„ Modely testovania a ohodnotenia hypotez “

Vladimír Marko
19.02.2015

analytické metódy v spoločensko-humanitných disciplínach
analytical methods in social sciences and humanities
APVV-0149-12 · www.amesh.sk

A hypothesis…

- A hypothesis is **an idea or hunch about the world**. It might concern a claim about an individual variable, such as the degree of pollution in the River Trent, or it might be concerned with the relationship between two or more variables... Most hypotheses are prompted by **a combination of inspiration** and **experience** of the world (an *empirical hypothesis)* (Phelan - Argument and Evidence, Critical Analysis for the Social Sciences 1996).

- “Hypotheses are single tentative **guesses, good hunches** –assumed for use in devising theory or planning experiments intended to be given a direct experimental test when possible”. (Eric Rogers, 1966)

- We use the word 'hypothesis' to refer to **whatever statement is under test**, no matter whether it purports **to describe** some particular fact or event or **to express a general law** or **some other, more complex, proposition**. (Hempel, Philosophy of Natural Sciences, 1966, 209p.)

- “An hypothesis is **a proposition which attempts to solve the problem** raised by the provocative fact, and thus is an explanation the scientist hopes is true…The purpose of an hypothesis is **to explain provocative facts and so to solve a problem**.” (Feibleman, The Logical Structure of the Scientific Method, 1959)

- “A hypothesis is **a conjectural statement** of the **relation between two or more variables**”. (Kerlinger, 1956)

- “Hypothesis is **a formal statement that presents the expected relationship between an independent and dependent variable**.”(Creswell, 1994)

- “**A research question** is essentially a hypothesis asked **in the form of a question**.”
- “It is **a tentative prediction** about the nature of the relationship between two or more variables.”
- “A hypothesis can be defined as **a tentative explanation** of the research problem, a possible outcome of the research, or an educated guess about the research outcome.”(Sarantakos, 1993: 1991)

- “An hypothesis is **a statement or explanation** that is suggested by knowledge or observation but has not, yet, been proved or disproved.” (Macleod Clark J and Hockey L 1981)

A hypothesis…

- “we examine the scientific ideas tested by experience namely **the guesses** called hypotheses, the hypotheses upgraded to laws, and the system of laws called theories.” ... hypothesis ≠ fiction (Bunge, M. – Philosophy of Science, Volume One, From Problem to Theory, 1998)

- A “hypothesis”— **an idea or postulate** that must be phrased as **a statement of fact**, so that **it can be subjected to falsification.** The hypothesis is constructed in advance of the experiment; it is therefore unproven in its original form. (David J. Glass and Ned Hall, A Brief History of the Hypothesis, Cell 134, August 8, 2008)

- The term *hypothesis* can appropriately be applied to any statement that is intended for evaluation in terms of its consequences (*John Earman and Wesley C. Salmon*, "The *Confirmation* of Scientific Hypotheses“)

- Ordinarily, when one talks about hypothesis, one simply means a mere assumption or some supposition to be proved or disproved. But for a researcher hypothesis is **a formal question that he intends to resolve**. Thus a hypothesis may be defined as **a proposition or a set of proposition set forth as an explanation** for the occurrence of some specified group of phenomena either asserted merely as **a provisional conjecture to guide some investigation or accepted as highly probable in the light of established facts**. Quite often **a research hypothesis** is a predictive statement, capable of being tested by scientific methods, that relates an independent variable to some dependent variable. (Kothari, C.R. - Research Methodology - Methods and Techniques, 1985, 2004, p. 184)

- At the most basic level, both hypotheses and theories are **types of scientific explanation**. A hypothesis **can be anything from a vague hunch to a finely detailed, though speculative, account of how or why something has come to be the case**. In general, however, the point of characterizing an explanation as a hypothesis is to note that it is **tentative and unproven**... Yet both claims must be regarded as hypotheses, since the exact details of what happened remain in question. (Carey, SS - Scientific Method, Beginners Guide 2001 (2004)

To sum up - A hypothesis is:

- premise (Bacon, …)
- an idea or hunch about the world
- the guess
- a provisional conjecture
- a combination of inspiration and experience of the world
- refer to whatever statement is under test
- a proposition which attempts to solve the problem
- a conjectural statement
- a research question
- a tentative prediction
- a tentative explanation
- a statement of fact
- a statement or explanation … not yet proved or disproved
- type of scientific explanation
- any statement that is intended for evaluation in terms of its consequences

- Different kinds of hypotheses
- working hypothesis
- research hypothesis
- theoretical hypothesis
- empirical hypothesis
- experimental hypothesis
- explanatory hypothesis
- predictive hypothesis
- ad hoc

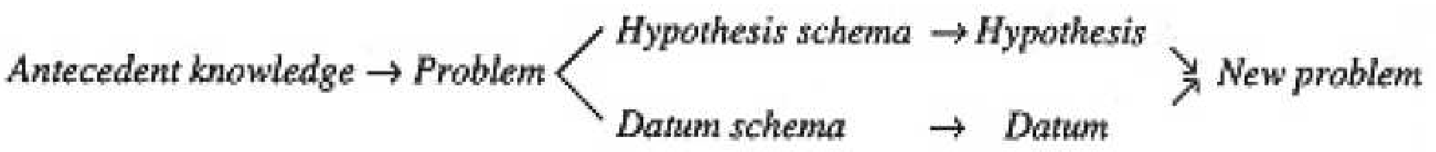
Generation of hypothesis

- hypothesis generation
  - Popper –insight, hypothesis formation is, or at least can be, a largely extra-scientific activity; hypotheses can be devised creatively, via philosophical reflection (from the armchair).
  - Today – HGF (Hypothesis generation framework) – The “End of Theory” conception
    - there is still a large gap between the huge amount of effort and resources invested in disease research and the little effort in harvesting the published knowledge. Hypothesis generation framework (HGF) finds “crisp semantic associations” among entities of interest - that is a step towards bridging such gaps. The proposed framework is fast, efficient, and robust in generating new hypotheses to identify factors associated with a disease. A full integrated Web service application is being developed for wide dissemination of the HGF. A large-scale study by the domain experts and associated researchers is underway to validate the associations and assertions computed by the HGF.

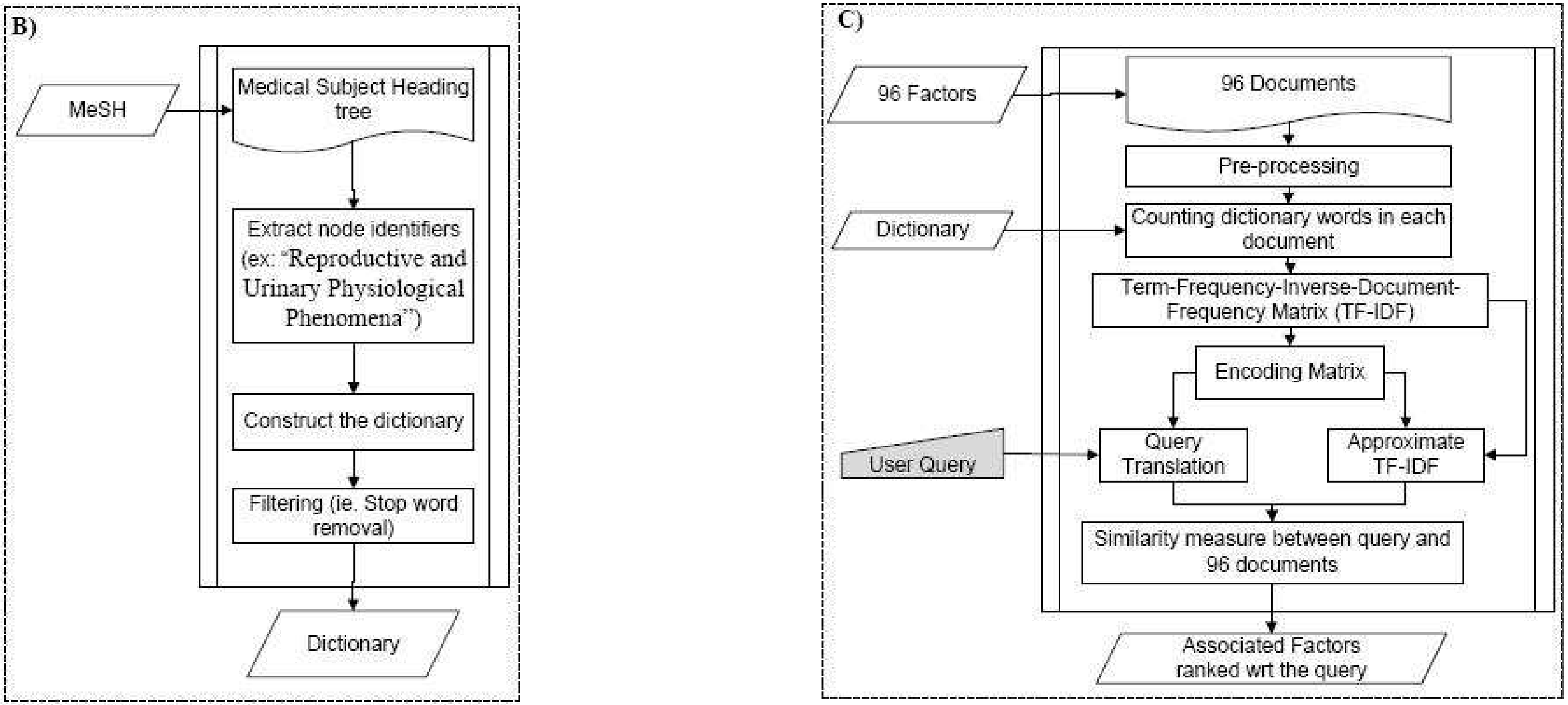
Generation of hypothesis

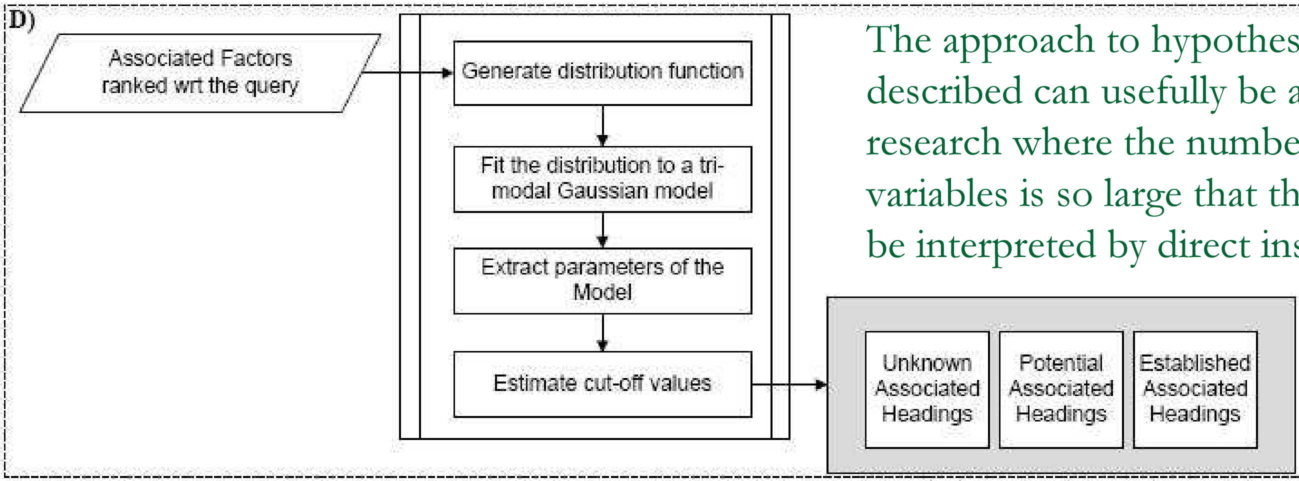
Generation of hypothesis is the way of data organization.

- The generator of a problem is a schema or prepositional function and the solution to it is a statement that is a datum or a hypothesis, depending on whether it overreaches the experience at hand and is corrigible, or not.
- Eliminate hypotheses and you are left with data of a comparatively uninteresting kind superficial, stray, unexplained as well as with minor problems, generated by data schemata.



The generator of hypothesis – medicine (HGF, "the end of theory")

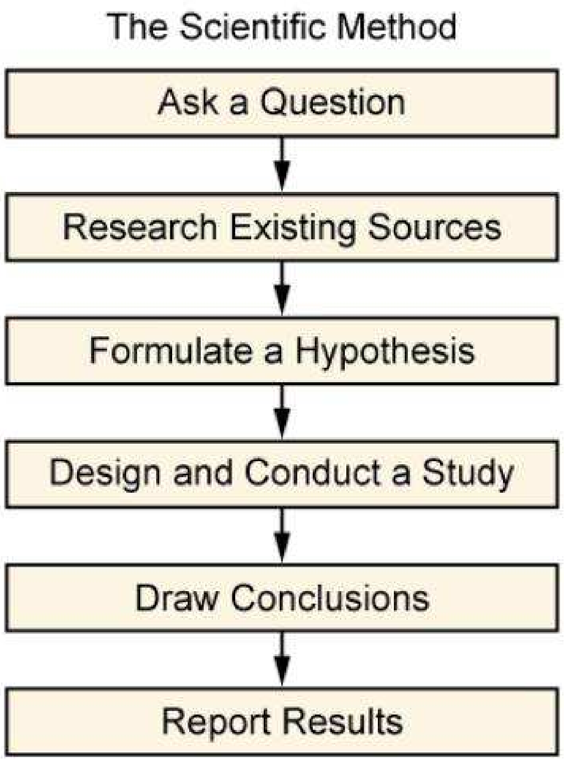


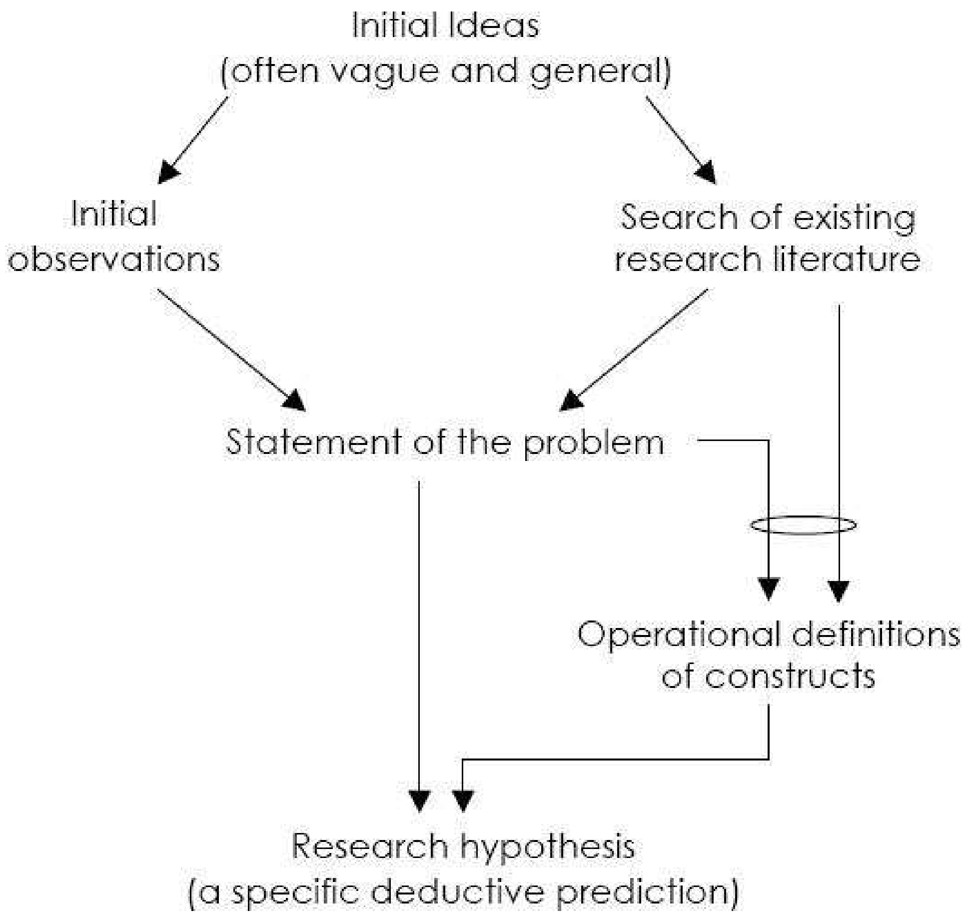


The approach to hypothesis generation just described can usefully be applied in any research where the number of objects and variables is so large that the data cannot easily be interpreted by direct inspection.

Generation of hypothesis – “grounded theory”

Forming hypothesis after data-mining





Working hypothesis (provisional, conductive, researching, “exploratory research”)

- **Working hypothesis**, a hypothesis suggested or supported in some measure by features of observed facts, from which consequences may be deduced which can be tested by experiment and special observations, and which it is proposed to subject to an extended course of such investigation, with the hope that, even should the hypothesis thus be overthrown, such research may lead to a tenable theory.
  Century Dictionary Supplement, v. 1, 1909, New York: The Century Company. Reprinted, v. 11, p. 616

- …"working hypotheses," which signal that **conceptualization is in its preliminary stages**. The working hypothesis is a pivotal concept in Dewey's (1938) theory of inquiry. Working hypotheses are a **"provisional, working means of advancing investigation"**; they lead to the discovery of other critical facts (Dewey, 1938, 142). This is the type of theory that Kaplan (1964, 268) would describe as **"provisional or loosely formatted."** Although the working hypothesis is preliminary, we still emphasize that it should, like all hypotheses, be in the form of a statement of expectations. Further, it must be possible to collect evidence that either supports or fails to support the expectation. Further, working hypotheses (like formal hypotheses) are **never proven**. They are supported with **empirical evidence**. Working hypotheses **direct inquiry** because they help to establish the connection between the research question and the types of evidence used to test the hypothesis.
  Shields & Tajalli, Intermediate Theory: *JPAE: Journal of Public Affairs Education*, vol. 12, no. 3, Summer 2006
- **A conductive hypothesis** is one which suggests lines of inquiry beyond the limits of its own verification.
  Feibleman, the Logical Structure of the Scientific Method, 1959

Microconceptual Framework	***Research Purpose***
I. Working Hypotheses	Exploration
2. Categories	Description
3. Practical Ideal Type	Gauging
4. Models of Operation Research	Decision Making
5. Formal Hypotheses	Explanation / Prediction

Hypothesis

- part of **the problem solving** activity (Bunge, Feibleman, …)
- an indicative **conditional**
- grounded in a **previous knowledge**
- expresses some **believes** (expectations…, in respect to a previous knowledge)

If testable (in respect of its **function** in the problem solving activity) it could be	**If not (yet) testable** (in respect of its **function** in the problem solving activity) it could be
❑ theoretical hypothesis ▪ (indirectly testable, by reduction to observational terms; or deductively inferable from other theoretical or empirical statements) ❑ empirical hypothesis ❑ experimental hypothesis ❑ explanatory hypothesis ❑ predictive hypothesis …	❑ working hypothesis or research hypothesis (related to previous knowledge; anticipation; insufficient conceptual and evidential background for better formulation; construed to advance [preliminary] research) ▪ directional or ▪ non-directional ❑ Ad hoc ▪ (partially related or not related to previous knowledge)

A **question** is not hypothesis (even could help in hypothesis formation where data are insufficient)

theoretical hypothesis – plus "saving the phenomena"

Giere, R. 1983. "*Testing theoretical hypotheses*." In J. Earman, ed., Minnesota Studies in the Philosophy, Vol. 10. Minneapolis: University of Minnesota Press

Model –as a set of objects that satisfies a given linguistic structure.

- **A system** – set of state variables and system laws that specify **the physically possible states of the system** and perhaps also its possible evolutions.
- **Theoretical models** have by themselves **no empirical content**—they make no claims about the world. But they **may be *used* to make claims about the world**. This is done by **identifying elements of the model with elements of real systems** and then claiming that **the real system exhibits the structure of the model**. Such a claim I shall call a *theoretical hypothesis.*
- The definition of a **model** amounts to the definition of a **predicate**. A theoretical hypothesis, then, has the logical form of predication: This is an X, where the corresponding definition tells what it is to be an X.
- **Goal of science** is the discovery of *true universal generalizations* of the form: All A's are B. Philosophers, beginning with Hume, have **reduced the concept of physical necessity to that of universality**, and scientific explanation has been analyzed in terms of derivation from a generalization.

- **A theory – a conjunction of universal generalizations.** This would mean that testing a theory is just testing universal generalizations… **Theory** – refer to **a more or less generalized theoretical hypothesis** asserting that **one or more specified kinds of systems** fit a given type of **model**. This seems broad enough to encompass all the sciences, including geology and physics.
- **The simplest form of a theoretical hypothesis** is the claim that **a particular, identifiable real system fits a given model** (geological models, models of natural selection)

- **Testing a theory**, then, means **testing a theoretical hypothesis** of more or less restricted scope. **The scope of a hypothesis** is crucial in any judgment of **the bearing of given evidence** on that hypothesis. Knowing what kind of thing we are testing, we can now turn to an analysis of empirical tests.

- **Saving the phenomena** *(traditional)* **dilemma** – *simplicity* (harmony) of the model or *the phenomena justified* regarding a hypothesis.
- **Another suggestion** (Clavius, Descartes, Leibniz, Huygens, Newton, Whewell and Jevons [Whewell vs. Mill; Peirce vs. Keynes]) – **true hypotheses are revealed by their ability to *predict* phenomena before they are known.**

THE LOGICAL BASIS OF HYPOTHESIS TESTING (Giere)

Scientific Inquiry: attempts to provide good justification for believing a hypothesis is true or false.

Theoretical Model: a generalized explanation of observed phenomena.

Theoretical Hypothesis: a contingent statement asserting that some real system corresponds to the theoretical model.

Hypothesis: contingent statement that is treated as the object of research. Hypotheses are logically justified by exhibiting them as conclusions of appropriate arguments.

premises:
Background
Initial **c**onditions (**IC**)
Outcome of experiment (**P**rediction or its negation)
conclusion: the **H**ypothesis

If (Not **H** and **IC** and **B**), then very probably Not **P**
P and IC and B
Thus (inductively), **H**

- An **argument** must meet 2 conditions to **justify** believing the conclusion (hypothesis):
  - (1) the premises (assumptions, experimental/initial conditions) must themselves be justified.
  - (2) there must be sufficient connection between the premises and conclusion
- Giere's 3 criteria of **good test** provides the basis for judging this connection.
  - 1) prediction is deducible from the hypothesis together with the initial conditions.
  - 2) prediction is improbable when considered out of context from hypothesis
  - 3) prediction is verifiable
- The **experiment** determines the truth and falsity of the **prediction**.
  - If the prediction is successful the hypothesis is justified.
  - If the prediction fails the hypothesis is refuted.

Problem of evidence - experience or knowledge?

How to define fact, datum, evidence, etc?

- Evidence – justified belief
- Davidson: evidence – (only) **causally related** with beliefs
- against scepticism and against the phenomenal conception:
  - observations, evidences, beliefs… have **propositional character -** *observation statements* (Chalmers, *What is this thing called Science*?
  - “we take a subject's evidence to consist of **all and only those propositions that the subject knows”** (Williamson, *Knowledge and Its Limits, 2000; …)*

- Evidence as Sign, Symptom, or Mark *(*Austin, *Sense and Sensibilia*)
  - Achinstein's (2001) concept of ‘**veridical evidence**’:

‘in this respect the present concept functions like “sign” or “symptom”. A rash may be a sign or symptom of a certain disease, or a good reason to believe the disease is present, even if the medical experts are completely unaware of the connection and so are not justified in believing that this disease is present, given what they know’ (p. 25)

- One's evidence consists exclusively of one's **current mental states**

(Conee, Earl and Feldman, Richard, 2004*. Evidentialism*)

- What is the role of evidence in testing some theory?
  - **Hempel** (1945, ‘Studies in the Logic of Confirmation’) **qualitative** (clasificatory) role of evidence – either evidence supports or does not supports a given hypothesis
  - **Carnap** (1950. *Logical Foundations of Probability)* – **quantitative** role of evidence – ***how much***:*‘Hypothesis H is confirmed by evidence E to degree R’*
  - **Achinstein** (1983, *The Concept of Evidence*; 2001, *The Book of Evidence*) - something can increase the probability of a claim without providing evidence for that claim.

Testing hypothesis - what deductive logic has to say in this regard?

- three elementary rules of inference involving *h*→ *t* and *t*.
  - *1. t*, *hence h*→ *t*, where *h* is an arbitrary statement.

    This inference pattern leads nowhere precisely because *h* is arbitrary. given *t*, logic will allow for **infinitely many hypotheses**.

  - *2. h and h*→ *t, hence t (modus ponens)*.

    But the *modus ponens* is useless for the empirical test of hypotheses, because our problem is now to find out the truth value of our posit *h* on no ground other than the truth value of a statement implied by it which, as shown by the previous inference rule, is a poor indication.

  - *3. -t and h*→ *t, hence -h (modus tollens)*.

  - *t and h*→ *t, hence h* (*fallacy of affirming the consequent*)

- Summary:
- deductive logic is powerless to validate initial assumptions
- deductive logic only helps to show which hypotheses are false

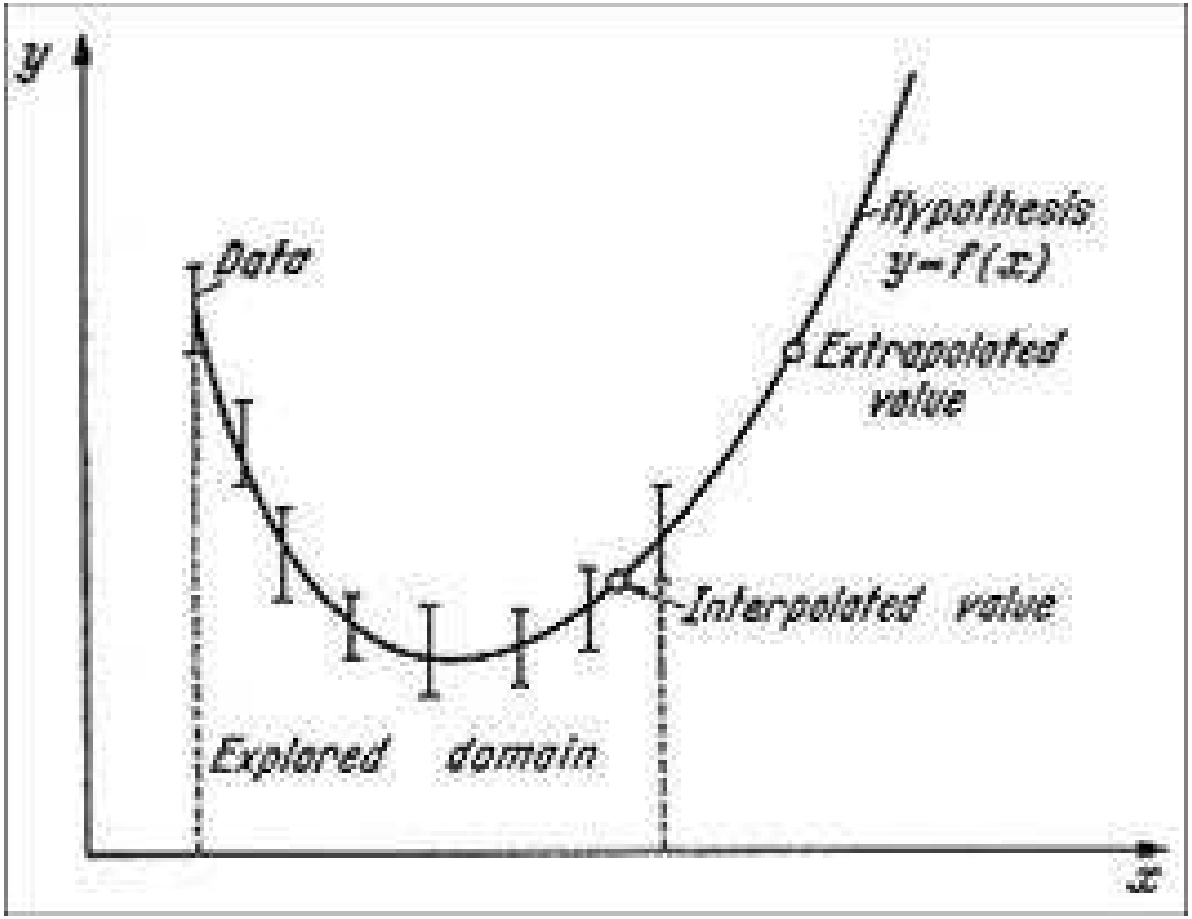
Bunge, 1998, 2, p. 342

All hypotheses are ***indirectly*** **testable** since, by definition, they contain non-observational concepts. Furthermore, they are ***incompletely*** **testable** by experience because, also by definition, they go beyond it.

1. Antecedent knowledge A unquestioned in the present research and relevant to a hypothesis h to be tested.	A
2. Problem: Is h true?	h?
3. Additional (test, auxiliary) hypothesis: From A and h derive "If h, then t". Unless t derives from A alone, h→t will have to be subjected to a further, independent test.	h→t
4. Translate t into t* for example, refer planetary positions, calculated in the sun's reference frame, to the earth. (simplification, Newton physics: “physical body” ⟶ “sun”)	t*
5. Experience: test for t*, i.e. find a set of data e comparable to t*.	e
6. Translate e into e* for example, express planetary position readings in geocentric coordinates. (simplification, Newton physics: “physical body” ⟶ “sun”)	e*
7. Compare t* to e* and decide whether they are epistemically equivalent or not; if they are, assert t*, if not negate t*.	
8. Inference concerning h: if t* was confirmed, declare h supported by t; otherwise conclude that h is undermined by t.	

- Step 3, i.e. the introduction of the **auxiliary hypothesis** " h→t " enabling us to make contact with experience.
- Steps 4 & 6, translation of non-observational to **observational concepts**

An "empirical curve" fitting a set of data and representing the hypothesized y=f(x):



Testing a hypothesis – verification, falsification, confirmation, …

- **Verificationism** (Hume, Comte...) – meaning of proposition **– only statements about the world that are empirically verifiable or logically necessary are *cognitively meaningful*** – Wittgenstein tradition, Carnap, Neurath, Weismann, Schlick –
- The thesis was elaborated in Waismann's "Theses" dated to "around 1930" which were presented as Wittgenstein's considered views. **"The meaning of its sentence is its verification"** in conversations with Schlick and Waismann on 22 December 1929 and 2 January 1930 (see Waismann 1967, *Wittgenstein and the Vienna Circle. Conversations Recorded*, ed. by B. McGuiness, transl. by J. Schulte and B. McGuiness, Oxford: Blackwell, repr. 1979.]
- **Schlick – all statements are meaningful for which it was logically possible to conceive of a procedure of verification** (concerned with constructed languages only) **1932**, "Positivismus und Realismus", *Erkenntnis*, 3: 1–31; transl. by P. Heath "Positivism and Realism" in Schlick 1979b, pp. 259–284, **1936** "Meaning and Verification", *Philosophical Review*, 45: 339–369, 1979b, *Philosophical Papers Vol.2 (1925–1936)*, Henk L. Mulder and Barbara van de Velde-Schlick (eds.), Dordrecht: Reidel.)
- **Outcome –** Wittgensteinian verificationist's criterion rendered **universally quantified statements meaningless**.
- **A J Ayer** - **two types of verification—*strong* and *weak*— weak** verification would be obtainable **when a proposition is rendered probable**. Ayer, Alfred J., 1936, *Language, Truth and Logic*, 2nd ed., London: Gollancz, 1946
- **Empiricist** demarcation: the **analytic/synthetic** and the **observational/theoretical** is tied together by *the verifiability criterion of meaningfulness*, according to which, in relation to a given language, L, a sentence *S* is meaningful **if and only if it is either analytic**-in-L or **synthetic**-in-L as an observation sentence or a sentence whose truth follows from a finite set of observation sentences.
- **Quine – Distinction: traditional demarcation of analytic/synthetic is untenable** "Two Dogmas of Empiricism", *Philosophical Review*, 60: 20–43, repr. in Quine *From a Logical Point of View*, Cambridge, Mass.: Harvard University Press, 1953, rev. ed. 1980, pp. 20–46.

Testing a hypothesis – verification, falsification, confirmation, …

- **Popper** urged **that verifiability be replaced with falsifiability** … *Logic of Scientific Discovery*, section 6, footnote 3
- **Carnap replaced verification with confirmation.** Rudolf Carnap, Testability and meaning, *Philosophy of Science* 3 (4):419-471 (1936)
- "If by verification is meant a definitive and final establishment of truth, then no (synthetic) sentence is ever verifiable, as we shall see. We can only confirm a sentence more and more. Therefore we shall speak of the problem of *confirmation* rather than of the problem of verification. We distinguish the *testing* of a sentence from its confirmation, thereby understanding a procedure-e.g. the carrying out of certain experiments-which leads to a confirmation in some degree either of the sentence itself or of its negation. We shall call a sentence *testable* if we know such a method of testing for it; and we call it *confirmable* if we know under what conditions the sentence would be confirmed. As we shall see, a sentence may be confirmable without being testable; e.g. if we know that our observation of such and such a course of events would confirm the sentence, and such and such a different course would confirm its negation without knowing how to set up either this or that observation."
- "... Therefore here also **no complete verification is possible** but only a process of **gradually increasing confirmation** ..."

- *Definition 16.* **A *sentence* S** is called ***confirmable*** (or completely confirmable, or incompletely confirmable) if the confirmation of S is reducible (or completely reducible, or incompletely reducible, respectively) to that of a class of observable predicates.

  [Note, 1950. Today I should prefer to replace Def. 16 by the following definition, based on Def. 18: A *sentence* Sis *confirmable* (or ...) if every descriptive predicate occurring in S is confirmable (or ...) .]

- *Definition 18.* **A *predicate* 'P'** is called ***confinnable*** (or completely confirmable, or incompletely confirmable) if 'P' is reducible (or completely reducible, or incompletely reducible, respectively) to a class of observable predicates.

Testing a hypothesis – verification, falsification, confirmation, …

- **Hempel** (1950, 1951), meanwhile, demonstrated that the verifiability criterion could not be sustained. Since it restricts empirical knowledge to observation sentences and their deductive consequences, scientific theories are reduced to logical constructions from observables.
  - `Studies in the Logic of Confrmation', was published in 1945, **right after first pioneer works in the area** (e.g. Hossiasson-Lindenbaum 1940), but **before Carnap's (1950, 1952)**
  - Hempel, 1950, "Problems and Changes in the Empiricist Criterion of Meaning" *Revue Internationale de Philosophie*, 41(11): 41–63.
  - Hempel, 1951, "The Concept of Cognitive Significance: A Reconsideration," *Proceedings of the American Academy of Arts and Sciences*, 80(1): 61–77.

- Three stages of empirical testing (Hempel [1945] 1965, 40-41):
  - First, we design, set up and carefully conduct scientific experiments, we try to avoid misleading observations, double-check **the data**, clear them up and finally bring them into a canonical form that we can use in the next stage.
  - In the second stage, these data are brought to bear on **the hypothesis** at stake - do they constitute supporting or undermining evidence?
  - Third and last, **the hypothesis is re-assessed** on the basis of a judgment of confirmation or disconfirmation: we decide to accept it, to reject it or to suspend judgment and to collect further evidence.

- In these three stages, **only the second stage is, or so Hempel argues, accessible to a logical analysis**: the first and the third stage are full of pragmatically loaded decisions, e.g. which experiment to conduct, how to screen of the data against external nuisance factors, or which strength of evidence is required for accepting a hypothesis. Evidently, those processes cannot be represented by purely formal means. That's different for the second stage which compares observational sentences (in which the evidence is framed) with theoretical sentences which represent the hypothesis or theory. This is the point where logical tools can help to analyze the relation between both kinds of sentences and to set up criteria for successful scientific confirmation.

Testing a hypothesis – verification, falsification, confirmation, …

Raven paradox (of confirmation) “Studies in the Logic of Confirmation” (1945)

Ravens are black ⇔	non-ravens are non-black
Ravens are black ⇔	shoes are white
	subst. non-ravens / shoes; non-black / white

- ***purely qualitative* predicates**
- But **paradoxical cases are confirmatory**: Hempel replied that even the existence of **a quantitative measure of evidential support** poses no challenge to his conclusion that the paradoxical cases—non-black non-ravens, such as white shoes—are confirmatory.
- *“Blite”*, “*Grue*” (Goodman 1955, *Fact, Fiction, and Forecast*)

Maher (1999, "Inductive Logic and the Ravens Paradox“): A non-raven (of whatever color) confirms that all ravens are black because

  - (i) the information that this object is not a raven removes the possibility that this object is a counterexample to the generalization, and
  - (ii) it reduces the probability that unobserved objects are ravens, thereby reducing the probability that they are counterexamples to the generalization.

- **Toward introducing auxiliaries**

1958, **Hanson** explained **that even direct observations must be collected, sorted, and reported with guidance and constraint by theory**, which sets a horizon of expectation and interpretation, how observational reports, never neutral, are laden with theory.

- **Glymour** solution – (1980, *Theory and Evidence*) embedded a refined version of Hempelian confirmation by instances in his analysis of scientific reasoning.
  - hypothesis *h* is confirmed by some evidence *e* even **if appropriate auxiliary hypotheses and assumptions** must be involved for *e* to entail the relevant instances of *h* (a *three*-place relation concerning **the evidence**, **the target hypothesis**, and **(a conjunction of) auxiliaries**).

What is the subject of the test: theory or isolated hypothesis?

- Duhem noted that **scientific hypotheses are tested in groups**.

- Kitcher - **Hypotheses are independently testable** (*without* recourse to the theory being tested). (Kitcher, "*Abusing Science*")

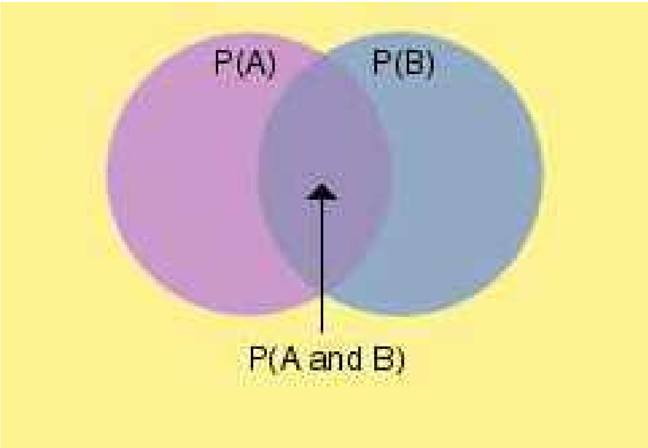
Testing hypothesis - Bayes' theorem

- Inputs:
  - data
  - prior probability ("a subjective degree of belief");
- |Bayes' theorem| → posterior probability
  - P(A|B) - conditional probability - probability of an event, based on another event happening, "the conditional probability of A given B", "the probability of *A* under the condition *B*"
  - P(A ∩ B) - joint probability;
  - P(A), P(B) marginal probabilities;
- Definition:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \mathrm{P(B)} > 0$$

- Analogously:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}, \mathrm{P(A)} > 0$$



$$\mathrm{P(B \cap A) \equiv P(A \cap B)} \rightarrow P(A \mid B) = \frac{\mathrm{P(B \mid A)P(A)}}{\mathrm{P(B)}}$$

Testing hypothesis - Bayes' theorem

- Let H be a hypothesis to be tested, and
- P(H) is called prior probability of H before new evidence E became available.
- P(E|H) is called likelihood - the probability of observing E *given* H.
- P(H|E) is the posterior probability.

$$P(H \mid E) = \frac{P(E \mid H)P(H)}{P(E)}$$

$$Posterior = \frac{\text{Likelihood} \times \text{Prior}}{\text{Normalizing factor}}$$

Testing hypothesis - Bayes' theorem (example)

http://stattrek.com/online-calculator/bayes-rule-calculator.aspx#report

- Marie is getting married tomorrow, at an outdoor ceremony in the desert. In recent years, it has rained only 5 days each year. Unfortunately, the weatherman has predicted rain for tomorrow. When it actually rains, the weatherman correctly forecasts rain 90% of the time. When it doesn't rain, he incorrectly forecasts rain 10% of the time.
- What is the probability that it will rain on the day of Marie's wedding?

- Event A1. It rains on Marie's wedding.
- Event A2. It does not rain on Marie's wedding
- Event B. The weatherman predicts rain.

- P(A1) = 5/365 = 0.0136985 [It rains 5 days out of the year.]
- P(A2) = 360/365 = 0.9863014 [It does not rain 360 days out of the year.]
- P(B | A1) = 0.9 [When it rains, the weatherman predicts rain 90% of the time.]
- P(B | A2) = 0.1 [When it does not rain, the weatherman predicts rain 10% of the time.]
- We want to know P(A1 | B), the probability it will rain on the day of Marie's wedding, given a forecast for rain by the weatherman.

$$P(A1 \mid B) = \frac{P(B \mid A1)P(A1)}{P(B \mid A1)P(A1) + P(B \mid A2)P(A2)}$$

$$P(A1 \mid B) = \frac{(0.014)(0.9)}{[\,(0.014)(0.9) + (0.986)(0.1)\,]} = 0.111$$

Testing statistical hypothesis (a sketch)

Frequentists: From Neymann/Pearson/Wald setup. An orthodox view that sampling is infinite and decision rules can be sharp. Hypothesis testing is the use of statistics to determine the probability that a given hypothesis is true.

The usual process of hypothesis testing consists of four steps.

- 1. Formulate the null hypothesis **H0** (commonly, that the observations are the result of pure chance) and the alternative hypothesis **Ha** (commonly, that the observations show a real effect combined with a component of chance variation).
- 2. Identify a test statistic that can be used to assess the truth of the null hypothesis.
- 3. Compute the P-value, which is the probability that a test statistic at least as significant as the one observed would be obtained assuming that the null hypothesis were true. The smaller the ***p***-value, the stronger the evidence against the null hypothesis.
- 4. Compare the ***p***-value to an acceptable significance value ***a*** (sometimes called an alpha value). If ***p≤a***, that the observed effect is statistically significant, the null hypothesis is ruled out, and the alternative hypothesis is valid.

Statistics (the same example)

- data

X	$\frac{1}{73}$	$\frac{72}{73}$
$P(X)$	$\frac{9}{10}$	$\frac{1}{10}$

MEAN

$$\mu = \sum x \cdot p(x)$$

$$\mu = \frac{1}{73} \cdot \frac{9}{10} + \frac{72}{73} \cdot \frac{1}{10}$$

$$\mu = \frac{81}{730} = 0.110958904109589$$

DEV.

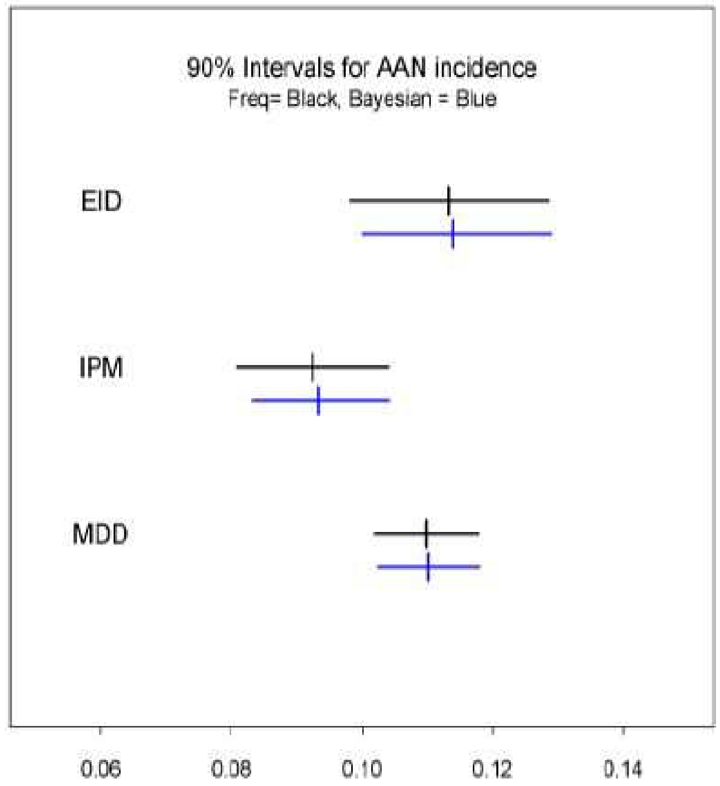
$$\sigma = \sqrt{\sum x^2 \cdot p(x) - \mu^2} = \sqrt{\frac{5193}{53290} - \left(\frac{81}{730}\right)^2} = \sqrt{\frac{45369}{532900}} \approx 0.2918$$

VAR.

$$\sigma^2 = \sum x^2 \cdot p(x) - \mu^2 = \frac{5193}{53290} - \left(\frac{81}{730}\right)^2 = \frac{45369}{532900} = 0.0851360480390317$$

Advantages of the Bayesian approach:

- ability to incorporate prior knowledge (prior beliefs) into the analysis
  - the Bayes factor will depend on the priors
- Bayesian assessment of detection and attribution is more explicit than in the frequentist framework.
  - frequentists claim detection when the H0 is rejected. It is not clear what it means in probabilistic terms.
  - attribution is claimed as failure to reject HA.
  - noniformative prior probability
- In Bayesian setting one must give, and justify, clear definitions for detection and attribution.
- Posterior probabilities can be used in a decision cost/benefit analysis to minimize expected costs and losses.



Evaluation (of rival hypothesis)

epistemic and cognitive virtues:

- testability
- explanatory power
- predictive power
- simplicity
  - syntactical, semantical, pragmatical
  - polemics on Goodman, “Safety, strength, simplicity” (1961)
- a priori (prior) probability
  - Salmon, The *Foundations* of Scientific Inference (1967); Toulmin; Carnap, Probability and Content Measure (1966)
- theoretical fruitfulness (fecundity, richness)
- conformity with background knowledge
- utility

Bunge, Hypothesis classification

- Formal (syntactical) properties of hypotheses
  - Structure of predicates
  - Range
  - Systemicity
  - Inferential power
    - specifiability (possibility of instantiation)
    - counterfactual force
- Semantical properties of hypotheses
  - *distributive (hereditary)* or global (nonhereditary)
  - order of semantical property
  - precision
  - Referent of predicates in respect of leading predicates - Mediate referent and Immediate referent
    - 1. Experience-referent hypotheses
    - 2. Experience and fact-referent hypotheses
    - 3. Fact-referent hypotheses
    - 4. Model-referent hypotheses

Bunge, Hypothesis classification

- Epistemological Viewpoint –focus on the inception, ostensiveness, and depth of hypotheses
  - Inception
    - 1. Analogically
      - (i) substantive analogy
      - (ii) structural analogy
    - 2. Inductively found hypotheses
      - (i) first degree induction
      - (ii) second degree induction
  - The degree of ostensiveness (or conversely of the degree of abstractness) of scientific hypotheses.
    - 1. Observational or low-level hypotheses contain only observational concepts
    - 2. Nonobservational hypotheses are those containing nonobservational concepts
  - Depth
    - 1. Phenomenological hypotheses
    - 2. Representational or "mechanismic" hypotheses

- Ground – more or less grounded on previous knowledge
    - 1. Guesses
    - 2. Empirical hypotheses
    - 3. Plausible hypotheses
    - 4. Corroborated hypotheses

Bunge, Hypothesis classification

- **Testability - the requirements an assumption must satisfy in order to pass for a scientific hypothesis are:**
  - (i) it must be **well-formed, self-consistent**, and have **maximal strength** with respect to the empirical evidence relevant to it;
  - (ii) it must **fit** the bulk of **relevant available knowledge**; and
  - (iii) it must, in **conjunction with other formulas**, entail consequences translatable into observational propositions.

  - **Untestable** formulas are those that cannot be subjected to test and cannot consequently be assigned a truth value.

  - **Test for factual truth** – attributions of factual truth can only be made on <u>**the strength of** empirical tests</u>.
  - **Empirically testable** means "sensitive to experience". Now a hypothesis sensitive to empirical data may be supported (confirmed) or undermined (disconfirmed) by them.
  - Testable hypotheses may then be
    - (i) purely confirmable,
    - (ii) purely refutable, or
    - (iii) both confirmable and refutable.

Bunge, Hypothesis classification

- **Strength**, a hypothesis h should be midway between any empirical information e relevant to it (weakest) and the body of knowledge A to which it may eventually be incorporated either as a theorem or as a postulate: $S(A) \geq S(h) > S(e)$.
  - An alternative way of stating the condition of maximal analyticity is to say that we want the body of antecedent knowledge A to give h a greater truth value than it would have if A were rejected: V(h|A)> V (h|– A). The larger the value of V(h|A) the better the theoretical validation of h. In the extreme case in which h follows from A we get maximal verisimilitude: (A|– h) –> [V(h|A)= 1. If, on the other hand, -h follows from A, then the prior (theoretical) truth value of h will be minimal: (A|– ~h) –> [V(h|A)= -1.

- **Strength** is **necessary but not sufficient** for testability – the stronger a hypothesis the greater its testability (K. R. Popper).
  - For example, at the beginning of our century several investigators suggested that the mechanism of nervous impulses was the release of chemicals; at the time (1903) this hypothesis <u>did not seem testable and was abandoned until, later on, adequate techniques were deliberately invented to test it</u>.
  - Moral: testability is not inherent in hypotheses but is a methodological property scientific hypotheses possess to varying degrees **relative to a body of empirical and theoretical knowledge**. Consequently, let us not write 'T(h)' for "h is testable" but rather 'T(h|AE)' for "h is testable relative to the antecedent knowledge A and the empirical procedures E".

- **the occurrence of empirical concepts somewhere along the line of the test process**: otherwise no experience would be relevant to our hypothesis.

Bunge, Hypothesis classification

- **No hypothesis is independently testable**, and this because no premise is by itself sufficient for the derivation of testable consequences. In the case of a theory we will need, in addition to the hypothesis itself, other assumptions of the theory and/or empirical data (since no hypothesis yields by itself empirical information. The question, then, is not whether a theory contains some assumption which is not independently testable and must therefore be regarded as suspect, but rather whether
  - (I) every one of the assumptions of the theory is actually necessary for the derivation of testable consequences,
  - (ii) the axioms make up a consistent whole having at least some empirically testable consequences (lowest-level theorems), and
  - (iii) there are no ad hoc hypotheses serving just to prop up some of the hypotheses and incapable of being tested independently of the hypotheses they protect.
- **Excluding parasitic untestable hypotheses** (plus ad hoc) clinging to an otherwise testable system and obtaining apparent support from the confirmation of the latter's low-level theorems.
- A **translation of theoretical into empirical statements** is not a purely linguistic affair: it consists in establishing certain **correspondences between conceptual objects** (e.g., "mass-points") **and empirical objects** (e.g., small bodies).
- In summary, the standard statement that experience tests hypotheses and theories is elliptical: propositions containing nonobservational theoretical concepts entail no empirical statements. What we do is to establish certain empirical models that can be more or less accurately matched with some low-level theoretical statements. **This is why we do not say that a hypothesis h entails its evidence e, but rather that h entails a testable consequence t that, when suitably translated, can be compared with an evidence e.**

Bunge, Function of hypotheses

Research consists in handling problems and every problem

- (i) is posed **within a body of knowledge involving hypotheses**, and
- (ii) is generated by a schema (**the problem generator**) that, when filled, becomes **either a datum or a hypothesis**.

Hypotheses are present in all steps of research, both in pure and in applied science, but they are particularly conspicuous on the following occasions:

- (i) when we attempt **to sum up and generalize the results of our observations**,
- (ii) when we try **to interpret the foregoing generalizations**,
- (iii) when we attempt **to justify (ground) our opinions**, and
- (iv) when we **design an experiment or plan a course of action** in order to either gather **further data** or **test an assumption**.

Ad hoc hypotheses

(the most tempting and the most distrusted by honest thinkers of all philosophical affiliations) **are unavoidable and welcome in science as long as** they contribute to enriching the foundation of important hypotheses and to ensure their consistency with other hypotheses. Ad hoc hypotheses are acceptable when they protect important ideas from rash criticism (such as the one based on the latest measurement), just as they are unacceptable if they prevent every criticism.

The main functions of hypotheses in science are the following:

1. **Generalizing experience:**

   summing up and extending available empirical data.

   A conspicuous subclass of this kind of hypotheses is the generalization, to an entire population, of "conclusions" (hypotheses) "drawn" from particular samples,

   Another conspicuous member of the same class is the so-called **empirical curve**, i.e. the continuous line that is made to join a set of points on a plane representing each an empirical datum.

   The extension beyond the set of data may be done, in this case, either by **interpolation** (assumption of values intermediate between those observed) or by **extrapolation** (assumption of values beyond the explored range, as in prediction).

   An "empirical curve" fitting a set of data and representing the hypothesized y=f(x).

2. **Inference starters**.

   Initial posits or wagers, trial hypotheses, work ing hypotheses, or simplifying assumptions serving as premises in an argument even if suspected to be false. Examples: (i) in an indirect demonstration we try the negate -t of the thesis of the theorem "h t" to be demonstrated and find out whether such a conjecture leads to a contradiction; (ii) the initial value assumed in the computation of a function by a method of successive approximations, or in the measurement of a magnitude, is likewise hypothetical; (iii) a grossly simplifying assumption that renders the application of a theory possible such as, e.g., that the earth surface is flat, or perfectly spherical over a certain region. Such hypotheses, when they are known to be strictly false, can be called pretences.

3. **Research guides:**

   exploratory guesses, i.e. more or less reasonable (founded) conjectures that are both the object of investigation and its guide. They range from precisely stated working hypotheses to rather vague conjectures of a programmatic kind.

4. **Interpretation:**

   explanatory hypotheses, or conjectures supplying an interpretation of a set of data, or of another hypothesis. Representational hypotheses are all interpretive since they enable us to interpret, rather than just generalize, data in terms of theoretical concepts. By contrast phenomenological hypotheses are of the generalizing type. For example, the electromagnetic field hypotheses (organized into a theory) explain the behavior of perceptible bodies of a kind.

5. **Protection of other hypotheses.**

   ad hoc conjectures the sole initial function of which is to protect or save other hypotheses from either contradiction with accepted theories or refutation by available data. For example, W. Harvey (1628) hypothesized the circulation of the blood, which is not an observable process and did not account for the difference between venous and arterial blood; to save the hypothesis he introduced the ad hoc hypothesis that the artery-vein circuit is closed by invisible capillary vessels, which were in fact discovered after him.

- Ďakujem za mimoriadne trpezlivú pozornosť