only=TRUE:

```
find factors value=risk operator=contains highlight_only=TRUE  
color factors field=found lo=white hi=#7FC97F
```

You can use any of the normal conditional formatting options, e.g. color factors field=found lo=blue hi=red - the hi colour will be used for factors which match the search, and the lo colour for those which do not.

This is a useful way to highlight factors which have the opposites symbol ~:

```
find factors value=~ highlight_only=TRUE  
color borders field=found
```

At the moment you can't use this to colour factor background by one criterion and colour factor borders by another (because they both use the hidden field called found).

Printout (for example, printout of quotes) option in tables

-
Do you want some preformatted quotes for your report?

Links

The screenshot shows a table interface with the following columns: link_id, from_label, to_label, statement_id, and quote. The table contains five entries:

link_id	from_label	to_label	statement_id	quote
1	High rainfall	Flooding	1	Recent periods of high rainfall have led to flooding.
2	Flooding	Damage to Property	1	The flooding has damaged property.
3	Flooding	Damage to Businesses	1	The flooding has damaged businesses.
4	Flood risk reduced	Increased investment into the area	1	Reducing the flood risk and increasing investment into the area.
8	Funds from Orgx	Flood risk reduced	1	Funds from Organisation X have reduced the flood risk.



Each table can also be shown as simple text for copying and pasting.

If you want quotes, go to the links table, click the button to display the printout, choose the columns you want. They will be shown in the Printout in the same order. Then just copy and paste.

Path tracing

If you want you can also leave either (but not both) factor selectors blank, e.g.

- to search from some specific factor or factors to "anywhere", or
- to search from "anywhere" to some specific factor or factors.

(However, at the moment you can not calculate robustness if either factor selector is blank.)

Calculating robustness by field, e.g. "by source"

Just because there is a lot of evidence for all the different parts of a path does not mean that any one source, or type of source, actually mentioned all those parts.

For this reason, `trace robustness` also lets you ask that question.

Trace robustness

This filter shows all the paths (up to a length which you can specify) from one factor or set in the Robustness tab.

Apply filter!

Keep this panel open?

Trace how many steps down?

0

3

Begin tracing from which factor(s)?

Start typing to select a factor, or factors, or part of a factor

End at which factors?

Start typing to select a factor, or factors, or part of a factor

For the robustness calculation, what field to use for counting? Leave blank for none.

Start typing to select a links field

For example you can put `source_id` in the box in order to count how many sources mentioned at least one entire path from the source to the target factor(s).

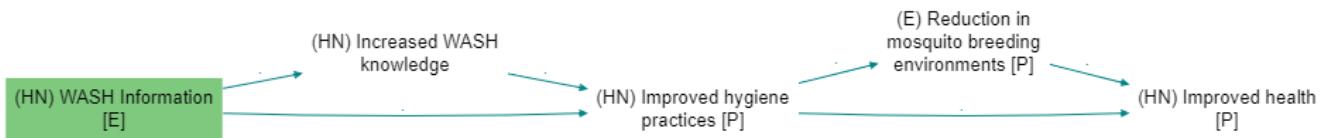
This is important because it's usually more credible to ask how many people mentioned all the bits of some path than to ask how many pieces of evidence are there altogether.

Or you could put `District` to find out, looking at each district individually, there is evidence (possibly assembled from different sources within that district) for the path.

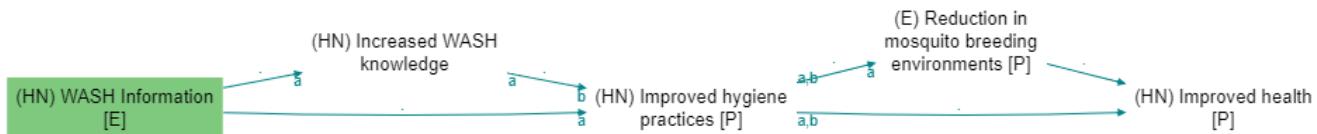
Mark links for continuity (Print View only)

This filter is experimental and the details will certainly change.

The transitivity trap can make it a challenge to interpret your maps. Some people (or somebody) said that improved hygiene led to reduction in mosquito environments, and some people said reduction in mosquito environments led to improved health, but was this the same or different people?



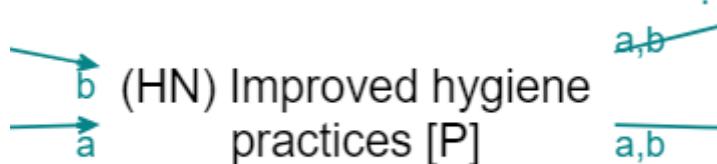
Add the filter `mark_links` (there is no button for it, you have to type it) provides the following diagnostics:



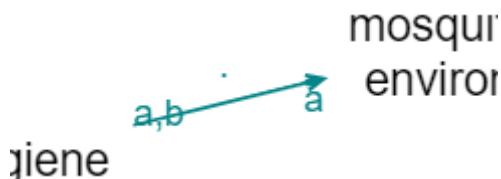
The incoming (bundles of) links to every factor (actually to every factor with outgoing links) are labelled a, b, etc, including when links are bundled e.g. by gender. Then the outgoing links are marked with say `a` if at least one of the sources who mentioned the outgoing link also mentioned link `a`.

So we can see that none of the people who said that improved hygiene led to reduction in mosquito environments also said that reduction in mosquito environments led to improved health: there is no label at all on the arrow going out of reduction in mosquito environments. There is no source continuity.

Note that the labels get re-used for each factor, so the `a`s and `b`s here are related:



but the `a`s here are not:



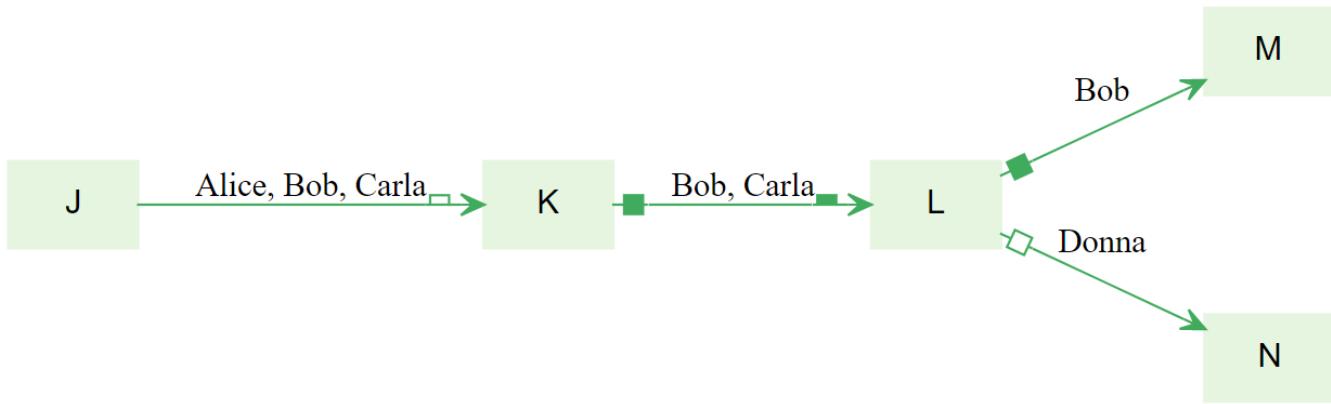
This also works with all the other fields, e.g. you can type `mark_links` field=statement_id in order to test statement continuity, which is a stricter test of continuity. `source_id` is default so you don't need to type it specially.

Yes, it is a bit difficult to communicate this in a report. But it is important for interpretation. Of course a chain without source continuity isn't an invalid chain per se, it's just something to be aware of.

We will probably also add a simpler metric for outgoing links which does not distinguish between the incoming links, something like "Percentage of sources who mentioned a link leaving factor F who mentioned any of the links entering F". This metric could be used to colour or scale the links, or perhaps be printed on the tail of the links.

Show continuity

Summary



Above, the links are labelled with the sources.

The □ open half-box at the end of the first link tells us that at least half but not all of these stories stop here: less than half the sources mentioned any link out of K.

The ■ filled box at the start of the second link tells us that all of these stories are continuations: all these sources mentioned some link into K.

The — filled half-box at the end of the second link tells us that at least half but not all of these stories continue: Bob mentioned some link out of L, but Carla did not.

The □ open box on the link from L to N tells us that this story is not a continuation: Donna did not mention any link into L.

There is no UI for this filter yet. You can just type

`show_continuity`

in the advanced editor.

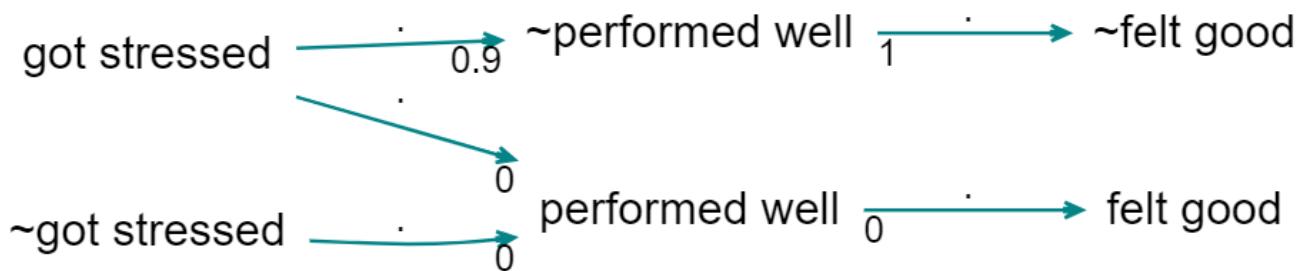
The four kinds of boxes are (possibly aggregated) indicators of continuity, with respect to sources, between stages in a path.

If you want to look at say statement continuity rather than source continuity (the default), type

```
show continuity field=statement_id
```

If you want to see numbers (see examples below) rather than symbols (see examples further below; symbols are the default) then type:

```
show continuity type=label
```



Here, the 0.9 says that 90% of the sources mentioning the link to ~performed well also mention the link from ~performed well. The 1 says that 100% of the sources mentioning the link from ~performed well also mention the link to ~performed well. And the zeros below say that there is no source continuity at all.

What this doesn't tell you is, when there are more than one incoming link, which of them have sources which continue to the outgoing link (that is what the bs and cs are for in `mark_links`). It's just an aggregate.

But what happens with filters which actually transform the map: zoom, bundle factors and combine opposites? Zoom can create its own version of the transitivity trap, if we have:

eating lemons --> health; no scurvy

and

health; fitness --> fast runner

Health; being physically fit → Ability to run fast

Eating lemons → Health; not getting scurvy

we should be very careful when concluding (when zooming)

eating lemons --> health --> fast runner

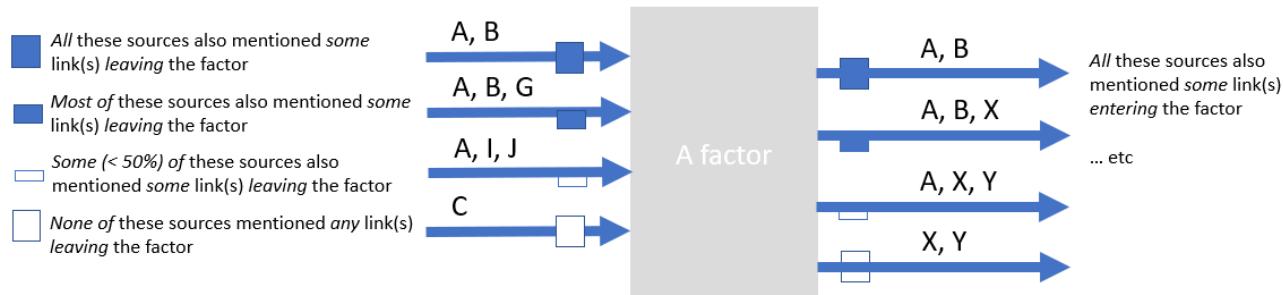
... and indeed, showing continuity highlights this error:



Showing continuity with arrowtypes

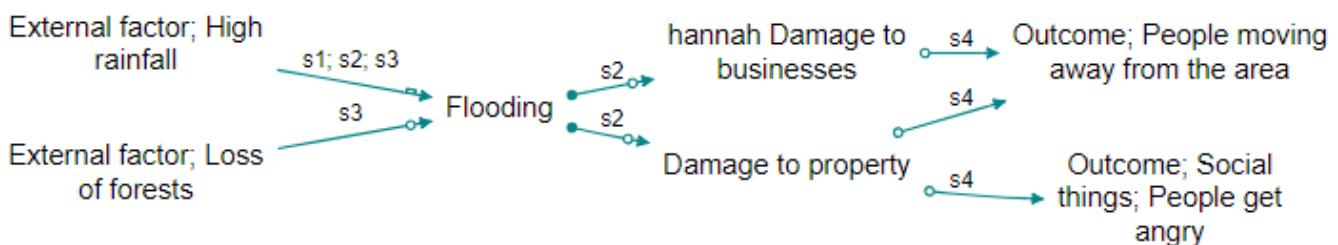
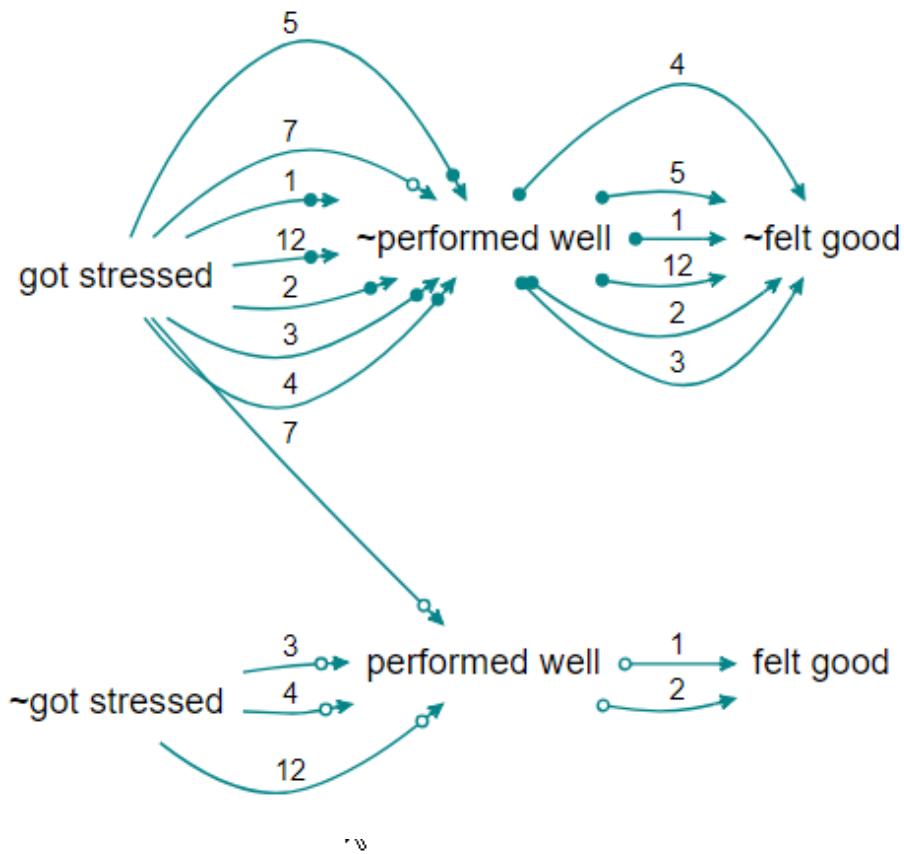
Printing actual numbers (from 0 to 1) on the arrows can be very confusing. So the default is to use symbols.

- white box: 0
- half white box: <= 0.5
- half full box: >.5
- full box: 1





{width=650}



Roundtripping with the statements table

Statements table on our Test server: the `statement_id` field now IS included when you download a file.

- This means you can delete rows and reorder rows (by dragging them around in Excel). Don't change the `statement_id`.

- Note: If you delete rows which had links, these links will become orphaned and will be displayed in the maps only when you do not filter by any statement_id. This probably isn't what you want -- you might want to delete the corresponding links too. If you choose to add new rows as well as deleting old rows, the statement_id of the new rows will be the row number, so make sure this is what you want -- will they then get "gifted" orphaned links?

!--chapter:end:88888-changelog.rmd--

(PART*) Final section {-}

!--chapter:end:90000-PART-final.rmd--

Acknowledgements

This is just the beginnings of a list of the people who have helped contribute to this Guide and to the Causal Map app ...

- Samuel Goddard, the amazing intern ...

The team at BSDR ...

Friends and colleagues around the world ...

- Randy Puljek-Shank
- Simone Ginsburg

!--chapter:end:90010-acknowledgements.rmd--

Appendix: Publications and links

Publications

Causal Map is quite new, but we're working hard publishing our ideas. Here is some relevant prior work.

Copestake, J., Morsink, M., & Remnant, F. (2019). Attributing Development Impact: the Qualitative Impact Protocol case book. March 21, Online.

Powell, S. (2019). Theories of Change: Making Value Explicit. Journal of MultiDisciplinary Evaluation, 15(32), 53–54.

Powell, S. (2019). Supplement to Theories of Change: Making Value Explicit. Journal of MultiDisciplinary Evaluation, 15(32), 53–54.

Powell, S. (2018). The Book of Why. The New Science of Cause and Effect. Pearl, Judea, and Dana Mackenzie. 2018. Hachette UK. JMDE Journal of MultiDisciplinary Evaluation, 13, 47–54.

Powell, S. (2019). Supplement to Theories of Change: Making Value Explicit. Journal of MultiDisciplinary Evaluation, 15(32), 53–54.

Presentations

- European Evaluation Society 2022 – Causal Mapping for evaluators
- European Evaluation Society 2018 – presentation on Value in Theories of Change
- UK Evaluation Society 2018 – presentation on Judea Pearl and Theories of Change
- American Evaluation Association 2017 – A graphical language for evaluation

Longer blog posts

- When an interview beats an RCT (with cartoon!)
- Scandal or yawn? Simpson's Paradox hits the IFRC
- Various posts on LinkedIn

ResearchGate

James Copestake

Fiona Remnant

Steve Powell

Resources and apps

- [Slides about theories of change and causal mapping](#)
- [The original Theorymaker app](#)
- [Help for the original Theorymaker](#)

Bibliography of selected causal mapping approaches

We also have a more complete Zotero library on causal mapping. Here are some of the key papers.

[Ackermann, F., & Eden, C. \(2004\). Using Causal mapping: individual and group; traditional and new. *Systems Modelling: Theory and Practice*, 127–145.](#)

[Ackermann, F., Howick, S., Quigley, J., Walls, L., & Houghton, T. \(2014\). Systemic risk elicitation: Using causal maps to engage stakeholders and build a comprehensive view of risks. *European Journal of Operational Research*, 238\(1\), 290–299. <https://doi.org/10.1016/j.ejor.2014.03.035>](#)

[Axelrod, R. \(1976\). The Cognitive Mapping Approach to Decision Making. In *Structure of Decision: The Cognitive Maps of Political Elites* \(Issue May 2019\).](#)

[Axelrod, R. \(1976\). The Analysis of Cognitive Maps. In *Structure of Decision: The Cognitive Maps of Political Elites* \(Issue May 2019\).](#)

[Axelrod, R. \(1976\). The cognitive mapping approach to decision making. *Structure of Decision*, 221–250.](#)

[Chaib-Draa, B., & Desharnais, J. \(1998\). A relational model of cognitive maps. *International Journal of Human Computer Studies*, 49\(2\), 181–200. <https://doi.org/10.1006/ijhc.1998.0201>](#)

[Copestake, J., Morsink, M., & Remnant, F. \(2019\). *Attributing Development Impact: the Qualitative Impact Protocol case book*. March 21, Online.](#)

[Goodier, C., Austin, S., Soetanto, R., & Dainty, A. \(2010\). Causal mapping and scenario building with multiple organisations. *Futures*, 42\(3\), 219–229. <https://doi.org/10.1016/j.futures.2009.11.007>](#)

[Jackson, K. M., & Trochim, W. M. K. \(2002\). Concept Mapping as an Alternative Approach for the Analysis of Open-Ended Survey Responses. *Organizational Research Methods*, 5\(4\), 307–336. <https://doi.org/10.1177/10944280223714>](#)

Krentel, A., & Aunger, R. (2012). Causal chain mapping: A novel method to analyse treatment compliance decisions relating to lymphatic filariasis elimination in Alor, Indonesia. *Health Policy and Planning*, 27(5), 384–395. <https://doi.org/10.1093/heapol/czr048>

Laukkanen, M. (1994). Comparative Cause Mapping of Organizational Cognitions. *Organization Science*, 5(3), 322–343. <https://doi.org/10.1287/orsc.5.3.322>

Laukkanen, M. (2012). Comparative Causal Mapping and CMAP3 Software in Qualitative Studies | Laukkanen | Forum Qualitative Sozialforschung / Forum: Qualitative Social Research. *Qualitative Research*, 13(2), volume 13. <http://www.qualitative-research.net/index.php/fqs/article/view/1846/3371>

Laukkanen, M., & Eriksson, P. (2013). New designs and software for cognitive causal mapping. *Qualitative Research in Organizations and Management: An International Journal*, 8(2), 122–147. <https://doi.org/10.1108/QROM-08-2011-1003>

Mapping, P. S., & Incentive, R. H. (2020). *Participatory Systems Mapping in action — Supporting the evaluation of the Renewable Heat Incentive*. 17.

Maps, T. C., & Elites, P. (1918). Chapter Title: Decision for Neoimperialism: The Deliberations of the British Eastern Committee in 1918 Chapter Author (s): Robert Axelrod Book Title: Structure of Decision: The Cognitive Maps of Political Elites Book Editor (s): Robert Axelrod Pub.

Markiczy, L., & Goldberg, J. (1995). A method for eliciting and comparing causal maps. *Journal of Management*, 21(2), 305–333.

Nadkarni, S., & Shenoy, P. P. (2004). *Nadkarni and Shenoy 2004 -A causal mapping approach.pdf*. 38(2).

Nadkarni, S., & Shenoy, P. P. (2001). Bayesian network approach to making inferences in causal maps. *European Journal of Operational Research*, 128(3), 479–498. [https://doi.org/10.1016/S0377-2217\(99\)00368-9](https://doi.org/10.1016/S0377-2217(99)00368-9)

Penn, A., & Barbrook-Johnson, P. (2019). *Participatory Systems Mapping: a practical guide*. 1–4. <https://www.cecan.ac.uk/sites/default/files/2019-03/PSM Workshop method.pdf>

Pidd, M. (n.d.). *Systems Modeling*. In *Elements*.

Srivastava, R. P., Buche, M. W., & Roberts, T. L. (2011). Belief Function Approach to Evidential Reasoning in Causal Maps. *Causal Mapping for Research in Information Technology*, 109–141. <https://doi.org/10.4018/9781591403968.ch005>

Sterman, J. (2001). System dynamics modeling. *California Management Review*. http://trafficlight.bitdefender.com/info?url=http%3A//www.systemdynamics.org/conferences/2002/proceed/papers/CAREVIEW/C2STERMA.PDF&language=en_US

Wallis, S. E. (2019). The missing piece of the integrative studies puzzle. *Interdisciplinary Science Reviews*, 44(3–4), 402–429. <https://doi.org/10.1080/03080188.2019.1680145>

Wallis, S. E. (2020). Understanding and improving the usefulness of conceptual systems: An Integrative Propositional Analysis-based perspective on levels of structure and emergence. *Systems Research and Behavioral Science*, February, 1–22. <https://doi.org/10.1002/sres.2680>

Wallis, S. E., & Wright, B. (2019). Integrative propositional analysis for understanding and reducing poverty. *Kybernetes*, 48(6), 1264–1277. <https://doi.org/10.1108/K-03-2018-0136>

!--chapter:end:99000-publications.rmd--

r if (knitr:::is_html_output()) '# References {-}'

!--chapter:end:99900-refs.rmd--

1. Any preset ending with (basic) is also the default filter that you will see when you first click onto that tab, making flicking between presets easy. ↵ ↵