Industrial Physcology - Research

1. Approaches to knowledge acquisition
   - → custom / superstition : lucky charms etc.
   - → intuitively → men think women are bad drivers
   - → respected source. eg. prediction by company director → int. rates will drop - people believe it
   - → reasoning - eg. trying a diet that worked in the past → reason it will work again
   - → experience → experience something nice → so it is nice eg places to live, see, eat etc.

Scientific → logic of investigation
- ↳ identify problem + formulate a hypothesis
- ↳ design the study or element
- ↳ carry out the study / experiment
- ↳ test hypothesis
- ↳ communicate results (journal / report)

Characters of scientific approach
- * maintain control = exclusion of foreign variables that might influence results
- * operationalise definitions
- * repeat investigations

exclude parts that could influence your study

define all the terms which can be used to measure them (eg good means specific qualifications

Test the experiment.

Repeat

2 types of research

↳ descriptive

↳ experimental

Descriptive research

* Secondary records - eg marriage records/census data
* Naturalistic observation → unobtrusive collection of data on a spontaneous behaviour eg hidden camera
* Case study - interviews/test results/ → less data orientated and more clinically orientated
* N=1 study - 1 person eg effects of therapy behavioural study
* Self reporting → introspection → one's own thinking feelings/perspectives
* Correlation study → degree of relationship between 2 measured variables - eg the correlation or relationship between motivation / productivity
* Ex post facto study - compare the effect of 2 or more variables - but researcher cannot control the variables → eg gender/age
* Longitudinal + cross-sectional studies.

→ Longitudinal - field studies that measure the same attributes of the same group of people at regular intervals

↳ Cross-sectionals - field study that measure the same attributes of individuals at the different age levels

* Surveys + Interviews → eg Questionaire from a representative Sample of the population

* Realiability - consistancy with what is being measured
  Validity - measure what is supposed to measure

* Participant observation - researcher actvly participates in the study while observing + recording the behaviour of others

* Meta - analysis → quantative technigire which is used to integrate + describe the results of a large number of studies of the same participant.

Experimental Research

→ seeks to determine casual (ie. cause + effect) relations between variables. It enables us to observe the result (effect) of systematic modification (manipulation) of one or more variables (cause) under controlled conditions.

→ Psychological experiment

→ objective observation of phenomena (behaviour/ actions/ responses to questionerce) etc. which occur in strictly controlled situations where one or more factors are varied (manipulated) and the others are kept constant

- advantages of experimental approach
  * One can infer casual relations
  * Researchers can manipulate one or more variable of their choice
  * Usefulness: experiments produced results that have stood the test of time, and led to other studies that offered solutions

- disadvantages of experimental approach
  * Lab findings are obtained in an artificial environment (no real life situations)
  * Time consuming
  * not adequate for the study of human behaviour
    └ yet sometimes it can be the only way to explain human behavior.

Experimental research - 2 types
  * field experiments - real life eg in organizations
  * Lab experiments - labs - controlled environment
    eg. sleep deprivation tests

Quasi experiments: influence of external variables can not be controlled: eg: testing the effect of longterm memory of a brain injury - cannot control the brain (internal) or assign random people to the group.
can't control the cause of the injury.

Internal + external validity

internal - extent ~~to~~ to which the observed effect of an experiment is caused by the experimental treatment condition, that is the independent variable.

eg students getting better grades because of private tutoring - But it could be that they are smarter + not necessarily the private tutoring → experimental treatment condition

External → the extent to which the results of an experiment can be generalized beyond the experimental situation to all the members of the same target population.

eg training techniques (study, lectures, multimedia) → selects 3 groups + each get a technique. → takes a test → multimedia group does better in the test → generalize that multimedia training should be used its most effective? external validity

1st priority in experimental research is to identify the treatment effect Once effect is verified by means of properly controlled, internal valid studies, external validity can be investigated.

P.2: StudyGuide for test questions.

2 Steps for Research

* 1 → identify research problem.
    * define problem clearly
    * purpose of the study / research

* 2 → Scope of research
    - * must be manageable -
    * consider practical / financial limitations.
    * urgency of problem will help determine scope (lenght) of research.

* 3 → identify practical Limitations
    * time
    * cost
    * size of sample
    * data collection techniques
    * processing techniques
    * access to information
    * nature of available information
    * form in which information is available
    * validity + reliability of - information
        - measuring instruments

* 4 → Evaluate the available research tools
    * knowledge
    * skills
    * statistical techniques

- computer equipment
- interpretation skills
- report writing skills

*5 → Survey of literature.

- collect your data
- survey/review existing data already in the field

*6 Research planning + design

- What is the problem being researched and how will it be researched?
- Data - what info is needed ? How will it be obtained? How much info is available?
- Criterion evaluation → What data can be used to measure the usefulness or success of the results.

*7 → Writing a research proposal.

- done before research project can be done.
- will be the basis of projects approval or denial.
  * Problem and context
    - state problem
    - sub problems
    - hypotheses
    - limitations
    - definitions of terms
    - assumptions.
    - importance of study

* Survey of relevant data
* data + processing + interpretation
  - primary / secondary data
  - criteria used to determine what data are to be used
  - research methodology
  - identification + handeling of subproblems
* Qualifications of researcher to conduct the ~~project~~ Project
* Description of proposed study
* A selected bibliography

* 8 → Data Collection.
  - what data
  - Why this data
  - nature of data.
  - how data will be ~~de~~ collected
  - on how many people
  - who will collect the data?
  - what will be done to ensure data collected is reliable

* 9 Data Processing
  * processing techniques
  * additional info collected
  * How will data be collected + made available for processing.

*10 → Step 10 Writing a research repat (feedback + recommendations)

Research Repeat:

Front Matter

* Title page
* Table of contents
* list of tables
* List of figures

Text

* Definition of problem and context / background
* Survey of relevant literature
* Description of population
* General procedures
* Results
* Summary, conclusion + recommendations
* Bibliography
* Appendices

→ Identify a problem + define it in words
→ Determine the scope of Research
→ identify practical Limitations
→ Identify practical Limitations
→ Graduate the available research tools
→ Survey of literature.
→ Research + planning
→ Writing a research proposal

→ Data processing
→ Writing the research project.

3/ numbers, variables + measures

Quantitive methods

* Efficiency - lots of data available in tables ect - makes it easier to draw conclusions

* Approximation / modelling of real world phenomena (Pictures)
  ↳ allows us to infer important dimension of similarity judgements. - if not used can lead to ~~long~~ sorting through a long set of verbal discriptors and we might not arrive at the dimentions as clearly.

* powerful & centuries - old language. eg weather maps
- deterministic claim - must be sure A will lead to B.
- Probabilistic claim - A might not lead to B (Chance expectation)

Functions of quantification:

* Administrative. eg monetary
* evidentiary aids or systems.

Variables + constants

When measure height → height is the variable.

Variables → measured entities (or attributes of entities) that can take on different values eg: height / weight.

3/

Constants are quantities that do not change but always have the same value eg: speed of light.

Continuous measures → take any value within the range defined as valid for a particular variable.

eg measure of something like food. - can be 1, 5, 2.5, 2.35 etc.

Discrete variables → certain values → order people complete a marathon.

Also known as categorical variables.

can take on values within a range 1. 2. 3. 4 NOT 1,5 etc.

⇒ 4 types of measurement scales

- x Nominal.
- x ordinal
- x interval
- x ratio.

Nominal variables → difference between categories of objects persons etc. Numbers are used here as labels to distinguish between the categories. eg: religion, sexes, etc.

Ordinal. → indicate categories different from each other + ranked or ordered in terms of an attribute.

eg. developing countries and developed countries. (one has more economic structure)

Grades A, B, etc.

Surveys I agree, I disagree, Strongly agree etc

3/

Interval variables -> true quantitive measures eg. temperature.
IQ score
- Tests

Ratio. -> all the properties of interval variables, but they also have a true zero value eg: 40 year old is 2x the age of a 20 year old.
Can't say that 80% in a test means twice as much as someone who got 40%

Magnitude -> property of mareness. - eg height can say 1 person is taller than another.

Equal intervals -> interval scale = equal intervals -> diff. between all points are the same (Nominal + ordinal does not posess the property of equal intervals)

Absolute 0 = when 0 is measured - nothing is measured
eg measure the wind a zero will mean that no wind is blowing.

	Magnitude	Property equal intervals	Absolute 0
Nominal	No	No	No
Ordinal	Yes	No	No
Interval	Yes	Yes	No
Ratio	Yes	Yes	Yes

3/

Independent + dependent variables

Independent → presumed to affect or determine other variables. eg: smoking → lung cancer.
IV. Predictor

Dependent → variables affected or determined by independent variables.
DV. Response / criterion.

Samples, populations, statistics and parameters.

Population → entire collection of elements or individual
eg: average income per capita → need everyone income. → add + divide by number of people

→ Sample: as we are normally unable to collect data from all. we collect a sample. (estimate.)

Parameter → numerical values summarising the population data.

Statistic → numerical values summarising sample data.

Statistical interference → act of generalising from a sample to a population.

-Statistics - we are often interested in hypothetical populations

3/ Problems with the quantitive approach

→ easy rationality of probabilistic inference has become an institutionalized canon, and has usurped other evidentiary forms

→

4. Self math tests

5/ Displaying Data.

Graphical + tabular displays of data allow us to "see" the distribution of scores of on a variable

→ frequency of a score refers to the number of times that given score appears within a dataset

→ frequency distribution is a tabular or graphical representation of a dataset indicating the set of scores on a variable together with their frequency.

→ types of data
variables that can take on a few values → integers → discrete values ⇒ goals scored in a match.
students in a class

⇒ continuous variables ⇒ height / mass of people
= infinitive number of possibilities

⇒ % frequency → percentage of observations that fall in each outcome category

Frequency Bar Chart

→ Values or scores of a nominal data are used for identification
→ does not indicate the amount of an attribute.
→ Bar chart is a graphical representation of nominal data in which a vertical bar reflects the frequency of each category on a discrete (or categorical) variable

table class interval smallest → biggest ↓

→ aspect ratio of a bar chart is the ~~ratio~~ ratio of the height to the width of the bar.

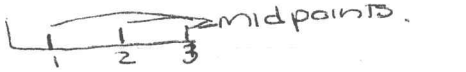
⇒ calculate % frequency → dividing the frequency for the category by the total number of subjects. (N)

$$f \div N = f\%$$

N = total amount of scores (f = total)

→ table do score lowest to high.

NB must have midpoints



"Have gaps between scores."

→ class interval is a division of category of scores on a grouped frequency distribution

→ grouped frequency distribution → tabular or graphical representation of ordinal, interval or ratio data. Scores are grouped into class intervals, for which frequencies are given

→ real upper limit + real lower limit refer to the true ~~new~~ boundaries of a class interval. They are found midway between

eg. class interval 10-19 → real upperlimit = 19.5 (½ between 19, 20)

20-29

→ real lower limit of 20-29 = 19.5

10-19 = 9.5

Cummulative percentage (%) frequency → % of items within a dataset that have value less than or equal to a specified score.

↳ add - frequency % with one below it.

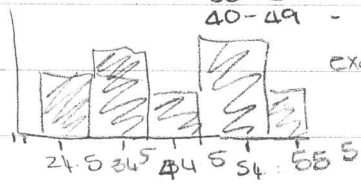
% frequency	Cummulative %
2.5	100
18.75	97.50
37.5	78.75
33.75	41.25
6.25	7.5
1.25	1.25

Cummulative frequency → from bottom up
add the frequency to the one below it.

The histogram.

→ midpoint of each class. if 1st class has nothing leave blank

class – 20-29 – midpoint = 24.5
30-39 " 34.5
40-49 - " 44.5



example midpoints

Midpoint of class interval
$= RLL + \left(\frac{RUL - RLL}{2}\right)$

→ no gaps between classes
→ indicate midpoint on horizontal AXIS.

Describing of frequency distributions

Shape of Distribution:

→ to describe the shape of a frequency comment on its
* symmetry, * modes
* skewness * kurtosis

Symetrical Unimodal Distribution

Symetrical Bimodal Distribution

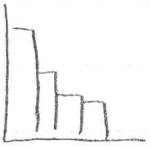
Symmetrical → often has centermost points in the middle of the distribution

Unimodal → peak in the middle

Bimodal → peak on the sides

negative peaked - right
positive peaked - left

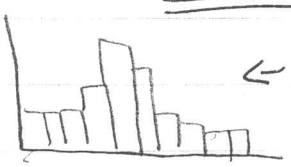
Kurtosis → peakness of a distribution
- ↳ flat (platykurtic)
- ↳ leptokurtic (peaked)
- ↳ mesokurtic (normal)



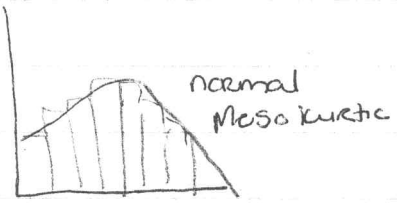
← positively skewed distribution

← negatively skewed distribution

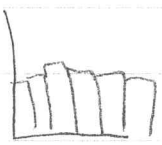
Kurtosis



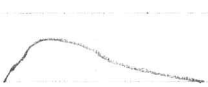
← Peaked distribution (leptokurtic)



normal Mesokurtic



← Flat distribution platykurtic



Percentile Rank – % of cases that lie at or below a specific point of the scale on which the data are measured

Percentile → point on the scale at + below which a specified percentage cases in a dataset falls.

$$\text{Percentile rank} = \%\text{ below} + \frac{\text{score} - \text{RLL}}{\text{class interval with.}} (\text{interval } \%)$$

% below = cummulative percentage of frequency of the class below in which the score falls

score = score in which we are trying to determine % rank

RLL = real lower limit of the interval in which the score of interest occurs

class interval width = width of class interval.

interval % = % of distribution that falls into the category of interest.

$$\text{Score at P} = \text{RLL} + \frac{\text{PR} - \%\text{ below}}{\text{interval } \%} (\text{interval width})$$

PR = Percentile Rank for which we wish to identify a score

Median (50^{th} %)

→ value that devides a distribution into 2 halves

Midpoint = average of upper + lower limits of the class interval

$$(UL + LL)/2$$

6/ Measures of central tendency

The Mean (arithmetic average of grap of numbers) $\bar{X}$

→ average of a grap of numbers.

$$\bar{X} = \frac{\Sigma x}{n}$$

$\bar{X}$ = mean

Σx = total of all scores

n = number of scores

Population mean = $\mu = \frac{\Sigma X}{n}$

Median

↳ middle score in a ranked distribution of scores

↳ not influenced by extreme scores

* Always use the formula first to compute the median location $\left(\frac{N+1}{2}\right)$
* Arrange data smallest → largest.
* Count positions up to the position then determine median

Mode

→ the most frequently ~~seen~~ occuring score in a distribution

7/

Measures of variability

Measures of variability indicate the degree to which the scores are dispersed or different from each other

3 commonly used measures of variation
- ↳ Range
- ↳ variance
- ↳ standard deviation.

notion of variation is central to many statistical concepts + procedures.

Range.

→ the distance between the bottom + the top of the dataset

* crude range ⇒ difference between the maximum + minimum scores
* extended range ⇒ difference + 1
* interquartile range → difference between 25th + 75th quartiles

∴ Range = highest score − lowest score.

w about
t don't need
to calculate
SS

Average deviation (Do NOT need to calculate, or define)

o only

→ average of the absolute distances of the individual scores from the mean of the distribution.

NB **Variance** (S^2) = sample variance

→ average of the squared distances of individual scores from the mean of distribution

→ smaller variance - closer the individual scores to the mean
→ greater the variance - more widely individual scores are dispersed around mean.

NB **Formula to be used to calculate standard variance**

$$S^2_x = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}$$

→ we use N-1 → gives a better estimate of population variance.

Variance → population
Do not use this formula to calculate sample variance

Variance of Population (info only)

$$\sigma^2 = \frac{\Sigma (x-\mu)^2}{N}$$

σ^2 = variance of population

Σ = sum of

x = each score

μ = Population mean $\mu = \frac{\Sigma x}{n}$

n = number of observations

Steps
* calculate difference $x - \mu$ each score
* calculate squared difference - do each score + add together.

Standard Deviation

- Square Root of the average of the squared distances of the individual scores from the mean of the distribution.

$$S_X = \sqrt{s^2 x}$$

Standard divation for population => info only

$$\sigma = \sqrt{\sigma^2}$$

All variances + standard deviations are possitive numbers. (if you get a negative number you made a mistake)

8 Correlation = r

Recap mean = $\bar{X}$

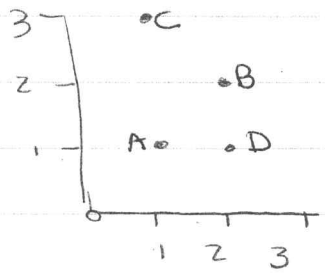
standard deviation = s

Scatter diagram is a graphic representation of correlation

correlation → relationship between X and Y.

correlation → r symbol.

example scatter diagram



* remember to name (identify) points.

Correlation coefficient. Pearson

$$r = \frac{N\Sigma XY - \Sigma X \Sigma Y}{\sqrt{[N\Sigma X^2 - (\Sigma X)^2][N\Sigma Y^2 - (\Sigma Y)^2]}}$$

Steps *Calculate ΣX = Sum of X - all scores added together

ΣY = sum of Y - all scores added together

ΣX^2 = calculate each X score^2 + then add

ΣY^2 = calculate each Y score^2 + add together

ΣXY = calculate X times Y for each score + then add together.

Correlation scores can only have a value between -1 and 1

If you get bigger than 1 or smaller than -1 you made a mistake.

Interpretation of the correlation coefficient

r = positive = 0 to 1

r = negative = -1 to 0

r = no relationship = 0

Book P 184 table

- nature of correlation (positive or negative) determined by direction of line.

examples positive correlation (both factors increase)

* more bread eat for lunch - heavier the person will be
* more years of teaing → higher salarey

examples negative correlation

* heaver the load of the truck → the less mileage kilometres the truck will drive increase in one weight load will cause decrease in mileage.
* more days staff is absent the lower the output of the factory

Strengh of the carelation

weak 0 → 0,39
moderate 0,40 → 0,79
strang 0,80 → 1,00

example :

- 0,95 = strang negative carelation
- 0,35 = weak possitive carelation
- 0,82 = strang possitive carelation
- 0,70 = moderate negative carelation
- 2,48 = IMPOSSIBLE must be between -1 and 1
- 0 = no carelation.

Deduction:

0,168 carelation → degree of job statisfaction and quality of working life

∴ modeatly possitive carelation and people who hae a high degree of job satisfaction will also hare a high quality of working life
(Both ↑ possitive → higher)

⇒ Stranght of the carelation is determined by the consentration of points → closer they are to a straight line, stranger the carelation

Camman Variance

* proportion of variance in $\bar{X}$ scores that is attributable to variance in $\bar{Y}$ scores → shared or camman variance between X and Y.

r^2 = proportion of common variance, calculated by squaring the correlation coefficient (r)

$r^2 \times 100$: persentage of common variance r^2

Only r^2 can be converted to an persentage not r

$r^2 = r \times r$

$\cdot \times 100$

$= r^2$

example $r = 0,40$

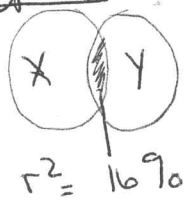
$r^2 = 0,40 \times 0,40$

$r^2 = 0,16$

$r^2 = 0,16 \times 100$

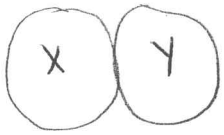
$= 16\%$

Graphic display



$r^2 = 16\%$

no correlation



100%

Correlations

→ correlation coefficient cannot be calculated as a % directly

→ refer to linear relations between 2 variables

↳ non linear if not fall into a straight line on a scatter diagram

→ correlation coefficients should <u>not</u> be averaged

→ does not imply a cause-effect relation between variables when interpretting correlations → <u>Avoid</u> using cause; influence and lead to

Factors that influence correlations

* outliers in a data set
* homogeneity of the population (make sure test groups are similar)
* restrictions in the range of variables
  ↳ make sure that variables don't have unrepresented small variances - restriction of range is a serious threat to the accurate interpretation of correlation coefficient

SG P43 - graphs → P44 - samples

SG P69 → walk through sample

SG P76 → walk though sample

SG P93 → 99

SG: 102 -