Theories of Change. Keeping it logical.

Steve Powell

Press your space bar or arrow keys to begin, or type o for an overview, f for fullscreen

Intro

These slides…

- present a simple and powerful way to construct and think about Theories of Change (based on DAGs, Pearl (2000))
- you don't have to stick to it. But working with it can help you sharpen your sense for logical distinctions which can enhance or ruin your own Theory of Change.

Feel free to use just one or two of these ideas in your Theory of Change.

There is no club to join or product to buy!

pogol.net/tocs.pptx

Questions: steve@pogol.net

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). Theories of Change: Making Value Explicit. Journal of MultiDisciplinary Evaluation, 15(32), 53-54.

Who’s suspicious of Theories of Change?

There are plenty of reasons to be suspicious of Theories of Change. Perhaps they:

- are unrealistic, too optimistic?
- are not flexible enough, don’t reflect how even the big features of plans change?
- don’t reflect the details of hands-on management and how the little features change?
- don’t reflect how new problems (and responses) can *emerge*?
- are too worried about measuring things in numbers?
- don’t account for complexity of different stakeholders with different visions?
- assume “everything is linear”?

These are good criticisms, but Theories of Change are broad enough to accommodate them.

These slides do *not* deal with the practical or political aspects (who creates Theories of Change, how and why) - even though these aspects are very important.

Variables and **arrows**

Theories of Change

There are many ways of drawing and understanding a Theory of Change.

Here is one useful (and dominant?) way based on DAGs. There are four basic ideas (variables, arrows, intervention & value), and we *derive* the others (“precondition”, “outcome”, “input/action”, “assumption”, “mechanism”, “impact”, “effect”, “effectiveness” from these basic ideas.)

A Theory consists of

- **variables** - things that could be *different*, which have a value or level you might in principle measure or assess…
- joined by **arrows** that mean “this causes that”; if you tweak *this*, it will change *that*.

Just use the best available information about what “what causes what” in an area you are interested in. Concepts like “assumption” or “precondition” don’t yet apply (thank goodness). If a factor is importantly involved, include it.

A Theory of Change is a Theory plus:

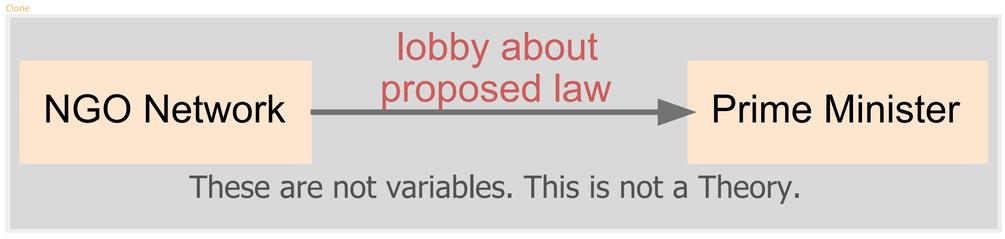
- information about what we **value** and what we can **do** to get it.

You can build almost any Theory of Change from these principles.

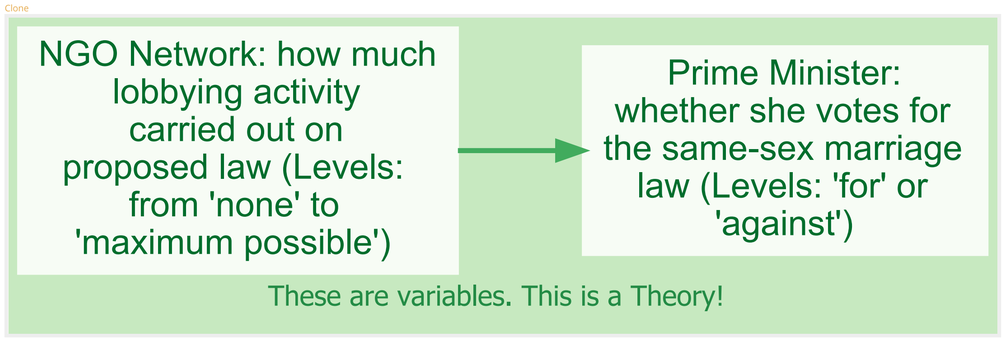
If you stick to them your Theories of Change should be *evaluable*, and you will avoid many logical headaches.

(This distinction is maybe the same as the one sometimes made between *Theory of Change* and *Theory of Action*)

Variables are not actors



This kind of network diagram can be very useful but these rectangles are *things*, not *variables*.

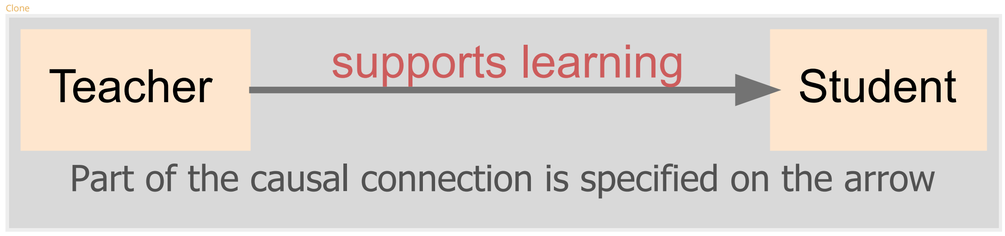


Now, these are *variables*. They **could be different**. They can take different levels or values:

- the “from” variable can be zero or very high, or anywhere in between
- the “to” variable can be only “for” or “against”

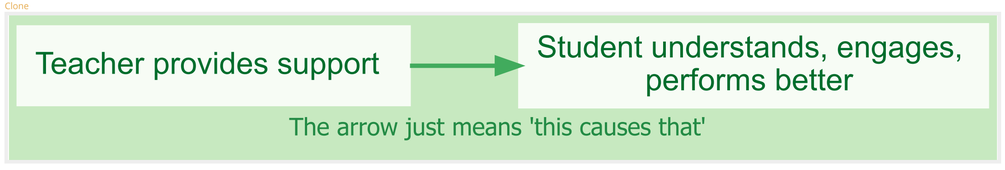
If you can’t directly “measure” them you could design indicators to estimate them. They don’t have to be numerical or clearly defined. But it should always make sense to ask “what value or level does this variable have right now?”

Arrows mean just plain causation



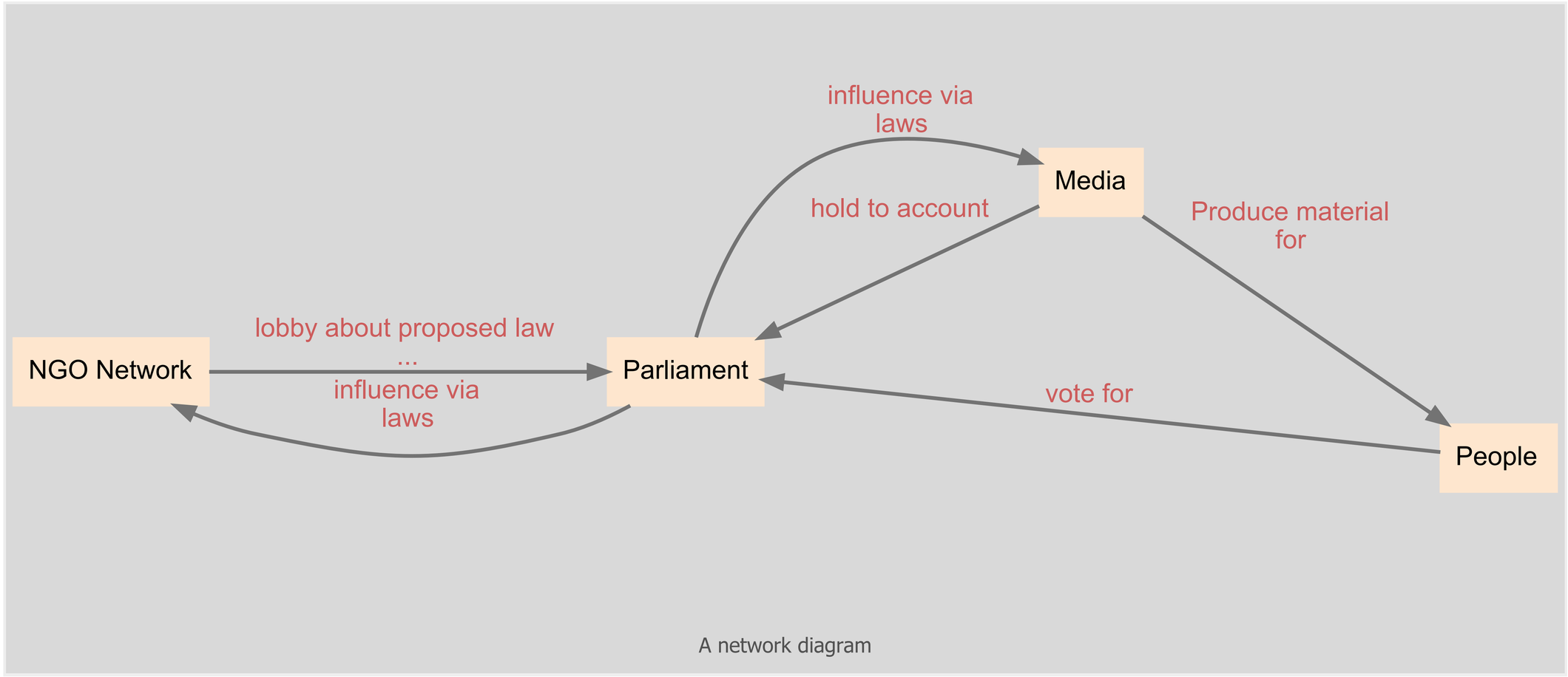
Not a Theory.

In a strict Theory of Change, arrows always just mean “causes”. All the rest of the information about *what causes what* has to go into the variables.



Variables: practice

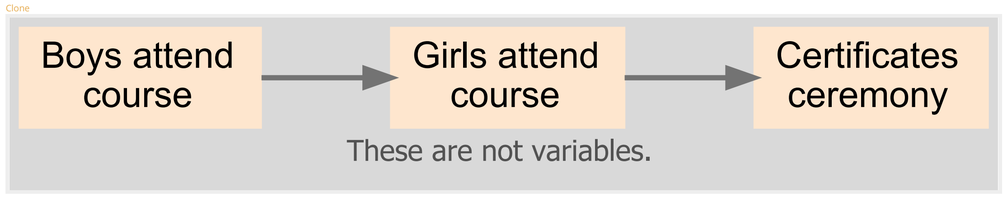
Clone



A network diagram

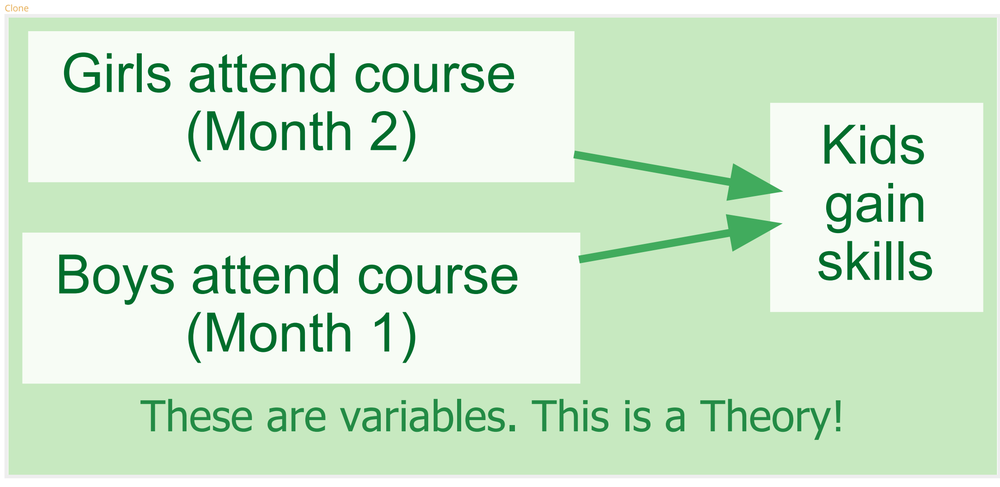
Convert this network diagram into a Theory of Change consisting only of variables?

Variables are not events

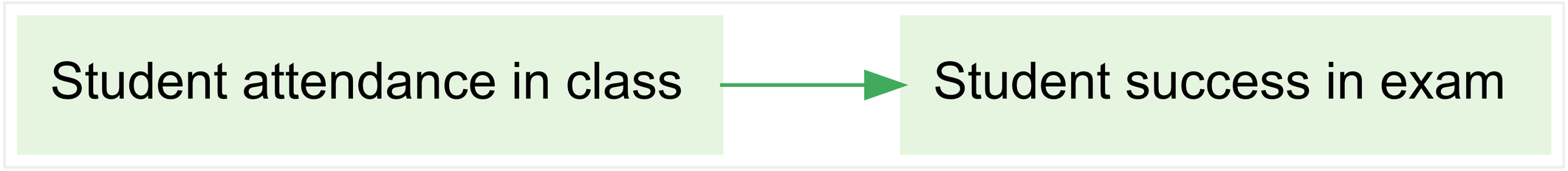


A Theory is not a Gantt chart.

The boys attending course 1 does not cause the girls to attend course 2. (It may be one of many influencing variables, if success of course 1 is a precondition for holding course 2.)



Do think about the kind of **levels** your variables have



- Are these variables yes/no or continuous?
- How would we measure or check them?
- Do we assume variables are always “false/true”?
- Does it matter if we dichotomise …
  - … variables?
  - … targets?

Different types of variable: different types of **level**

Note

A variable label can contain information about the different levels the variable can take, by actually listing them like this: "***Levels:*** red, orange, green" or in some other way, e.g. "Levels: counting numbers" or by using symbols. This isn't compulsory, but can help clarify what *kind* the variables are.

How many kinds of variable can you think of?

Clone

Percentage of legislators supporting the law (Levels: percentages)

Public's confidence in the President ◪

Public's confidence in the President (Levels: lo-hi)

The colour of this traffic light (Levels: red, orange, green)

Enabling law on same-sex marriage is passed at Federal level (Levels: false, true)

Enabling law on same-sex marriage is passed at Federal level ◨

Number of signatures on petition (Levels: counting numbers)

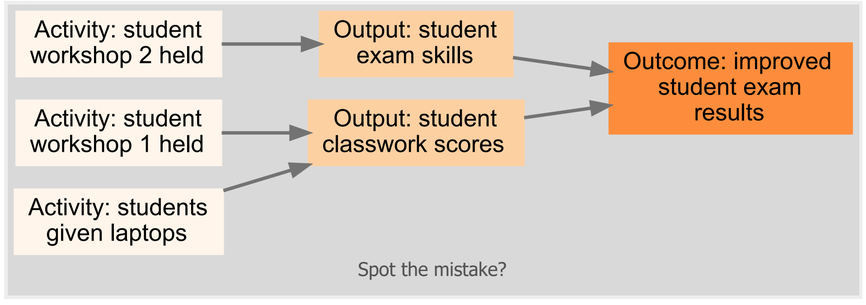
Specifying what kind a variable is - can help the reader understand what is meant

Either all or none of your variables should mention facts/targets/changes

Indicators should never mention targets (DFID 2011). But often it is very useful if our *variables* do.

A target is a *Difference made*. This may not always be a difference from the baseline.

Clone

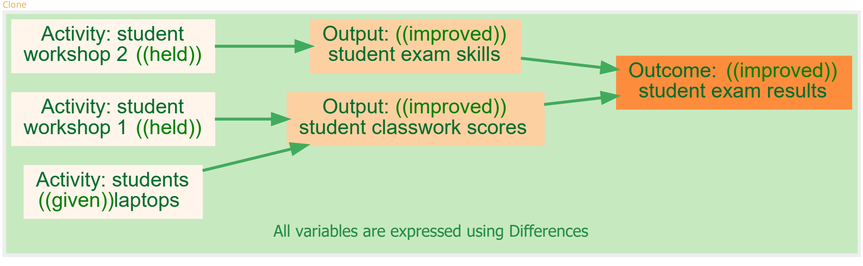


The problem is that the two intermediate variables are just shown with variable names whereas the others are also expressed in terms of Differences.

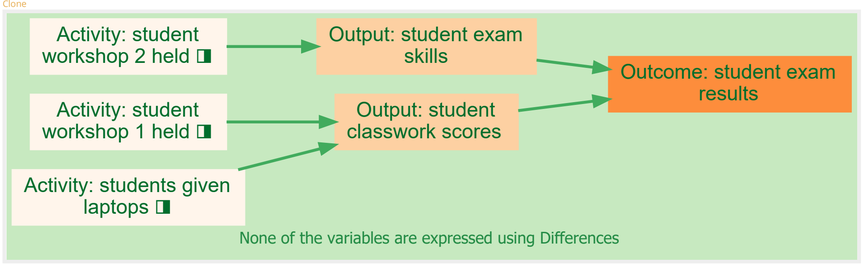
If you take this at face value, this will give you endless headaches down the line.

Two solutions:

Clone



Clone



Simple Theories: using arrows to show the causal influences *from* one or more variables *to* another.

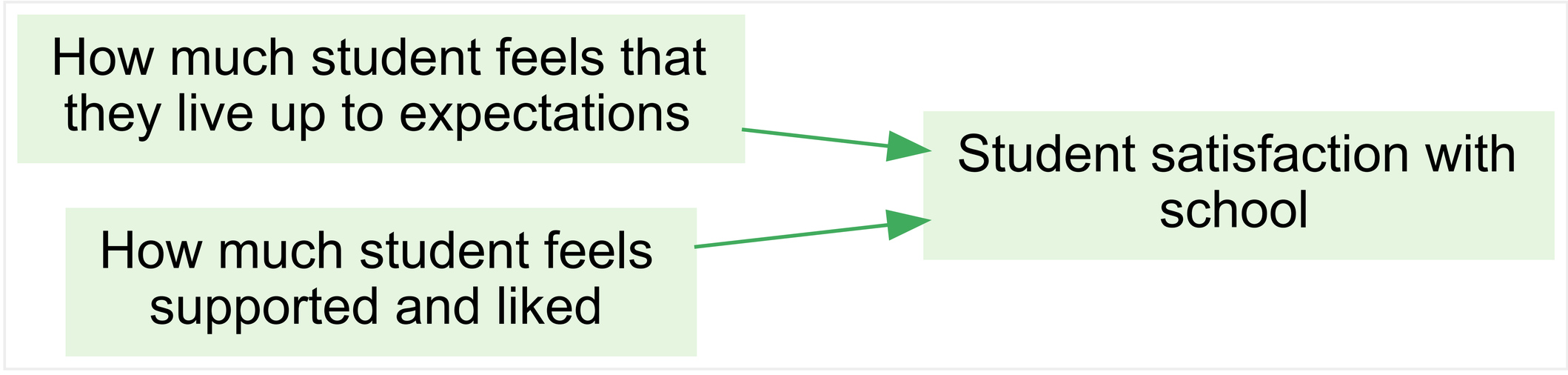
We use arrows to link one or more variables to another.

Meaning: if you manipulate the “from” variables, this will make a difference to the “to” variable.

Note

We can call the combination of these variables (and the arrows joining them) a ***simple Theory***.

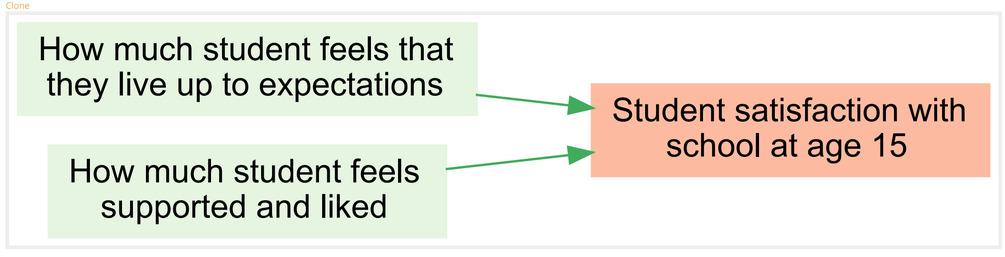
Clone

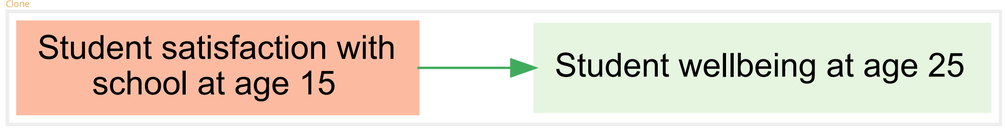


“How satisfied is the student? It depends on their feeling of support, of living up to expectations , …”

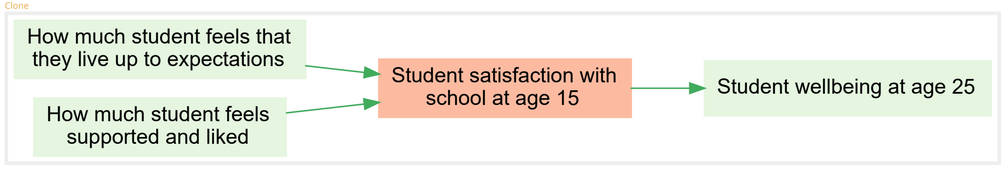
Combining simple Theories

If we have more than one simple Theory and they share some variables …





… we can join them together into a composite Theory of Change.

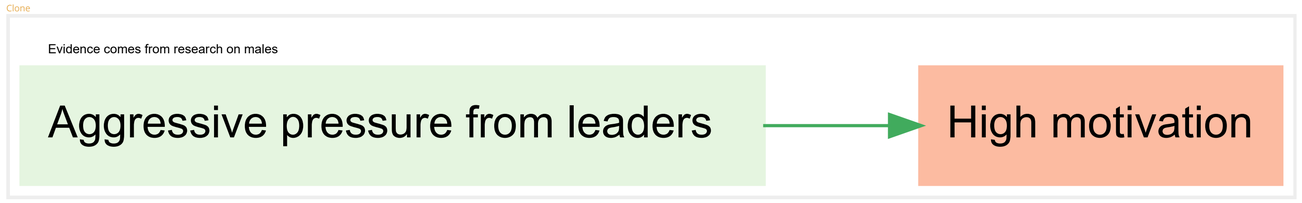


Ideally, each simple Theory within a larger one makes sense on its own and has its own evidence.

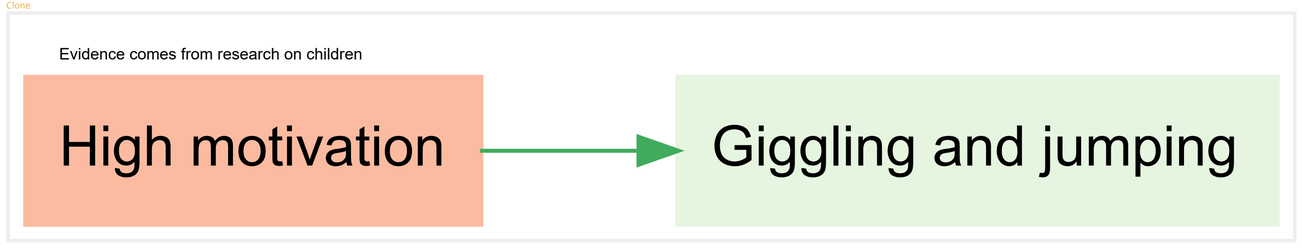
- But snapping together theories like this may result in loops, so be careful.

Combining theories: **Don’t** mix up **contexts**

Clone



Clone



Snap these together, giving a new Theory.

Clone

Practice: does this happen in your field of work? ℹ

Tip

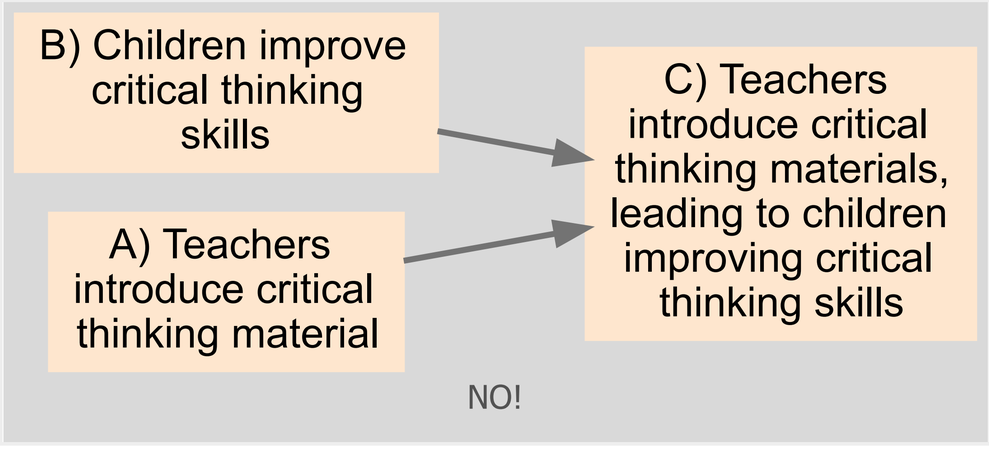
Combining theories: **Do** note the **evidence**

- Remember and note down where the evidence comes from
- Ask whether what we know from one context is transferrable to our context

Tip

Don’t try to sell a causal link as a variable!

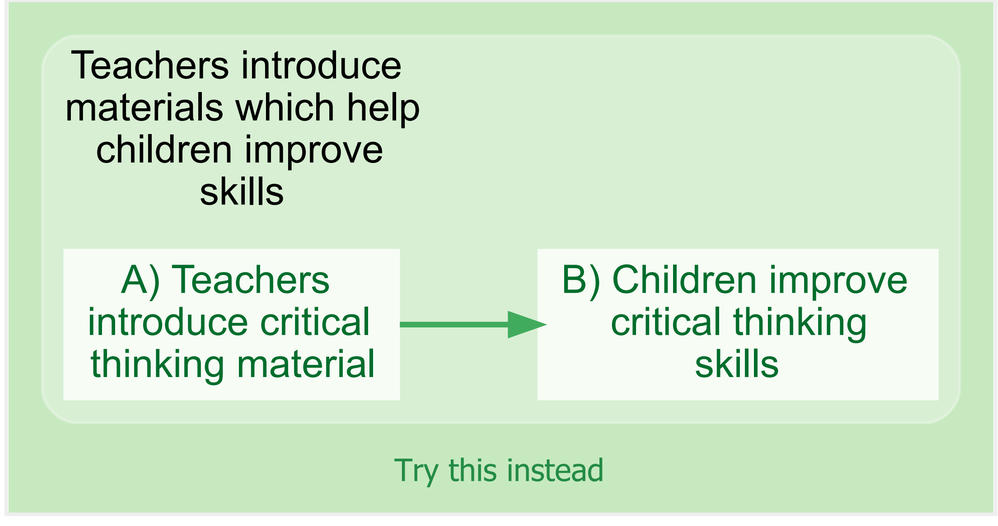
Clone



Breaks at least two rules:

- The arrows imply that A and B influence C; in fact A influences B and C just summarises that fact.
- The variables in a Theory of Change should be free of causal words like “through” or “via” or “because”; they should have a value or level you might in principle measure or assess.

Clone

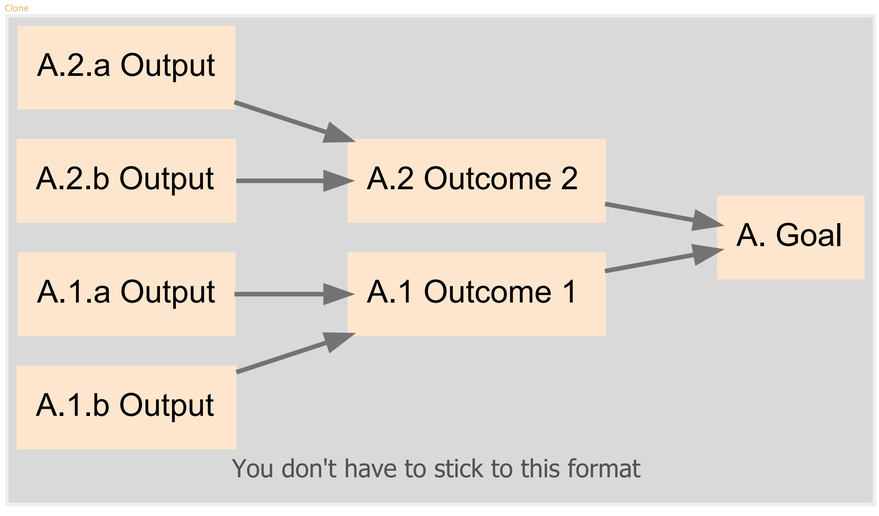


The causal story stands for itself.

We also used a grouping box to label the causal story (optional).

Do let your variables influence more than one other variable

Clone



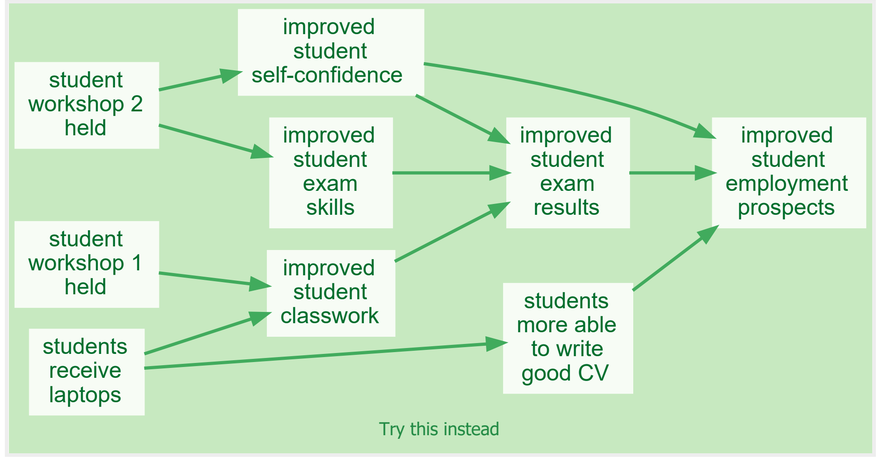
In lots of Theories of Change, especially those based on Logical Frameworks, every variable only has one child variable which it influences.

The “one child rule”

- is convenient for administrators
- allows us to number the variables in our project hierarchically
- a convenient straight-jacket

But in real life … Of course a variable can influence more than one other variable!

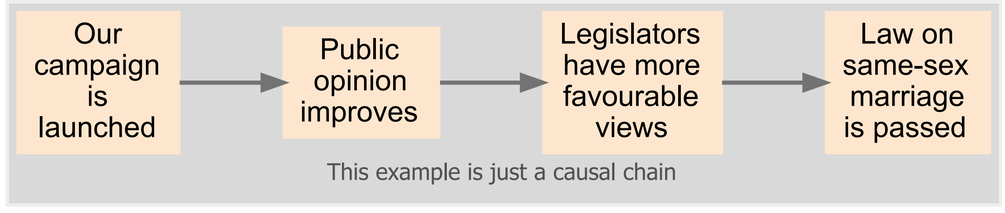
Clone



ToCs: what we can **do** to get things we **value**

Show which variables we intervene on: ▶

Clone



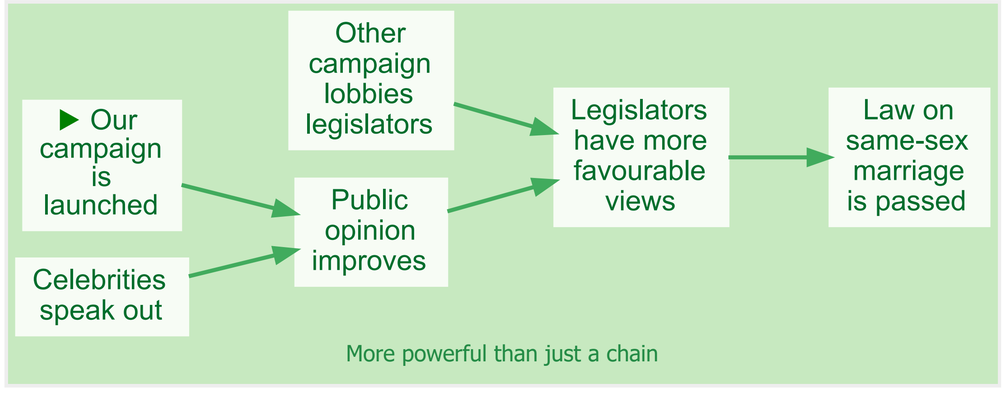
In many Theories, the "no-parent" variables (the ones with no influence variables contributing to them) are assumed to all be under the control of our project.

But of course the world isn't like that.

Sometimes it can be really useful to show where we intervene and where we don't.

Make it clear which variables you are intervening on, by marking them with an *intervention symbol*: ▶ and also making sure it is clear what the intervention does – for example, training 20 teachers (rather than training the usual number, 10, or maybe none).

Clone



Usually, when we include intervention variables, the downstream variables are also expressed in such a way that the consequences of the intervention can also be read off them, e.g. "Law on same-sex marriage is passed (rather than not passed)" – as opposed to, say, "Whether the law on same-sex marriages is passed".

The *other* no-parent variables in the diagram are not under our control. In this example, they have also been expressed as Statements ("Celebrities speak out" – rather than not speaking out.) The difference made to the downstream variables such as "Legislators have more favourable views" take these Statements or assumptions into account. You can call them "assumptions" or "preconditions". You can leave them out of your diagram and just make a note.

Show which variables we value: ☻

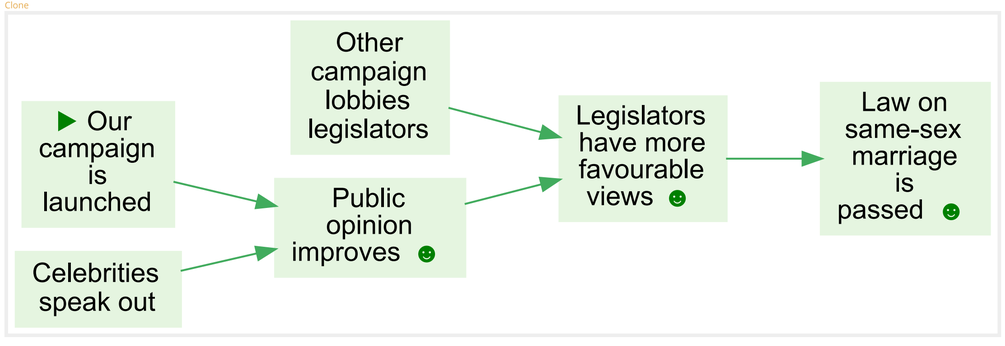
In many Theories, the “no-child” variables (the ones with no consequence variables) are all assumed to be ones we *care about* (and the *only* ones we care about).

But of course the world isn’t like that.

Note

Sometimes it can be really useful to show which **variables we value**, e.g. by marking them with a heart or smiley ☻ . (We can also use a “frowny” symbol: ☹ for things we don’t want.)

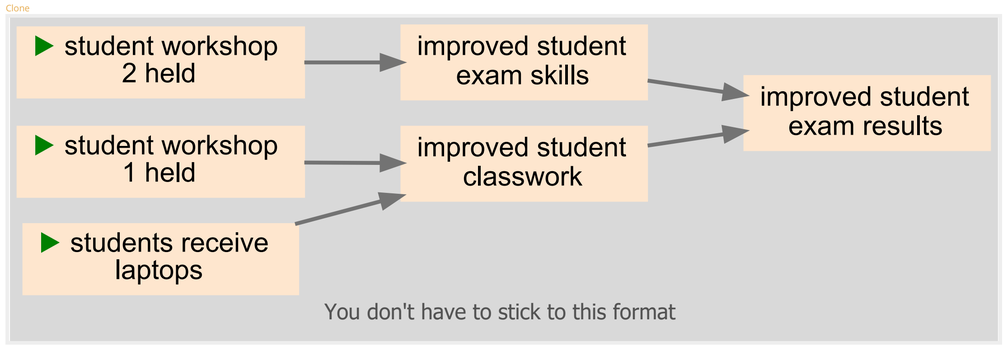
Clone



In this case, we value the legislators and the public having a more positive attitude to same-sex marriage not only because they are more likely to pass the law, but for other reasons we aren’t mentioning here.

It’s also possible to write *more than one* symbol, for example: ☻ ☻ ☻ , to distinguish, say, between things we value a little and which we value a lot. Or just write e.g. “Value=high” or “Value=low”.

Allow yourself more than one *valued variable*

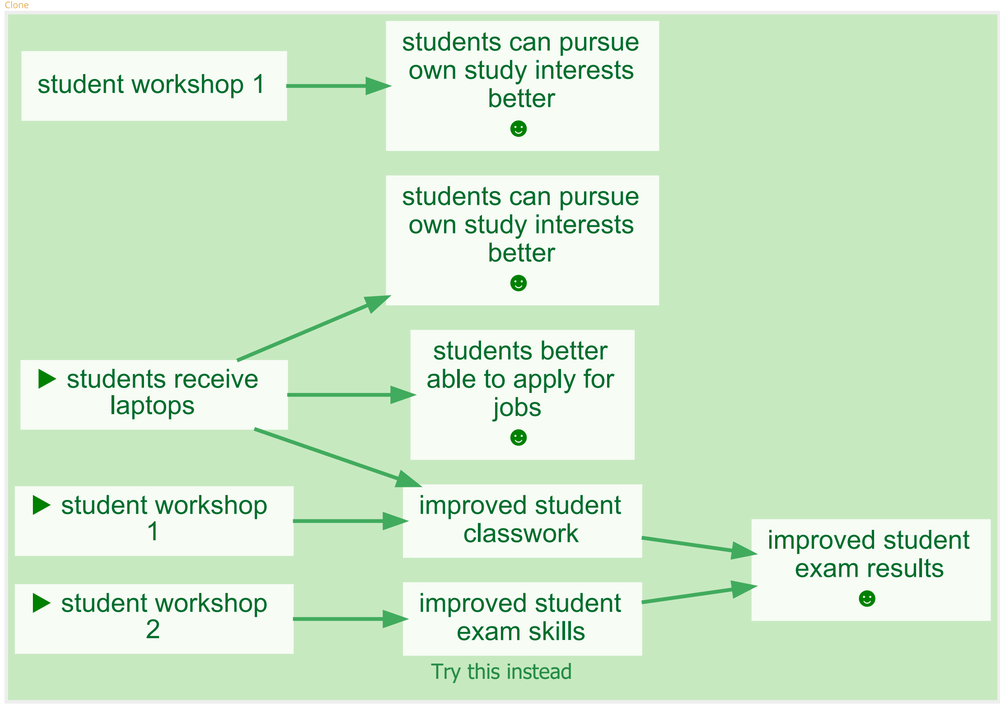


Lots of Theories of Change have only one valued variable,

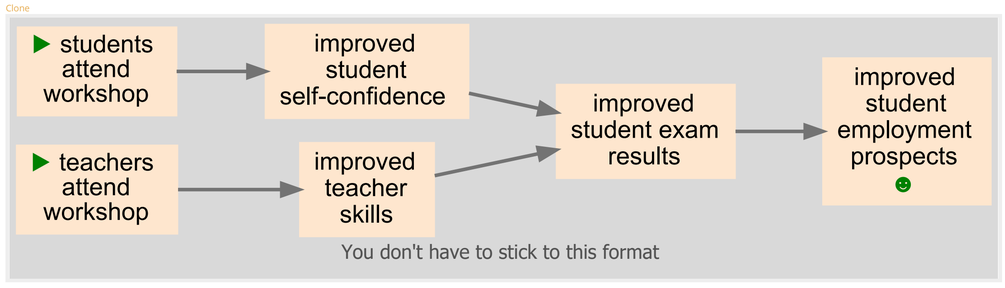
- the last one(s) in the chain
- often called “Top-level Outcome”, “Goal” etc.

But in the real world, you (or maybe other stakeholders) often value more than one variable.

The laptops might have other valuable consequences for the students, so mark them as valuable, e.g. “☻ ” or even “☻ ☻ ”.

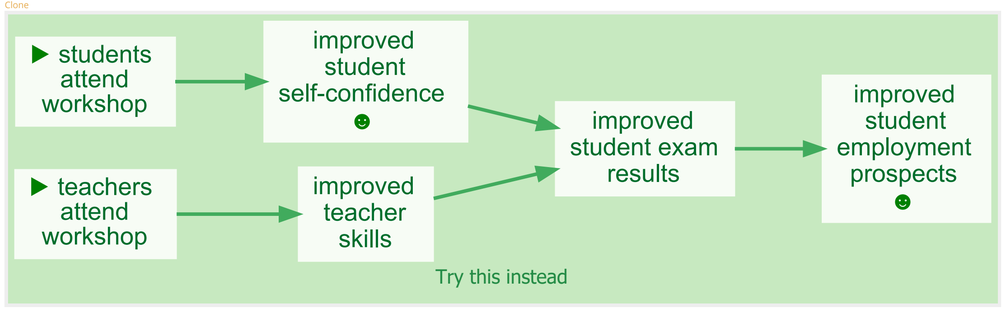


Allow yourself to value intermediate variables too



In the real world, you (or maybe other stakeholders) might value variables which are *not* at the end of a chain. If so, mark them as valued too, e.g. with a " ☻ ".

Maybe you would still be quite pleased with the project if it made a big difference to student self-confidence regardless of its effect on exam results:

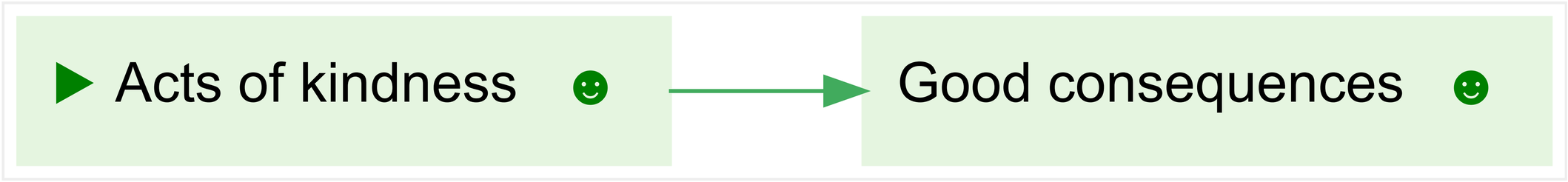


Some intermediate variables (like attending workshops) are really only means to an end.

But some are valuable in their own right: don't undersell your project if it really does produce additional value! Add more " ☻ " for variables you value more!

Allow yourself intrinsic value

Clone

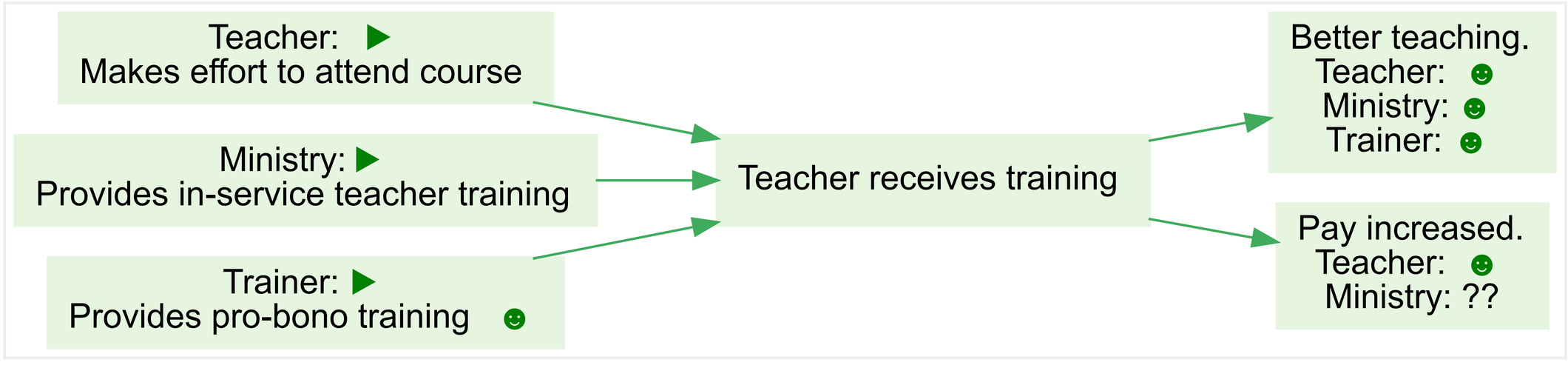


We already noted that a valued variable does not have to be at the end of a causal chain. In fact, it can even be at the beginning of one.

This kind of thinking is very important for instance to understand spiritual and religious motivation. It is related also to the idea of the “humanitarian imperative”. From this point of view, the consequences can be more or less important, but certainly the process is important.

Multiple Actors: Who acts on what, who values what?

Clone



When different stakeholders or "Actors" interact, they are usually able to intervene on different variables. This can be done by using different colours, or by writing the name of the Actor *before* the ▶ symbol, as we do here. We can use ☻ symbols, preceded by the name of the Actor who values this particular variable, in the same way. ℹ

Previously we have only seen Value ☻ and Control ▶ symbols used globally to show what "we", the makers of the Theory, value and control. Now we are zooming out and taking a meta-perspective.

This kind of presentation can be really useful for realising that different Actors might do different things, or put different a different emphasis on the same things, because their motivations are different.

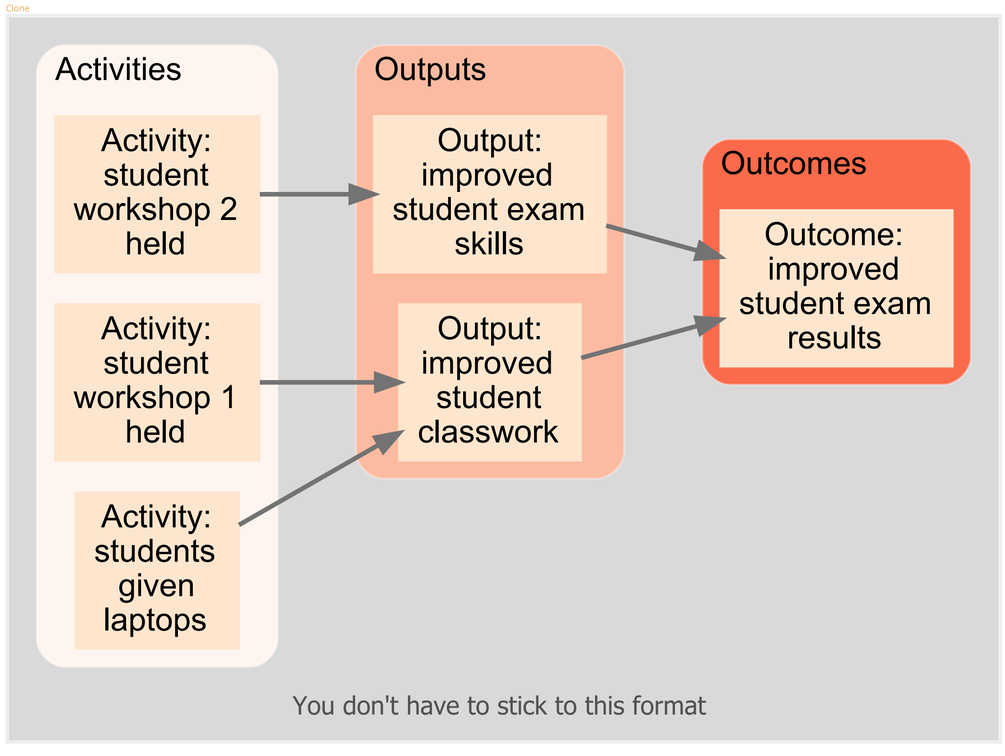
In this example, part of the trainer's motivation is simply the act itself - of providing pro-bono training. All the Actors value the main outcome, improved teaching, whereas the increased pay to which the teacher is entitled after completing the training is a positive motivation for the teacher but a negative one for the Ministry. ℹ

In this case, the Actors differ on two of the dimensions of the Theory of Change - what they control and what they value - but share the third aspect, namely the Theory itself, the one shown here. More generally, they might not.

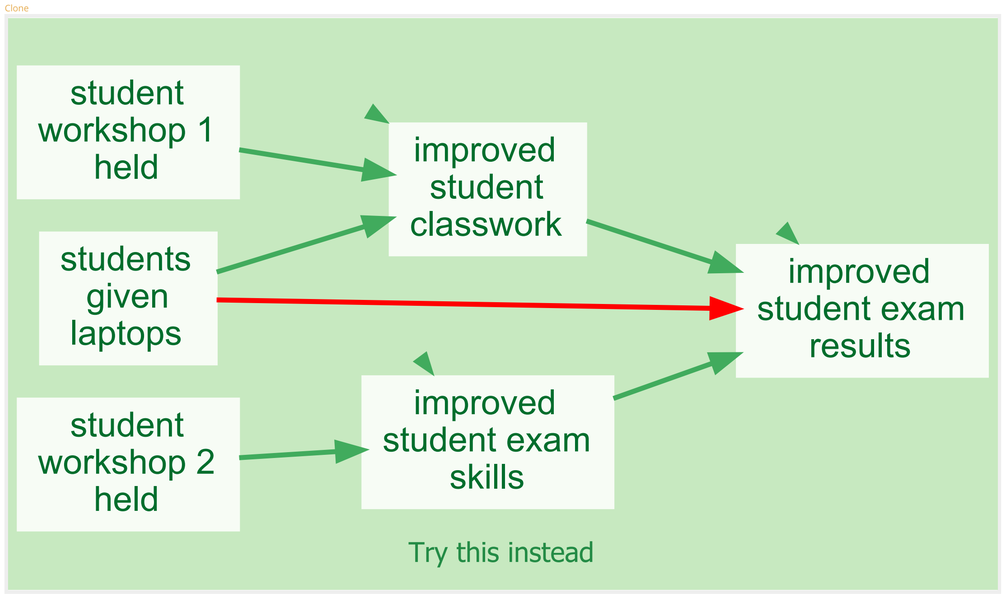
See also (Hansen and Vedung 2010) on evaluation with multiple stakeholders, and (Van Ongevalle et al. 2012) for "actor-focused" evaluation approaches.

Slices and **phases**

Is it useful to divide a Theory into **Slices** like “Outputs” & “Outcomes”?



In real life, variables do not fit neatly into slices. Sometimes it is useful to pretend that they do. Sometimes not.



… the laptops also directly contribute to improved exam results, perhaps because the students can revise better: add an arrow - it is shown red here.

Students receive laptops is **both one step *and* two steps away** from the final outcome. **So the “slices” structure breaks down**.

Slices: practice

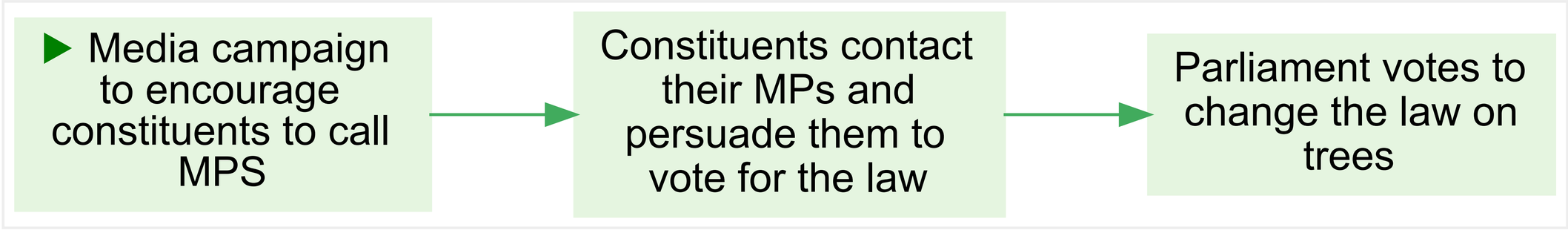
- Find a Theory of Change which is arranged in neat slices and see if you can make it more realistic but less tidy
- OR Find a Theory of Change which is NOT arranged in neat slices and see if you can “tidy” it!

Advantages and disadvantages?

- Can “slices” dictate programming? Should they?

Slices: you don’t have to intervene at arm’s length

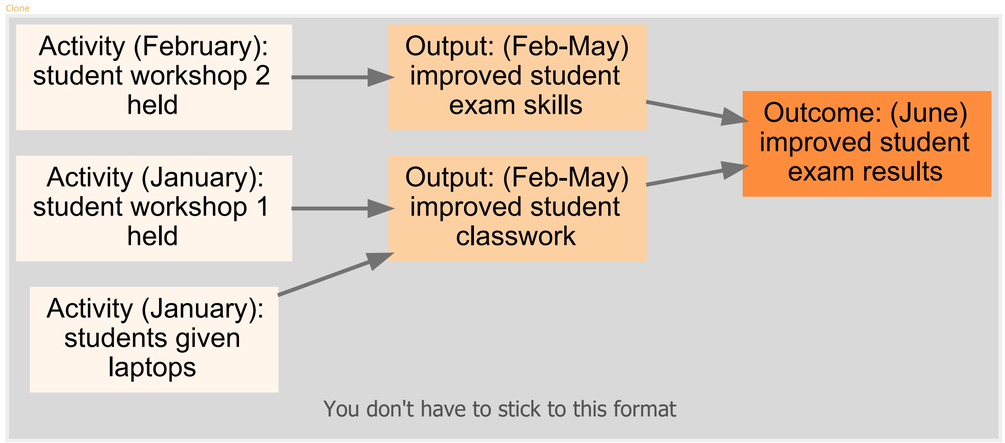
Clone



Does all campaigning have to work “at arms length”?

Slices: you don’t have to intervene at arm’s length (2)

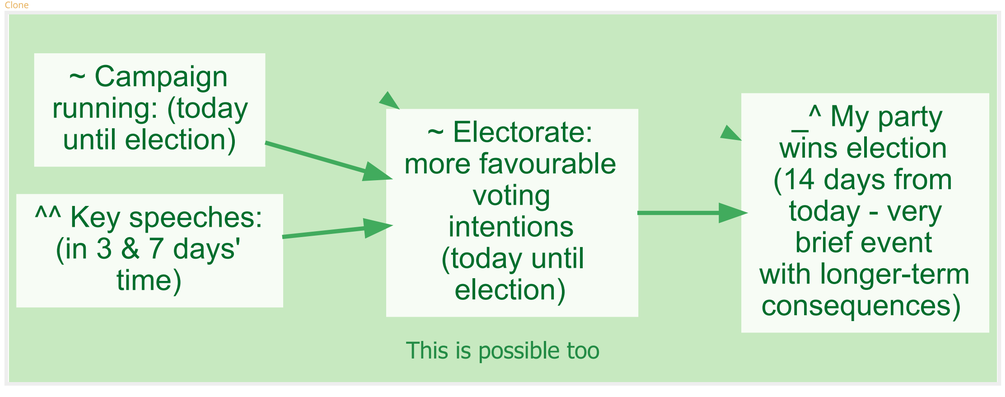
Is it useful to divide a Theory into **Phases**?



Here, the activities in each slice are supposed to take a specific amount of time and each completes before the next starts: the *slices* are *phases*.

Sometimes useful, sometimes a painful straightjacket.

The Prime Minister calls snap elections in two weeks, my party has to launch a campaign to influence a whole nation's behaviour in 14 days:



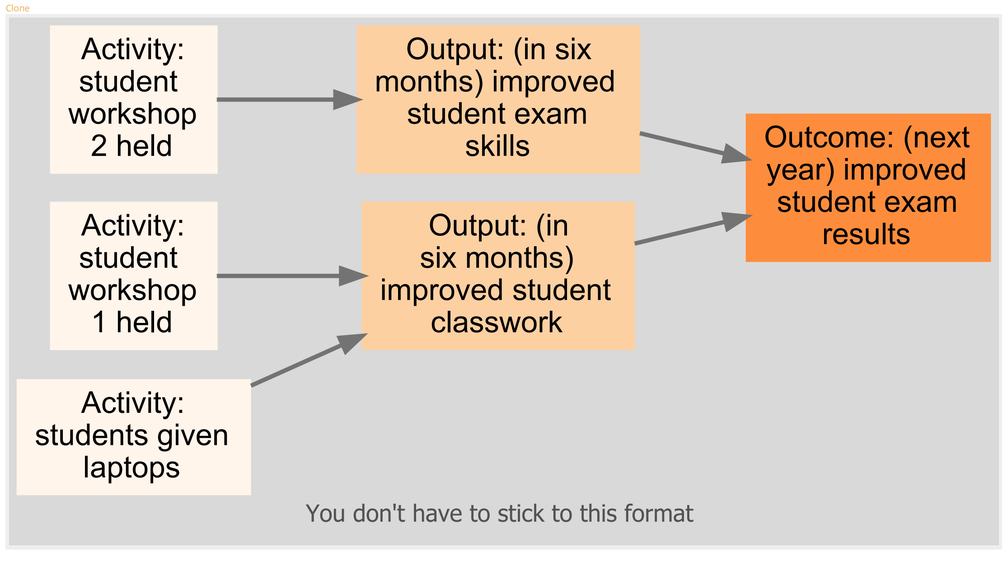
Continuous variables can stretch over time, and discrete variables can repeat, and their timings can overlap one another in different ways.

Phases: Practice

- How do we indicate phases graphically?
- Can phases dictate programming? Should they? i
- Find a Theory of Change which is arranged in phases and see if you can make it more realistic but less tidy
- OR Find a Theory of Change which is NOT arranged in phases and see if you can “tidy” it into phases!

Advantages and disadvantages?

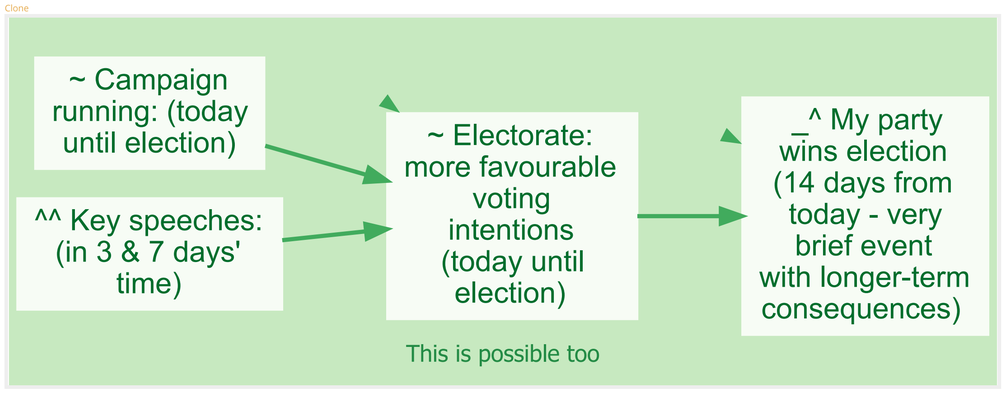
Position may tell you nothing about duration



Downstream variables such as “final Outcome(s)”:

- do they always take longer to achieve?
- do they always sustain longer?

Sometimes they do, sometimes they don’t.



Making a difference can take minutes or centuries, and the Difference can then last for minutes or centuries. Depends on many things but not a variable’s position in a Theory.

Sustainable social change is hard but not because it is so many links away in some Theory.

Loosen up your logframe

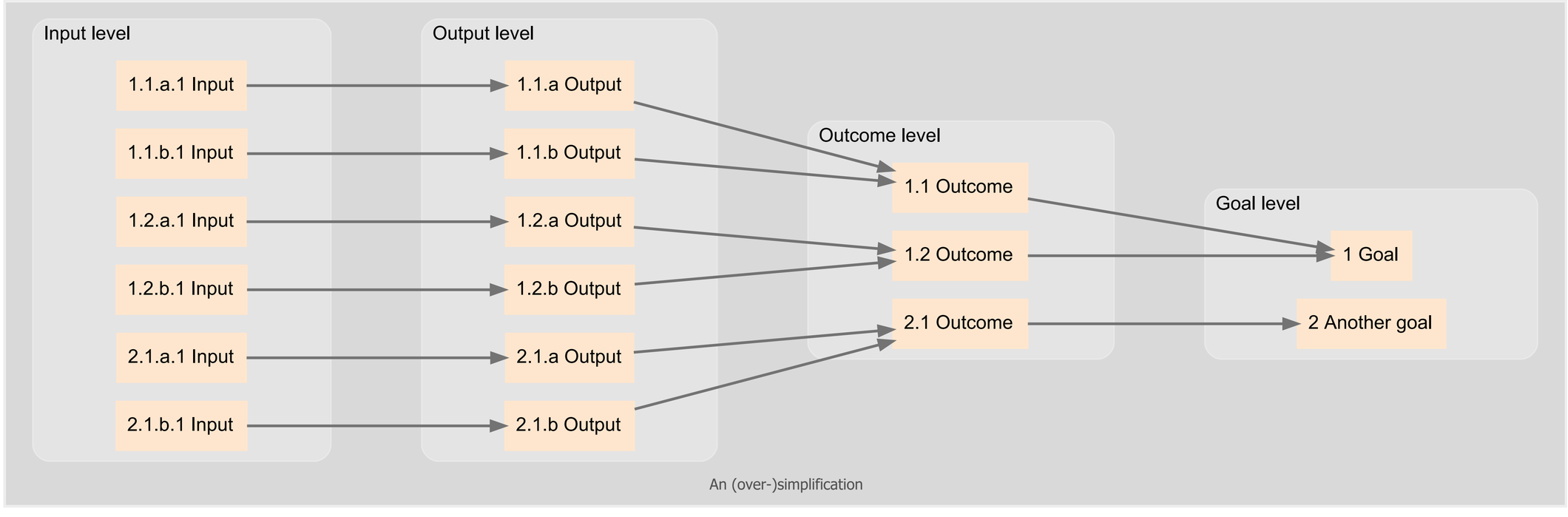
In Logical Frameworks, variables are divided up into slices: Inputs, Outputs, Outcomes etc. This is a very powerful simplification:

- The first slice is just the variables we can control (called “Activities”, etc)
- The final slice is just the variable(s) we really value (called “Goal”, etc)
- The slices “happen” one after another …

… But it’s usually just too much of an oversimplification.

Sometimes we have to stick to such a rigid format, sometimes we choose to, sometimes we don’t.

Clone



An (over-)simplification

Variables linked **by definition**

Be careful with variables *defined* in terms of others

Often, variables overlap each other *conceptually*.

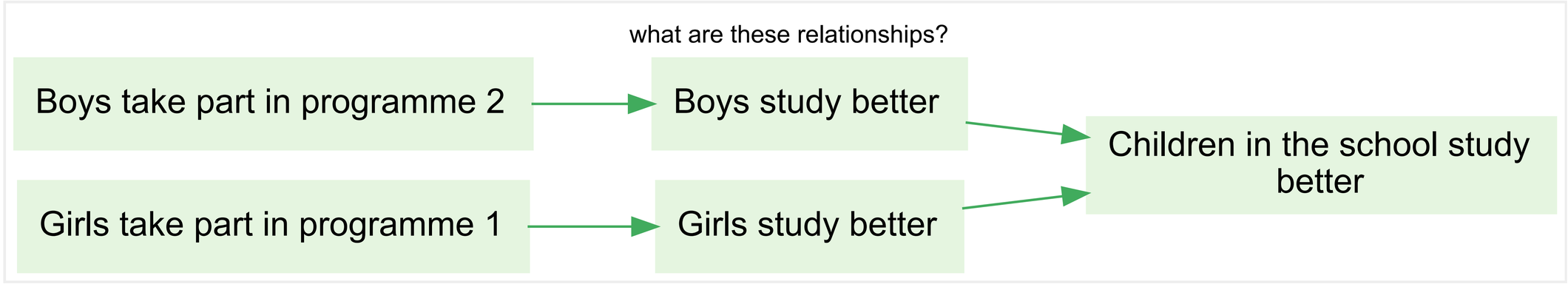
They are not separate entities. They are related to one another through their definitions.

Note

Very common, very often unnoticed, very challenging.

Spot the definitional relationship here?

Clone

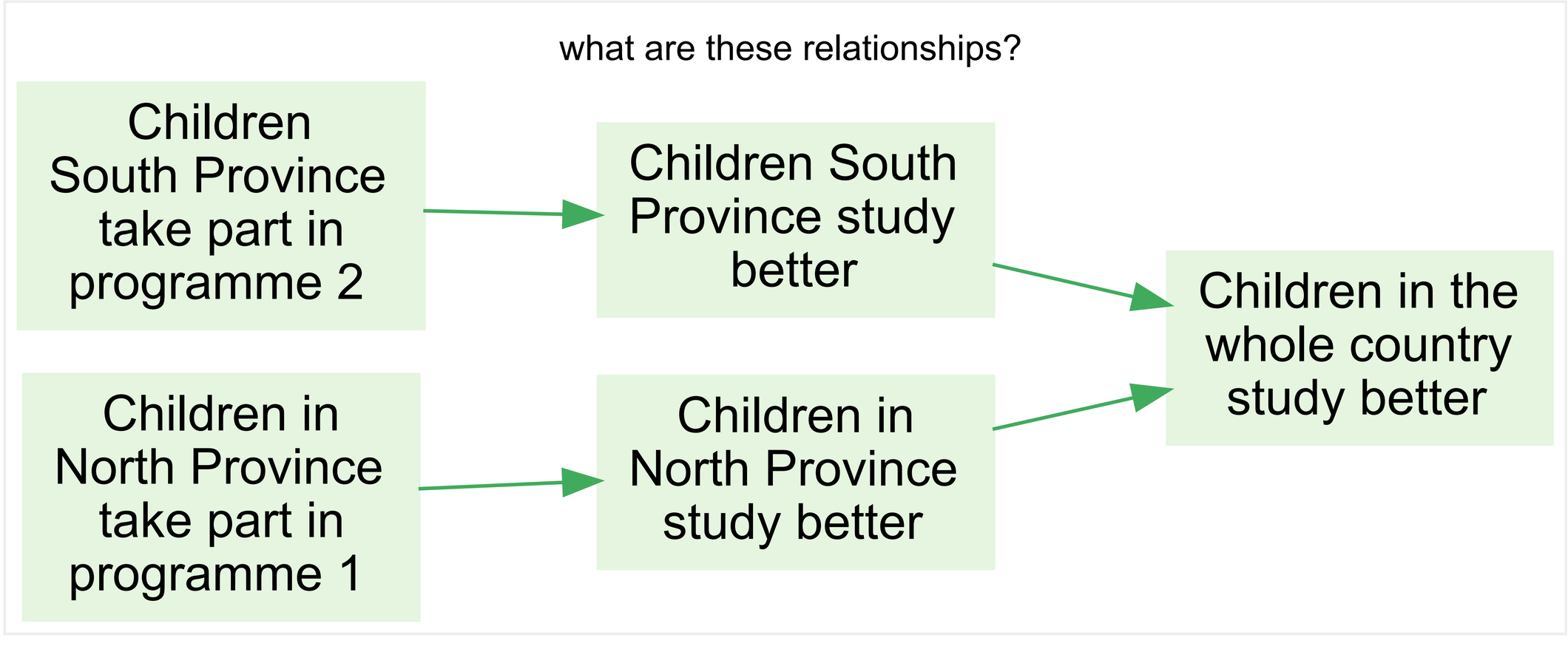


Tip: if two variables are related by definition, measuring one means (partially) measuring the other! A real headache for results-based management etc.

Variables *defined* in terms of others: more practice

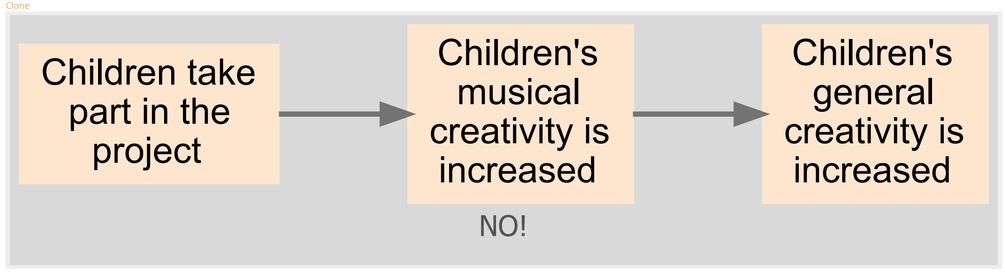
Spot the definitional relationship here?

Clone



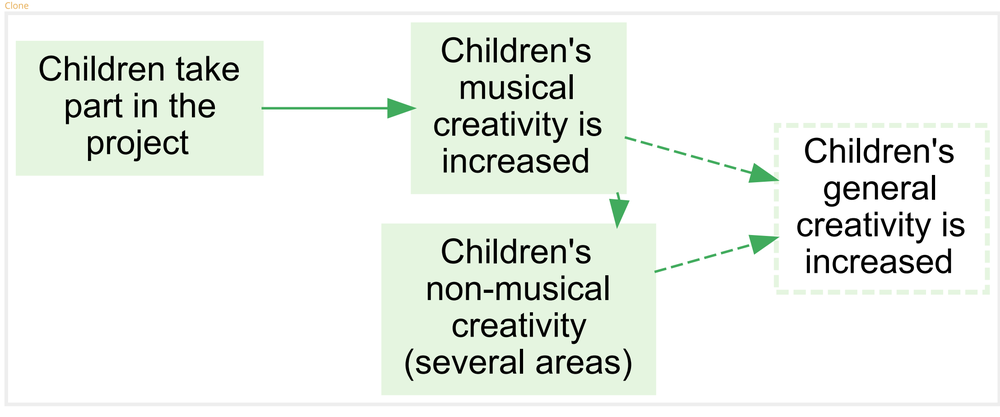
Be even more careful with variables which *partially define* others

A project's activities lead to an increase in children's musical creativity; the management want to show how that is a benefit for the donor's broader focus, which is children's creativity in general (defined as the average of creativity in several different domains, one of which is music):



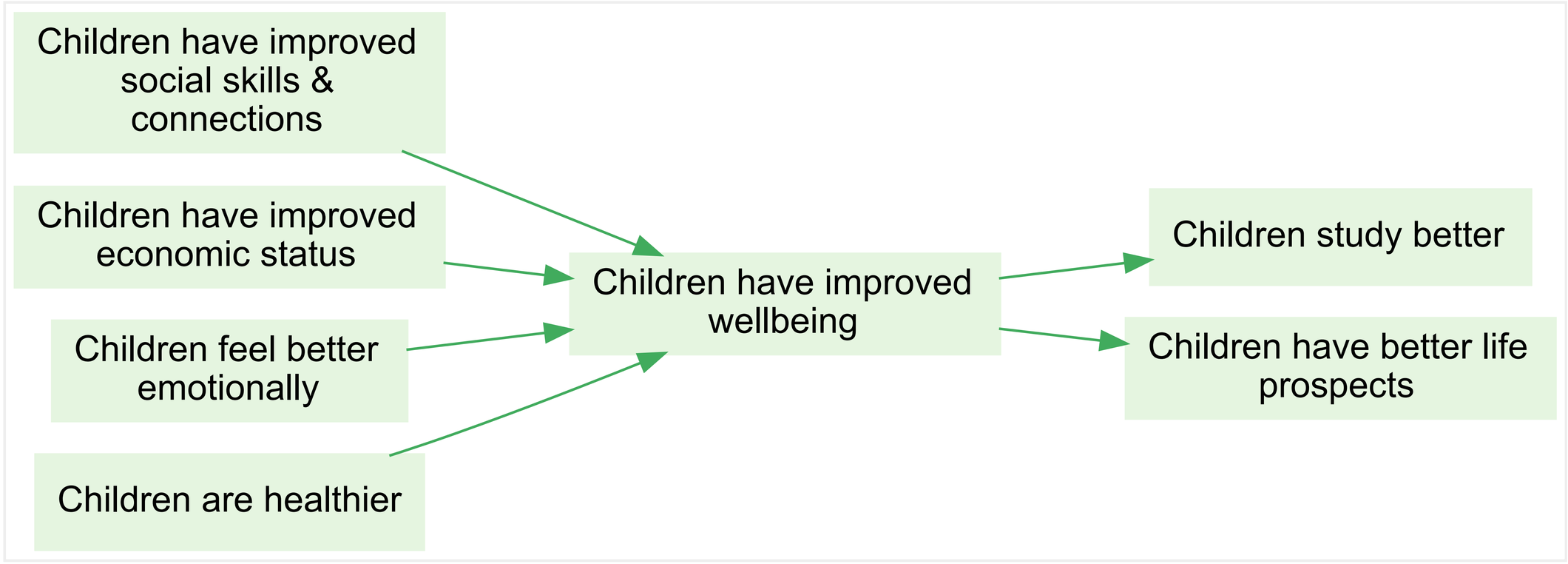
General creativity is (partially) defined in terms of musical creativity, so a *dashed* arrow should be shown along with other variable(s) too, in order to complete the definition …

But at the same time, the project claims that more musical creativity also *causes* an increase in general creativity, which would seem to result in a link which is both causal and definitional at the same time: not a good idea. Instead, we can show a link from musical creativity to the *other* part(s) of the definition, like this:



Variables *partially defined* in terms of others: more practice

Spot the definitional relationship here?



Grouping boxes

Practice

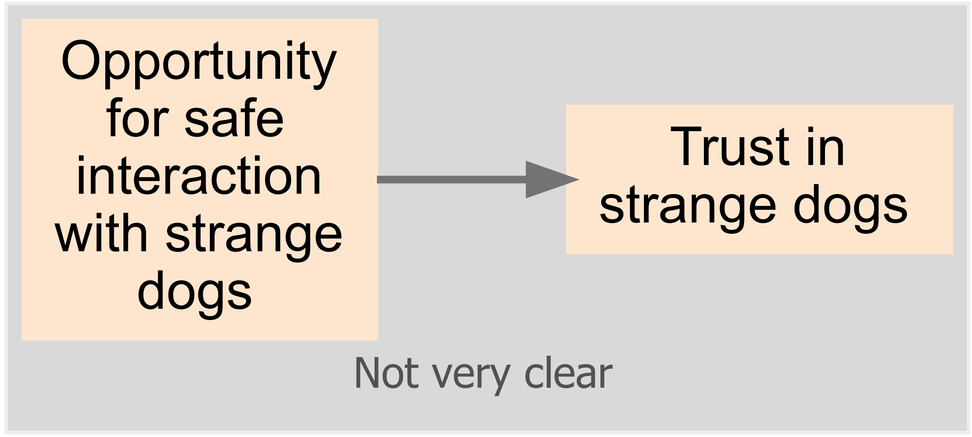
Can you find examples of Theories of Change using larger boxes which group together sets of variables?

Do show specific **features** of variables, like who (or what) they *belong to*, like this: “Child: …”

Note

We can specify *who or what* a variable *belongs to*, by **mentioning the ‘*owner: …*’ at the start of the variable name**. We can do the same for any other feature which a variable might have, e.g. time-point, country, county … and which might distinguish it from other variables.)

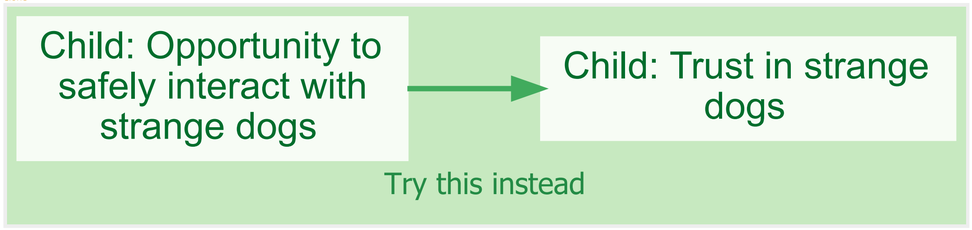
Clone



This suggests that somehow, trust in strange dogs depends on opportunities for interaction.

But *who* has this trust? *Who* has these opportunities? We can clarify like this:

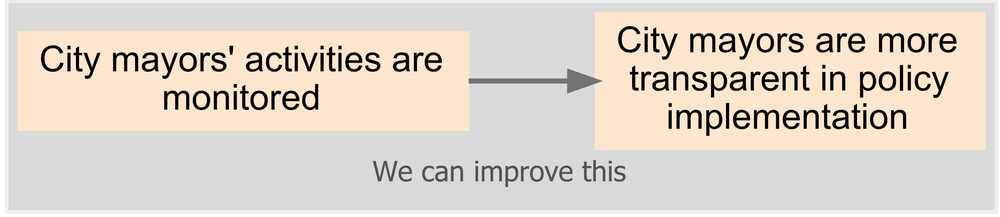
Clone



Tip

Do say who does what, and avoid passive sentences!

Clone

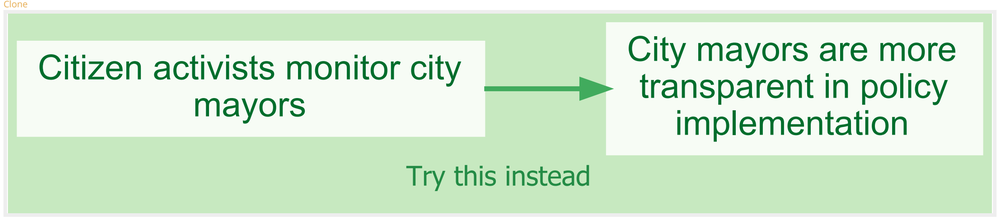


… the two variables above look like they belong to the City mayors.

But they don’t.

This is better:

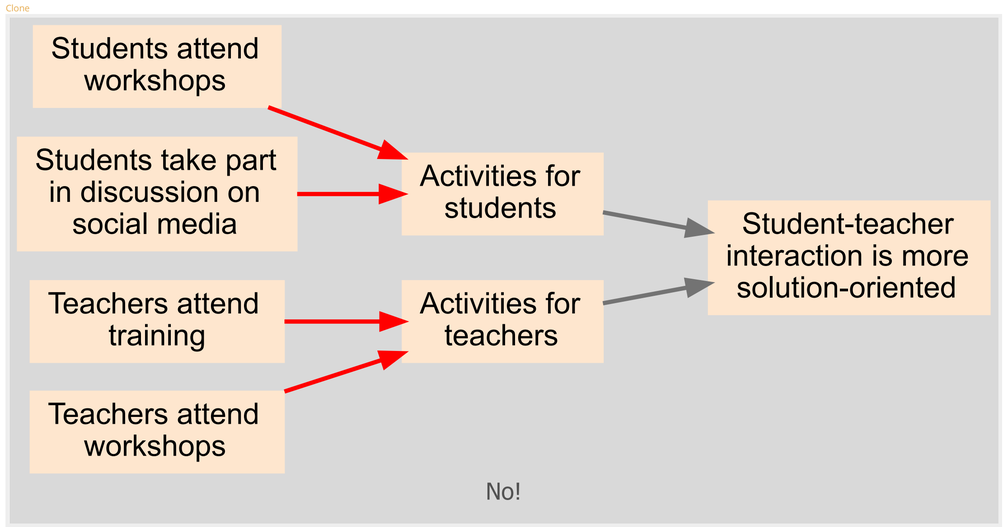
Clone



We’d first go to the *activists* to gather data on the variable.

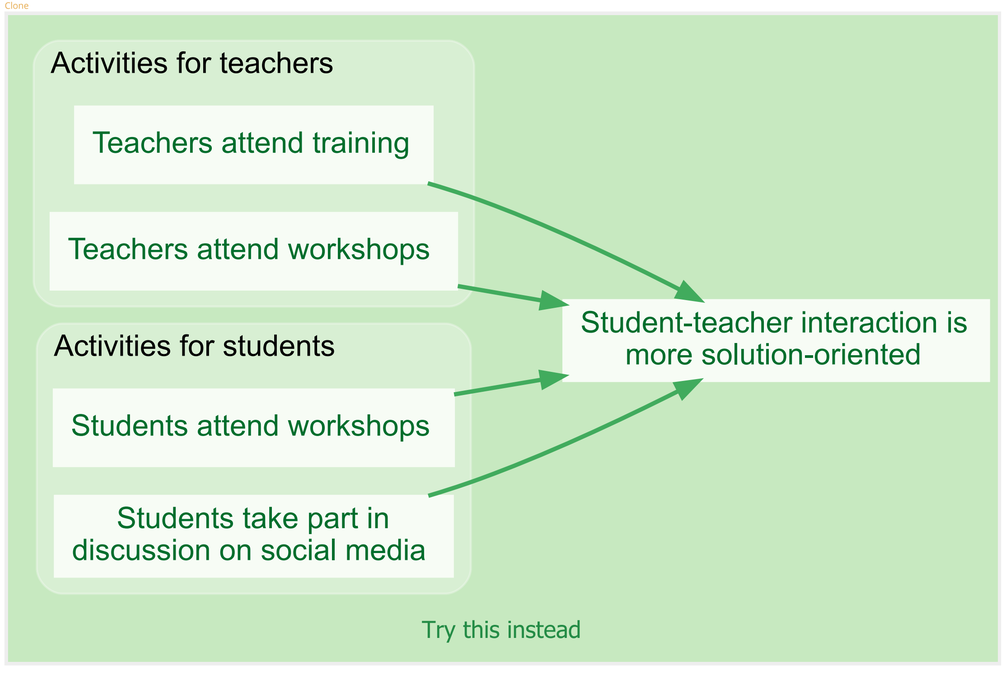
But perhaps we don’t know who is going to do the monitoring, and that’s why we want to use the passive. We could say: “*Somebody* is monitoring city mayors”. But that is a “hard-to-search” variable. Where would we go, what would we do, to gather data on that?

Don’t use a variable to group project parts



This Theory fragment tries to use “Activities for teachers” and “Activities for students” as a way to organise the diagram. But this is wrong, because all arrows in a Theory of Change should show causal influence, and here the red arrows do not.

Just just eliminate the original variables and use two grouping boxes instead:

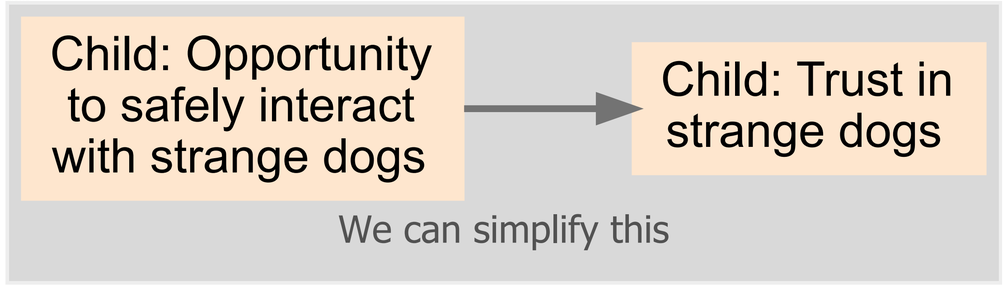


Do use **Grouping boxes** for variables with similar features

Note

When the labels of several variables begin with the same *feature* or *owner* e.g. “Child:”, they can be surrounded by a ***grouping box***; that text is then deleted from those variables and added to the box. It’s the same thing, it just saves a few words and helps structure the diagram.

Clone

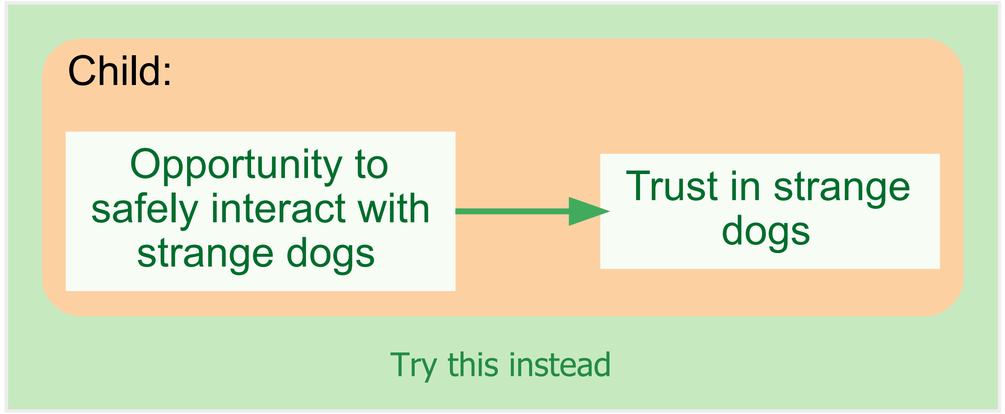


This Theory says that trust in strange dogs depends on opportunities for safe interaction.

Both variables have a special feature (`Child:...`) which tells you who or what the variable belongs to.

Other kinds of box are summarised later.

Clone



The grouping box also reminds us that these two variables belong to the *same* child. Useful.

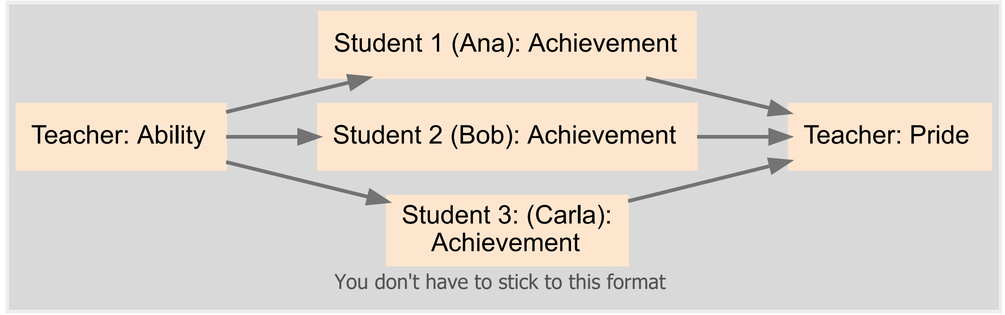
One variable **for each** person in a group

Note

The words "***For each***" at the start of a variable label, followed by a (explicit or implicit) list of people, time-points etc., says that this is actually a set of variables, one for each person or thing in the list.

The same teacher teaches three different students - his ability contributes to their achievement, and their achievement in turn influences his pride.

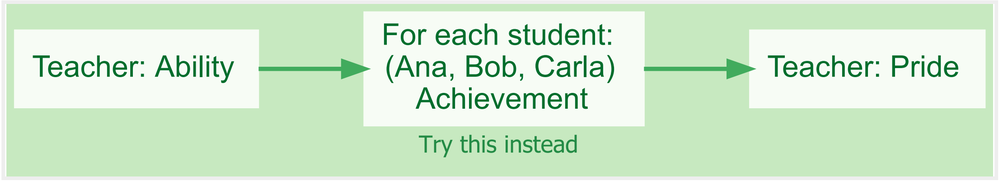
Clone



Sets of variables like this (similar to one another but belonging to different people or time points, or countries, or firms …) are very common.

Some people, especially statisticians, use the word "variable" only for *sets of variables* like these. That can make things simpler in statistics, but in Theories of Change we often have individual variables as well as sets of them. We can call these sets "for-each variables": treat them like *one* variable but remember, really it is a *set*. (This is not an established practice but should be self-explanatory.)

Clone



Theories of Change with **nested grouping boxes**

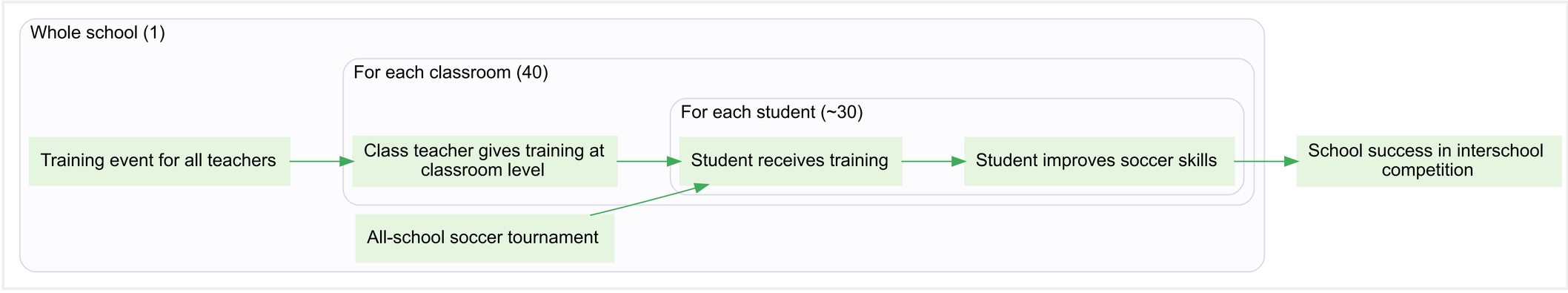
We can use *nested* grouping boxes to show quite complicated ideas very simply.

Here there is 1 interschool competition, maybe 40 classrooms, 40 teachers, 30 children in each classroom …

The nested diagram gives you a clue how to monitor or evaluate this programmel

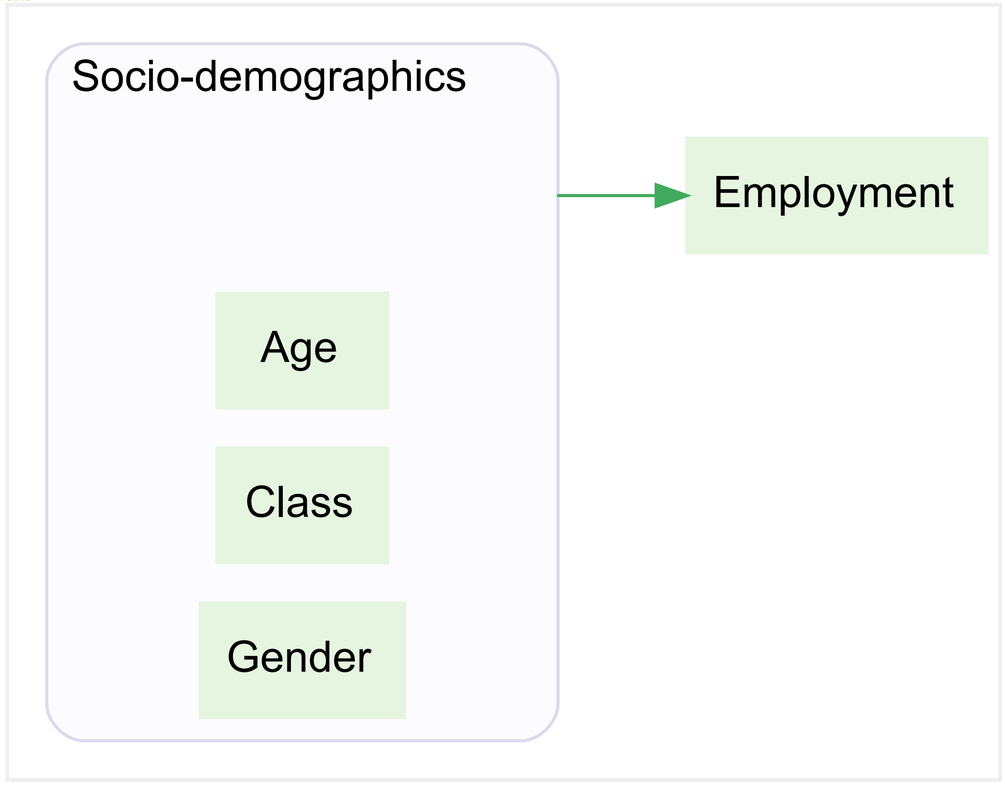
- If you are using questionnaires, how many will you need?
- If you are setting target levels for training received per student, how would you do it? For each teacher?

Clone



Arrows *from* grouping boxes

Clone

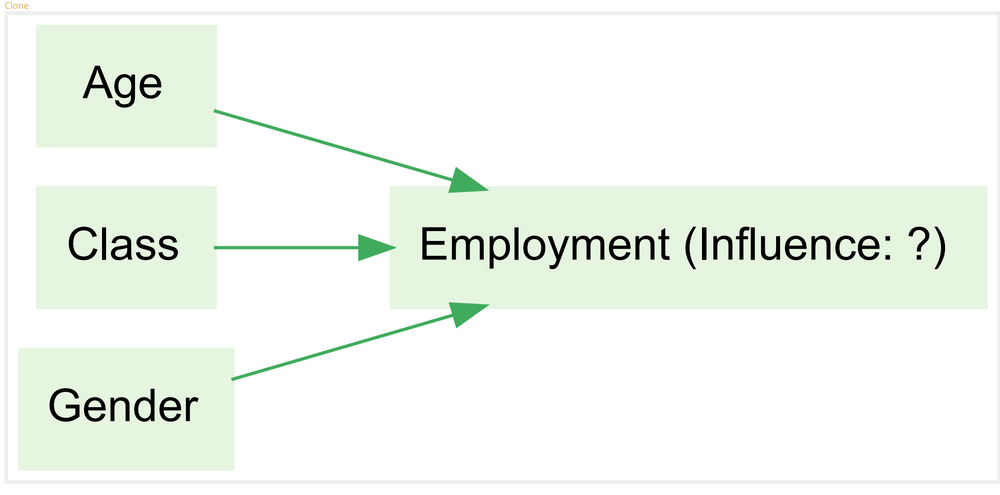


We often see this kind of diagram in the wild.

An Arrow *from* a Box is usually a kind of a shorthand which replaces Arrows from the variables it contains…

… so it is equivalent to something like this:

Clone



… but … what does this mean?

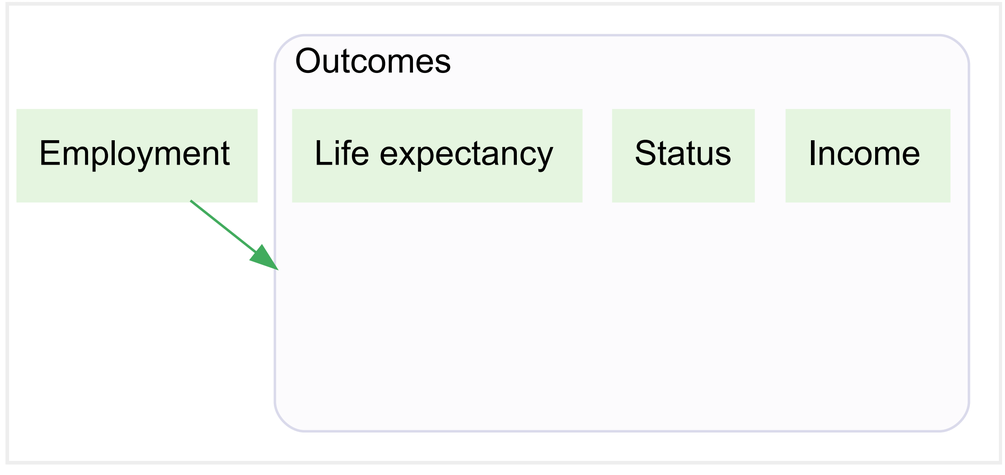
The minimal interpetation is this:

Note

An arrow from a box to a variable V means that at least *some* of the variables in the box have *some* kind of influence on V.

Arrows *to* grouping boxes

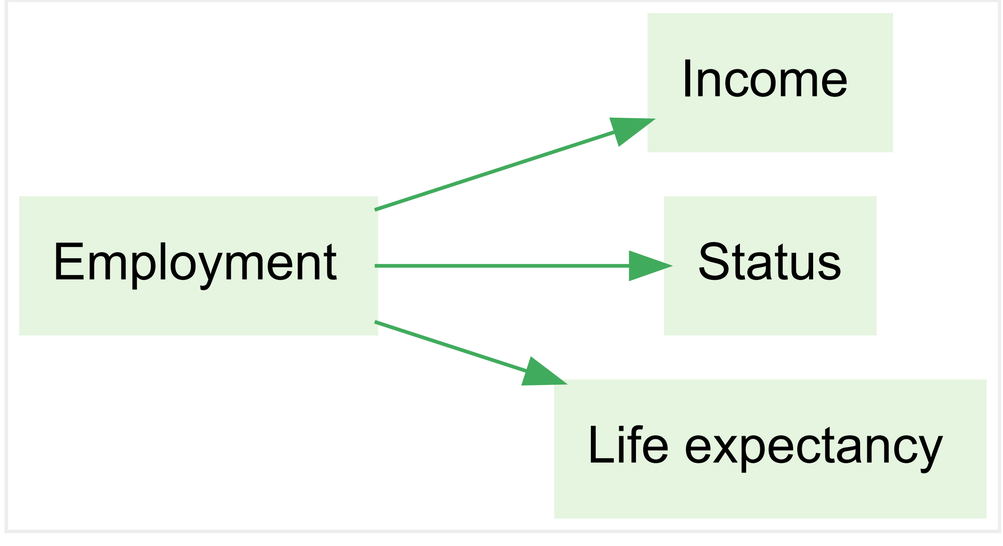
Clone



An Arrow *to* a Box is usually a kind of a shorthand which replaces Arrows to the variables it contains…

… so it is equivalent to something like this:

Clone



… but … what does that mean?

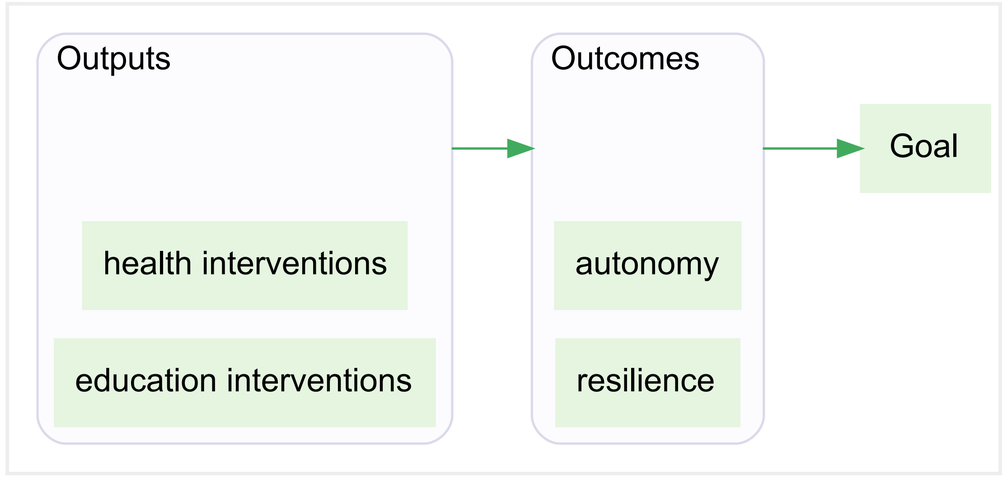
The minimal interpetation is this:

Note

An arrow to a box means that at least *some* of the variables within it are influenced in *some* way.

Arrows *between* grouping boxes

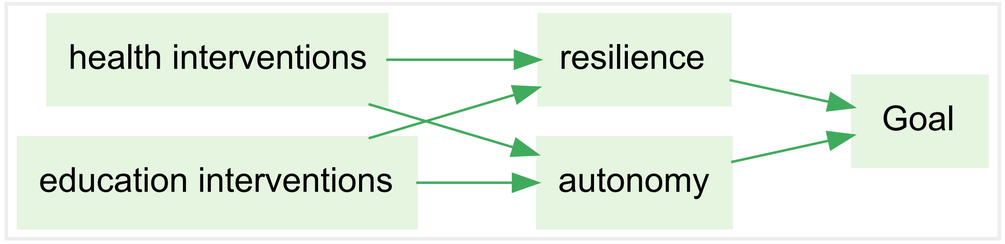
Clone



An Arrow *between* boxes is usually a kind of a shorthand, perhaps just to avoid drawing too many arrows …

… so it is equivalent to something like this:

Clone



… but … what does that mean?

The minimal interpetation is this:

Note

An arrow between boxes means that at least *some* of the variables in the first box influence at least *some* of the variables in the second, in *some* way.

Outro

Thanks!

steve@pogol.net

causalmap.app

References

DFID. 2011. “How to note: Guidance on using the revised Logical Framework.” *Department for International Development*, 1–33.

Hansen, Morten Balle, and Evert Vedung. 2010. “Theory-Based Stakeholder Evaluation.” *American Journal of Evaluation* 31 (3): 295–313. https://doi.org/10.1177/1098214010366174.

Pearl, Judea. 2000. *Causality: Models, reasoning and inference*. Cambridge Univ Press.

Van Ongevalle, Jan, Huib Huyse, Cristien Temmink, Eugenia Boutylkova, Anneke Maarse, and Double Loop. 2012. “Dealing with complexity through ‘actor-focused’ Planning, Monitoring & Evaluation,” no. November. http://www.pso.nl/files/images/Praxis Paper 26 - Dealing with complexity through PME{_}1.pdf.

App