MODULE 4: DECISION-MAKING AND EVALUATION

CT ACADEMY | DATA ACTION LAB

10. EVIDENCE-INFORMED DECISION-MAKING

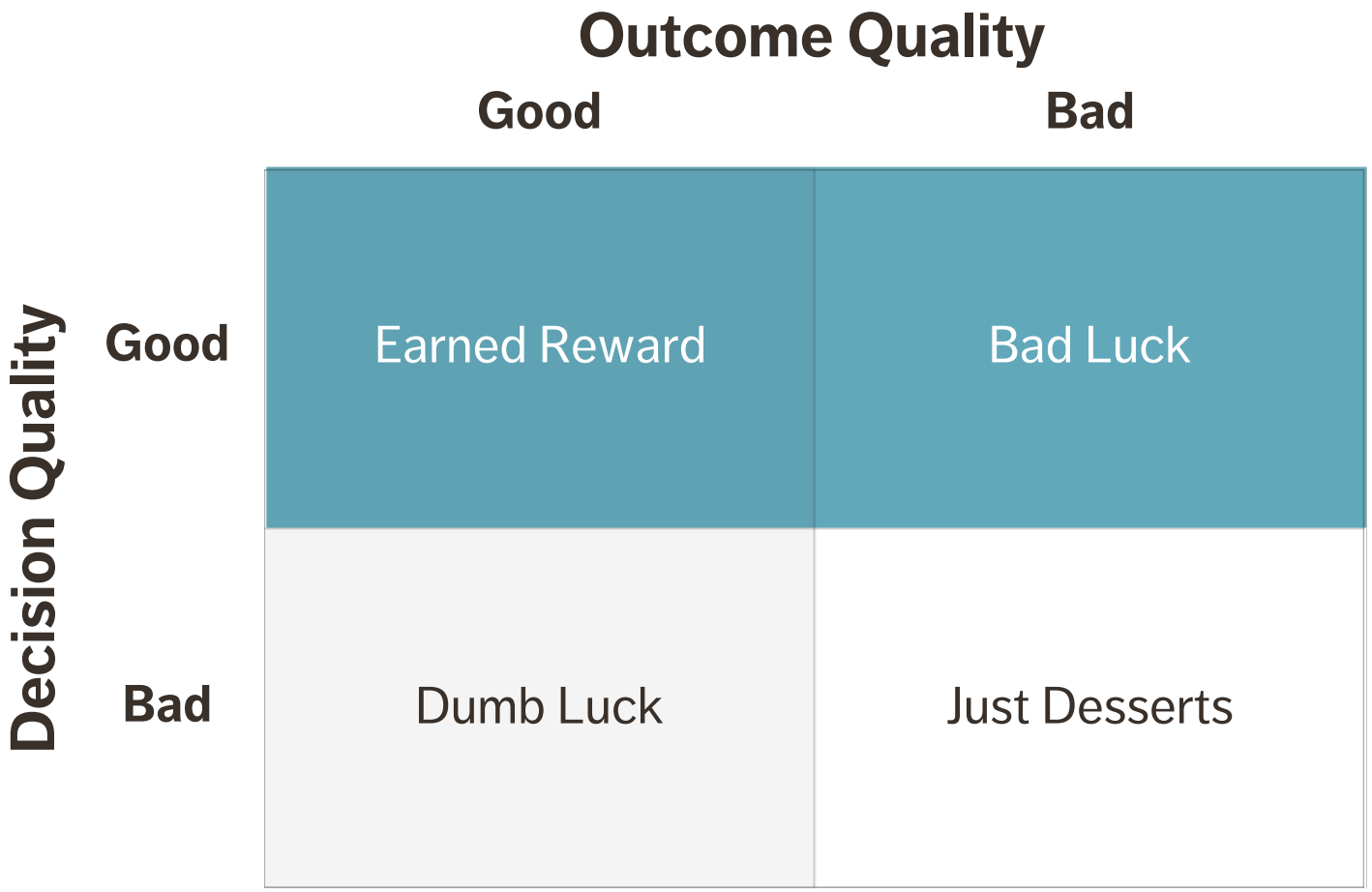
DECISION-MAKING AND EVALUATION

LIFE TURNS ON TWO THINGS

Luck
(outside your control)

Decisions
(within your control)

RESULTING



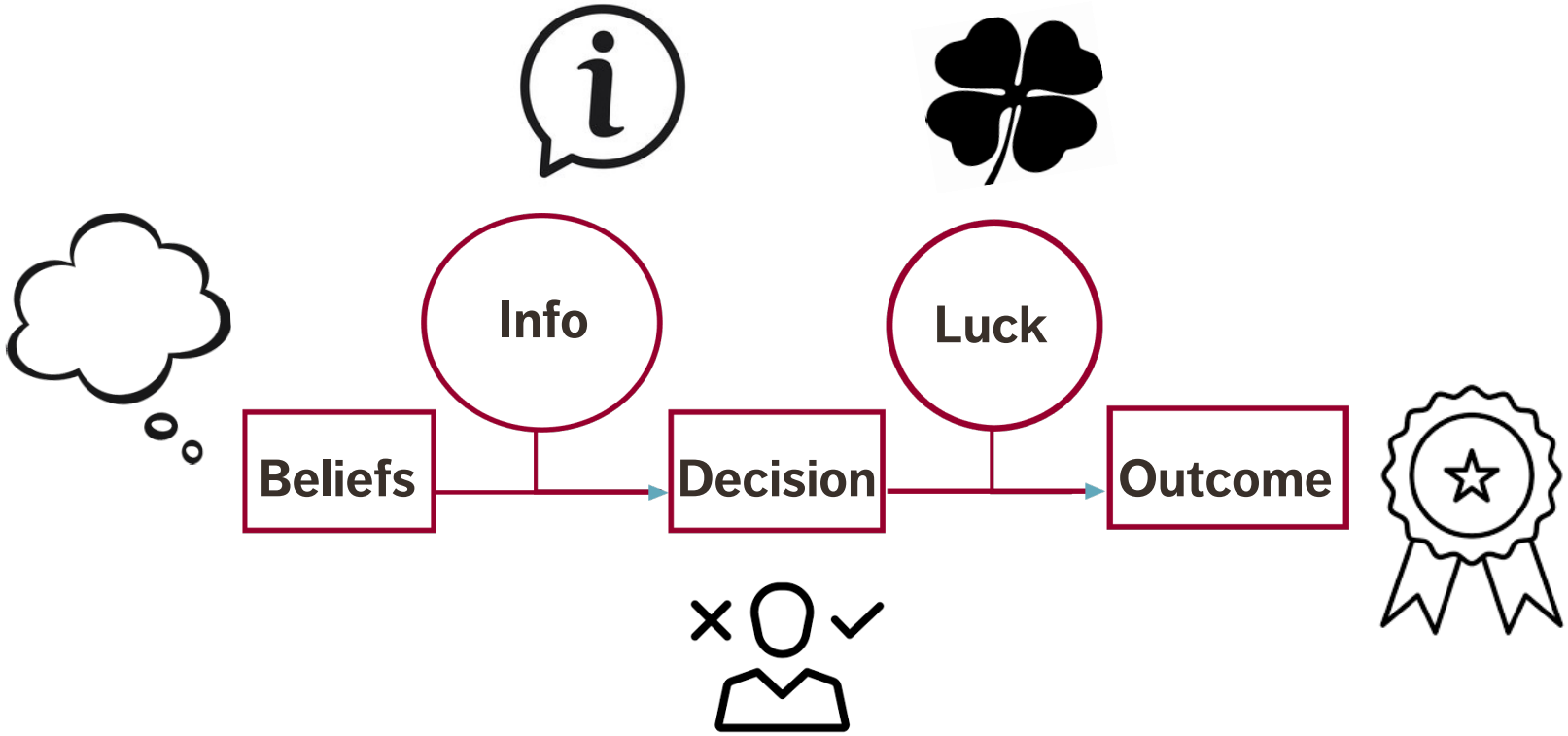
Poker players warn of "**resulting**": assessing the quality of a decision based **solely on its outcome**.

Problem: "resulting" makes us lack compassion for ourselves and for others.

Bad outcomes do not necessarily equate to **poor decision-making**.

Exercise: find examples for each quadrant.

LUCK AND INFORMATION (REPRISE)



WHAT IS ANALYSIS IN THE GOC?

Drawing conclusions?

Gathering and presenting evidence (pivot tables)?

Providing options?

Providing opinions/hypotheses/beliefs/recommendations?

Pushing your agendas?

TYPICAL ANALYSIS ACTIVITIES

Analysis is an activity done *to* something.

We analyze the **situation** or the **problem:**

- Gathering facts and evidence
- Summarizing the facts
- Reviewing and evaluating facts
- Combining facts
- Generating new statements or hypotheses
- Breaking down concepts into simpler concepts

Common theme: facts!

- Building up more complex concepts from simpler concepts
- Defining concepts
- Using reasoning to derive new facts
- Determining if statements are true (facts) or false
- Determining how confident we are about a statement being true or false

CRITICAL THINKING

Critical thinking (supported by analysis, reasoning, inference) is important.

Using **rigor** and **methods:** also important.

This is not a course on logic, BUT…

ultimately reasoning activities are all about getting at the (a?) **truth** – having enough true facts at your fingertips to keep you from making bad decisions.

SYSTEM OPTIONS (REVISITED)

System 1: automatic decisions (“gut-feeling”)

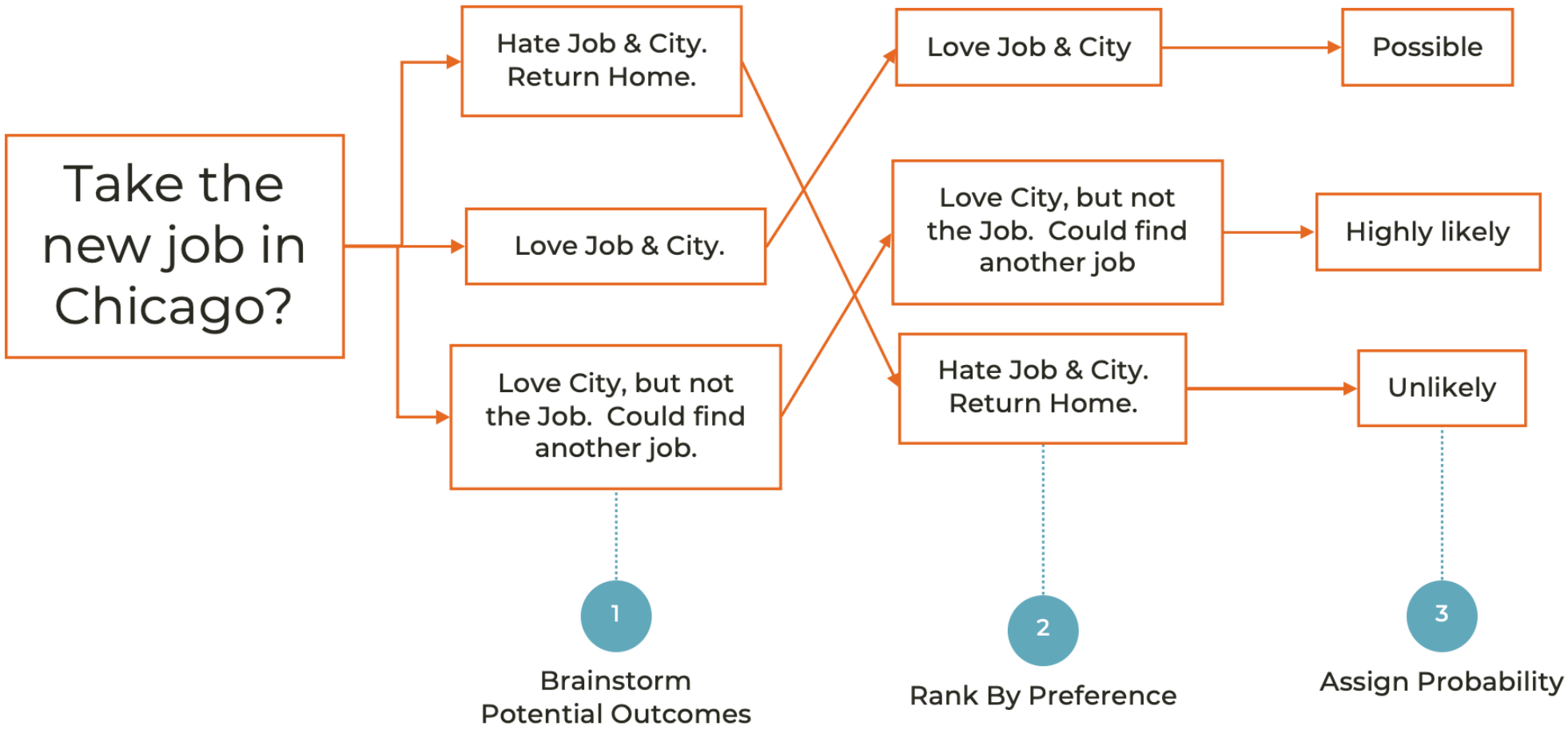
System 2: scientific method (“controlled environment”)

But also…

System 2: everything else (i.e., your job)

We need to use all reasoning types, with emphasis on what is **plausibly** true.

ANNIE DUKE'S GUIDE TO DECISIONS



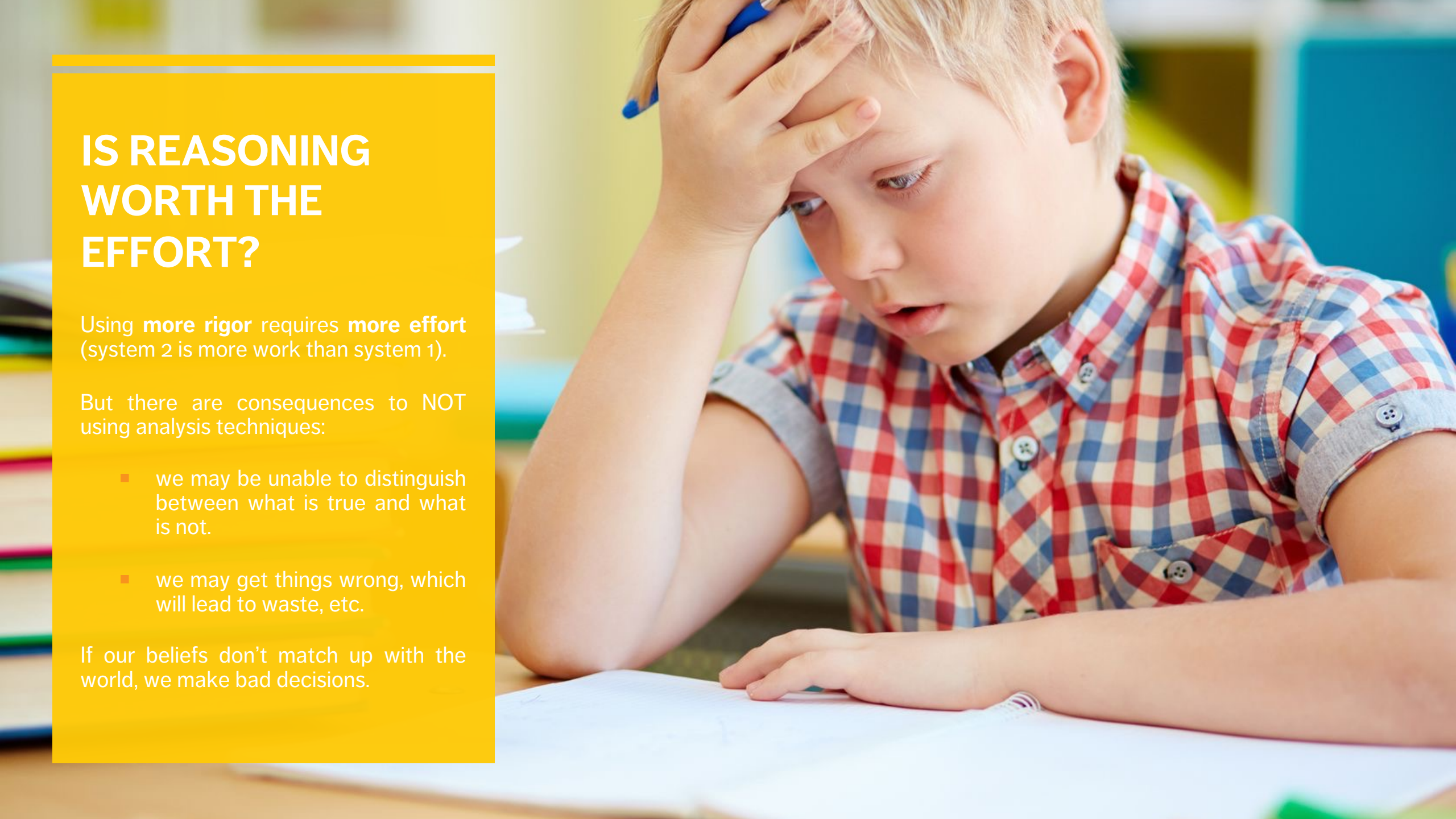
IS REASONING WORTH THE EFFORT?

Using **more rigor** requires **more effort** (system 2 is more work than system 1).

But there are consequences to NOT using analysis techniques:

- we may be unable to distinguish between what is true and what is not.
- we may get things wrong, which will lead to waste, etc.

If our beliefs don't match up with the world, we make bad decisions.

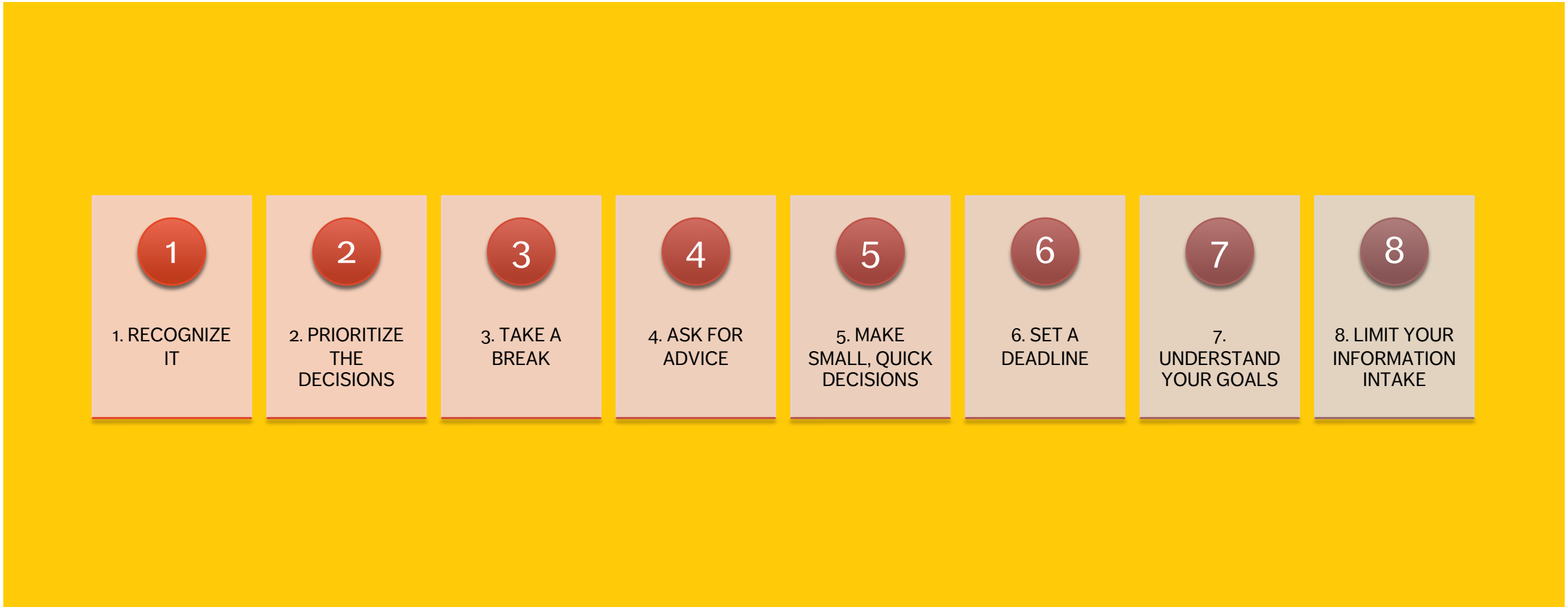


ANALYTICS: ANALYSIS PARALYSIS

We may need to make a decision with **less than complete** information. What is the risk of **not deciding** vs. the risk of making a less-than-perfect decision?

Analysis paralysis is caused by overthinking a situation and worrying about the outcome at the expense of decision-making. It is **perfectionism**, taken to an extreme (not good).

“It doesn’t matter in which direction you choose to move when under a mortar attack, just so long as you move. Decisions are never final for the simple fact that change is never absolute. Rather, change is ongoing. To stay competitive and progress at the rate of change requires adaptive decisions that can be iterated and improved upon on the fly.” [Jeff Boss, Forbes]



ANALYTICS: AVOIDING ANALYSIS PARALYSIS

WHAT IS YOUR ANALYSIS GOAL?

Do you want to:

- carry out actions based on **what is in your data**?
- gain a deeper understanding of something **specific**? (specific individual(s)? specific group(s)?)
- come to **general** conclusions that extend beyond the specific?

Local vs. Global

Here vs. Everywhere

Past/Present vs. Future

Situational Awareness vs. Contingency Planning

TYPICAL SEQUENCE OF REASONING

1. Start with premises: knowledge/assumed true beliefs

2. Carry out reasoning

3. Reach conclusions: new knowledge/potentially true beliefs

This approach can also be used to generate a **logical argument**

YOU ALREADY DO THIS, INFORMALLY

Suppose I pause at the top of a set of stairs with an armful of stuff. What argument might be playing out **unconsciously**...?

- IF I have too many things in my hands, THEN I can't hold on to the railing going down the stairs
- IF I don't hold onto the railing, THEN I might stumble
- IF I stumble, tTHEN I might drop my stuff to stop myself falling down the stairs
- IF I drop my stuff, THEN some of it might break
- IF my stuff breaks, THEN then I'll be sad

CONCLUSION: I currently have too many things in my hands.

YOU ALREADY DO THIS, INFORMALLY

How do we act on such a conclusion?

- Because I have too many things in my hands, I might drop them on the stairs and break some of them.
- This would make me sad :(
- Instead, I could choose to make two trips so I can hold on to the railing.
- If I make two trips instead of one, this doesn't mean I won't drop something and break it, but it does increase my confidence that I won't drop something.

Decision and Action: "I will split the load into two parts and make two trips." or "Nah, that's not likely to happen; I'll tough it out and make one trip."

REASONING VULNERABILITIES

Conspiracy theories **mindset:** individuals jumping to invalid conclusions because they cannot reason and/or recognize bad evidence.

Is it **plausible** that there are microchips in the COVID vaccine? How would you gauge the degree of plausibility?

Thought exercise: you are given a stable of deductive logicians and a stable of debaters to help you make decisions. Which would you chose? Is any of them of use to you?

FORMAL RIGOROUS REASONING

Mathematicians and philosophers developed **formal methods** to bring **rigour** to informal reasoning – we can think of these as reasoning tools.

Using these tools increase our chances that we will **end up with true statements**, in which we can feel confident (if not always 100% so).

Without rigor, we can succumb to **biased reasoning**, which prevents us from reaching **true conclusions** or **justified conclusions**.

FORMAL REASONING TECHNIQUES

INDUCTIVE, PLAUSIBLE, DEDUCTIVE, ABDUCTIVE, ANALOGICAL REASONING

FURTHER SPECIALIZED TECHNIQUES: SCIENTIFIC METHOD, STATISTICAL REASONING, MATHEMATICAL AND COMPUTER MODELLING

EVIDENCE-BASED ANALYSIS, WHICH MAY BE MORE-OR-LESS TECHNICAL

FORMAL REASONING TECHNIQUES

Reasoning strategies:

- deducing new facts from existing facts (**deductive** reasoning)
- generalizing from examples (**inductive** reasoning)
- reasoning to the best explanation (**abductive** reasoning)
- using analogies and models (**analogical** reasoning)

These last three techniques are examples of **plausible reasoning** – you are not guaranteed to reach the truth, but you are increasing your level of certainty.

PLAUSIBLE REASONING

Consider the following scenario [Jaynes, 2003]:

- you are walking down a deserted street at night;
- you hear a security alarm, look across the street, and see a store with a broken window, and
- someone wearing a mask crawls out of the broken window with a bag full of smart phones.

What might a first system 1 conclusion be?

What might a system 2 conclusion be?

PLAUSIBLE REASONING

Say we concluded that the person crawling out of the store is stealing merchandise from the store. How do we come to that conclusion? It **cannot** come from a logical deduction based on evidence.

Indeed,

- the person crawling out of the store **could** have been its owner who,
- upon returning from a costume party, realized that they had misplaced their keys
- just as a passing truck was throwing a brick in the store window,
- triggering the security alarm, after which
- the owner then went into the store to retrieve items before they could be stolen,
- which is when you happened unto the scene.

The original conclusion is not **deductive**, but it is at least **plausible**.

DEDUCTIVE VS. PLAUSIBLE REASONING

Plausible reasoning:

If A is true, then B is more plausible

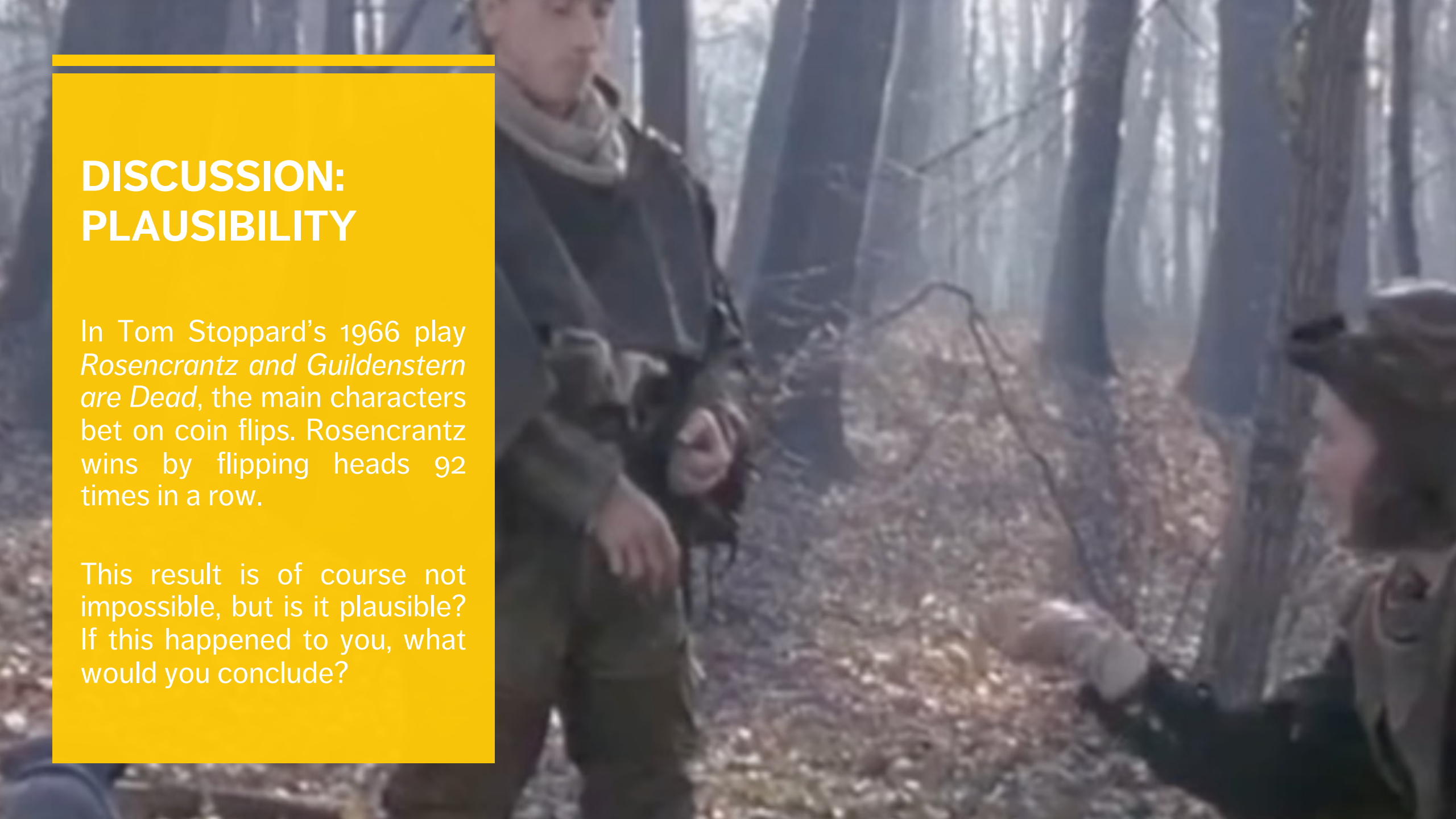
B is true

A is more plausible

If "the person is a thief" (A is true), we would not be surprised to "see them crawling out of the store with a bag of phones" (B is plausible).

We do "see them crawling out of the store with a bag of phones" (B is true).

Thus, we would not be surprised if "the person were a thief" (A is more plausible).

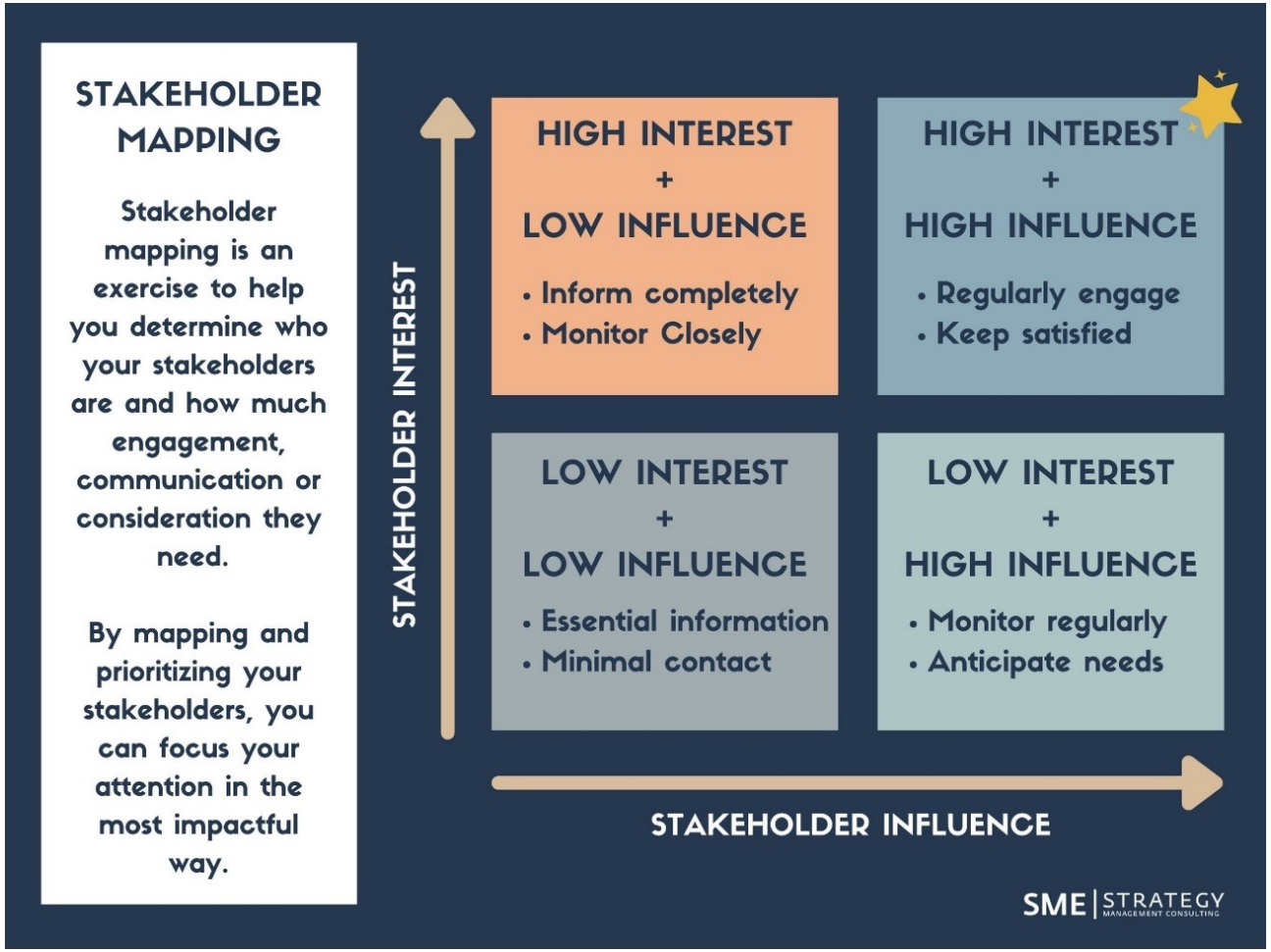


DISCUSSION: PLAUSIBILITY

In Tom Stoppard's 1966 play *Rosencrantz and Guildenstern are Dead*, the main characters bet on coin flips. Rosencrantz wins by flipping heads 92 times in a row.

This result is of course not impossible, but is it plausible? If this happened to you, what would you conclude?

BUILD RELATIONSHIPS WITH STAKEHOLDERS



11. EVALUATING OUTCOMES

DECISION-MAKING AND EVALUATION

WHAT ARE OUTCOMES?

When we talk about "**evaluating outcomes**", we need to be specific about what each word means.

The GoC has policies and directives that identify both words (see the next few slides); to be precise, our definitions will borrow from the approach used to create a "**logic model**" (more details on that soon).

Evaluation activities must be **well-defined** and grounded in **high-quality** data.

DATA STRATEGY REFRESH – MEASURING OUTCOMES

Specific high-level desired outcomes are identified in the GoC 2023-2026 Data Strategy Refresh:

- effective, equitable, ethical, and inclusive services, programs, and policy
- trusted and accountable government
- greater public value from data
- enhanced evidence-informed decision-making
- support for Indigenous data sovereignty

Lesson: outcomes from initiatives on which we are working should align with one or more of the items above.

OUTCOME MANAGEMENT GUIDE & POLICY ON RESULTS

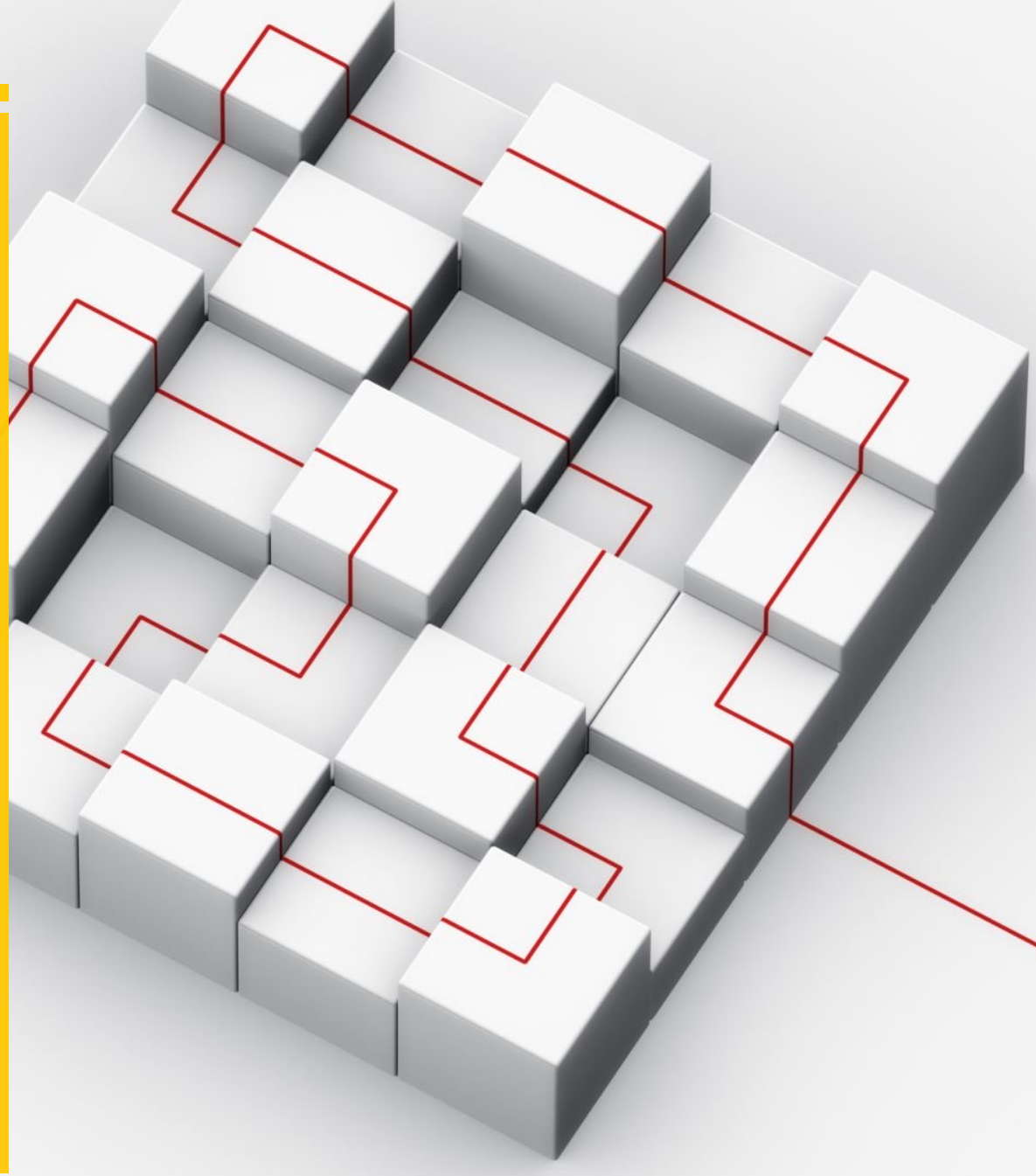
TBS's "Outcome Management" guide is aimed specifically at **projects** and **portfolio management**. Good guidelines are referenced in the document, including the use of **logic models**.

The GoC also has a "Policy on Results", which includes the "Departmental Results Framework". The purpose of this document and program is to provide departments with guidance on how to **best measure the success** of **programs**, **initiatives**, **projects**, and **ongoing operations**.

LOGIC MODEL

These references don't really provide us with a sensible way of defining outcomes for our own use, however.

It is easier to use an approach called a **logic model**, which provides a framework to logically think though various aspects of our work and to lay out a series of sensible **connected evaluations**.



LOGIC MODEL

A **logic model** (LM) is a graphical representation that outlines the sequence of **actions**, **resources**, and **intended outcomes** of a program, demonstrating how activities are **connected** to the results they are supposed to achieve.

The key components of a logic model are:

- **inputs:** resources, contributions, and investments that go into the program, e.g., time, money, staff, volunteers, equipment;
- **activities:** what the program does with the inputs to fulfill its mission, e.g., services provided, programs conducted, products developed;
- **outputs:** the direct products of program activities, e.g., number of workshops held, participants served;
- **outcomes:** the specific changes in program participants' behavior, knowledge, skills, status, and level of functioning, e.g., increased knowledge of a subject, improved skills, and
- **impact:** the broader changes that occur as a result of the program, e.g., long-term effects on the community or system.

LOGIC MODEL

In such a model, outcomes need to be **measurable** and quantitively "**link**" outputs to impacts.

What does this look like in a financial context? For example, a health program might invest funds into public education to stop smoking.

In this case we would have:

- **outputs:** program smoking cessation advertising in print and online;
- **outcomes:** a measurable % of a city population reading the advertising, and
- **impact:** a measurable decrease in smoking (measured through surveys & cigarette sales).

LOGIC MODEL

Inputs	Activities	Outputs	Outcomes	Impacts
▪ $500k ▪ 2 FTE ▪ Project oversight	▪ Advert design ▪ Liaise with retail ▪ Website and design ▪ Campaign design	▪ Print adverts ▪ Online adverts (delivered through Google advertising)	▪ Retail establishments putting up print adverts ▪ Adverts delivered online to relevant search criteria	▪ Fewer people smoking within city boundaries

WHAT IS EVALUATION?

Evaluation is the **systematic process** of assessing the **effectiveness** of an activity. With the LM as a **reference point**, we see that evaluation means **different things at different stages** of the process.

Logic Model Stage	Example of Evaluation
inputs	Do we have all the required inputs? Are the inputs of sufficient quality?
activities	Are we able to initiate activities with the inputs that have been defined? Are the activities able to transform the inputs effectively into outputs?
outputs	Have we produced enough of the outputs with sufficient quality?
outcomes	Have we induced the appropriate level of positive change in the stakeholder behaviour, knowledge, skills or in an environments level of functioning?
impact	Have we made long-term positive impacts on the community or system?

WHAT IS EVALUATION?

In the smoking cessation example, we could have the following evaluation.

Logic Model Stage	Example of Evaluation
inputs	$500k and 2 FTE budget to work with service providers. Once work was performed, actuals came in at $498k and 1.5 FTE. Inputs were evaluated as “effective”.
activities	Activities (ad design for print and online) were completed but took longer than expected (1 month over schedule). Activities were evaluated as “partially effective”.
outputs	The ads were designed to specifications and passed all quality control checks (5 out of 5). Outputs were evaluated as “effective”.
outcomes	Outcomes were measured over time (number of posters put up on retail walls and number of on-line ads interacted with). The targets for both (50% of retail establishments and 10% click rate on website) were met but took longer than anticipated (by 3 months). Outcomes were evaluated as “partially effective”.
impact	Over the course of 1 year, a statistically significant drop in cigarette sales was measured within the city boundaries and all surveys (150) contained interview data that identified that between 15% and 20% of respondents had at least attempted smoking cessation. Impacts were evaluated as “effective”.

WHAT IS EVALUATION?

Note that in the previous example:

- a **methodology** was defined for evaluation (reduction in sales, survey results);
- data was **available** to perform the evaluations at each step ($ and FTE budget);
- the evaluation steps were **related** through defining logical statements and assumptions (the higher the quality of the ads, the more likely it was to be engaging with the target audience, etc.).

In the last point, "**correlation does not imply causation**". We CAN conduct causal analyses (very carefully), but they require fair amount of forethought (and need to be modified on a case-by-case basis).