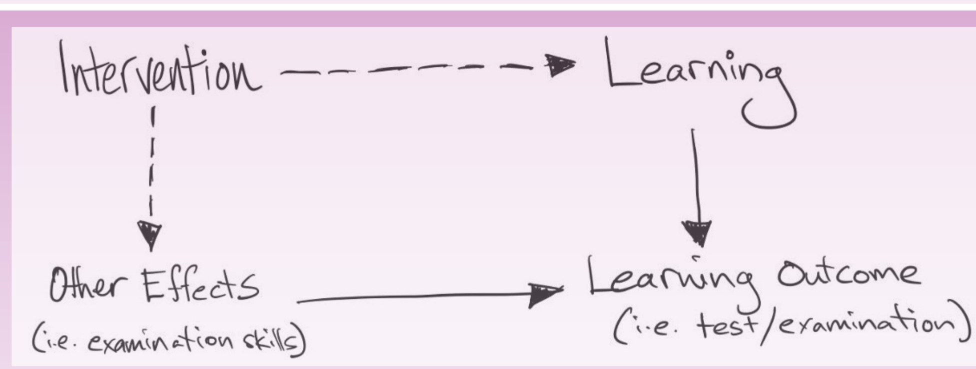
Participatory Causal Modelling of Learning Systems

Ben Hicks : Connected Intelligence Centre : University of Technology Sydney

Learning Analytics (LA) aims to improve the learning process. This necessitates a *causal* interpretation of observational data. One way to model causal structure is by using causal **Directed Acyclic Graphs (DAGs)**. The visual formalism of the model requires little technical knowledge to engage with, providing an opportunity for non-technical experts to remain engaged deep into the crafting of critical statistical assumptions about the learning system, including the importance of latent variables.



A **DAG** is a graphical representation of how variables influence each other in a system. If we think changes in A result directly in changes in B we draw $A \rightarrow B$. This forms a **D**irected **G**raph (**DG**). If it is also **A**cyclic (no loops) then we have a **DAG**, which has a precise mathematical translation to the joint probability distribution of the data which can be leveraged in a several ways.

Constructing DAGs as a collaborative thinking process

How can I collaborate with stakeholders?

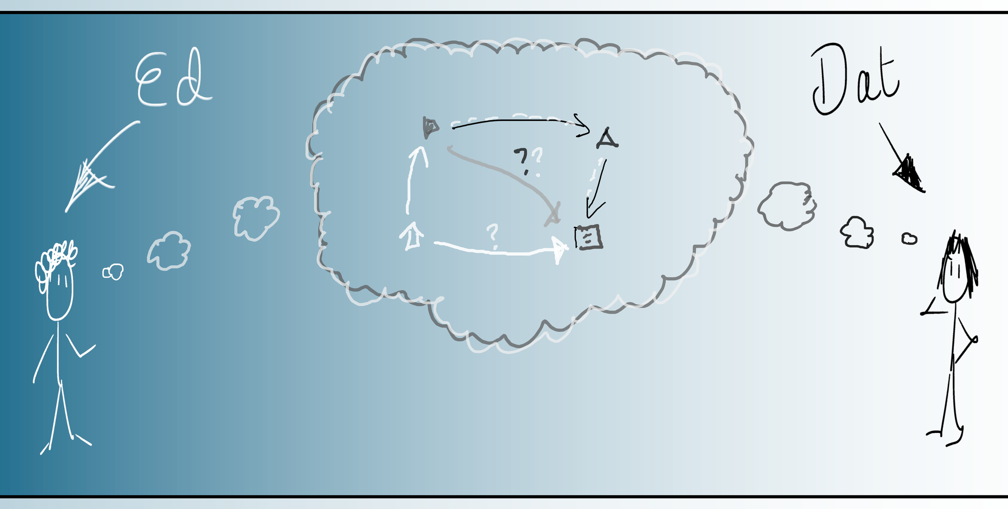
Affordances of DAGs for Learning Analytics

How could these these models help LA?

Can I just have a go at drawing one with you now?

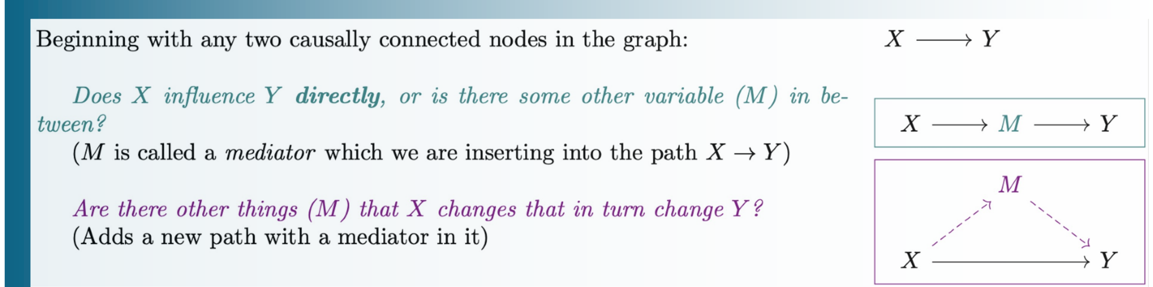
Abstract but accessible

Constructing a causal **DAG** requires minimal technical knowledge. This means it offers a way for non-technical **Ed**ucation experts to describe the important aspects of a learning system so that it can be leveraged by **Dat**a experts.



Interrogating the model, together

The graphical model can be interrogated at any point using prompts (such as the example below). This helps facilitate precision in thinking about what is important in the system.



Abstraction allows comparison

Models built from a range of stakeholders can be compared. The formal requirements of constructing a causal DAG allow for easier comparison – the price to achieve this is paid through reducing the complexity of the system so it can be modelled.

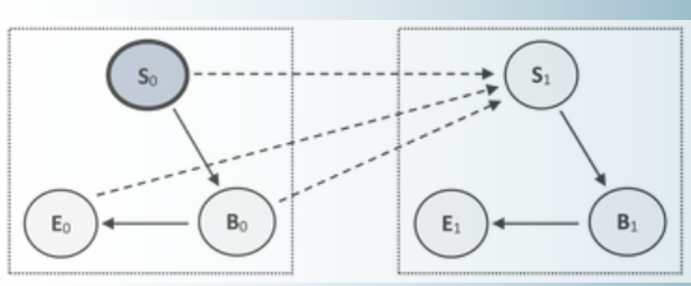
From graph to causal claim

Any given causal DAG directly translates to how we can decompose a joint probability distribution, using *do*-Calculus (Pearl, 2009). For instance, the graph $X \leftarrow Y \rightarrow Z$ describes the factorization $P(X,Y,Z) = P(X|Y)P(Y)P(Z|Y)$ and with it the conditional independence $X \perp Z|Y$.

What is important here is that each DAG describes how best to understand the **causal** relationships between variables. Typically, this is used to find causal relationships from observational data by understanding which variables to control for to minimise bias (Weidlich et al., 2022). This set of variables, for a given relationship, is known as the **adjustment set**.

Testable learning theories

Kitto et al. (2023) outline a way to use a causal DAG of a learning theory to generate a collection of implied conditional independence relationships. These can be used build tests to examine to what degree the theory holds in the observed data.



One of three Self-regulated Learning causal DAGs postulated in Kitto et al. (2023)

DAG informed dashboard design

An adjustment set of variables has implications beyond statistical models. If a dashboard shows a comparison between X and Y then an unbiased view of the data should compare these variables within levels of the adjustment set. This could be implemented with filters, slicing or aggregation.

PhD publication plan

This poster describes a slice of the work towards my PhD. Feedback welcome, and cases to take part in collaborative modelling *very* welcome!

A paper (not yet written) to help frame this kind of thinking.

Describing bounds for locally linear causation

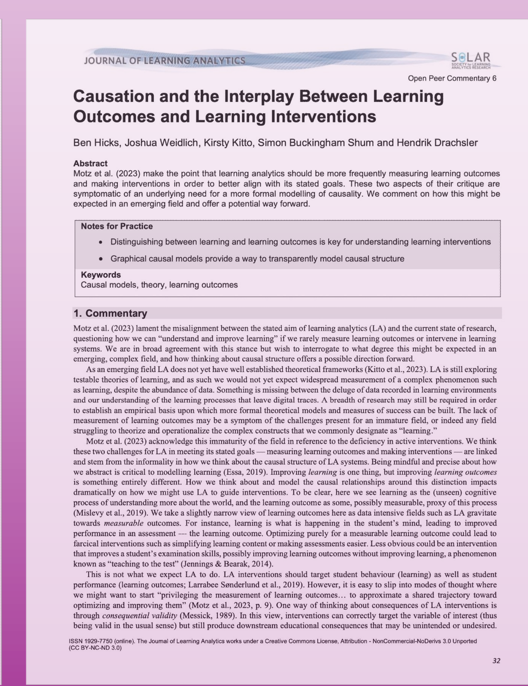
When we model a system using a DAG we are describing a system with linear causation – effects are additive. In reality most learning systems will exhibit non-linear causation if we look closely enough.

This (more philosophical) paper will argue that the key is to describe the boundaries within which the system is sufficiently describe with linear causation (such as a DAG).

A paper and a commentary introducing using DAGs as a way to think about education system.


Thinking with causal models: A visual formalism for collaboratively crafting assumptions

Ben Hicks, Kirsty Kitto, Leonie Payne, Simon Buckingham Shum


JOURNAL OF LEARNING ANALYTICS

Causation and the Interplay Between Learning Outcomes and Learning Interventions

Ben Hicks, Joshua Weidlich, Kirsty Kitto, Simon Buckingham Shum and Hendrik Drachsler

Collaborative Causal Modelling with Learning Experts

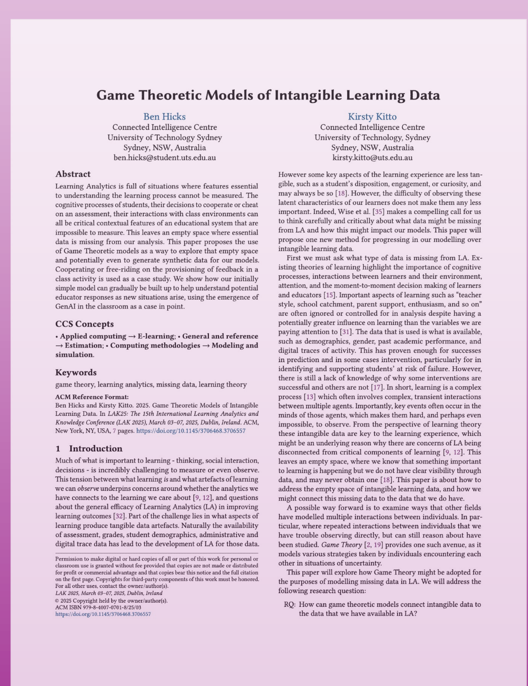
A paper (in progress) building on the work presented in this poster. I hope to find out:

RQ1: How well do these models represent an experts understanding of a learning system?

RQ2: What are the affordances of graphical causal models to develop thinking and a shared understanding?

Interested in modelling a system for this work? e: ben.hicks@student.uts.edu.au

An idea of what to do when non-linear causation is required.


Game Theoretic Models of Intangible Learning Data

Ben Hicks, Kirsty Kitto

References

Ferguson, R., Khosravi, H., Kovanović, V., Viberg, O., Aggarwal, A., Brinkhuis, M., ... & Yan, V. X. (2023). Aligning the goals of learning analytics with its research scholarship: An open peer commentary approach. *Journal of Learning Analytics*, *10*(2), 14-50. – "Causation and the Interplay Between Learning Analytics and Learning Interventions" is part of this.

Hicks, B., Kitto, K., Payne, L., & Buckingham Shum, S. (2022, March). Thinking with causal models: A visual formalism for collaboratively crafting assumptions.

In *LAK22: 12th international learning analytics and knowledge conference* (pp. 250-259).

Hicks, B., Kitto, K. (2025, March). Game Theoretic Models of Intangible Learning Data. In *LAK25: The 15th International Learning Analytics and Knowledge Conference.*

Kitto, K., Hicks, B., & Buckingham Shum, S. (2023). Using causal models to bridge the divide between big data and educational theory. *British Journal of Educational Technology*, *54*(5), 1095-1124.

Pearl, J. (2009). *Causality* (2nd ed.). Cambridge: Cambridge University Press.

Weidlich, J., Gašević, D., & Drachsler, H. (2022). Causal inference and bias in learning analytics: A primer on pitfalls using directed acyclic graphs. *Journal of Learning Analytics*, *9*(3), 183-199.