**Design Thinking**

1. **Quantitative Method: -**
2. **Module 1: Intro**
3. **Module 2: Origins**

* ***Origins of Scientific Method***
* In science we try to resolve these questions by using a set of principles and procedures called the scientific method.
* ***Short Story***
* **Classic –**
* Ancient Greek thinkers like Thales, Pythagoras, and Democritus sought natural explanations instead of divine ones, but Plato and Aristotle were the first to deeply consider how we gain knowledge.
* Plato believed that reality consists of ideal forms that exist independently of the physical world. Since the physical world is an imperfect reflection of these forms, true knowledge can only be gained through reasoning, making Plato a rationalist.
* Aristotle, while also a realist, believed that reality is the physical world. He argued that knowledge comes from sensory experience and observation, making him an empiricist.
* Aristotle developed formal logic, particularly the syllogism, as a tool for building knowledge. However, its reliability depended on the truth of the premises, which he believed could be established through observation.
* Some of Aristotle’s fundamental premises were incorrect due to selective observation, such as believing that insects have four legs or that men have more teeth than women.
* Plato’s and Aristotle’s ideas dominated for almost 2000 years, with little scientific advancement during the rise and fall of the Roman Empire and the early Middle Ages.
* Arab and Persian scholars like Ibn al-Haytham and Al-Biruni in the 10th century emphasized systematic observation and experimentation, setting the foundation for modern science.
* Thinkers like Roger Bacon and Grosseteste further refined scientific methods by advocating both inductive and deductive reasoning.
* The invention of the printing press allowed scientific ideas to spread widely, creating the conditions for the Scientific Revolution.
* Copernicus proposed a heliocentric model, but it was Galileo who defended it as a true representation of reality, directly challenging the Catholic Church and marking the beginning of the Scientific Revolution.
* **Enlightment –**
* Galileo is considered the father of modern science because he explicitly separated science from philosophy, ethics, and theology, which were under the Catholic Church’s control.
* He disproved Aristotle’s claim that heavier objects fall faster than lighter ones through logical reasoning and experimentation.
* Galileo provided observational evidence, such as Venus’s phases, proving that planets orbit the sun, directly challenging the geocentric model supported by the Church.
* Galileo’s assertion that heliocentrism was physically real led to his trial by the Inquisition and house arrest until his death.
* Descartes, despite rejecting many of Aristotle’s ideas, emphasized reasoning as the primary path to knowledge, famously stating, "Cogito, ergo sum" (I think, therefore I am).
* Francis Bacon championed inductive reasoning, arguing that knowledge should be built purely on observations and experience, making him the father of empiricism.
* Hume took empiricism to the extreme, arguing that since observations don’t guarantee future occurrences, we can never be certain of any scientific claim.
* In response to Hume’s skepticism, German Idealists proposed that reality is a mental construction and that reasoning should be the focus of scientific inquiry.
* As scientific and technological advances surged during the Industrial Revolution, scientists lost interest in idealist metaphysical debates.
* At the start of the 20th century, a new movement, logical positivism, emerged, advocating a return to strict empiricism and verifiable scientific statements.
* **Modern Science –**
* After WWI, the Vienna Circle (Schlick, Neurath, Carnap) rejected German Idealism and metaphysical questions, proposing Logical Positivism, which defined science as the study of meaningful, verifiable statements.
* Statements must be verifiable to be meaningful, either as **analytic (a priori, tautological, e.g., "all squares have four sides")** or **synthetic (a posteriori, dependent on observation, e.g., "all cats are born with tails").**
* Statements referring to unobservable entities were initially dismissed as meaningless, which led to problems in explaining concepts like electrons, gravity, and mental states.
* Logical positivists realized that scientific statements could never be fully verified, only confirmed, as future observations might contradict them.
* To address issues with verification, Logical Positivism evolved into Logical Empiricism, which accepted indirect references to unobservable entities if they played a role in scientific explanations.
* Karl Popper criticized verificationism, arguing that science should focus on **falsifiability** instead. A statement is meaningful if it can be disproven, not just confirmed.
* Quine argued that hypotheses cannot be tested in isolation since scientific theories rely on background assumptions, meaning falsification is not always straightforward.
* Thomas Kuhn introduced the idea that science operates within paradigms, which only change when accumulating contradictory evidence leads to a **scientific revolution.**
* Due to critiques from Popper, Quine, and Kuhn, Logical Positivism collapsed, giving way to more flexible approaches in the philosophy of science.
* Today’s science follows the **hypothetico-deductive method**, combining induction, deduction, and falsifiability, with **constructive empiricism** (Van Fraassen) suggesting that theories should aim for empirical adequacy rather than absolute truth.
* ***Philosophy of Science***
* **Epistemology –**
* Before accepting the hypothetico-deductive method, one must consider **ontology** (the nature of reality) and **epistemology** (how knowledge is acquired).
* Two major epistemological views exist: **rationalism**, which emphasizes reason and innate ideas, and empiricism, which prioritizes sensory experience as the basis of knowledge.
* **Rationalism (Plato & Descartes) –** Rationalists believe that at least some concepts are **innate** and that logical reasoning alone can reveal truths about the world.
* **Empiricism (Aristotle) –** Aristotle, the first empiricist, argued that **observation and sensory experience** are the foundation of knowledge, rejecting innate ideas and proposing the "tabula rasa" (blank slate).
* **Moderate Empiricism (Galileo)** – Galileo valued both **observation and logical reasoning,** stating that “the book of nature is written in the language of mathematics,” and he accepted thought experiments and unobservable entities.
* **Strict Empiricism (Hume & Logical Positivists)** – Later empiricists insisted that only **sensory experience** leads to true knowledge, dismissing statements about **unobservable** entities as meaningless.
* Bacon promoted empirical methods and inductive reasoning, laying the groundwork for the **scientific method**.
* **Constructive Empiricism (van Fraassen) –** This modern version of empiricism allows for **theoretical terms** while maintaining that only **observable** phenomena can be considered true.
* Constructive empiricism acknowledges that knowledge is **tentative** and subject to revision if new evidence contradicts current understanding.
* The balance between rationalism and empiricism continues to shape **scientific methods**, influencing how hypotheses are formulated, tested, and validated.
* **Ontology –**
* Ontology explores the nature of reality, focusing on two key questions: (1) Does reality exist independently of human thought? (2) What is the ontological status of particulars (specific instances) and universals (general properties like love or gravity)?
* Reality exists **only in the mind**; everything we perceive is a **mental construct**. Both particulars (e.g., a cat sleeping) and universals (e.g., love, gravity) exist, but only as mental representations.
* Reality exists **independent of human thought**, but **everything is made of matter**. Consciousness, thoughts, and emotions are just **by-products of physical interactions** in the brain.
* **Realism –** External reality **exists independently** of human thought, and **universals (e.g., love, gravity) are real**. Different forms of realism define how universals exist.
* **Platonic Realism –** Universals exist **independent of human perception** on a separate, abstract plane, as argued by Plato.
* **Scientific Realism** – Universals like gravity and love are considered r**eal** because they are useful and necessary for making **scientific claims** about observable phenomena.
* **Nominalism** – Accepts that reality exists **independent of human thought**, but **denies the existence of universals.** Terms like love and gravity are just convenient **labels** for patterns in the world, not real entities.
* Particulars are **specific instances** of observed phenomena, while universals are **general properties** that may or may not exist independently of those instances.
* These are **opposite views**: Idealism holds that reality is **mental**, while Materialism claims reality is **purely physical**.
* The debate between these views influences **scientific inquiry, metaphysics, and epistemology**, affecting how we define knowledge, perception, and reality.
* **Approaches –**
* Initially focused on **natural sciences**, the scientific method was later applied to **social sciences**, raising questions about whether a **realist-positivist** view should also apply to psychological and social phenomena.
* **Objectivism** – Maintains that social and psychological phenomena (e.g., intelligence, social cohesion) **exist externally** and independently of human perception, similar to natural sciences.
* **Constructivism** – Argues that social reality is **constructed** by social actors, meaning concepts like happiness or femininity **depend on cultural, historical, and social contexts** rather than existing as fixed external entities.
* **Interpretivism** – The epistemological counterpart to constructivism, which emphasizes **understanding social reality from the perspective of the people involved** rather than relying on objective measurement.
* **Hermeneutics** – Focuses on **interpreting people’s behaviour within their social context**, recognizing that actions can only be understood in relation to the cultural and historical environment.
* **Phenomenology** – Studies how people **experience** the world and the **meanings they attach** to their experiences, requiring researchers to **eliminate preconceived notions** to grasp subjects' perspectives.
* **Verstehen (Weber) –** Refers to **empathetic understanding** of social phenomena, requiring researchers to **immerse themselves in the subject’s culture** to understand their worldview and actions.
* (1) **Layered interpretation** increases the risk of **misinterpretation**; (2) **Lack of comparability** makes universal theories difficult; (3) **Differences in frame of reference** can prevent researchers from fully understanding their subjects' perspectives.
* **Qualitative research** (interviews, participant observation) aligns with **constructivism**, while **quantitative research** (statistical analysis, large datasets) aligns with **objectivism.**
* **Mixed Methods Approach** – Recognizing the strengths of both **qualitative and quantitative** methods, researchers increasingly use a **combined approach** to provide a more comprehensive understanding of social phenomena.
* **Goals –**
* Science aims to gain knowledge, but research can be categorized as either **universalistic** (general explanations) or **particularistic** (specific cases).
* **Universalistic Research** – Seeks to explain **broad, generalizable patterns** that apply to all people, groups, or societies. Example: studying the general effect of **violent video games on aggression**, regardless of game type or player demographics.
* **Particularistic Research** – Focuses on **specific contexts, groups, or time periods**. Example: investigating the impact of **raising the drinking age in the Netherlands** on teenage alcohol poisoning rates.
* **Fundamental research** seeks knowledge **for its own sake**, while **applied research** aims to **solve real-world problems**.
* If loneliness is linked to depression, applied research might introduce **a cat therapy program** to help lonely, depressed individuals feel better.
* A large-scale study on the **correlation between loneliness and depression** aims to understand the relationship but does not develop a treatment.
* While fundamental research is typically **universalistic**, it can also be **particularistic** when studying specific populations (e.g., violent video games and aggression in **particularistic** privileged young offenders).
* Applied research is often, but it can also have **universalistic elements**. Example: studying whether **human interaction or mere presence** of a pet reduces loneliness.
* Applied research can generate **new fundamental insights**, and fundamental research can **inform better applications.**
* **universalistic and particularistic**, as well as **fundamental and applied**, research approaches are valuable and can complement each other to improve knowledge and real-world solutions.

1. **Module 3: The Scientific Method**

* ***The Method***
* **Empirical Cycle –**
* A structured process in scientific research that follows a **hypothetico-deductive approach**, involving five key phases: **Observation, Induction, Deduction, Testing, and Evaluation.**
* **Observation Phase** – The starting point where a researcher notices a **pattern, event, or phenomenon** that sparks curiosity. Observations can come from personal experiences, others’ accounts, or prior research.
* **Induction Phase** – Specific observations are generalized into a **hypothesis or rule**. Example: If several people have bad experiences with their **mother-in-law**, one might hypothesize that **all mothers-in-law are horrible**.
* **Deduction Phase** – From the general hypothesis, **specific predictions** are made. For instance, if the hypothesis is true, then most people surveyed should rate their mother-in-law as horrible.
* The deduction phase also involves designing the study, defining concepts, selecting measurement tools, determining sampling methods, and planning data collection procedures.
* **Testing Phase** – New **empirical data** is collected and compared against predictions. This usually involves **statistical analysis** to determine whether the hypothesis holds up.
* If **8 out of 10** people rate their mother-in-law as horrible but **2 rate her as neutral**, the hypothesis is **partially refuted**, as the prediction was not fully met.
* **Evaluation Phase** – The results are interpreted to determine whether the hypothesis is **supported, adjusted, or rejected.** Hypotheses are rarely discarded outright but are often refined.
* **Iterative Nature** – The cycle **repeats**, as new observations from one study feed into future research, refining theories and improving understanding over time.
* **Scientific Progress** – By continuously refining hypotheses through this cycle, empirical science **gradually builds more accurate knowledge** about how the world works.
* **(Dis)confirmation –**
* Scientific hypotheses are tested using **empirical data.** In this case, the hypothesis states that **all mothers-in-law are horrible**, based on personal observations.
* A research setup is designed, where t**en colleagues rate their mother-in-law** as **likeable, neutral, or horrible**. If the hypothesis is true, all should choose “horrible.”
* If all ten colleagues rate their mother-in-law as horrible, the hypothesis is **provisionally supported**, but not **proven**, as future studies may contradict it.
* No empirical hypothesis can ever be **definitively proven**; the best we can do is **accumulate strong support** from multiple studies.
* If only **8 out of 10 colleagues** rate their mother-in-law as horrible, the prediction is false. Logically, the hypothesis **should be rejected**, but this is rarely the case in social sciences.
* Disconfirmation often leads to questioning **research design** rather than rejecting the hypothesis outright. Errors in **measurement, sample size, or procedural mistakes** can be responsible.
* If two colleagues rated their mother-in-law as **neutral**, it could be due to a procedural issue (e.g., they were uncomfortable giving negative responses). In such cases, the study setup, not the hypothesis, is reconsidered.
* New findings may lead to **refining** rather than rejecting a hypothesis. For example, if **only daughters-in-law** rated their mother-in-law as horrible, the hypothesis could be adjusted to reflect this pattern.
* Social science rarely sees **radical shifts** in theories. Instead, hypotheses are **gradually refined** based on small modifications and better study designs.
* Science progresses by **repeating studies, refining hypotheses, and improving research methods**, rather than through **absolute proofs or sudden discoveries**.
* **Criteria –**
* A confirmation does not prove a hypothesis, and disconfirmation does not automatically reject it. Studies are judged based on **reliability and validity**.
* A study is **replicable** if independent researchers can repeat it, and **reliable** if repeating the study yields **consistent results**.
* A study is **valid** if its conclusion about the hypothesized relationship **accurately reflects reality**.
* **Construct Validity** – Ensures that the study **measures and manipulates** the properties **it intends to**, avoiding errors in measurement or unintended influences.
* **Threats to Construct Validity –** If a depression questionnaire **measures social exclusion instead of depression**, or if **caring for a cat increases self-worth instead of reducing loneliness**, the study lacks **construct validity**.
* **Internal Validity (Causality)** – A study has **high internal validity** if the observed effect is **truly due to the hypothesized cause**, not **other factors**.
* **Threats to Internal Validity – Alternative explanations** lower internal validity. For example, in the cat study, **seasonal mood changes** or **increased physical activity** from caring for the cat could explain lower depression.
* **External Validity (Generalizability)** – A study is **externally valid** if its results apply to **other groups, settings, and times** beyond the original study.
* **Threats to External Validity** – If a study is conducted only on **elderly people and cats**, it may not generalize to **teenagers, different cultures, or other animals.**
* **Construct validity** ensures proper measurement, **internal validity** ensures the correct cause is identified, and **external validity** ensures results apply broadly.
* **Causality –**
* **Four Criteria for Causation** – To establish causation, four conditions must be met: (1) **The cause and effect are connected**, (2) **The cause precedes the effect**, (3) **The cause and effect occur together consistently (covariation)**, and (4) **Alternative explanations can be ruled out.**
* Giving lonely, depressed people a cat to take care of could reduce loneliness and, in turn, lower depression. To establish causation, researchers must ensure the effect follows the cause and happens consistently across cases.
* The hardest part of proving causation is ruling out **other possible causes**. For example, in the cat study, reduced depression might result from i**ncreased physical activity,** not reduced loneliness.
* Alternative explanations weaken **internal validity**. A key aspect of research methodology is designing studies that minimize these threats.
* **Correlation ≠ Causation** – Just because two things **happen together** doesn’t mean one causes the other. For example, children who play violent video games might be aggressive **for other reasons**, like upbringing or personality, rather than the games themselves.
* ***Threats to Internal Validity***
* **Participants –**
* **Maturation Threat –** Natural changes over time (e.g., depression improving on its own) may be mistaken for treatment effects. Use a control group.
* **Selection Threat** – Pre-existing differences between groups (e.g., physical fitness) can influence results. Use random assignment.
* **Selection by Maturation Threat** – Groups may differ in how quickly they naturally change (e.g., open-minded people recovering faster). Use random assignment.
* **Control Groups Reduce Bias** – Including a control group helps separate natural changes from actual treatment effects.
* **Randomization Strengthens Validity** – Randomly assigning participants prevents systematic differences, improving study accuracy.
* **Instruments –**
* Low Construct Validity – Occurs when instruments measure the wrong construct or contain systematic bias. Use well-validated measurement and manipulation methods.
* Instrumentation Threat – Happens when measurement tools change during the study (e.g., switching from one questionnaire to another). Maintain consistent measurement instruments throughout the study.
* Testing (Sensitization) Threat – Taking a pre-test can influence participants’ behaviour, creating an alternative explanation for results. Use a control group to account for pre-test effects.
* Unequal Pre-test Sensitization – If only the experimental group recognizes the study’s purpose, they may alter responses to align with expected outcomes. Include experimental and control groups that do not take a pre-test.
* Overall Strategy for Prevention – Maintain valid and consistent instruments, avoid changing measurement tools, and use careful experimental designs that account for testing effects.
* **Artificiality –**
* Experimenter Expectancy Effect – Researchers' unconscious expectations can influence participants' responses, leading to biased results. Use an Experimenter Blind Design, where the researcher does not know the expected outcomes.
* Demand Characteristics – Participants alter their behaviour based on their assumptions about the study’s purpose, especially if they believe the treatment should help them. Keep participants unaware of the study's true purpose.
* Rosenthal's Rat Study Example – Demonstrated that experimenters’ beliefs can unconsciously influence outcomes, even with animals, by affecting their treatment of subjects.
* Double-Blind Design – Both the experimenter and participants are unaware of group assignments, reducing bias from both researcher expectations and participant behaviour changes.
* Cover Stories & Debriefing – Researchers use plausible alternative explanations for the study’s purpose to minimize demand characteristics, followed by debriefing participants after the study.
* **Research Setup –**
* Ambiguous TemporalPrecedence – It may be unclear whether the hypothesized cause actually comes before the effect. Do violent video games cause aggression, or do aggressive children choose violent games?
* Solution to Ambiguous Temporal Precedence – Introduce or manipulate the hypothesized cause to ensure it happens before the effect. Assign children to play violent games and observe whether non-aggressive children also become aggressive.
* History Effect (External Events) – Unexpected events during the study can influence results. A news event about a violent crime by a minority group could affect attitudes in a study on stereotype reduction.
* History Effect (Small-Scale Events) – Events that occur within the study itself can bias results. If an experimenter faints in the control group, it might make participants more negative, skewing results.
* Solution to History Effects – While large-scale events are unavoidable, testing participants separately can prevent small events from affecting an entire group.
* Mortality (Dropout Bias) – Participants dropping out of a study can create an alternative explanation for results. A depression drug causes severe flatulence, leading to 80% dropout, leaving only participants who tolerate it.
* Uneven Dropout Rates Between Groups – If one group has a much higher dropout rate than another, the groups are no longer comparable.
* Solution to Mortality Threats – Document dropout reasons and analyse their potential impact on the study’s conclusions.
* If only those who tolerate side effects stay in a study, a drug may falsely appear more effective than it actually is.
* Careful experimental design, participant tracking, and awareness of external events help minimize these validity threats.
* ***Relevant types of Variables***
* **Variables of Interest –**
* A construct is an abstract concept (e.g., loneliness, depression), while a variable is a measurable or manipulable expression of that construct (e.g., loneliness measured by the UCLA loneliness scale).
* Constructs need to be defined in specific, measurable ways. Loneliness can be measured with a questionnaire or manipulated by giving participants a pet cat.
* A variable must show variation; otherwise, it is a constant. If all participants score the same on a depression test, depression is a constant and cannot be analysed.
* Independent Variable (IV): The cause or predictor (e.g., loneliness, manipulated through pet ownership).
* Dependent Variable (DV): The effect or outcome (e.g., depression, measured through a questionnaire).
* When studying cause-and-effect relationships, researchers control the independent variable to observe its impact on the dependent variable, ensuring clear conclusions.
* **Variables of Disinterest –**
* Confounders (Lurking Variables) – A variable related to both the independent and dependent variable, potentially explaining the observed effect instead of the independent variable. In a study on loneliness and depression, physical activity (from caring for a pet) may reduce depression rather than reduced loneliness itself.
* Confounders can be controlled by keeping them constant across all participants and measuring them and turning them into control variables for analysis.
* Control Variables – Measured variables that could influence the results, allowing researchers to account for their effects and verify if the main relationship still holds. Measuring physical activity levels to check if depression reduction is due to loneliness or exercise.
* Background Variables – Variables like age, gender, ethnicity, and socioeconomic status that do not directly affect the study’s main relationship but help assess how generalizable the findings are to a larger population.

1. **Module 4: Research Designs**

* ***Ways to Compare***
* **True Experiments –**
* **Manipulation** – The researcher controls the **independent variable** to ensure the cause precedes the effect, eliminating ambiguity in causal direction.
* **Comparison (Control Group)** – A **comparison group** without the cause helps determine if the effect happens naturally, reducing threats like maturation.
* **Random Assignment** – Participants are randomly assigned to conditions (e.g., experimental vs. control), ensuring **no systematic differences** between groups other than the independent variable.
* **Eliminating Alternative Explanations** – Randomization helps balance factors like age, gender, and pre-existing traits, reducing confounds that could explain the results.
* **Replication** – While not a formal requirement of true experiments, **repeating studies** helps confirm that randomization and results are reliable over time.
* **Factorial Designs –**
* **Multiple Independent Variables (Factors)** – Factorial designs investigate **two or more independent variables simultaneously**, allowing for a more complex understanding of their effects.
* **Main effects** examine the impact of each independent variable separately, while **interaction effects** explore how factors combine to influence the dependent variable.
* A study on migraine treatments could analyse **dosage (low, medium, high)** and **gender (men vs. women)** to see how these factors individually and jointly affect migraine attacks.
* **Higher-Order Interactions** – Additional factors (e.g., **diet**) increase complexity, allowing for **two-way** (e.g., dosage × gender) and **three-way** interactions (e.g., dosage × gender × diet).
* **Advantages of Factorial Designs** – They provide deeper insights by **examining multiple variables at once,** helping researchers **understand not just isolated effects but how variables interact** in real-world scenarios.
* **Repeated Measures –**
* In **between-subjects designs**, different participants experience different levels of the independent variable. In **within-subjects designs**, the same participants experience all levels.
* A **between-subjects design** would assign participants to **only one** dosage level, while a **within-subjects design** would have each participant try **all** dosage levels over time.
* Factorial designs can include **both between and within factors** (e.g., **dosage as a within factor, gender as a between factor**) to compare effects across groups.
* **Repeated measures designs** involve **short-term repeated testing** with manipulated independent variables, while **longitudinal designs** track variables over **months or years**, usually in **correlational studies**.
* **Fewer participants** are needed for statistical power, as each participant acts as their own control, making these designs **more efficient and practical**.
* ***It’s all about Control***
* **Manipulation –**
* In a true experiment, researchers control the independent variable by assigning participants to different levels (e.g., playing a violent video game for 0, 2, or 4 hours) to examine its effect on the dependent variable.
* The control group experiences no manipulation of the independent variable, while the experimental group undergoes a manipulated condition. Not all variables (e.g., age, gender) can be manipulated, as they are individual differences variables.
* Researchers verify whether the intended manipulation was successful by measuring the independent variable's effect, but this check should occur after measuring the dependent variable to avoid bias.
* External variables should be kept constant across conditions (e.g., room colour) to ensure that differences in the dependent variable are caused only by the independent variable, following the **ceteris paribus** principle.
* Unlike physics, social scientists struggle to control all variables, especially individual differences like age or education. Randomization and matching help manage these uncontrolled variables.
* **Lab vs. Field –**
* Experimental studies require control over the independent variable and extraneous variables, which is why many are conducted in lab settings that maximize **internal validity** by providing a controlled environment.
* Lab studies may lack **ecological validity** (real-world resemblance), but they can still maintain **experimental realism**, meaning they effectively capture psychological effects even in artificial conditions.
* A lab experiment on low self-confidence and salary negotiations uses artificial settings (e.g., computerized decisions), but it can still demonstrate causal relationships despite lacking real-world complexity.
* Field studies observe behavior in natural settings and may involve experiments, but they lack the control of lab studies, making them prone to uncontrolled variables.
* Lab studies prioritize **internal validity** (control over variables), while field research enhances **external validity** (real-world applicability). Both approaches are valuable and can complement each other in research.
* **Randomization –**
* Random assignment ensures that there are no systematic differences between experimental and control groups, eliminating alternative explanations for differences in outcomes.
* While randomization typically balances characteristics across groups, small sample sizes may lead to unequal group distributions due to chance, requiring randomization checks.
* Blocking ensures equal group sizes, while stratified randomization balances specific characteristics (e.g., gender) across groups, though it becomes complex with multiple variables.
* It’s impractical to stratify for all possible subject characteristics, and controlling too many variables becomes difficult.
* In within-subjects designs, conditions should be counterbalanced to prevent order effects (e.g., learning or fatigue influencing results).
* **Experimental Designs –**
* Participants are randomly assigned to either an experimental or control condition, with the dependent variable measured after exposure to the independent variable.
* Adds a pretest before exposure to the independent variable to assess baseline differences and measure changes over time, useful when maturation is a threat to internal validity.
* Combines the two-group design and pretest - post-test design by running both with and without a pretest, allowing researchers to control for pretest sensitization effects.
* All participants experience all conditions of the independent variable, with random assignment to order of exposure to control for order effects.
* ***What if we have no Control?***
* **Matching –**
* When Random Assignment is Impossible due to pragmatic, ethical, or inherent reasons (e.g., studying the effect of sex on political conservatism). Leads to a selection threat to internal validity, making systematic differences between groups harder to rule out.
* Matching as a Solutionwhere participants are matched on relevant background variables (e.g., age, education) to reduce selection bias. Helps exclude alternative explanations for differences between groups.
* Risk of Undermatching occurs when matching is done on a variable measured with error and related to both the independent and dependent variables. Leads to regression to the mean, where extreme values on the pretest tend to move closer to the mean on a post-test.
* Disadvantaged toddlers with higher-than-average cognitive scores are matched with advantaged disadvantaged toddlers with lower-than-average scores.
* On a second test, disadvantaged toddlers tend to score lower (closer to their group mean), while advantaged toddlers tend to score higher (closer to their group mean). This creates an illusion that Sesame Street is harmful to disadvantaged children, even if it benefits both groups equally.
* When Undermatching Doesn't Occur, this issue does not apply to variables that can be measured precisely, like age, sex, or educational level.
* **Quasi-Experimental Designs –**
* Lack of random assignment is the key feature. Can also lack manipulation and comparison groups. Still aims to demonstrate causal relationships, but with more threats to internal validity.
* Ethical or practical constraints prevent random assignment (e.g., parental choice in violent game study). Individual differences variables (e.g., sex, trauma) can't be manipulated. Sometimes called natural experiments when variables occur naturally.
* Static Group Comparison Design: Post-test only, with no random assignment; Selection threat is a major concern.
* Pretest-Post-test Non-Equivalent Control Group Design: Adds a pretest to assess pre-existing differences; Helps clarify causal direction and estimate maturation effects.
* One-Group Pretest-Post-test Design: Establishes temporal precedence, but threats to validity remain.
* Interrupted Time-Series Design: Uses multiple measurements before and after treatment to track natural changes; A sudden change right after treatment suggests an effect.
* Replicated Interrupted Time-Series Design: Adds a second group with the same measurements but no treatment; If both groups show the same pattern, a history threat may be responsible; A second group receiving the treatment at a different time helps confirm causality.
* Without proper controls, we might falsely conclude that violent games cause aggression.
* Multiple measurements and comparison groups help rule out history effects.
* **Correlational Designs –**
* Correlational designs do not manipulate the independent variable. Some studies don’t even identify an independent variable, as they don’t test causal relationships.
* Measures a large group at one point in time to study associations between variables. Often uses surveys, but survey method not equal to research design.
* Measures one individual multiple times at fixed intervals to observe patterns over time. Often confused with longitudinal studies, which involve groups instead of individuals.
* Panel design: Follows the same group over multiple time points. Cohort design: A type of panel study where participants share a common starting point (e.g., same birth year or job entry).
* Longitudinal: Any study (correlational, experimental, quasi-experimental) tracking participants over time. Repeated Measures: Experimental design where the same participants experience different conditions. Time-Series in Quasi-Experiments: Measures groups repeatedly before and after an intervention.
* **Other Designs –**
* Case Studies focuses on one person or group (quantitative or qualitative). Single-subject, time-series, or N=1 research fall under this category. Negative Case Analysis: Searches for contradictory evidence to test hypotheses, useful for all-or-none claims.
* Evaluation Research investigates the effectiveness of a policy or program. Summative: Assesses outcomes (was it effective?). Formative: Assesses processes (how does it work?). Usually non-experimental due to ethical/practical constraints.
* Experimental studies are typically experimental and focused on individualtreatments. Common in clinical psychology (e.g., cognitive behavioural therapy for depression). Differs from evaluation research, which targets broader societal or educational programs.
* Validation Research assesses the quality of measurement instruments (e.g., surveys, questionnaires, tests). Uses statistical analysis to check internal consistency. Ensures that related constructs show expected patterns of association. Important in psychometrics and sociometric.

1. **Module 5: Measurement**

* ***What is it?***
* **Operationalization –**
* Variable represents a construct (e.g., political power) in a more abstract form. Operationalization is the specific, concrete procedure used to measure or manipulate a construct (e.g., using "influence in parliament" as a variable).
* Political power could be operationalized in different ways, such as counting bills passed or asking analysts to rate influence. Love of animals: Could be operationalized by observing interactions with pets or using a self-report questionnaire.
* Exposure to animals can be operationalized in various ways, such as through direct care, visiting a zoo, or watching animal documentaries.
* Operationalizations may not capture the entire construct. Self-reports may show more positive attitudes, while observed behaviors may reveal less affection toward animals.
* The operationalization might only measure a part of the construct, so it's important to define which aspect of the construct is being assessed to avoid misleading conclusions.
* **Measurement Structure –**
* Measurement represents relationships between objects or groups using numerical values to capture certain properties (e.g., body length).
* Inequality: Different objects or persons have different values. Order: Numbers represent the rank or order of objects (e.g., tallest to shortest). Equal Differences: Differences between objects are the same and reflected in equal numerical differences.
* Ratios compare the scale or magnitude between objects, such as one being twice as tall as another, with corresponding numbers reflecting that ratio.
* Using cardboard to visually measure body length helps demonstrate how measurements reflect inequality, order, differences, and ratios. This process is laborious but becomes easier with tools like a tape measure.
* Assigning numerical values to properties allows researchers to aggregate data and use statistical tools for analysis and drawing conclusions.
* **Measurement Levels –**
* Nominal Level only allows for the differentiation of values with no meaningful order (e.g., nationality, sex, pet preference).
* Ordinal Level allows for differentiation and ordering of values (e.g., test scores). Differences and ratios cannot be meaningfully interpreted.
* Interval Level allows differentiation, ordering, and interpretation of differences between values (e.g., temperature). The zero point is arbitrary, so ratios don't hold meaningful meaning.
* Ratio Level allows differentiation, ordering, interpretation of differences, and ratios. The zero point is non-arbitrary and meaningful (e.g., length, age).
* By categorizing a ratio-level variable (e.g., age), the exact measure is lost (e.g., age categories instead of exact years), leading to a loss of ratio and interval information.
* **Variable Types –**
* Nominal is unordered categories (e.g., animal preference, nationality).
* Ordinal is ordered categories with no meaningful differences or ratios (e.g., test scores indicating rank).
* Binary/Dichotomous is exactly two categories (e.g., male/female, smoker/non-smoker).
* Polytomous is categorical variables with more than two categories. For categorical variables, we focus on frequencies, not means (e.g., animal preference counts).
* Quantitative Variables: Interval and Ratio allow meaningful differentiation, ordering, and comparison of differences and ratios (e.g., weight, temperature); Mean Comparison is for quantitative variables, calculating the mean and comparing groups makes sense.
* Continuous can take any value within a range (e.g., weight, height).
* Discrete is limited set of values (e.g., number of pets, can only be whole numbers).
* Relevance is the discrete/continuous distinction matters less for interpretation than categorical/quantitative classification.
* ***How do we assess its Quality?***
* **Measurement Validity –**
* Construct Validityrefers to whether an instrument accurately measures the intended construct (e.g., fondness for cats).
* Face Validity is expert opinion on whether the instrument seems to measure the intended construct. Predictive (Criterion) Validitydemonstrates that the instrument can predict a related outcome, but doesn't guarantee it's measuring the right construct. Convergent Validity ensures the instrument correlates with other measures of the same construct (e.g., a questionnaire on cat fondness correlates with observations of behaviour around cats). Discriminant Validity ensures the instrument doesn't correlate with unrelated constructs (e.g., no correlation between cat fondness and pizza fondness).
* Multi-Trait Multi-Method Matrix is a systematic approach to test convergent and discriminant validity by using multiple methods to measure different traits and examining their correlations.
* Validating social and psychological constructs is difficult because there's no direct, undisputed measure like a tape measure for physical properties. Thus, indirect methods like convergent and discriminant validity are essential.
* **Measurement Reliability –**
* Reliability refers to the consistency or stability of an instrument. A reliable instrument should produce similar results when measuring the same property repeatedly.
* Test-Retest Reliability is administering the instrument twice and checking consistency between results. It's useful when the property being measured is stable over time. Internal Consistency is checking consistency between different parts of the instrument at one time. It’s used for instruments with multiple questions that measure the same construct (e.g., split-halves reliability). Intra-Observer Consistency is when one observer rates the same behaviour multiple times, and the results are compared. Inter-Observer Consistency is comparing ratings from different observers to ensure consistency.
* Systematic Error is if an instrument measures an additional construct (e.g., positive attitude alongside cat fondness), it can result in consistent but misleading measurements, making the instrument less valid but still reliable. Random Error is random fluctuations lead to different measurements each time, reducing reliability.
* An instrument must be reliable before it can be valid. Low Reliability & Low Validity is high random and systematic error, making it hard to distinguish true scores. Low Reliability & High Validityis high random error but little systematic error; the right property is measured, but imprecisely. High Reliability & Low Validity is high precision but measures the wrong property. High Reliability & High Validity is the best case, where the instrument measures the right property with high precision.
* Statistical Modelling ispsychometricians and sociometricians use statistical methods to separate random and systematic errors from the true score.
* ***How do we measure?***
* **Survey, Questionnaire, Test –**
* Survey is a general term for a set of questions covering various topics like biographical information, opinions, behaviours, etc. Questionnaire is focused on a single construct or related constructs, typically used for measuring psychological traits, emotional states, or attitudes. Test is designed to measure an ability, such as intelligence or math proficiency.
* Clear instructions are essential for proper understanding (e.g., how to round numerical answers in a math test). A cover story can prevent participants from guessing the study's purpose and potentially distorting their responses.
* Interviewer-administered Surveys is the direct interaction between the interviewer and respondents (more expensive). Self-Report Surveys is for respondents to complete the survey independently, typically using paper or online platforms. Online Surveys is convenient for respondents, offers additional controls for researchers (e.g., question order, required responses, identifying unusual patterns). However, a downside is a low response rate.
* Items are the individual questions, statements, or words participants respond to. These can be accompanied by discrete response options or continuous scales. Stems are the core part of the item (question, statement, or word) that respondents react to.
* Using multiple items to measure a construct helps to reduce random errors and assess internal consistency. More items generally increase reliability.
* Items is a set of items related to fondness for cats (e.g., petting a cat, watching cat videos, etc.). Scale and Sum score items are scored (e.g., "agree," "neutral," "disagree"), and the scores are summed to form a scale indicating the person's level of fondness for cats. Recoding some items might need to be reversed (e.g., negatively worded items) to align with the scale.
* Subscales - questionnaires can have different subscales measuring various dimensions of a construct (e.g., academic motivation could include intrinsic motivation, extrinsic motivation, and fear of failure). Statistical Methods exist to determine whether different dimensions of a construct are distinguishable based on the responses. This summary outlines how surveys, questionnaires, and tests are used to measure psychological constructs and the importance of clear instructions, validity, and reliability in their design.
* **Scales and Response Options –**
* Likert Scale is a common scale that measures a single property using comparable statements. Respondents indicate their level of agreement with each statement. Items should be monotone, meaning that respondents who have a greater degree of the property should consistently agree more strongly with the items. After recoding, items should be monotone. A higher score indicates a higher level of the property being measured (e.g., higher cat fondness).
* Non-monotone items, like "cats are a necessary evil to get rid of mice," are problematic because both cat lovers and extreme cat haters might disagree, making it hard to interpret the scores. Other scales like differential and cumulative scales exist but are used less frequently. Likert items usually have 3 to 7 response options, but a graphic rating scale can also be used, where respondents mark a line between two extremes.
* The process questionnaire construction starts with clearly defining the property to be measured. Different dimensions of a construct may be identified, each measured with specific items. Identifying dimensions requires in-depth analysis. For example, academic motivation can be intrinsic (love of the subject) or extrinsic (desire for a better job). Items should describe specific situations or expressions of the property. Vague or general items lead to misinterpretation.
* Items should be short, simple, and unambiguous. Avoid double negation, unfamiliar words, or overly complicated formulations. Avoid ambiguous formulations, like asking if cats are funny without clarifying whether it means "make you laugh" or "are strange." Double-barrelled questions should be avoided. For example, the question "I love to pet and pick up cats" is confusing for people who like one but not the other.
* Items should be neutral, not suggestive, to avoid influencing respondents' answers. Ensure that all respondents can answer the items. For example, include a filter question if a person has never interacted with a cat. Extreme wording (e.g., "never" or "always") should be avoided as it makes it difficult for respondents to provide moderate answers. Response options need to be clear, exhaustive, and mutually exclusive. For example, age categories should be appropriately defined so that all respondents can choose an option without overlap.
* There are many other factors in constructing scales, items, and questionnaires, including item order, handling sensitive questions, and combining self-report with others' ratings. This is just a surface-level overview of the process.
* **Response and Rater Bias –**
* Acquiescence is the tendency to agree with all statements regardless of content. It can be detected by including negatively worded items. For example, in a cat fondness questionnaire, someone agreeing with both positively and negatively phrased items shows inconsistency.
* Social Desirability is a bias where respondents answer questions in a socially acceptable way, presenting themselves favourably. This can be detected by adding items like "I've never stolen anything" or "I've never lied." Strong agreement with such statements suggests bias in the responses.
* Extreme Response Style is where respondents choose the most extreme answer without carefully considering the strength of their agreement. Unlike acquiescence, responses are consistently extreme, but not necessarily accurate.
* Bias Toward the Middle is the tendency to avoid extreme responses and consistently choose the middle option. This can be detected by including very extreme items (e.g., "cats are purely evil creatures"). If respondents select the middle option for all items, their responses are inconsistent.
* Halo Effect is a positive or negative evaluation of one characteristic influences ratings of other characteristics. For instance, attractive people might be rated as more intelligent or better at their job.
* Generosity Error (Leniency Effect) where raters are overly generous or lenient in their evaluations, leading to consistently high ratings.
* Severity Error is the opposite of generosity, where raters are overly strict or negative in their evaluations, leading to consistently low ratings.
* These biases are hard to detect but can be addressed by using multiple raters with clearly defined coding schemes and checking inter-rater reliability. However, bias shared by all raters is particularly difficult to identify.
* **Other Measurement Types –**
* Physical Measures is common in biology, medicine, and psychology. Examples include electrical skin conductance (for arousal), eye tracking (for attention), EEG and FMRI (for brain activity), and reaction times (for cognitive ability).
* Observational Measurement is used in sociology, psychology, and educational sciences. Systematic observation involves registering specific behaviours through coding schemes that define categories, behaviour criteria, timing, and circumstances. It also requires training and calibration of observers to ensure high inter-rater reliability.
* Trace Measurement involves assessing behaviour indirectly through physical evidence, such as counting used tissues after a therapy session to measure depression.
* Archival Data uses pre-existing data collected by others, such as census data. Example: Investigating whether higher income correlates with more votes for conservative parties.
* Content Analysis is a method that combines features of observational, archival, and trace measurement. It involves structured coding of texts like newspaper articles or blogs. This method can be used to analyse political discourse, such as identifying emotional and rational language in political speeches.
* Structured Interviews is where questions, question order, and response options are predetermined, similar to surveys. They have higher response rates but can be challenging for sensitive topics.
* Unstructured (Open) Interviews is a qualitative method where the conversation can flow freely based on the respondent’s answers. This makes it harder to compare and aggregate data but provides deeper insights.
* Other qualitative methods include case studies, focus groups, and participatory observation, though these are not discussed in detail here.

1. **Module 6: Sampling**

* ***What is it?***
* **External Validity Threats –**
* **Sampling Concepts –**
* ***What types are there?***
* **Probability Sampling –**
* **Probability Sampling (Simple) –**
* **Probability Sampling (Complex) –**
* **Non-Probability Sampling –**
* ***Random and Systematic Error***
* **Sampling Error –**
* **Non-Sampling Error –**
* **Sample Size –**

1. **Module 7: Practice, Ethics & Integrity**

* ***It starts with record-keeping***
* **Documentation –**
* **Data Management –**
* ***Ethics towards Participants and Society***
* **Unethical Studies –**
* **Ethics Towards Participants –**
* ***Sloppy vs. Solid Science***
* **Research Integrity –**
* **Questionable Research Practices –**
* **Peer Review Process –**
* **Dissemination Problems –**

1. **Module 8: Catch Up**
2. **Module 9: Exam**