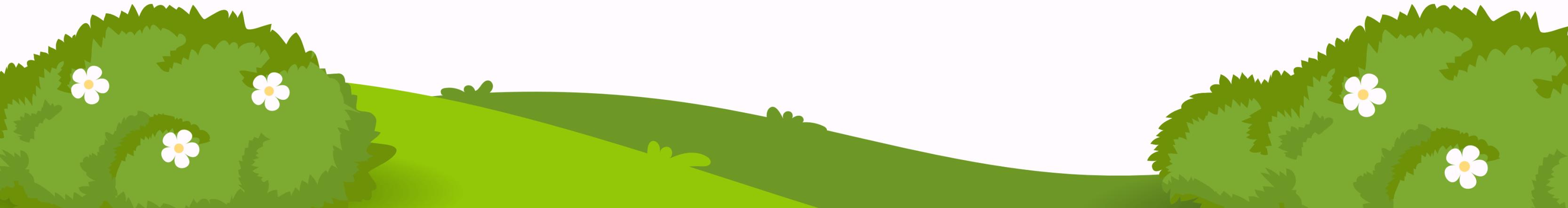
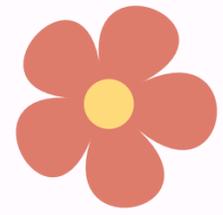


RELIABILITY, VALIDITY & TRUSTWORTHINESS



in quantitative & qualitative research





GROUP MEMBERS



NUR SAHIRA TANJUNG
059348



NIK NUR LIANA
SYAHIRAH
059537

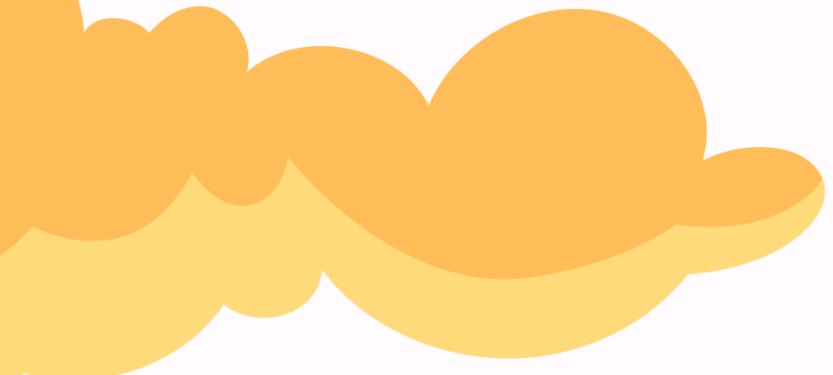


AINA HANIM SYAKIRAH
061610



NUR FARHANA
060511





RELIABILITY

- An important aspect of research that involves dependability, consistency, and replicability of results.
 - Concerned with precision and accuracy in measurement.
 - Some variables can be measured precisely (ex: height), while others cannot (ex: musical ability).
 - For research to be considered reliable, it should demonstrate that similar results would be obtained if the study were replicated with a similar group of participants in a similar context.
- 



RELIABILITY

- The concept of reliability is often associated with positivist research, but it is not exclusive to that approach.
- Qualitative research also needs to be reliable, although the criteria and methods for achieving reliability may differ from quantitative research.
- Relevant to both quantitative and qualitative research and should not be overlooked in either approach.



RELIABILITY IN QUANTITATIVE RESEARCH

- Can be categorized into three main types:
 1. Stability
 2. Equivalence
 3. Internal consistency
- Reliability refers to the consistency of results and is influenced by:
 1. Research situation
 2. Factors affecting the researcher or participants
 3. The measurement instruments used



RELIABILITY AS STABILITY

This type of reliability measures consistency over time, similar samples, and the uses of the instrument in question.

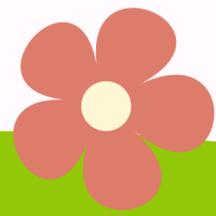




1.

Time



- A reliable instrument produces consistent data from similar respondents over time.
 - If a test is conducted and then repeated within an appropriate timeframe without any changes, similar results should be obtained.
 - The researcher needs to determine an appropriate length of time to avoid memory effects or external influences that may distort the data.
 - Correlation coefficients can be calculated to assess the reliability of pre- and post-tests.
- 
- 
- 



2.

Similar samples and the uses of the instrument in question

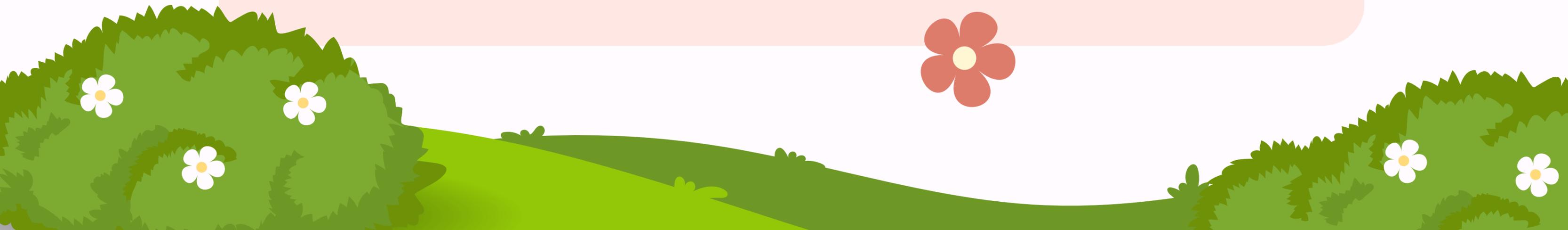
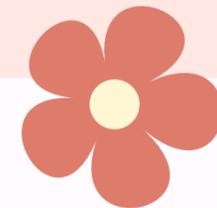


- If a test or questionnaire is administered simultaneously to two closely matched groups, who share significant characteristics (e.g. age, gender, ability etc.), similar results or responses should be obtained.
 - The correlation coefficient can be calculated for the whole test or specific sections of the questionnaire (by using a correlation statistic or a t-test as appropriate), and should be high to be considered reliable.
 - This type of reliability is particularly useful in piloting tests and questionnaires.
- 
- 
- 



When using the test/re-test method, certain considerations should be taken into account, including:

- The time period between the test and re-test should not be too long to prevent changes in situational factors.
 - The time period should not be too short, as participants may remember the first test or intervention effects may be too strong (e.g., Hawthorne effect or immediacy effect).
 - Participants may have developed an increased interest in the field and may have pursued additional information between the test and re-test periods.
- 





RELIABILITY AS EQUIVALENCE

Can be achieved through two main approaches, which are
Equivalent forms and Inter-rater reliability.





1.

Equivalent forms (a.k.a 'alternative forms') of a test or data-gathering instrument



- If an equivalent form of the instrument yields similar results, it indicates this form of reliability.
 - For example, in an experiment, the pre-test and post-test are designed as alternative forms to measure the same issues.
 - This type of reliability can also be shown when equivalent forms of a test or instrument produce consistent results when applied simultaneously to matched samples, such as two random samples in a survey.
 - Reliability can be measured through statistical tests (e.g. t-tests or Mann-Whitney U tests), by examining high correlation coefficients, or by comparing means and standard deviations between the two groups.
- 





2.

Inter-rater reliability



- 
- When multiple researchers are involved in a study, it is important to ensure agreement among them by following consistent data entry procedures.
 - This is especially relevant in cases where structured observational or semi-structured interview data are collected, and researchers need to agree on how to categorize and enter the data.
 - At simple level, inter-rater reliability can be calculated as a percentage:

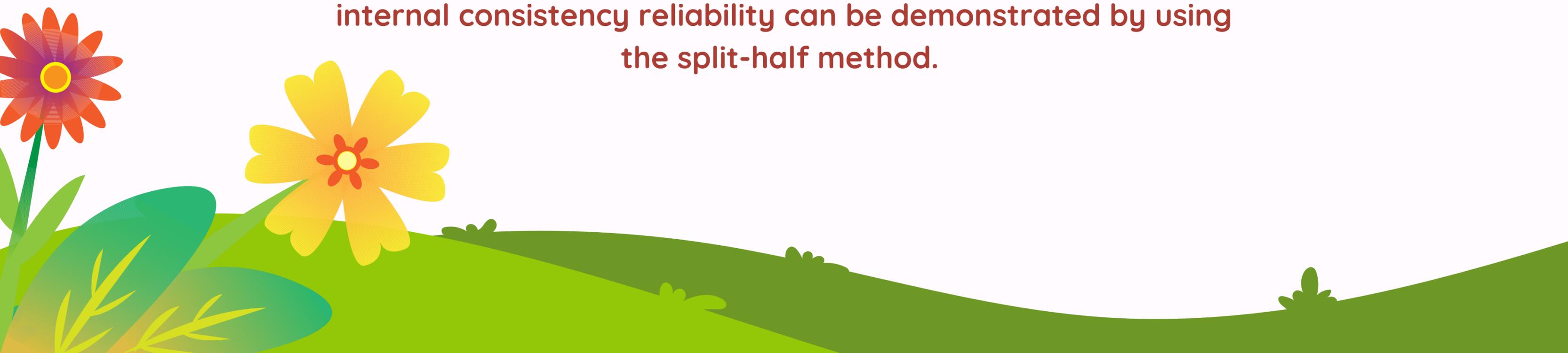
$$\frac{\text{Number of actual agreements}}{\text{Number of possible agreements}} \times 100$$





RELIABILITY AS INTERNAL CONSISTENCY

Unlike the test/re-test and equivalent forms methods, which involve conducting tests or using instruments twice, internal consistency reliability can be demonstrated by using the split-half method.





Split-half method:



Example: A researcher wants to measure self-esteem and creates a scale with 8 items (questions).

1.

Data Collection:

The researcher designs a self-esteem scale with 8 items. Each item is rated on a scale from 1 to 5, with 1 representing low self-esteem and 5 representing high self-esteem.

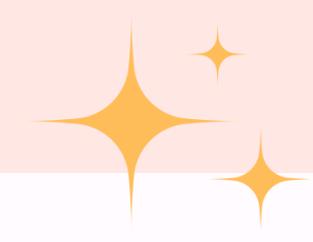
2.

Splitting the Scale:

The researcher randomly divides the 8 items into 2 equal groups: Group A and Group B.

Group A: Item 1, Item 3, Item 5, Item 7

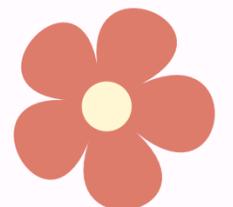
Group B: Item 2, Item 4, Item 6, Item 8





3.

Scoring:



Each participant's responses are scored separately for Group A and Group B. For example, if a participant rates Item 1 as 4, Item 3 as 3, Item 5 as 5, and Item 7 as 2, their score for Group A would be $4 + 3 + 5 + 2 = 14$. Similarly, their score for Group B would be calculated based on their responses to Items 2, 4, 6, and 8.

4.

Analysis:

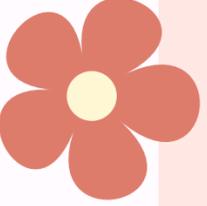
The researcher calculates the correlation coefficient between the scores of Group A and Group B. Let's assume the correlation coefficient is found to be 0.85.





5.

Reliability Estimation:



The correlation coefficient of 0.85 indicates a strong positive relationship between the two halves of the scale. It suggests that the items within the scale consistently measure the same construct of self-esteem. However, since we only used half of the items, the reliability estimate needs to be adjusted.

6.

Adjusting the Reliability Estimate:

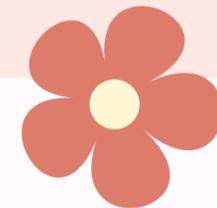
The researcher applies a reliability adjustment formula, such as the Spearman-Brown formula or Cronbach's alpha, to estimate the reliability of the full scale based on the split-half correlation coefficient. Let's assume that the reliability adjustment formula provides an estimated reliability of 0.90 for the full scale.





Factors influencing the reliability of a data-collection instrument:

- the length of the data-collection instrument (e.g. a test)
 - the heterogeneity of the group being investigated (the greater the heterogeneity, the greater the reliability)
 - the abilities of the participants
 - the methods of testing for reliability
 - the nature of the variable that is being measured or investigated
- 



To improve reliability, researchers can:



1.

minimize any external sources of variation by standardizing and controlling data collection conditions

3.

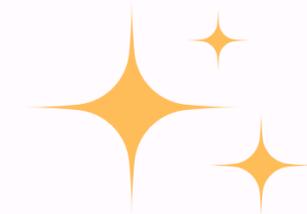
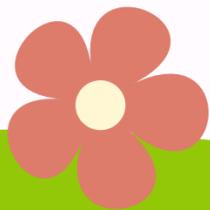
widened the number of items on a particular topic

2.

ensure consistency among researchers (inter-rater reliability) through training

4.

exclude extreme responses during data analysis (e.g., outliers)



VALIDITY

1.

CONTENT VALIDITY

2.

INTERNAL VALIDITY

3.

FACE VALIDITY



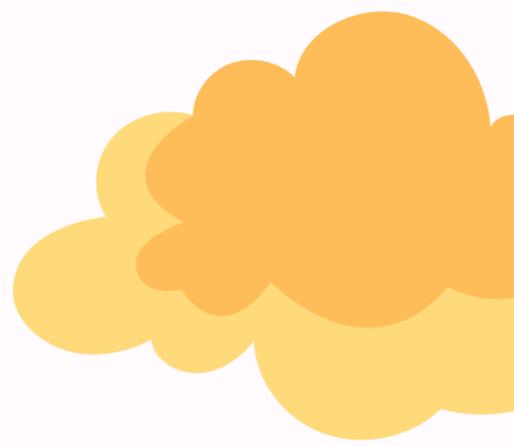
VALIDITY

Validity refers to how well a study measures what it claims to measure and how accurately it draws conclusions.



VALIDITY FEATURES

- controllability (ability to control variables)
- replicability (ability to replicate the study)
- consistency (consistency of results)
- predictability (ability to predict outcomes)
- generalizability (ability to apply findings to a larger population)
- randomization of samples (random selection of participants)
- neutrality/objectivity (lack of bias)
- observability (ability to observe and measure phenomena)

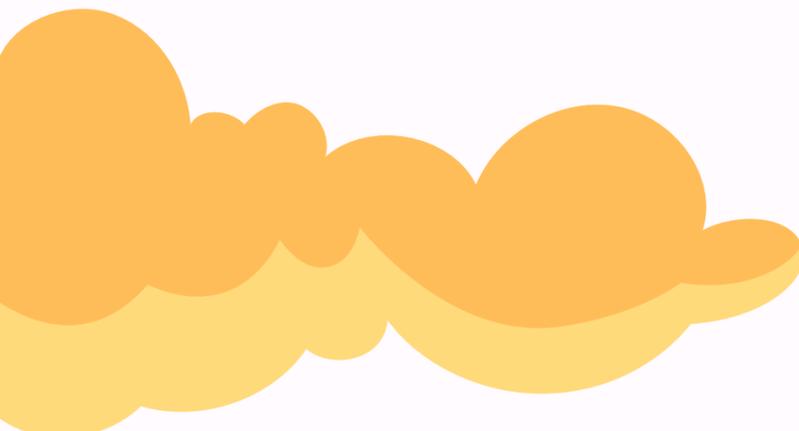




ASPECTS IN A STUDY: VALIDITY

- staying true to the assumptions underlying the statistical methods used
- ensuring the construct and content validity of the measurement tools employed
- conducting careful sampling
- avoiding various threats to internal and external validity





SHADISH ET AL. (2002)

Statistical conclusion (Shadish et al., 2002) validity, which refers to the accuracy of statistical conclusions.



COMPROMISED BY FACTORS:

- low statistical power
- violation of statistical assumptions (e.g., normal distribution of data, linearity, sample size)
- measurement errors
- limited range of data from the measurements
- inconsistent procedures for the treatments/interventions
- extraneous variables
- wide variability in outcome measures
- errors in statistical formulas
- false assumptions of causality.



CONTENT VALIDITY

Content validity refers to the extent to which an instrument covers all the relevant aspects or items of the topic it is meant to measure. It is important for the instrument to represent the broader issue being investigated and for the selected sample items to be representative and comprehensive.

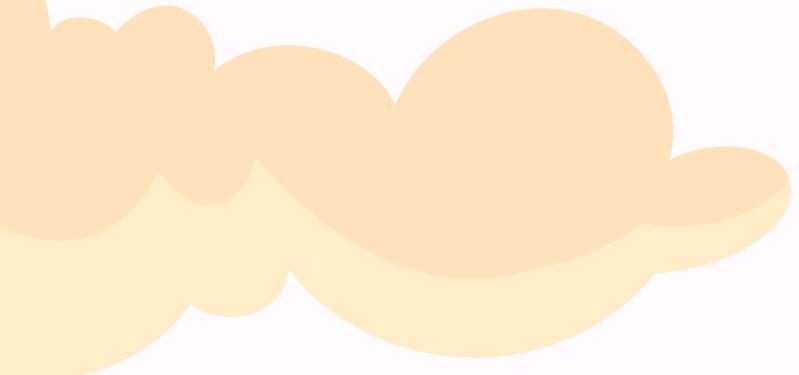




To demonstrate content validity, the instrument must show that it fairly and comprehensively covers the domain or items that it purports to cover (Carmines and Zeller, 1979, p. 20) because it is unlikely that each issue will be able to be addressed in its entirety simply because of the time available.

The researcher must ensure that the elements of the main issue to be covered in the research are:

- a fair representation of the wider issue under investigation (and its weighting)
 - the elements chosen for the research sample are themselves addressed in depth and breadth
 - careful sampling of items is required to ensure their representativeness
- 
- 
- 

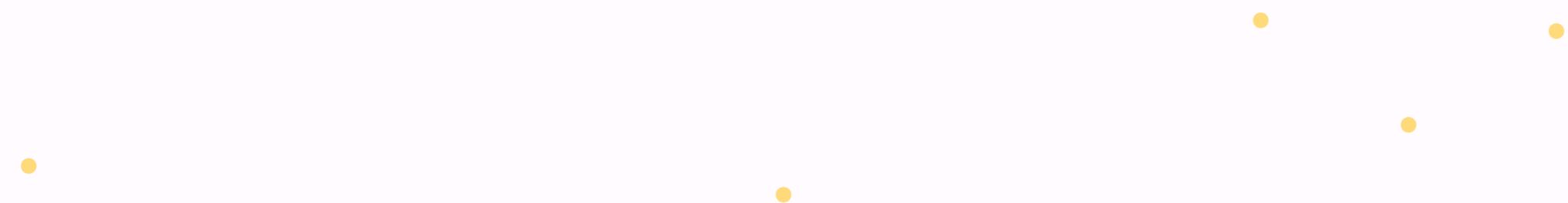


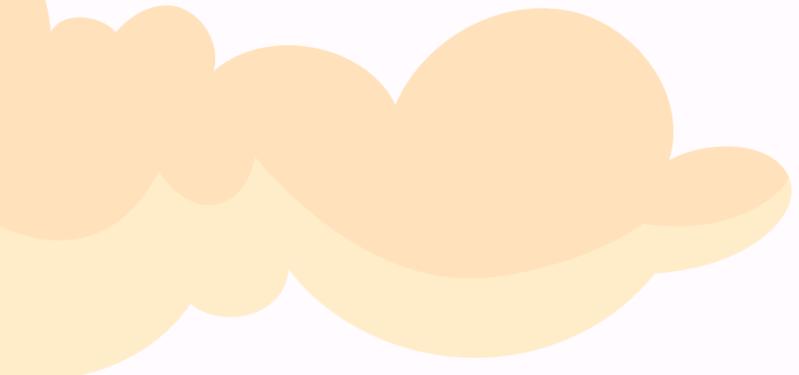
🌸 EXAMPLE 🌸

If a researcher wants to assess how well a group of students can spell 1,000 French words, but decides to use a sample of only fifty words for the spelling test, the test must include fifty words that fairly represent the range of spellings found in the 1,000 words.

This can be achieved by including all spelling rules or covering possible spelling errors in proportions similar to those occurring in the 1,000 words.

The researcher must ensure that the 1,000 words cover all the aspects of spelling they are interested in and then randomly select fifty items to check if they adequately represent the 1,000 items.

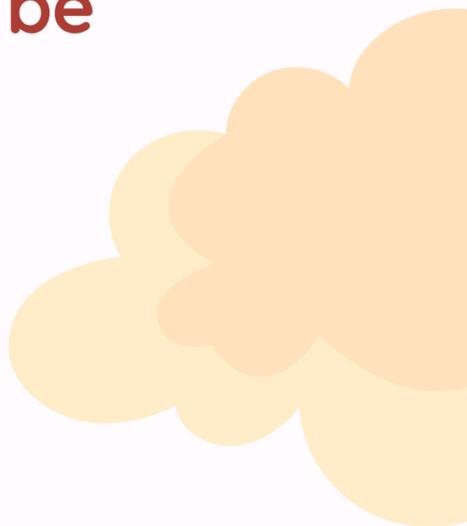




CHALLENGES

- Identifying the characteristics required in the population (such as people or spelling items).
- Defining the universe of content from which the sample will be drawn.

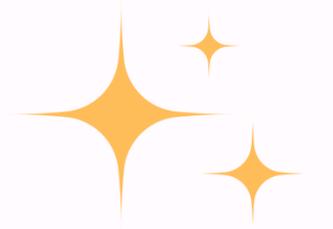
TIPS: Expert opinion or jury validity can be helpful in this process.





INTERNAL VALIDITY

- Refers to the extent to which a study accurately measures the cause-and-effect relationship between variables.
- It focuses on whether the experimental treatments truly make a difference in the specific experiments being examined and if the research is free from errors or validity violations.
- On the other hand, external validity considers how the observed effects can be generalized to other populations or settings.



11 THREATS

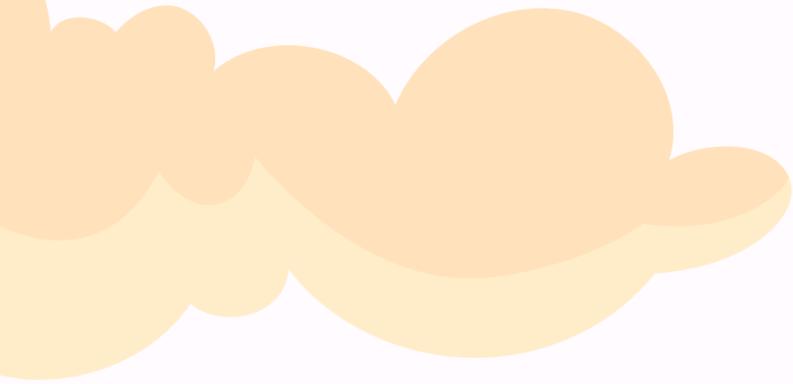
History	Other events occurring between pre-test and post-test observations can mistakenly be attributed to treatment differences.
Maturation	Subjects naturally change over time, which can produce differences independent of the research.
Ambiguous temporal precedence	It's important to determine which variable is the cause and which is the effect.

11 THREATS

Statistical regression	Subjects who initially score extremely high or low on a pre-test are likely to move closer to the average on a post-test, leading to a regression to the mean.
Testing	Pre-tests can produce effects unrelated to the research, such as sensitizing subjects or practice effects.
Instrumentation	Unreliable tests or instruments can introduce errors, particularly if there are changes in observer skills or instrument calibration.

11 THREATS

Selection	Differences in subject selection for comparison groups can introduce bias and interact with other factors.
Experimental mortality	Subjects dropping out of long-running research can confound the effects, as those who remain may differ from the initial sample.
Instrument reactivity	The data collection instruments themselves may influence the behavior of the participants.



11 THREATS

Selection-maturation interaction	Confusion between the research design effects and the effects of the variable being studied.
Type I and Type II errors	Type I error is a false positive, while Type II error is a false negative. Adjusting the level of significance can address these errors.

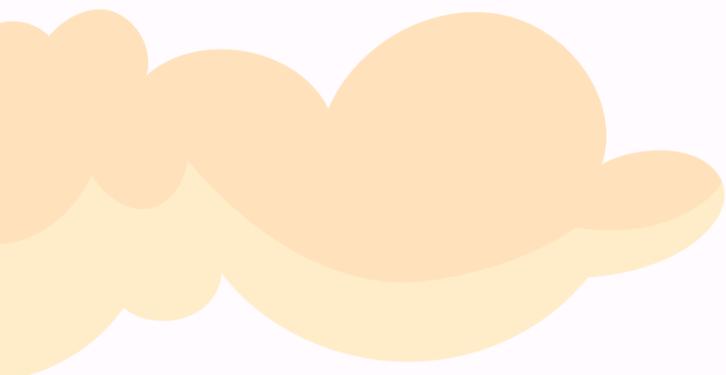


THREATS: ARY ET AL. (2002)



- "Construct underrepresentation" refers to the inadequate representation of a concept or factor in the instrumentation or data collection. This could occur when the measurement is too narrow or selective, not fully capturing the construct being studied.
- "Construct-irrelevance variance," which refers to the influence of other unrelated factors on the factor or process being investigated.





TYPE I & II ERROR

Type I error happens when,

- we mistakenly conclude that there is a significant effect or relationship in our research when, in fact, there isn't one.
- To address this error, we can make our research more stringent by setting a higher level of significance.
- For example, instead of considering a result significant if the probability (represented by p) is less than 0.05, we can set it at 0.01.

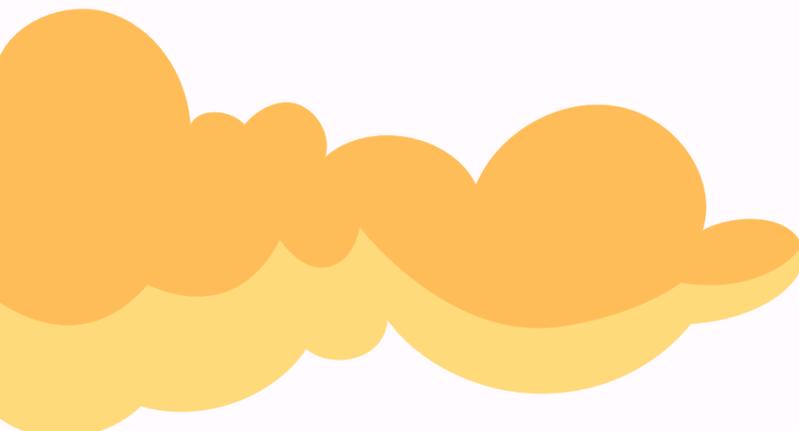
Type II error occurs when,

- we fail to detect a significant effect or relationship that actually exists.
- 
- 

TYPE II ERROR FACTORS

According to Boruch (1997), several factors can contribute to a Type II errors including:

- Inadequate validity of the measurement of the response to the intervention
- Irrelevance of the measurement of the intervention
- Low statistical power of the experiment
- Selecting the wrong population for the intervention.
- To address a Type II error, we can reduce the level of significance.
- For instance, instead of using $p < 0.05$ as the threshold for significance, we can use $p < 0.20$ or $p < 0.30$.



FACE VALIDITY

Refers to a subjective assessment of whether a measurement or research instrument appears to measure what it is intended to measure on the surface. It involves a common-sense evaluation of whether the items or questions in a measurement tool seem to be relevant and appropriate for capturing the concept or construct being studied.

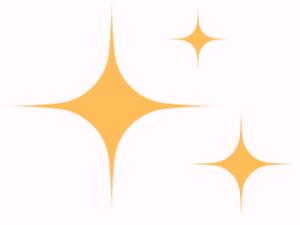




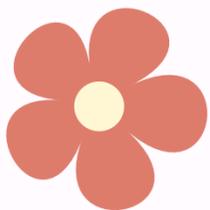
EXAMPLE



Suppose a researcher is interested in measuring the level of job satisfaction among employees in a company. They develop a questionnaire that includes items like:

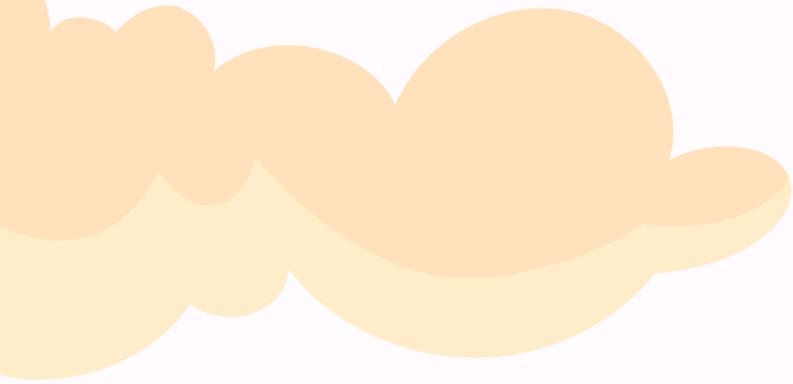
- "I am satisfied with my salary,"
 - "I enjoy the tasks I perform at work,"
 - "I feel valued by my superiors."
- 

Before using this questionnaire, the researcher may conduct a face validity assessment.

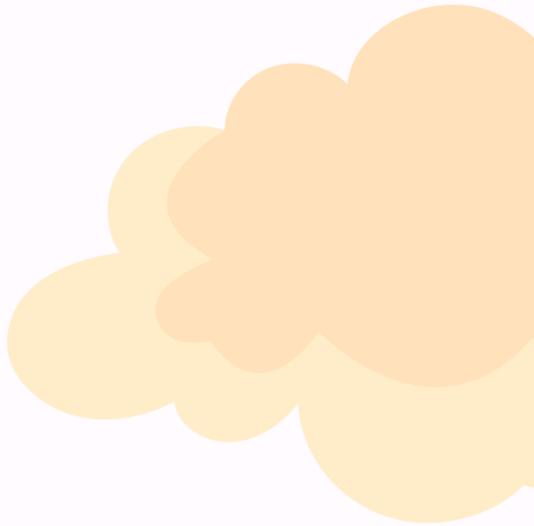


DURING THE FACE VALIDITY EVALUATION

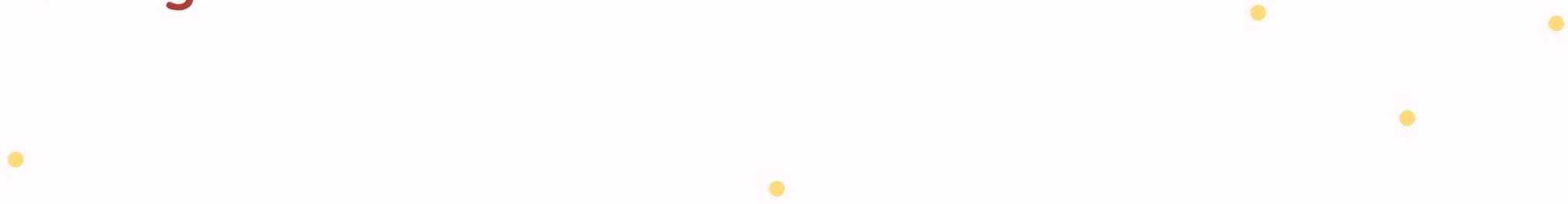
- The researcher and other experts or individuals familiar with the topic of job satisfaction would review the questionnaire items.
- They would ask themselves if the items appear to be directly related to job satisfaction and if they make intuitive sense.
- For example, they might agree that questions about salary, task enjoyment, and supervisor appreciation are relevant indicators of job satisfaction.
- If the questionnaire items are deemed appropriate and relevant, they would be considered to have face validity.



IMPORTANT NOTE



It's important to note that face validity is a:

- Subjective judgment
 - Does not provide definitive evidence of whether the measurement truly captures the intended construct.
 - It simply suggests that, based on a superficial examination, the items appear to measure what they are intended to measure.
 - Face validity can serve as an initial step in the development and evaluation of research instruments, but it is not a substitute for more rigorous forms of validity testing.
- 

The background features a bright, cheerful scene with a sun in the top right, orange clouds in the top left, and a green landscape at the bottom with a purple flower, a yellow flower, and some bushes with small white flowers.

TRUSTWORTHINESS

(QUALITATIVE RESEARCH)

The extent to which the research findings can be considered credible, transferable, dependant, and confirmable.



COMPONENTS OF TRUSTWORTHINESS

1.

CREDIBILITY

2.

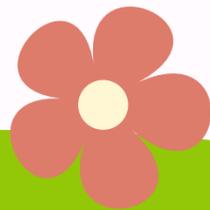
DEPENDABILITY

3.

CONFIRMABILITY

4.

TRANSFERABILITY





CREDIBILITY

THE 1ST COMPONENT OF
TRUSTWORTHINESS (QUALITATIVE)



CREDIBILITY

(SIMILAR TO THE CONCEPT OF INTERNAL VALIDITY)



- refers to how well the findings of a study align with reality
- **subjective judgment** that relies on individual perceptions

Strategies to ensure credibility in qualitative research:

1. Member checking

allows participants to review research write-ups and provide feedback on the accuracy of the data and interpretations

3. Prolonged engagement

spend a significant amount of time actively involved in the research setting, e.g. for a complete cycle or a specific period

2. Peer debriefing

seeking reactions and feedback from colleagues or co-researchers who are not directly involved in the study

4. Triangulation

using multiple methods, sources, or researchers to collect and analyze data

TYPES OF TRIANGULATION

	Type of Triangulation	Explanation	Example
1.	Methodological triangulation	Using more than one method to collect or analyze data.	<p>A study on the effectiveness of a new teaching method in a school.</p> <p>Tools: Classroom observations, interviews with teachers, and analysis of student performance data.</p>
2.	Data triangulation	Using different types of data to support findings.	<p>Researchers investigating the impact of a training program on employee satisfaction to gather a comprehensive understanding.</p> <p>Tools: Surveys, interviews, and company records.</p>

TYPES OF TRIANGULATION

Type of Triangulation	Explanation	Example
3. Investigator triangulation	Multiple researchers to independently analyze findings and compare their interpretations.	A research team studying the effectiveness of a study strategy class. Have each member evaluate the data from different sources and share their conclusions for further analysis.
4. Theoretical triangulation	Employing multiple theoretical perspectives to understand findings or guide the research.	Investigating the factors influencing consumer behaviour in a particular market. Utilize different theoretical perspectives to gain a comprehensive understanding of consumer decision-making.

TYPES OF TRIANGULATION

Type of Triangulation	Explanation	Example
5. Environmental triangulation	Conducting the study in different situations or contexts to explore the intended focus.	<p>Examining the effectiveness of a social program targeting at-risk youth.</p> <p>Can be conducted in various locations or communities to understand how the program's outcomes might vary in different contexts.</p>



DEPENDABILITY

THE 2ND COMPONENT OF
TRUSTWORTHINESS (QUALITATIVE)



DEPENDABILITY



- refers to the stability and consistency of research findings over time

Practices to ensure dependability in qualitative research:

1. Peer debriefing

having another researcher review and provide feedback on the researcher's field notes, including their interpretations

2. Audit trail

the documentation and record-keeping of the research process, including all the steps taken, decisions made, and changes implemented throughout the study

3. Anticipation of peer review

researchers are aware that their work and its products will be inspected by a peer, which prompts researchers to be cautious about what they record as factual observations and what they set aside as their own interpretive comments

4. Bracketing

involves separating the raw observations from the researcher's interpretations

BRACKETING

- Bracketing acknowledges the reflexive analysis that researchers engage in, which affects the research process.
- Bracketing doesn't devalue the researcher's thoughts; it simply recognizes that different processes are at play.
- However, researchers also need to be aware of the influence of their values and passions and actively monitor them.
- This self-awareness and transparency about their involvement in the research process is known as reflexive auditing.
- It involves describing how the researcher's values and decisions shaped the research project.



CONFIRMABILITY

THE 3RD COMPONENT OF
TRUSTWORTHINESS (QUALITATIVE)

CONFIRMABILITY



- the degree to which the findings are based on **objective evidence** rather than the researcher's biases or values
- research findings can be verified or confirmed by others

Strategies to ensure confirmability in qualitative research:

1. Reflexivity

acknowledging and examining the researcher's own values, beliefs, and biases and how they may affect the research process and findings

3. Peer review

other researchers review the research findings, then help identify any biases or assumptions that may have influenced the research and provide an objective assessment of the research findings

2. Audit trail

a detailed record of the research process that can be reviewed by others to verify the research findings

CONFIRMABILITY

Study on the experiences of ethnic minority students in higher education

The researcher conducts interviews and collects data from a diverse group of ethnic minority students studying in various universities across Malaysia. The study aims to understand their challenges, perceptions of discrimination, and coping strategies within the higher education system.

EXAMPLE

How to ensure confirmability in qualitative research?

The researcher need to:

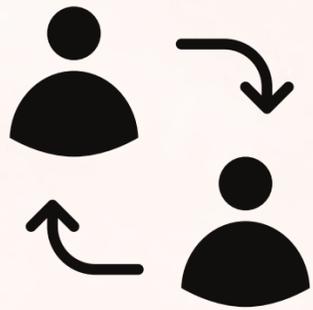
1. maintain a reflexive journal throughout the research process to document their own biases, assumptions, and personal values
2. maintain an audit trail to record all aspects of the research process
3. engage in peer review to review the methods used and provide feedback on the research findings



TRANSFERABILITY

THE 4TH COMPONENT OF
TRUSTWORTHINESS (QUALITATIVE)

TRANSFERABILITY



- the degree to which the findings of a study **can be applied to other settings** or contexts beyond the immediate research setting
- findings of a qualitative study can be generalized to other populations, contexts, or settings.

Strategies to ensure transferability in qualitative research:

1.

Use purposive sampling (selecting participants based on specific criteria that are relevant to the research question)

2.

Provide detailed description of the sociodemographic characteristics of the participants



TRANSFERABILITY



Study on language attitudes and code-switching among multilingual individuals in Kuala Lumpur.

The researcher conducts interviews and collects data from participants who represent various language backgrounds, including Malay, English, and Chinese speakers. The study explores the participants' perceptions of code-switching, their language preferences in different contexts, and the sociocultural factors influencing their language choices.

How to ensure transferability in qualitative research?

1.

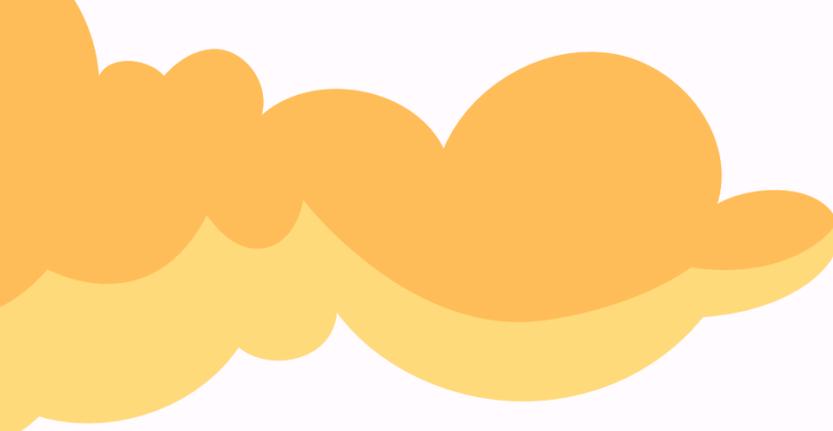
Share the detailed descriptions of the linguistic landscape in Malaysia

Status of Malay as the national language, the widespread use of English in education and business, and the presence of multiple Chinese dialects. The researcher also accounts for the multicultural and multilingual nature of Kuala Lumpur, highlighting the diverse linguistic practices and language policies in the city.

2.

Reference for future linguistic researchers

They can examine the similarities and differences in language attitudes, language use patterns, or sociocultural factors between Kuala Lumpur and their specific research site to determine the relevance of the findings.



ETHICAL CONSIDERATIONS

TO ENSURE THAT RESEARCHERS RESPECT THE RIGHTS, NEEDS, VALUES,
AND DESIRES OF THE INDIVIDUALS PARTICIPATING IN THE STUDY

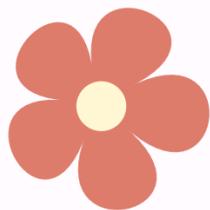
To ensure the protection of participants' rights, the following safeguards will be employed:

1. The research objectives will be articulated verbally and in writing so that they are clearly understood by the informant (including a description of how data will be used).
2. Written permission to proceed with the study as articulated will be received from the informant.
3. A research exemption form will be filed with the Institutional Review Board.
4. The informant will be informed of all data collection devices and activities.
5. Verbatim transcriptions and written interpretations and reports will be made available to the informant.
6. The informant's rights, interests and wishes will be considered first when choices are made regarding reporting the data.
7. The final decision regarding informant anonymity will rest with the informant.



CONCLUSION

QUANTITATIVE RESEARCH	QUALITATIVE RESEARCH
Validity and reliability	Trustworthiness <ul style="list-style-type: none">· Credibility· Dependability· Confirmability· Transferability
BENEFITS	
<ul style="list-style-type: none">• crucial for determining the quality and credibility of research findings• essential for future research in the respective research fields	





THANK YOU!

GROUP 12

