🌻 The transitivity trap

📅 **22 Aug 2025**

> From (Powell et al. 2024)

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

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

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

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

(Thanks to Gary Goertz for remembering this one!)

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

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

from

> B → C

and

> C → E

we can only conclude

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

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

For example:

If

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

and

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

can we deduce

> G → C → H

and therefore

> G → S

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

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

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

If

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

and

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

can we deduce

> A → C → D

and therefore

> A → D

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

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

Not just a problem for causal mapping

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

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

Can we mitigate the trap with careful elicitation protocols?

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

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

Transitivity trap, or identity trap?

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

References

Bollen (1987). *Total, Direct, and Indirect Effects in Structural Equation Models*. JSTOR.

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