How do mechanisms and phenomena relate to one another?

Tudor Baetu

mechanisms and their discovery

how are putative mechanistic components identified and how do researchers know that they are indeed parts of the mechanism producing the phenomenon under investigation?

this is largely a matter of experimental research

"In the life sciences, investigators developing explanations often (1) begin by **identifying** the mechanism responsible for a specific phenomenon to be explained, (2) proceed to **decompose** the mechanism into its parts and the operations they perform, and (3) finally **recompose** the mechanism to show how, as a result of the organized parts orchestrating their operations, the mechanism generates the phenomenon."

(Bechtel 2017, 256)

mechanisms and their discovery

1) a system responsible for producing the phenomenon of interest is identified
   - Bechtel & Richardson 2010: **locus of control**
     - "Before developing a mechanistic explanation of a particular phenomenon, one must identify which system is responsible for producing that effect" (p. 39)

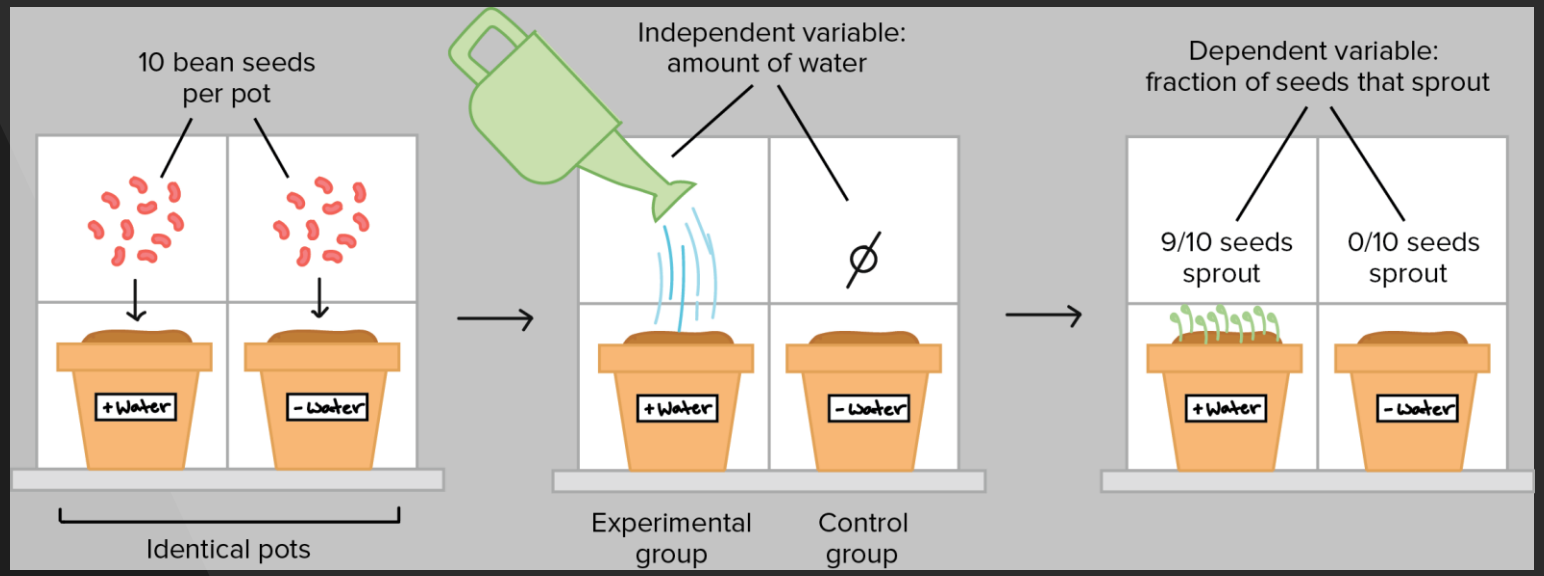
   - this requires the characterization of an **experimental setup** in the context of which **a phenomenon can be consistently reproduced**
     - *e.g., for T-cell activation, a typical experimental setup = cell model of an immune response = T-cells extracted from the blood of a human donor or precancerous/'immortalized' T-cell lines are grown in an artificial medium and stimulated by the addition of lipopolysaccharides*

mechanisms and their discovery

2) variables describing the experimental setup are targeted by experimental interventions in the hope of demonstrating that changes in the experimental setup and the physical systems of which it is composed result in changes in the phenomenon under investigation
   - Craver (2007, Chs. 2-3): **causal relevance** plays a crucial role in the identification of putative mechanistic components
     - causal relevance is demonstrated by means of **controlled experiments**, as analyzed by Woodward's (2003) interventionist account of causation
     - the causal relevance of a factor X in respect to target phenomenon Y is established as a result of experiments in which interventions on X result in changes in Y
     - *e.g., the IκB inhibitor was shown to be part of the regulatory mechanism of T-cell activity in light of experiments demonstrating that mutations in the sequence of IκB result in a prolonged activation of T-cells*

controlled experiments

- make a comparison between outcomes in two situations
  - a **test** experimental setup in which a factor under investigation is present
  - and a **control** setup in which it is not
- any **difference in outcomes** must be **caused** **by a difference between the situations**



controlled experiments

- in order to conclusively demonstrate that the factor is causally responsible for the difference in outcome, alternative explanations must be ruled out:
  a) reverse causation & common cause
     - only if there is a causal pathway linking factor and outcome as upstream cause to downstream effect, **interventions** on the factor result in changes in outcome
  b) effects of confounders
     - ensure that test and control conditions are comparable in all causally relevant respects minus the factor manipulated in the experiment
  c) accuracy of the intervention
     - target only the factor under investigation and no other factors which may contribute to differences in outcome

mechanisms and their discovery

3) mechanism is '**recomposed**', physically or conceptually, in order to show how the mechanism generates the phenomenon (Bechtel 2011)
   - physical/in vitro = reconstitution experiment
   - conceptual/in silico = computer simulation or narratives and diagrams such as those illustrated earlier
   - goal is to demonstrate that components organized, acting, and having the properties described in the mechanistic explanation are indeed sufficient to produce the phenomenon under investigation (= complete explanation?)

how do mechanisms and phenomena relate to one another?

- a mechanism is said to be **responsible for** a phenomenon in the sense that it
  - **produces** or
  - **underlies** that phenomenon
- two metaphysical interpretations
  - **etiological interpretation** → captures the intuition that mechanisms should be causally relevant to phenomena given that the discovery of mechanisms relies primarily on experiments demonstrating causal relevance
  - **constitutive interpretation** → captures the notion that phenomena are behaviors of systems explained by referring to the behavior of their parts, but raises a puzzle about how evidence for causal relevance can justify a metaphysical interpretation postulating a non-causal relationship between mechanisms and phenomena

the **etiological** interpretation

- the fact that mechanisms are elucidated by means of experiments designed to demonstrate causal relevance speaks in favor of a causal interpretation of the relationship between mechanisms and phenomena
  - → immediately suggests an **etiological interpretation**
  - chains of antecedent causes terminating in the phenomena to be explained
  - … but in the mechanistic literature, phenomena are seldom construed as outcomes
  - they are usually depicted as regularities, patterns or behaviors associated with biological systems
  - suggests a correlation among several variables
    - in contrast, a time-point outcome is just the change in the value of one variable
- mechanistic explanations are not aimed at identifying antecedent causes, but rather at elucidating the mechanisms ‘underlying’ input-output behaviors

the **constitutive** interpretation

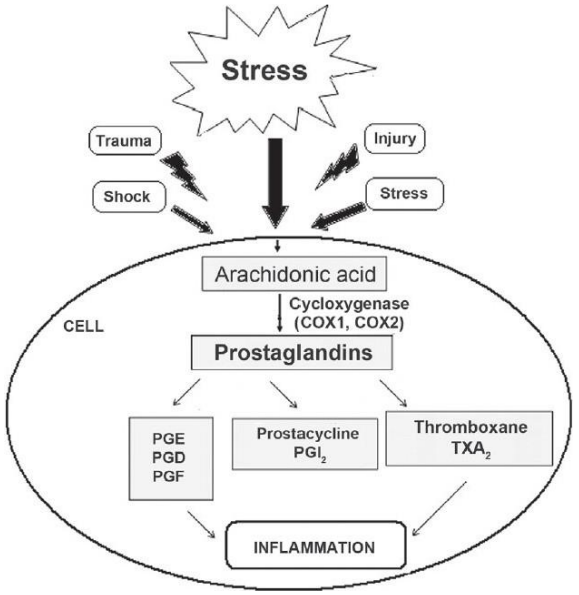
- one way of understanding this underlying-relationship is in terms of **constitution**:

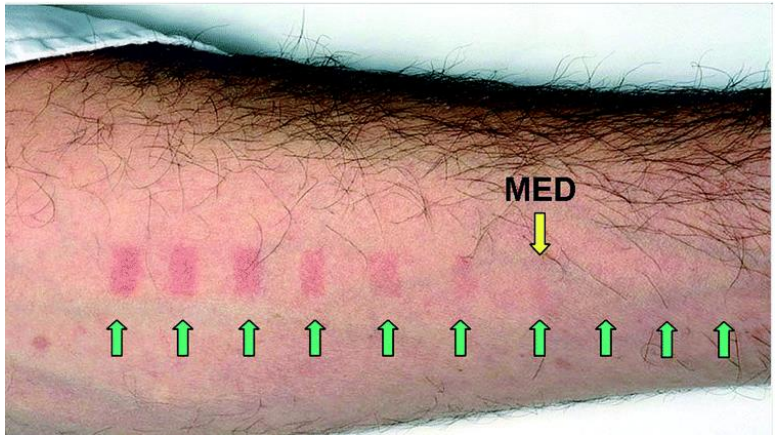
  "At the highest level, the eye transduces light into a pattern of neural activities in the optic nerve. This process can be decomposed into lower-level components and their activities. […] The conversion of light into patterns of neural activity by the retina can itself be decomposed into different components: in particular, the rods and cones that change their electrical state depending on specific features of the light stimulus (such as wavelength and intensity). Another level down, rod cell activation is also sustained by a mechanism. Light is absorbed by and activates rhodopsin, which then stimulates G-proteins. […] Each new decomposition of a mechanism into its component parts reveals another lower-level mechanism until the mechanism bottoms out in items for which mechanistic decomposition is no longer possible." (Craver and Bechtel 2007, 549)

- bottom line: the light transducing eye ***is not caused by***, but instead ***is made of*** rhodopsin signaling cells

the **constitutive** interpretation

- immediate difficulty = how can constitution relationships can be inferred given that experiments only provide evidence for causal relevance?
- Craver's solution = **mutual manipulability**
  - "a component is relevant to the behavior of a mechanism as a whole when [… the] two are related as part to whole and they are mutually manipulable […]:
  - (i) X is part of S;
  - (ii) in the conditions relevant to the request for explanation there is some change to X's ϕ-ing that changes S's ψ-ing; and
  - (iii) in the conditions relevant to the request for explanation there is some change to S's ψ-ing that changes X's ϕ-ing" (2007, 153)





example: the UV-induced erythema assay

- ***whole S** = guinea pig/human organism*
- ***S's behavior ψ** = the phenomenon of inflammation (sunburns)*
- ***part X** = prostaglandins*
- ***X's behavior φ** = the binding of prostaglandin receptors (which triggers a signaling cascade leading to the expression of several gene products involved in inflammatory responses)*
- ***bottom-up intervention** on a part having an effect on the whole = prostaglandin injections resulting in an inflammatory response*
- ***top-down intervention** on the organism-whole = ultraviolet exposure resulting in increased prostaglandin levels*

how do two causal relationships add up to constitution?

emphasis on causal relevance suggests that the relationship between mechanism and phenomenon is strictly one of **causal dependency** (Leuridan 2012)

… nevertheless, Craver & Bechtel insist that the **requirement for part-whole relationships** has several unpalatable consequences for a causal interpretation

chief among which is the fact that cause and effect are no longer distinct events (Craver 2007, 153-54; Craver and Bechtel 2007, 552-54)

the **constitutive** interpretation

- the constitutive interpretation is committed to the view that phenomena are macro-level sates consisting of micro-level mechanisms
  - constitution may be understood as ‘supervene on’, ‘are realized by’, ‘are identical with’ or ‘are made of’
  - whatever it is, **constitution** is definitively **not causal**

how do two causal relationships add up to constitution?

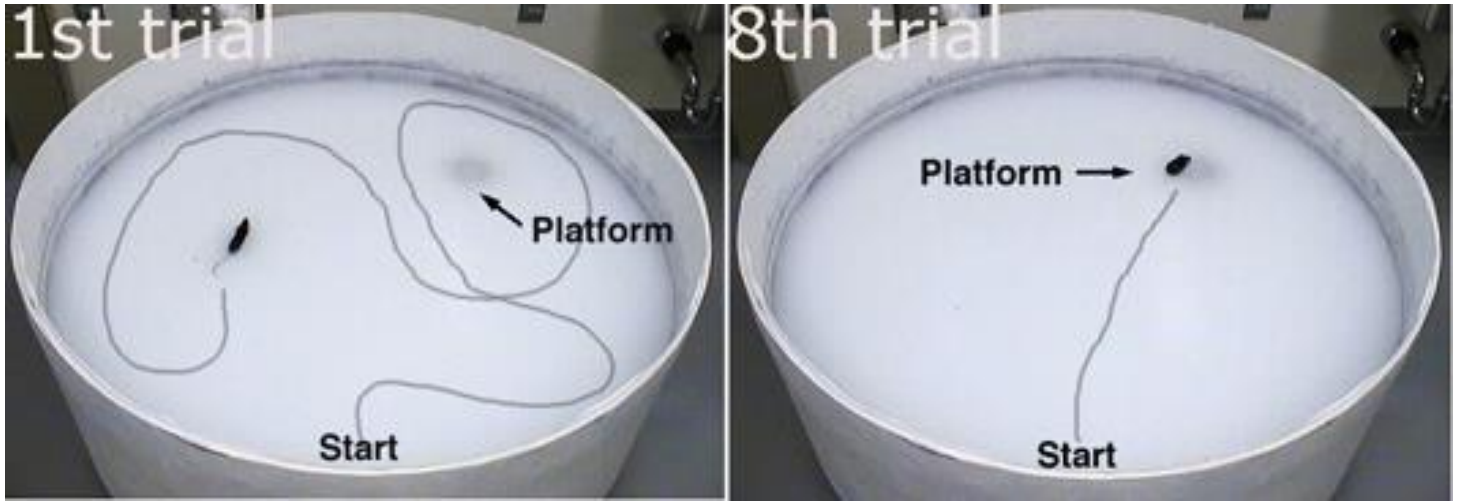
- … yet if it is indeed the case that two variables stand in part-whole relationships, then it cannot be conclusively demonstrated that interventions on a variable have an effect on the other variable
  - → it is impossible to demonstrate mutual manipulability in the first place

- if lower levels supervene on higher ones, as implied by the part-whole constituency requirement, then top-down interventions invariably have an effect on the behaviors of both wholes and their parts (Baumgartner and Gebharter 2016; Romero 2015)
- → several variables are targeted at the same time → the accuracy of the intervention is compromised (condition c) → it cannot be concluded that the effects observed in the dependent variable are indeed due to experimental interventions on the independent variable and not to direct interventions on the dependent variable

how do two causal relationships add up to constitution?

- the more general problem is a **failure to correctly identify the independent variable**

- Craver (2007, 166-70) describes the training of rats in a Morris maze as a top-down experiment in which the behavior of rat-wholes is manipulated (the independent variable, corresponding to the factor tested for causal relevance) and effects on the behavior of parts, in this case the long term potentiation of certain synapses, are measured (the dependent variable)



how do two causal relationships add up to constitution?

… yet that which varies between test and control conditions (the independent variable) are not the rats or their behaviors, which are assumed to be comparable at onset between test and control condition, but rather the maze, more specifically the variable ‘place of escape platform’, which takes the values ‘constant’ or ‘random’

Morris mentions that a rat was excluded from the experiment because it was found to be different from the other rats in a respect relevant to the measured outcomes ‘escape latency’ and ‘directionality of tracks’: *“One rat was found to have difficulty in swimming during the first Pretraining session and was replaced with another animal”* (1981, 242)

same incongruency is present in the inflammation example: what varies between test and control conditions are not the guinea pigs or their behaviors, but the intensity of UV radiation, which is the **input condition** of the phenomenon of ultraviolet-induced erythema

how do two causal relationships add up to constitution?

- Craver seems to be aware of this caveat when he argues that *"[o]ne intervenes on S's ψ-ing by intervening to provide the conditions under which S regularly ψs. Top-down experiments intervene in this way"* (2007, 146)
- methodological incongruity persists
  - the **phenomenon** under investigation is a behavior amounting to an **association of variables**
    - *e.g., inflammation is just an occurrence of symptoms, but a consistently reproducible set of symptoms in response to certain stimuli*
    - *spatial memory, which is not measured solely as the time it takes a rat to navigate a maze, but the time it takes a rat to navigate a maze given prior exposure certain environmental clues*
- in contrast, the **experiments** involved in the elucidation of the mechanisms responsible for these phenomena **target individual variables**
  - *e.g., UV exposure, erythema, the position of the escape platform in a water maze, directionality of tracks*

how do two causal relationships add up to constitution?

- it would seem that talk about bottom-up and top-down interventions plays on an ambiguity whereby the manipulation or measurement of a variable involved in the description of a phenomenon is arbitrarily equated with the manipulation or measurement of a phenomenon/behavior of the whole

levels of mechanisms

- **levels of mechanisms**
  - Craver and Bechtel introduce **two notions of constitution**
  - one **between physical systems**, which is not causal in nature
    - *e.g., the eye being composed of cells*
  - the other **between behaviors**, or phenomena, which is amenable to a causal mediation interpretation
    - *e.g., rhodopsin signaling pathway being part of light-to-neural-activity transduction*

levels of mechanisms

- further assume that behavioral and physical constitution are systematically aligned
- → the behavior of a system is hierarchically composed of–and experimentally decomposable into–the behaviors of its parts
  - *e.g., the light transducing eye consists of rhodopsin signaling cells*

- finally, they argue that the physical constitution element of this hierarchy trumps the causal relationship between mechanisms and phenomena
  - *e.g., the light transducing eye is **not caused by**, but rather **made of** rhodopsin signaling cells*

levels of mechanisms

- why assume that behavioral and physical constitution are systematically aligned?

- examples demonstrating a lack of alignment between behavioral and physical constitution
  - *e.g., UV-induced erythema is constitutive of UV-induced inflammation, yet both behaviors are documented in the same experimental setups/biological systems*

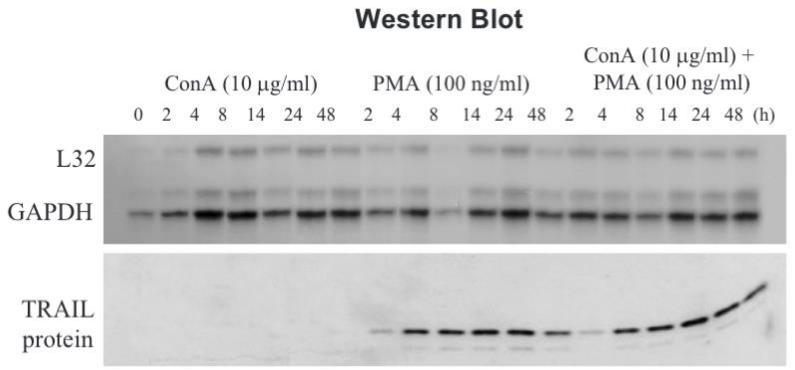
levels of mechanisms

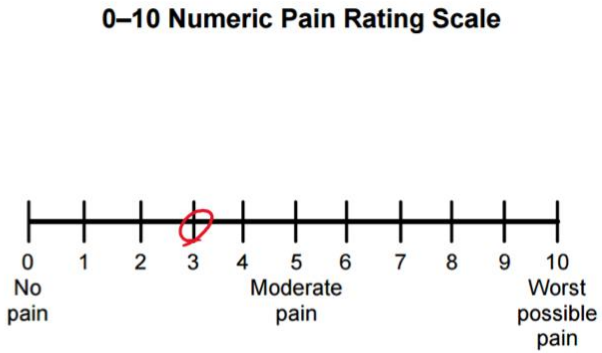
- why is a part-whole relationship an essential requirement for constitutive relevance?

- researchers didn't infer that **gene expression is part of the mechanisms responsible for inflammatory responses** *because gene expression occurs in cells and cells are parts of organisms*, but rather **because receptor binding and gene expression are causal intermediaries along the causal pathway linking exposure of an organism to a harmful stimulus and the ensuing inflammatory response**

the **causal mediation** interpretation

- experimental methodology assumes a **causal** **interpretation of measurements** (e.g., Trout 1998)
  - data consists of **physical effects**
  - these effects are **informative of the causal structure of the world**
    - i.e., differences in measured values reflect differences in the causal structure of the world





the **causal mediation** interpretation

- e.g., **standardization**
  - = measurement techniques are carefully replicated from one measurement to the next & designed in such a way as to generate perceptually unambiguous outputs meant to eliminate the potential for disagreement among observers
  - → any particular instantiation of the technique and any particular observer can be exchanged with any other without affecting the data
  - → safe to infer that differences in data outputs are due to reasons other than differences in the measurement system

  - thus, given the same measurement techniques, differences in measurements reflect differences in the causal structure of the world

the **causal mediation** interpretation

- worth considering the problem of constitutive relevance from an experimental standpoint

- the inflammation example
  - the phenomenon to be explained = ‘ultraviolet-induced erythema’ = the consistently reproducible induction of erythema in response to exposure to ultraviolet radiation
  - outcome variable measured = severity of erythema
    - *e.g., 1/barely visible to 4/purple with edema scale*
  - differences in measured values of erythema are attributed to differences in the causal structure of the experimental setup
    - *e.g., differences in UV intensity*

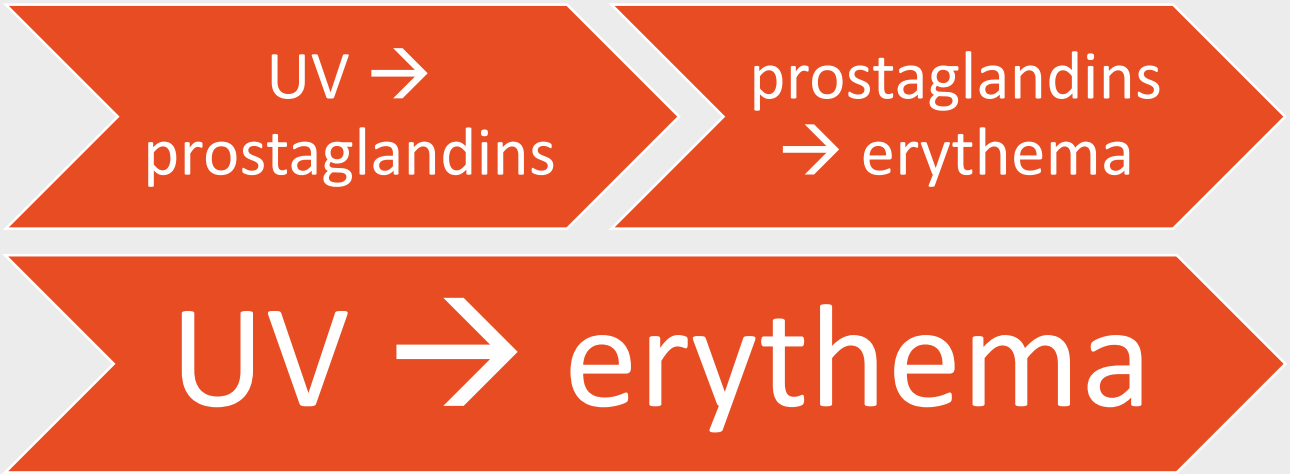
the **causal mediation** interpretation

- worth considering the problem of constitutive relevance from an experimental standpoint

- the inflammation example
  - so-called ‘bottom-up’ intervention = the phenomenon of ‘prostaglandin-induced erythema’ = the consistently reproducible induction of erythema in response to an increase in the tissue concentration of prostaglandin
  - so-called ‘top-down’ intervention = the reproducible phenomenon of ‘ultraviolet-induced increase of local prostaglandin levels’

the **causal mediation** interpretation

- relevant question
  - = are the causal structures involved in ultraviolet-induced prostaglandin synthesis and prostaglandin-induced erythema are constitutive, or parts, of the causal structure linking ultraviolet exposure to the development of erythema
  - causal pathways being parts/constitutive of other causal pathways



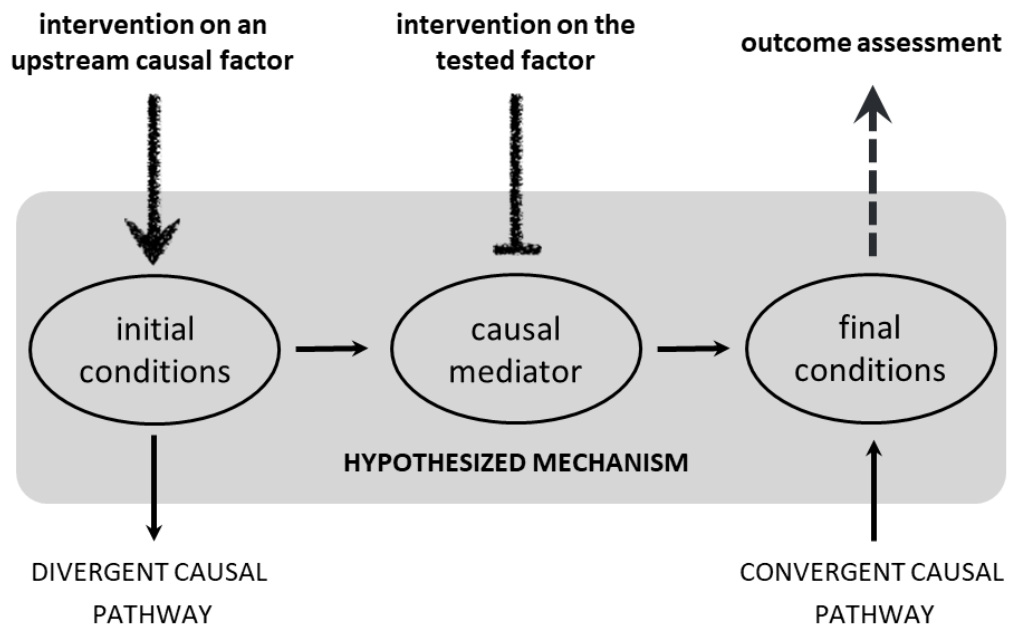
the **causal mediation** interpretation

- question can be reframed in terms of **causal mediation** (Baetu 2012; Harinen 2014)
  - to ask whether prostaglandin synthesis is constitutively relevant to the phenomenon of ultraviolet-induced erythema
  - = to ask whether prostaglandin synthesis is a **causal intermediary** **along a causal pathway** linking the ultraviolet exposure-input to the erythema-output

UV → prostaglandins → erythema

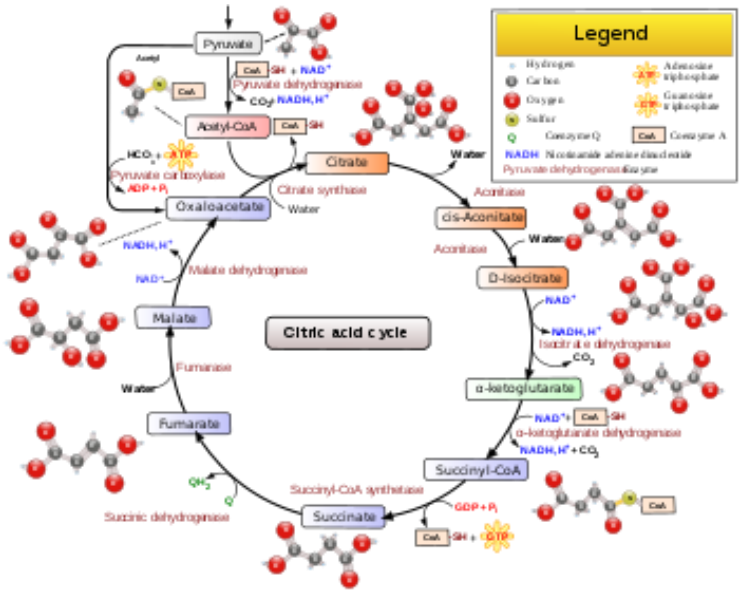
the **causal mediation** interpretation

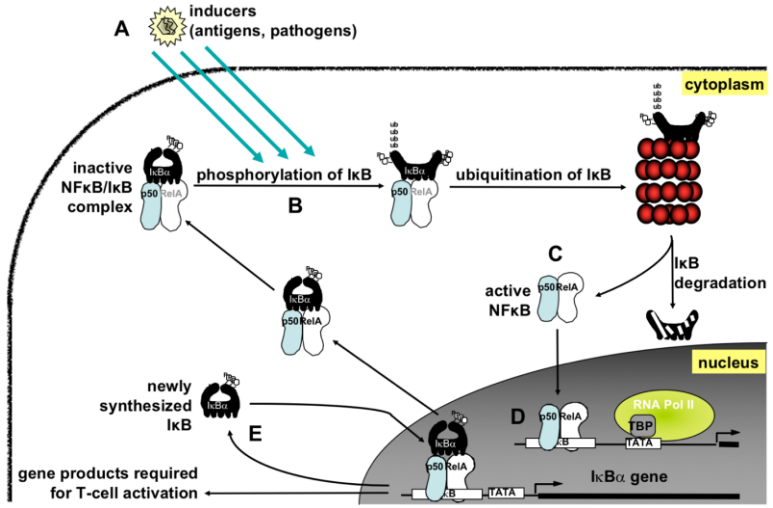
- there is a well-established experimental methodology for demonstrating causal mediation
  - common strategy = conduct a **knockout-type experiment** whereby two factors, usually the initial conditions and a putative mechanistic component, are simultaneously manipulated on an independent basis and the effects on a third variable, usually the output conditions, are observed
  - *e.g., when prostaglandin synthesis is blocked, ultraviolet exposure fails to cause erythema* (Langenbach et al. 1999)



the causal mediation interpretation

- under this interpretation
  - the mechanism = a **causal structure** **linking input and output conditions**
    - ‘linking’ is understood here in the experimental sense of **reproducible**, or not allowed to vary from one iteration of the experiment to the next
  - to be **constitutively relevant** to the mechanism = to be a **causal intermediary** along this causal structure

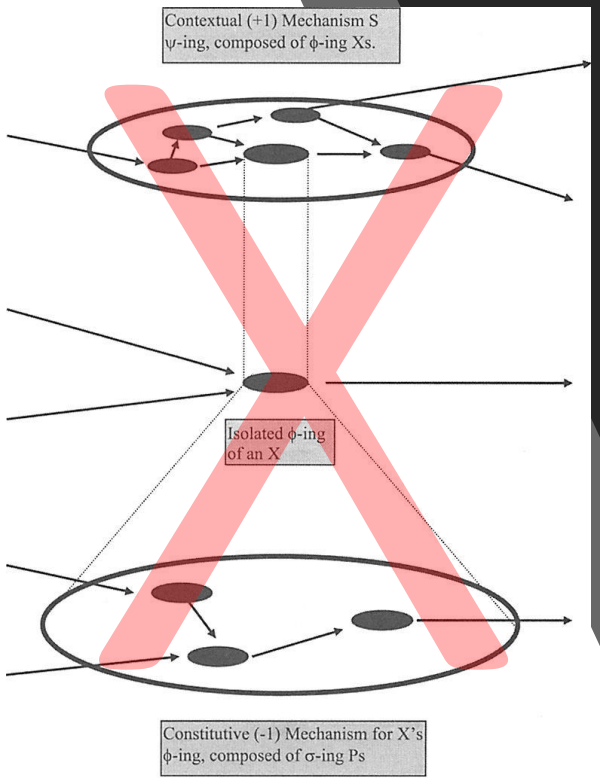




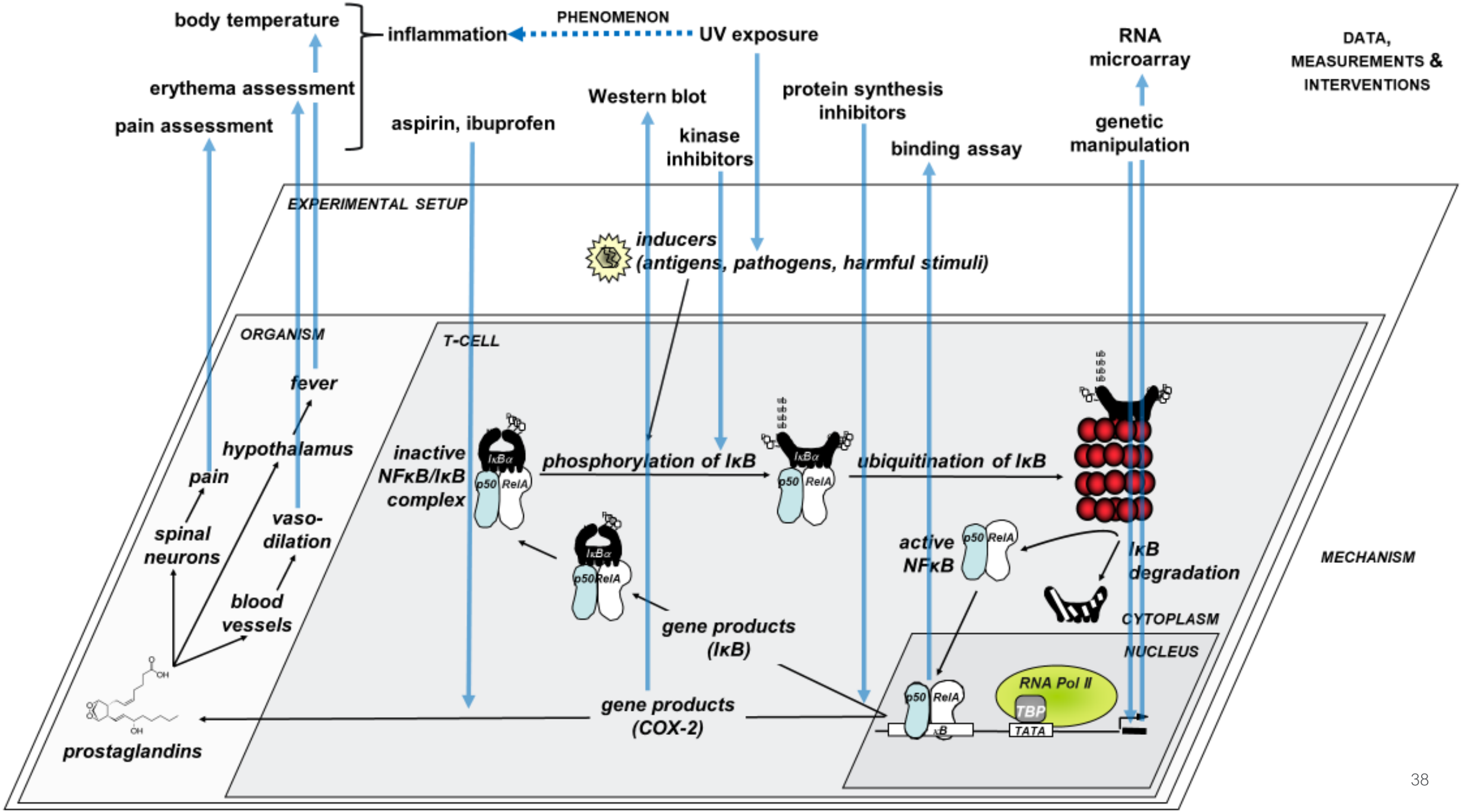
the **causal mediation** interpretation

- note: this does not entail that mechanisms are linear causal chains
- e.g., in the absence of measurements of intermediary stages, a **circular metabolic pathway**, such as Krebs' cycle, appears as a linear input-output phenomenon
  - 2 acetyl-CoA, 6 NAD^+, 2 FAD, 2 $ADP+P_i$ → 4 CO_2, 6 NADH, 6 H^+, 2 $FADH_2$, 2 ATP, 2 CoA
- the same applies to inflammation: the NF-κB molecular mechanism of involves **negative feedback loop** ensuring that inflammatory responses eventually shut down after being triggered by stimuli

a **level-free** conception of mechanisms



- implication
  - **no hierarchy/levels of mechanisms**
  - only a causal structure probed by an increasing number of experiments gradually revealing more and more causal intermediaries
  - → both the 'higher-level' **description of a phenomenon** and that of its 'lower-level'/'underlying' **mechanism** refer to the **same causal structure**
    - the difference =
      - a phenomenon is described in light of interventions and measurements probing very few aspects of this structure → the inner workings, or intermediary causal stages remain unknown and are often described as 'black boxes'
      - a mechanistic description is based on additional interventions and measurements targeting variables characteristic of intermediary stages → by disrupting the inner workings of the mechanism in ways that affect downstream variables used to test for causal relevance, causal factors initially hidden in the black-box are revealed


PHENOMENON
body temperature
inflammation
UV exposure
erythema assessment
pain assessment
aspirin, ibuprofen
Western blot
kinase inhibitors
protein synthesis inhibitors
binding assay
RNA microarray
genetic manipulation
DATA, MEASUREMENTS & INTERVENTIONS
EXPERIMENTAL SETUP
inducers (antigens, pathogens, harmful stimuli)
ORGANISM
T-CELL
fever
hypothalamus
pain
spinal neurons
vaso-dilation
blood vessels
prostaglandins
inactive NFκB/IκB complex
IκBα
p50
RelA
phosphorylation of IκB
ubiquitination of IκB
active NFκB
IκB degradation
gene products (IκB)
gene products (COX-2)
CYTOPLASM
NUCLEUS
RNA Pol II
TBP
TATA
MECHANISM

how do mechanisms and phenomena relate to one another?

- **causal mediation interpretation**
  - a phenomenon is analogous to an incomplete, or low-resolution, measurement-mediated representation of a mechanism
  - causal interpretation of measurements dictates that the **mechanism causally determines the phenomenon** for which it is responsible
    - although this causal relationship should not to be understood in the etiological sense of bringing about changes in a variable
    - … but rather in that of gluing together the values of multiple variables in relationships of association and causal dependency
  - levels/part-whole composition is not essential