Summary of AK2030 Theory and Methodology of Science

Yashar Honarmandi
yasharh@kth.se

January 12, 2021

Abstract

This is a summary of AK2030 Theory and Methodology of Science.

This course, rather than teaching any particular aspect of science, will focus on general methodological and epistemological issues when approaching problems in science.

Contents

1 General Philosophy of Science

Definitions Definitions explain the meaning of words. They consist of a definiendum (what is to be defined), a definiens (that which defines the definiendum) and a defining connective (which describes how the two are connected. Words such as “is” as well as logical comparisons are in this category). In addition, there may be delimiters which explain that the definiendum is an aspect of something else. For instance, in the definition “A system is spherically symmetric if it looks the same from all directions,” “spherically symmetric” is the definiendum and “A system is” is the delimiter.

Lexical Definitions A lexical definition is a definition based on a common understanding of a word or phrase.

Stipulative Definitions A stipulative definition is a definition based on reasons and arguments for choosing it in a particular way.

Ambiguity Ambiguity is the presence of two possible meanings of the same word.

Vagueness Vagueness is the unclarity of the limits of what is within the scope of a definition.

Narrowness and Broadness A definition is too narrow if it does not apply to something that it should, and too broad if it applies to something that it should not.

Knowledge Knowledge is true, justified belief. These three conditions are necessary, but not always sufficient. We will nevertheless proceed with this definition.

This definition is compatible with uncertainty, which is important to us as recognizing uncertain knowledge is vital in science. Naturally there are degrees of uncertainty, and our idea of knowledge should capture that.

In science, the justification of belief is central in the social aspects of its practice. An ideally justified belief requires consideration of all relevant reasons for that belief, and science is publically defended exactly to adhere to this ideal.

Inference Inference is the logical process of going from a set of facts and/or evidence to some conclusion.

Inductive and Deductive Inference Inductive reasoning is the process of going from a set of observations to a general claim. Such reasoning is an attempt at extending our knowledge beyond what can be directly observed. Deductive reasoning is the process of explicating or extracting knowledge about something particular from previously possessed knowledge.

Inductive reasoning is fallible - even the best of efforts at induction might still overlook something for which one has no evidence. Deductive reasoning, on the other hand, is infallible - deduction necessitates that you have sufficient knowledge, and proper reasoning based on sufficient knowledge preserves the truth of your statements.

Examples of Deduction and Induction A few ways of inductive inference are

- Direct inference, where a property is extrapolated from a finite set of observations.
- Projection, where a sequence or trend is used to extrapolate a future observation.
- Generalization, where a hypothesis about a population is formed based on a finite set of observations.

A few ways of deductive inference are

- Modus ponens, in which a set of assumptions A is combined with the idea that if A is true then their implications B are true to infer that B is true.
- Modus tollens, in which a conditional $A \to B$ is combined with the observation that B is false to infer that A is false.

Ill-Formed Inferences An ill-formed inference is a logically inconsistent inference.

The Goals of Science The goals of science are to predict, explain and design phenomena. A phenomenon X is

- predicted if one can know at what time or in what way X will occur.
- explained if the causes of X are known.
- designable if it is known to satisfy certain properties.

Instrumentalism and Realism Do scientific theories constitute claims that are true or false? Realists say yes, whereas instrumentalists say that theory only orders observations and do not in themselves contain truth.

Hume's Problem As there exists an infinite set of inference rules, in particular inductive inference rules, the conclusions of any particular inference are only justified if the choice of inference rule is justified. Hume's problem discusses the justification of inference rules.

Hume's argument was the following:

1. Assume that inference is either deductive or inductive.
2. Assume that in order to justify an inference rule, it has to be inferred from some set of principles.
3. Assume that inference rules cannot be deduced, as past and previous inference need not be connected,
4. Thus, inference rules are arrived at inductively.

The implication of the latter statement is that there must be a set of inference rules allowing you to infer the validity of the rule you want use. But this set of inference rules is not necessarily valid in itself. The result is a so-called infinite regress of the same problem of justification, which according to Hume could only be resolved if no inference rule is justified.

Solving Hume's Problem Hume's problem would imply that the practice of science is itself irrational. This would be a complete disaster, and must therefore be resolved.

One way to solve this came from Karl Popper, who denied the premise that science rests on using inference rules. This will be discussed later.

Another solution is to deny the impact of Hume's argument by studying what offers justification, in the hope that this is not destroyed by the infinite regress. Foundationalism is one attempt at this that asserts that there exists a foundation upon which justification is built. Coherentism is an alternative, which desires that claims be justified from some coherent system containing the set of previously accepted claims. In this view, inference rules are good if they are coherent with the rest of scientific knowledge. The result is that inference rules are merely manifestations of coherence and a way to connect different practices in a coherent manner.

The Hypothetico-Deductive Method The hypothetico-deductive method consists of

1. Formulating a hypothesis.
2. Deducing consequences of the hypothesis.
3. Testing whether the consequences are true.
4. If the consequences are false, reject the hypothesis deductively. Otherwise, strengthen your confidence in the hypothesis inductively.

Quality Criteria for Hypotheses A good hypothesis must be

- Either true or false.
- Not a tautology, i.e. something that is necessarily true or false by definition.
- Contain some generalization or discussion of the unobservable.

A good set of consequences of a hypothesis must be

- Observable.
- Follow from valid deduction.
- Relevant for the hypothesis. This part informs what kinds of experiment you want to do. For instance, if the hypothesis is A implies B, then you would need to find cases where A is true and B is false at the same time to reject the hypothesis. In other words, studying cases where A is false or B can at best be a tool for verifying the validity of the deductions made.

Falsifiability and Falsification A theory or hypothesis is falsifiable if it has implications the truth or falsehood of which can be observed. The act of observing an implication to be false is called falsification.

Popper's Falsificationism Popper attempted to circumvent Hume's problem by rejecting the ability of science to strengthen hypotheses, leaving it only with the ability to deductively falsify them. While this seems to solve the problem, some critiques of this solution are:

- It provides no way to give hypotheses confidence, in contrast to what seems reasonable and what is scientific practice.
- Along with an observation comes auxillary hypotheses about the validity of the experiment. A particular falsification can only falsify the conjuction of the fundamental hypothesis and the auxillary hypotheses, limiting our ability to perform modus tollens on the fundamental hypothesis. This is called the Duhem-Quine thesis.
- Hypotheses may be modified ad hoc solely to fit particular observations. Popper countered this by claiming that modifications that reduce falsifiability are ad hoc and should not be used, in contrast to other modifications, which might be necessary.

Confidence Confidence in a hypothesis may be described either qualitatively or quantitatively.

Bayesianism and Frequentism A possible qualitative approach is to assign a probability of truth to a hypothesis and try to find it from observations combined with the probability of these observations given the hypothesis. This way of assigning probabilities is characteristic of the Bayesian approach to probability.

The frequentist view of probability, on the other hand, is that it is the frequency with which some event occurs. As hypotheses are not events, it follows from this view that probabilities cannot be meaningfully assigned to hypotheses, limiting the possibility of quantitatively strengthening confidence.

Strengthening Confidence How can our confidence in a hypothesis be strengthened, and why does observations of consequences of a hypothesis increase our confidence in the hypothesis itself?

One answer might be found in the compatibility of the hypothesis and its consequences. However, this is not strong enough on its own. Relevance is also needed - otherwise, many observations might seemingly verify far-reaching conclusions.

The issue of under-determination is also prominent, as multiple hypotheses might be compatible with the same observations. Given the observation of some consequences, the question then becomes which hypotheses should have strengthened confidence. One solution to this, often used in science, is Occam's razor, where the simplest hypothesis is selected. Another solution is to assign some relevance to the consequences of a hypothesis and only strengthen one's confidence upon the observation of relevant consequences.

Another requirement, stronger than the one previously presented, is that the conclusions be unlikely if the hypothesis is false. A test such that a given consequence is unlikely if the hypothesis is false is called severe testing. While this is an improvement, one must first of all avoid double-counting evidence - if you use some piece of evidence to construct a hypothesis, you cannot use that evidence to verify your hypothesis. In addition, if there is low confidence in the hypothesis to begin with, severe testing will not sufficiently strengthen it.

Observation Loosely speaking, observation is the event of experience through the senses. It is key to all scientific practice. Observations can be divided into categories:

- Direct observation, which is unaided sensory experience.
- Aided direct observation, which uses instruments to amplify the senses.

- Indirect observation, in which an event is not directly observed and the event is experienced through its effect on its surroundings.

Empiricism Empiricism is the thesis that sensory knowledge is the ultimate basis for knowledge.

Logical Empiricism and its Refutation Logical empiricism is the position that theory and experiment can be separated, that is that theory does not inform experiment. If this were true, then any scientific theory would rest on a solid and independent foundation of experimental observation.

This position has been refuted. Its refutation is as follows: First of all, indirect observation always depends on theory, as theory dictates how to interpret how the indirect observations pertain to the underlying phenomena. Next, aided direct observation sometimes requires theory as instruments may distort the phenomena they are used to observe, and separating between instrumental effects and actual phenomena requires theory. Thus theory and observation cannot be separated.

The Problem of Nomic Measurement The circular relation between theory and observation is formulated in the problem of nomic measurement. Its statement is as follows:

- Assume that you want to measure property X, which is not directly observable.
- Assume that X may be inferred from Y, which is directly observable.
- For the inference to be performed, one needs some relation $X = f(Y)$. This relation cannot be obtained experimentally obtained as it would require X and Y to be known at the same time, violating the assumption.
- As f cannot be identified from measurement alone, it must come from theory. Thus theory and experiment are inextricably connected.

Operationalization Direct observation, while powerful in a sense, is generally error-prone and can often only give qualitative descriptions, hence indirect observational methods are usually preferred. The process of linking properties (a so-called property of interest) to directly observable effects is called operationalization of that property. If the two are linked by some stable relation (a so-called hypothesized causal chain), that guarantees that when the property is present, so is the observed effect. Using operationalization we may infer information about the property from observing the effect.

Quality Criteria for Operationalizations The quality criteria for an operationalization are

- That the underlying property be well-defined in order to allow for proper inference.
- That the relation between the property and the effect is valid.
- That the relation is sufficiently stable.
- That the effect is publically observable with sufficient precision.

Operationalism Operationalism is the position that everything is defined through the operations by which we observe or measure them. Its main appeal is to remove the issue of theory-dependence.

Criticisms of Operationalism The criticisms of operationalism are:

- Operationalism prohibits two differently performed measurements to pertain to the same underlying concept.
- Operationalism makes it impossible to criticize a measurement for not properly capturing an underlying concept.

Measurement The process of measurement consists of the following:

1. Define the concept that you would like to measure.
2. Operationalize the concept.
3. Specify a measure and define units of comparison.
4. Represent the results with numbers.

Quality Criteria for Measurements The quality criteria for a measurement are:

- There must be a unit of comparison.
- The unit must be sufficiently stable.
- Anyone must have access to the same measure.

Scales The representation part of a measurement implies the need for scale. There are five kinds of scales:

1. Nominal scales, in which samples are given ID numbers.
2. Ordinal scales, in which objects are ordered according to some qualitative measure.
3. Interval scales, which allow the comparison of distances.
4. Ratio scales, which allow the comparison of both distances and ratios.
5. Absolute scales, which allow the comparison of absolute numbers, i.e. where the absolute numbers have a significance by themselves.

Different scales of the same type made to represent the same property must represent the same empirical structure. Each set of scales is defined by the set of allowed transformations, and combined with the previous this implies that what can be inferred from a measurement is limited by the allowed transformations. The allowed transformations are:

- Transformations that preserve uniqueness for nominal scales.
- Positive monotone transformations for ordinal scales.
- Positive linear transformations for interval scales.
- Positive scaling transformations for ratio scales.
- None for absolute scales.

Note that any scale also allows the transforms below where it appears in the list.

Measurement Error The measurement error is defined as the difference between the measured and true value of some property. It comes in two forms: Systematic and random error.

Before introducing the two, we introduce the notions of precision and accuracy. Precision is the absence of variation, and accuracy is closeness to the true value. Systematic errors are thus inaccuracies, and random errors are imprecision.

Convergent and Divergent Validity Convergent validity is achieved if different methods with causally independent operationalizations measure the same target under the same conditions and yield the same result. Divergent validity is achieved if an operationalization applied to substantially different targets under different conditions yield different results.

Experiments An experiment is a controlled observation in which the observer manipulates the real variables that are believed to influence the outcome, both for the purpose of intervention and control. Its purpose is to justify accepting or rejecting a hypothesis. Its characteristics are:

- manipulation.
- intervention with the independent variables.
- control of disturbing factors.
- observation.

Mill's Method of Difference Mill's method of difference describes the logic and procedure of experiments. It has the following steps:

1. Ask what causes a phenomenon E.
2. Conjecture that C is the cause.
3. Produce situations S_1 and S_2 in which neither C nor E occur and such that all relevant causal factors are the same.
4. Activate C in S_1 only.
5. Observe E in S_1 only.
6. Assert that something causes E in S_1 and that nothing causes E in S_2.
7. As the two only differ by C, conclude that C causes E.

Baked into these steps are a number of assertions and ideas which might be difficult to realize - for instance, it might be hard to know what factors are causal. This necessitates good design.

Repeatability, Reproducibility and Replicability If an experiment is described such that it contains enough information that others can repeat it, the experiment is repeatable. If a competent repetition of the experiment yields the same result, the experiment is reproducible. If a competent independent experiment using independent data, methods and experimental infrastructure yields the same results, the experiment is said to be replicable. These categories will be important when looking for errors in experimental design.

Internal and External Validity If the relation between intervention and observed effect inferred from an experiment is true and not clouded by uncontrolled background factors, it is internally valid. If such an inference also holds for different targets and not just the same experiment, it is externally valid.

Models Models are important tools in science used to help us describe the world. They have different aspects, and may be approached through any of them. These are:

- Representations, where the models attempt to represent some target in the real world to some extent. This is useful when studying the real-world target might be impossible or infeasible, morally or ethically prohibited or cognitively difficult.
- Idealizations, where the model contains only the relevant aspects of the real-world target. Models are generally analogies of the real-world target. These may be positive (containing similar aspects), negative (containing idealizations) or neutral (containing descriptions of things that cannot be known in the target).
- Purpose-dependent tools - some models describe certain properties better than others.
- Things to be manipulated. The kind of analogy provided by the model limits the available set of manipulations - if some property is idealized away in the model, then investigating that property in the model has no bearing on the target. Manipulations may also be used to reveal neutral analogies as either positive or negative.

Quality Criteria for Models Models do not necessarily have quality criteria that apply equally well in all contexts, and in fact these often trade off with one another. Nevertheless, some criteria are:

- Accuracy. A model is accurate if it is similar to the target with respect to relevant properties.
- Robustness. A model is robust with respect to some condition if changing it does not change the result.
- Parameter precision. A model has higher parameter precision than another if the specification of the former implies the specification of the latter.
- Simplicity. A model is simpler than another if it contains fewer variables and parameters than the other.
- Tractability. A model is tractable with respect to some principles if the results can be obtained by applying the set of principles to the model.
- Transparency. A model is transparent if the user is capable of understanding how the results are produced.

Models and Experiments Models and experiments are similar in that they both contain a set of variables and parameters, as well as both entailing the manipulation of something and the observation of the effect. However, they differ in the errors that can be made. The greatest issues with experiments concern internal validity, whereas the greatest issues with models concern whether the relevant analogies of the model hold.

The Structure of Explanations Explanations consists of an explanandum - what is to explained - and one or more explanans - statements that increase understanding. An important ingredient in explanations is contrasting the explanandum to some other scenario, and this contrast might affect the answer you seek or obtain.

Scientific Knowledge and Explanations The process of explanation is vital in fulfilling the goals of science - namely:

- making predictions by providing reasons for expecting a phenomenon to occur in a particular way or at a particular time.
- explaining phenomena by providing lawful reasons why a phenomenon occurs.
- designing phenomena by providing reasons for expecting why a certain manipulation satisfies certain functions.

Our description of explanation seemingly makes it very similar to prediction - one looks to the future, the other to the past.

Understanding Understanding a phenomenon can be taken to be the ability to answer questions about what would happen if things were different. The process of understanding thus involves tracing so-called productive relations, which describes what features of a system produces what others.

Causation We try to clarify the notion of causation by defining some associated terms.

X is a direct cause of Y with respect to a set of background variables V if and only if there exists an intervention on X that changes Y with V held constant. Built into this notion is the idea that X causes Y with respect to some background variables, which is a model. Therefore one cannot say that X causes Y generally.

X is a contributing cause of Y with respect to a set of background variables V if and only if there exists a causal chain comprised of direct causes extending from X to Y.

The Deductive-Nomological Account The deductive-nomological account is way of providing an explanation. It answers the question of why some particular phenomenon occurred by deducing it from a set of natural laws and circumstances.

Critiques of the Deductive-Nomological Account If we modify our concept of explanation to be the process of providing understanding, we see that explanation must entail identifying productive relations. This is a better notion of explanation because it helps us distinguish between explanation and prediction. The deductive-nomological account does not distinguish between the producer and what is produced, and is therefore insufficient to provide explanations. Furthermore, the deductive-nomological account does not identify the relevance of the different circumstances.

Another critique comes from considering so-called singular causal explanations. These are explanations based purely on recounting a sequence of events. In many practical contexts such explanations are sufficient, but according to the deductive-nomological account they are not. Hence the deductive-nomological account does not cover everything that we would like to consider an explanation.

Causal Explanations and their Quality Criteria Causal explanations are explanations that identify difference-making contributing causes to an explanandum.

Quality Criteria for Causal Explanations The quality criteria of causal explanations are:

- Accuracy - the ability of the explanans to describe the state or properties. The explanans only needs to identify the difference-making contributions.
- Precision - the more precisely the explanandum states a contrast, the better.
- Difference-making - the explanans must identify all contributing causes.
- Non-sensitivity - the explanans must have low sensitivity to background causes.
- Cognitive salience - the explanation should be fit to its audience. This typically means making it simpler.

2 Ethics

Morality Engaging with morality involves asking oneself questions about what one may or ought to do. The answers may be permissions, prohibition and obligations.

The descriptive view of morality describes it as a code of conduct accepted by a society, group or individual. This view is limited, however, as it carries no natural distinction from, for instance, etiquette, law and instrumental rationality. Furthermore, it carries a high degree of relativism.

The normative view of morality, which will be the focus of the rest of the discussion, views morality as a code of conduct that all rational persons would endorse.

Ethics Ethics is a theory that offers normatively valid reasons for rationally endorsing some code of behaviour. Depending on the choice of ethical theory, people might agree on a particular choice but disagree on why this is the correct choice.

Consequentialism Consequentialism is an ethical framework in which moral assessment is based on the consequences of the action. This begs the question of how to assess the consequences of an action. Subtheories of consequentialism specify this further, utilitarianism being a noteworthy example.

Some problems with consequentialism include:

- It is extremely demanding, requiring all possible consequences to be evaluated. As a consequence of the complexity, all actions are seemingly either required or forbidden.
- It permits some intuitively wrong actions.

Deontology Deontology is an ethical framework in which an action is moral if it fulfills relevant rules or duties.

Compared to consequentialism, deontology is simpler, includes the morally gray and the amoral and incorporates the preference of the agents into ethical considerations. However, its issues include:

- The possibility of allowing disastrous consequences in the name of principle.
- The fact that rights and duties are categorical, limiting flexibility.

Virtue Ethics Virtue ethics asserts that morality consists of having and exemplifying certain good character traits, or virtues.

Compared to consequentialism, it lacks the overdemanding nature, and compared to deontology, it explains the motivation for people to be moral. This system has problems too, however, including:

- Providing no clear guidance for how to act in any particular situation.
- Making it hard to resolve conflicts.

Morals and Experimental Design In experimental design moral questions often arise, in particular when the subjects are human. What is important to maintain the morality of the study and how important different aspects are will naturally vary according to different ethical frameworks.

Informed Consent Informed consent of a human subject to being experimented upon is consent that is

- informed - the subject has received all relevant info about the purpose of the project, how it will be carried out and the effects on themselves.
- voluntary - not forced and sufficiently free from negative influence.
- decisionally-capacitated - the subject can assess the provided information, appreciate how it concerns them and make and communicate their decision.

According to a consequentialist, informed consent might be important because individuals know best what is good for them and because it is good for science as an institution to establish and maintain trust. According to a deontologist, one should treat others as autonomous beings and not lie. According to a virtue ethicist, it is good to show sincerity, reflexivity and respectfulness.

Scientific Misconduct Scientific misconduct may manifest as

- fabrication - making up data or results.
- falsification - manipulating materials, equipment or process, or changing or omitting data or results, such that the research is not accurately presented.
- plagiarism - appropriation of other people's ideas, processes, results or words without credit.

A consequentialist might argue that misconduct undermines the trust of the institution of science or wastes resources. A deontologist might argue that lying violates some fundamental principle. A virtue ethicist might argue that the moral path would be to show sincerity and humility, and thus not perform such acts.

Whistleblowers A whistleblower is someone who exposes scientific misconduct.

Authorship A person is assigned authorship of a work if they substantially contributed to the conception of the work, drafted the final work, gave final approval and agreed to be held accountable to all aspects of the work.

Why is the proper assignment of authorship important? A consequentialist might argue that the impartiality of science, public trusti in science, incentives for authors and the saving of resources are important consequences of this. A deontologist might argue that proper authorship preserves intellectual property rights and properly assigns responsibility. A virtue ethicist might argue that proper authorship reflects and allows authors to display sincerity, humility, resoluteness and reflexivity.

Ghost Authors and Gift Authors A ghost author is someone who contributes to a work according to the authorship criteria but is not listed as an author. The opposite is a gift author, or honorary author.

The Precautionary Principle The precautionary principle states that if an activity raises threats to human health or the environment, precautionary measures should be taken even if all causal relations are not established.

Consequentialism and Risk Consequentialism has a few approaches to the issue of risk. One is the actualist approach, which measures the goodness of an action solely based on the actual outcome. This approach is clearly insufficient. Another is the maximization of expected utility, defined as the probability,weighted average of the utility.

Criticisms of the consequentialist approach include:

- one might take it as important to avoid disproportionate disaster, but the consequentialist approach provides no obvious way to take this into account.
- one might want to avoid disproportionately exposing certain individuals, but this does not emerge naturally from consequentialism.

Deontology and Risk When assessing risk using deontology, a fundamental issue is how small a probability of a risk violation actually counts as a violation. This issue is extremely complicated, some complications including distinctions between intentional and unintentional risks, voluntary and imposed risks, self-produced and externally produced risks and the existence and absence of benefits.

3 Engineering

The Roots of Science and Engineering The natural sciences are rooted in natural philosophy, with the goals of explaining and understanding the world. The knowledge was formulated in terms of abstract reasoning and theory and communicated with books and lectures.

Contrast this with engineering, which is rooted in the crafts. The goal of the crafts is the creation of so-called technical artefacts. The knowledge was specified and local, and communicated to training on the job by masters. From this point of view the distinction between the two is more apparent and can be traced.

Tacit Knowledge Tacit knowledge is a kind of knowledge that is difficult to transfer verbally. Crafts typically deal significantly in this kind of knowledge.

Modern Engineering Modern engineering has changed the engineering of the past, turning the tacit explicit, the local general and the trial-and-error approach into a systematic approach to attaining knowledge.

Technical Artefacts Technical artefacts are:

- Existing material objects.
- Purposeful.
- Objects that realize their function through their physical properties.

Inherent in this is a duality between the physical properties and function of an artefact - both enter into the notion of a technical artefact, but neither can be reduced to the other. An artefact may be characterized by describing its properties, the intended use by the designer and the application by a user.

Function The function of an object can be ascribed by a user or assigned by a designer. The assignment by a designer is performative, and if successful, the designer has described a new instance of a functional kind.

Devices may:

- Malfunction - lose the physical property necessary for their function.
- Be misused - engaged in a way that does not establish the intended function or any alternative function.
- Fail - be unable to support he function.

The latter implies that only assignments backed up by expertise establish function.

The assignment of function by a designer is a normative process in a different way than the ascription by a user is - it establishes normative claims about what the artefact ought to do.

The Design Process In general the design process involves the chain starting at a set of needs, moving trough a functional description and a physical description to create an artefact. The core description of design may be taken as the development of a physical description corresponding to a functional description. A creation-centered description of design is the creation of physical objects that robustly satisfy a functional description. A client oriented description of design is the creation of physical objects that robustly satisfy a client's needs.

Involved in the design process are multiple steps, each requiring knowledge. When moving from needs to a functional description, what functions must the object satisfy to fulfill the given needs? When moving from a functional to a physical description, which structures might satisfy the necessary functions? When moving from the physical description to the object, how does one produce objects with the desired physical properties? The knowledge required for each step is not necessarily contained within science or engaged with by scientists, again highlighting the distinction from engineering as simply applied science. In particular, you might find operational principles, normal configurations, device-restricted simplifying theory, phenomenological theory and quantitative assumptions.

Design Methodology In the above chain there are a number of issues of a methodological nature.

The first is the determination of appropriate functional requirements when going from the needs. Determining these should be done in such a way as to

- be implementation-independent - refrain from specifying the physical implementation.
- be complete - describe all the desired effects.
- be adjusted to the use-context - takes user behaviour into account.
- have quantitative criteria.

The next is in the determination of the physical description. This can be done by

- functional decomposition.
- blind variation within the confines of previous knowledge and theory.
- weeding out possibilities through modelling or testing.

Next there is the issue of deciding between possible solutions. One way to do this is by optimization, involving an exhaustive search and complete evaluation of possibilities. This might work, but is a lot of work. Another possibility is so-called satisficing, where one starts by finding a minimal function threshold and searches until that threshold is met. This is often more close to practice.

The final issue is validation, which may be performed in many ways. Some examples are

- testing.
- proof-of-concept.
- prototypes.
- direct trials.
- usability testing.

Definitions of Risk There are many different defintions of risk in use, each reflecting different aspects of the notion. Some definitions are:

- an unwanted event that may or may not occur.
- the cause of such an unwanted effect.
- the spread of possible outcomes of some action.
- the probability with which some unwanted effect occurs.
- the severity of some possibly occurring unwanted effect.
- the expected value of some uncertain event.

Primary and Secondary Prevention Primary prevention is eliminating the cause of some unwanted effect. Secondary prevention is providing means of mitigating the unwanted event itself.

Conditions of Decision Making Decisions can be made under

- certainty, where there is only one outcome.
- risk, where there are several possible outcomes and the results of all outcomes and probabilities with which they occur are known.
- ignorance, where some probabilities are unknown.
- deep uncertainty, where some unknown possibilities exist.

Risk Assessment and Management Risk assessment is the collection of information in order to estimate risk. RIsk management is making a judgment based on the collected information about how to handle the risk.

Expected Value-Based Risk Some arguments for using an expected value-based approach to risk include:

- It allows for including risk in optimization calculations, assuming that all involved factors are given in or can be converted to the same currency.
- It keeps science value-free by more easily allowing for the separation of risk assessment and management.

There are some problems with this approach as well, including:

- the existence of risk-adversity and risk-loving, i.e. the preference of possible gains or possible losses. The expectation value tells you nothing of how to make a choice if you lean strongly towards either category. One might counter this by introducing a notion of utility and optimizing this instead.
- the issue of expected value as defining risk, as opposed to merely measuring it. For example, variations might be an interesting to parameter to include in one's assessment, but if expectation value defines risk, variations could not possibly have any relevance.
- the presence of other subjective evaluations. For example, psychological studies show that humans consider involuntary risks worse than voluntary ones, views risks as worse when they lack control, consider novel risks worse and generally underestimate large risks and overestimate small risks, although the importance of these considerations is debated.

Decision Making Under Ignorance If one must make decision under ignorance, some strategies to proceed include:

- transforming the problem to a format where expectation value or utility optimization is possible. One common way to do this is to appeal to the principle of insufficient reason and assign all outcomes the same probability.
- employing the maximin strategy, where the focus is put on extreme outcomes, primarily the worst ones, and decisions are made only based on those.

Decision Making Under Deep Uncertainty If one must make decisions under deep uncertainty, some strategies to proceed include:

- working with safety factors. Engineers often do this both to account for quantifiable risks such as high loads and worse material, and for inquantifiable factors such as imperfect failure theory, unknown failure mechanisms and human errors.
- appealing to the precautionary principle.

Reasons for Using Non-Quantitative Risk Assessment One might want to use non-quantitative risk assessment to avoid high computational costs, obtain a simple risk analysis, avoid the lack of accuracy possibly involved in a quantitative assessment or for security purposes not sufficiently accounted for in quantitative assessments.

4 Methodology

Methodology is the study of the design of methods for gaining knowledge. It is different from general philosophy of science in that it does not contain, for instance, comparison, choice and justification of methods.

Choice of Method To answer the question of what methods should be chosen, three approaches are typically found:

- The conventional approach, in which you choose the same methods as your teachers or peers.
- The outcome-oriented approach, in which you choose the method that gives the best results.
- The reason-based approach, in which you choose the method for which you have the overall best reasons. Reasons here means considerations of the method with respect to a set of goals that you want to achieve.

Their advantages and disadvantages may be summarized as in table 1.

Approach	**Advantages**	**Disadvantages**
Conventional	• Makes choices easy	• Leaves you less open to correcting methodological mistakes and being critical of your own results • Makes it hard to collaborate with others due to inflexibililty
Outcome-oriented	• Oriented towards an intelligent choice of method	• Is too vague - what are the best results, for instance, and who judges this? • Science often involves long-term planning, so the optimal choice of method might be hard to know a priori
Reason-based	• Flexible	• Makes choices hard

Table 1: Advantages and disadvantages of approaches to method choice.

Experimental Control Experimental control consists of identifying relevant features that might interfere with your experiment and being able to influence these features such that alternative explanations can be ruled out.

A few methods for implementing experimental control are:

- Dividing your subjects into treatment and control groups.
- Keeping other factors constant by finding or creating situations with the same background.
- Eliminating background factors altogether by creating special circumstances. One way to do this is to blind your study.
- Separate out external factors by registering and accounting for their contribution.

Non-Experimental Procedures The types of experiments described by its definition are in principle limited to a lab environment, but there are other kinds of non-theoretical activities that are important in science, but are not experiments.

The first is observational studies, where you study subjects with no ability to manipulate, intervene or control the subjects.

The second is natural experiments, in which you have no ability to manipulate, but circumstances are equivalent to intervention and control having been done.

Next there are field experiments, in which you have the ability to manipulate, but some background factors cannot be controlled through manipulation.

A final one is simulation, where you have the ability to both manipulate, intervene and control, but only on a representation of your system.

Design Errors Erroneous design is design that fails to account for some background factor(s). It may manifest as:

- Failure to control for all relevant factors.
- Confirmation bias(/interpretation problem) in observations.
- The observer effect(/influence problem) spoiling your results.
- The placebo effect interfering.
- Selection bias in your treatment and control groups.

To detect such errors, one might employ what knowledge one possesses, of either theory or experimental practice, to identify them, or one might investigate the repeatability, reproducibility and replicability of the experiment to find out where it might have gone wrong.

Randomization Randomization is the process of dividing samples into treatment and control groups using randomness. Experiments set up this way are called randomized control trials.

Randomness is desired because it might eliminate selection bias, help convince others that your experiment is not rigged and can be helpful in blinding the identity of the treatment from both investigators and subjects. However, randomization is not the only way to achieve these goals.

Randomization does not help to equally distribute background factors between the treatment and control groups. Imbalances may instead be avoided by being checked and controlled for. One might also use so-called stratified randomization, where background factors are included in the randomization process.

Randomization has no bearing of the broadness of your sample, meaning that it adds nothing to the generalizability of your results to other populations than that which you have sampled.

Blinding Blinding is the process of obscuring the identity of the control and treatment groups, with the purpose of removing observer effect. There is single blinding, where the subjects do not know whether they are in the treatment or control group, and double blinding, where neither the subject nor the experimenter knows to which group the subject belongs.

Reducing Error Random error can be reduced by repeating measurements and performing new operationalization in such a way that you obtain more precise measurements.

To reduce systematic error, one must identify the causes. One can only be sure that no systematic errors are present if one has a collection of measurement processes that both exhibit divergent validity individually and convergent validity collectively. If this is not found, the search begins, with few general guidelines. A few examples of places to look are background variables, observer effects and the stability and standardization of measurements.

The control of error represents a set of auxillary hypotheses which enter in the hypothetico-deductive method.

Statistics Statistics is a set of mathematical methods for dealing with quantitative data. It is divided into descriptive and inferential statistics.

Lying With Statistics The practice of lying with statistics involves taking good data and using proper statistical methods to produce false or misleading claims.

Methodology and Statistics The methodological aspects of statistics entail choosing the right statistical tool for some particular task, justifying the choice and choosing a proper representation of the data.

Why Use Statistical Inference? Statistical inference might be useful for hypothesis testing if:

- The hypothesis has stochastic implications.
- One wants to quantify error or confidence.

Note that the statistical treatment generally weakens individual accounts of confirmations and falsification if the hypothesis is deterministic.

Fisher's Significance Testing Fisher's significance testing is a framework for rejecting hypotheses. Its steps are as follows:

1. Specify a hypothesis H.
2. Devise an experiment to test H and specify the possible outcomes of the experiments.
3. Determine the distribution of the test statistics.
4. Observe the outcome of the experiment.
5. Calculate the p-value, defined as the probability of observing a result as least as extreme as the observed one given H.
6. If p is smaller than some threshold value, termed the significance level, reject H.

p-Value Abuse Fisher's significance testing leaves room for exploitation. A few examples of how this can be done are:

- Modifying the hypothesis so as to mislead readers about what you have shown.
- Change the experiment or sample size until you get the desired result.
- Change how outcomes are partitioned.
- Use other sample distributions.
- Change the threshold.

Neyman-Pearson Testing Fisher's framework only leaves room for considering one hypothesis at a time. The Neyman-Pearson framework allows you to test two mutually exclusive and jointly exhaustive hypotheses. Labelling one of them as H_0, the outcome is designated according to the table below.

	H_0 is true	H_0 is false
H_0 is accepted	Correct	Type-II error
H_0 is rejected	Type-I error	Correct

Table 2: Designation of outcomes in Neyman-Pearson testing.

The typical approach is to set the desired type-I error rate to some acceptable value and perform power analysis on your test given this. The power of the test is defined as the probability of rejecting H_0 given that it is true, and depends on the set type-I error rate, the magnitude of the effect of interest and the sample size. Once that is done, you compare the obtained p-value to the set type-I error rate and use that to accept or reject H_0.

Bayesian Statistics The Bayesian approach to statistics provides another way of using statistics for testing a hypothesis. In this approach, probabilities of hypotheses being true are interpreted as subjective confidence in hypotheses. The steps are:

1. Formulate a set of competing hypothesis.
2. Determine prior probabilities of each hypothesis being true.
3. Collect data not used for informing your assignment of probabilities.

4. Determine the probability of the data given the hypothesis.

5. Compute the probability of the hypothesis given the data using Bayes' theorem.

6. Update the prior probabilities using the above.

Some issues with this approach include:

- The issue of assign prior probabilities. This may be solved by approaching the issue in a subjectivist manner, where one realizes that as long as the priors are not 0 or 1 everyone will agree in the end, or an objectivist approach, where one starts the process by dividing one's belief equally between the hypotheses.

- The issue of old evidence, which may not be reused.

- The issue of uncertain evidence, as the Bayesian approach builds on letting the probability of the evidence occuring going to 1.

Modelling Strategies I here introduce two modelling strategies. Their approaches reflect how we learn from models.

The first is to consider models as mirrors of the world. In this view you desire your model to be precise, a feature that you typically gain at the cost of simplicity, tractability and transparency. You will still not gain enough precision to avoid issues of external validity, i.e. to guarantee that the model describes the target.

The other view is to view models as isolations of relevant features of a system. This view requires that the target be divisible, at least to some extent, in such a way. It is also difficult to validate, as the target also has interaction effects.

Strategies for Identifying Causality While correlations are not sufficient or even necessary for causal relations, they are an important part of the evidence for certain causal relations. One strategy for identifying causal relations is Mill's method of difference, which has the drawback of requiring experiments. Another is instrumental variable analysis, defined by the following steps:

- Observe a correlation between X and Y.

- Find a variable Z that is known to affect X, but not Y.

- Use Z instead of X when estimating the effect of X on Y.

This strategy lets you work out causal relations only by observing correlations, but requires some causal knowledge in order to be implemented.

Interpretation of Behaviour The interpretation of human behaviour is an important part of qualitative studies. The process of interpretation may be explained using two frameworks. The first is simulation theory, where one's own cognitive mechanisms are used to simulate other people's mental state. The other is theory theory, in which a theory of the mind is constructed using evidence and hypotheses.

The Belief-Desire Explanation One way of explaining human behaviour is the following:

- Assume i desires outcome X over Y.

- Assume i to believe that action A would bring about X and not Y, while action B would bring about Y and not X.

- Infer that this explains why i chose A over B.

Methodological Behaviourism Methodological behaviourism is the rejection that mental states can be inferred, favouring instead the discovery of regularities between stimulus and and response.

Interpretation as Adopting the Intentional Stance Another way to interpret behaviour, human or otherwise, is to adopt the so-called intentional stance, described as follows:

- Treat the object the behaviour of which is to be explained as rational.
- Figure out a set of beliefs and desires it ought to have given its place in the world and purpose.
- Decide what it ought to to given the attributed beliefs and desires.
- Predict that it will do what it ought to.

Construct Validity In observing human behaviour, we need to assume that there is a relationship between the mental attitude of the subject and its response to the investigation of the mental state. If this exists, the observation has construct validity.

Quality Criteria for Qualitative Methods The quality criteria for qualitative methods are:

- repeatability.
- control of observer effect.
- reflection on observer bias.
- robustness, often ensured by the use of multiple investigation methods.

Interpretation Quality Interpretation is fallible. Notably, for a single input the process of interpretation involves many assignments of mental states, and many models provide plausible explanations for the same behaviour.

One solution is to avoid interpretation as a whole if possible, adopting methodological behaviourism.

Another solution is to take various steps to improve the quality of your interpretations. This can be done by

- minimizing the scope of the interpretations.
- find more evidence to help in uncovering mental states and processes.
- ensure the competence of the observers.
- check for coherence in the interpretation.
- perform intersubjective checks with a colleague.

Case Studies A case study is a study characterized by a low number of cases, all chosen for specific reasons, displaying the same outcome with no experimental manipulation performed. The goal of case studies is to understand the specific case. This is in contrast to multivariate analyses, which involve many cases chosen randomly or representatively where a dependent variable varies and experimental manipulation is performed or mimiced. The goal of multivariate analyses is to identify and understand patterns.

There are many potential purposes for case studies. The first is simple falsification of some claim or hypothesis.

Another is the establishment of analytical narratives. This is done by

- making a research question.
- selecting cases.
- using a narrative to elucidate the principal players, their preferences, key decision points and possibilities.
- the evaluation of the model through the testing of its implications.
- the consideration of alternative narratives.

Yet another possibility is qualitative comparative analysis, which is an analysis of a set of cases performed in order to gain a better understanding of the necessary causes.

Finally there is process tracing, which investigates cause-effect relations in a single case. The involved steps are:

1. Formulate hypotheses about any mechanisms in the case.
2. Work out differences in the empirical implications.
3. Check the case data for mechanistic fingerprints.
4. Conclude which mechanism applies.

Quality in Case Studies Steps that can be taken to improve case studies include:

- Clearly stating research questions.
- Clearly specifying theory background and hypotheses.
- Clarifying what constitutes a case unit.
- Making case selection explicit.
- Specifying what makes different cases comparable.
- Investigating all implications of one's theoretical hypothesis.
- Investigating rival hypotheses.