

Evaluation measures

Antonio Reyes Pérez (PhD)

Facultad de Lenguas y Letras

Universidad Autónoma de Querétaro

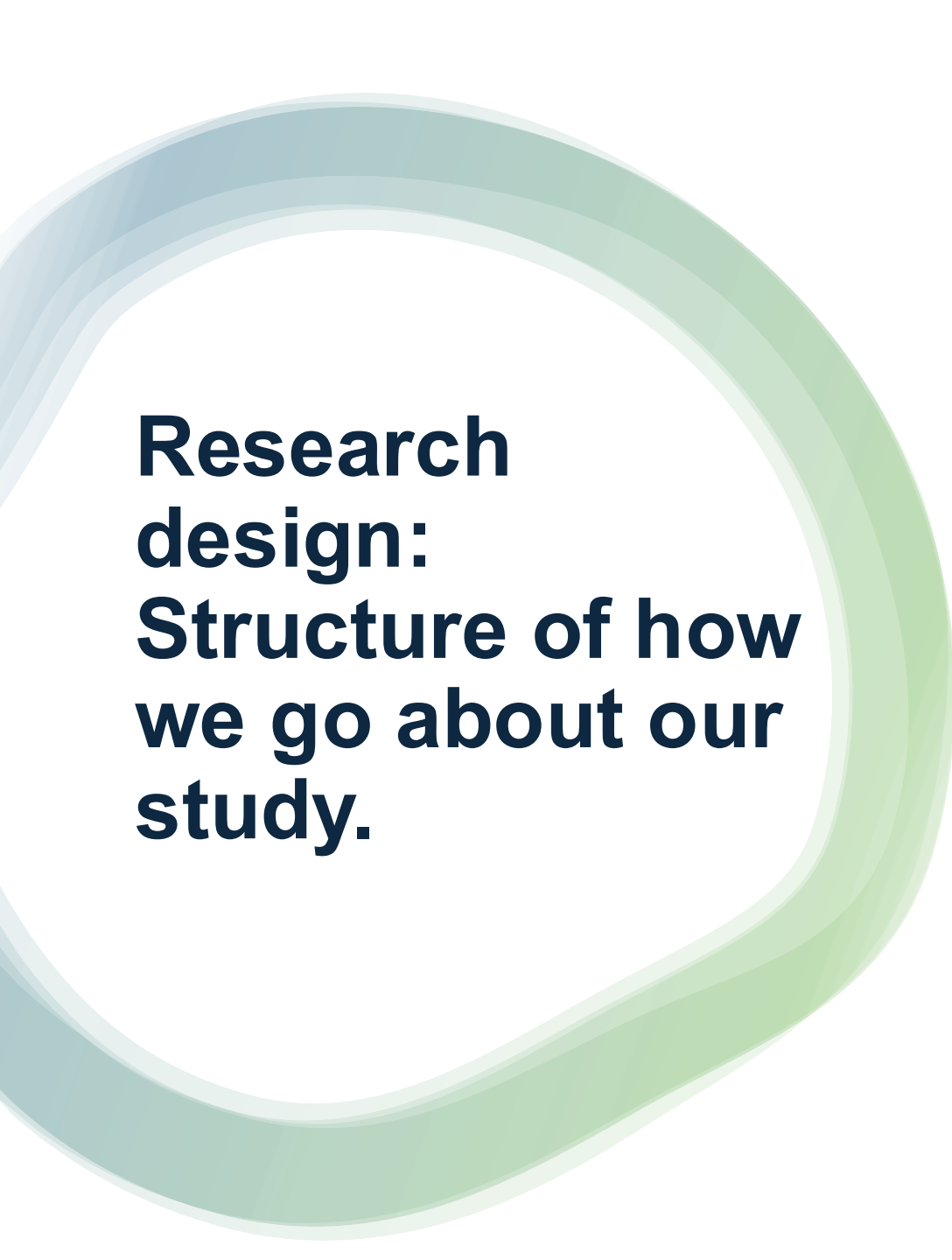
Data Processing Workshop

April 2025



Introduction

- Qualitative research is concerned with structures and patterns, and how something is.
 - Inductive: theory is derived from the results of our research.
- Quantitative research focuses on how much or how many there is/are of a particular characteristic or item.
 - Deductive: based on already known theory we develop hypotheses, which we then try to prove (or disprove) in the course of our empirical investigation.




Research design: Structure of how we go about our study.

- Framework around which we organize our study, but it does not refer to the actual tools we use to carry out our research (questionnaires, recordings, etc.).
- Not only do we need a set of tools to get our data
- We also need to think about a coherent and solid framework around which we organize our data collection.
- Order is very important > RQ
- We could have data, but is this data useful to answer our RQ?



Two typical designs

- Cross-sectional design (transversal): we collect a comparatively large amount of data at one point in time.
- Longitudinal design (longitudinal): they are based on the repeated collection of data over a longer period of time.
- Research design, theoretical background and actual methods used are inseparably linked and form the overall framework for our study,
- How data is collected.



Some notes about quantitative methods

- **Variables:** features, measurable > data > RQ
- Whenever we want to quantitatively measure something, that is, assign a variable value to a particular case, we need to thoroughly think about a reliable way to make this decision.
- Need a set of clear and objective definitions for each category or outcome.
- Measures should be designed in such a way that it comprises as many cases as possible.
- Example:
 - The number of different realizations of the vowels:
 - How often does /e/ occur in words such as peck, and how often is /e/ replaced by /æ/ making it homophone to pack.



Operationalization

- It is important to establish a clear set of rules as to what constitutes an /e/ and what counts as an /æ/.
- Anything up to x Hertz counts as an /e/ and everything below as an /æ/.
- Why X is X.
- Eliminate/reduce ambiguity

Measures

- **Reliability** refers to our measure repeatedly delivering the same (or near same) results:
 - if we use the same measure with the same people under the same conditions, our measure should give us the same result.
- **Validity** refers to our measure actually measuring what it is supposed to measure:
 - if we get significantly different results from previous research, under very similar circumstances, we should carefully look at our method before we get too enthusiastic about our results.

1. ¿Eres un niño o una niña?

☐ Niño ☐ Niña

2. ¿En qué grupo estás?

☐ 5º A ☐ 5º B

☐ 6º A ☐ 6º B

3. ¿Cuántos años tienes?

_____ años

4. ¿Cuánto pesas?

_____ kg.

5. ¿Cuánto mides?

_____ cm.

6. ¿Acostumbras/sueles desayunar en las mañanas?

☐ Sí, todos los días

☐ Sí, pero sólo algunos días

☐ No, nunca desayuno

7. ¿Has desayunado algo HOY antes de venir a la escuela?

☐ Sí ☐ No

8. ¿Qué desayunaste HOY?

☐ Leche

☐ Chocolate (tipo Chocomilk, Nesquick o parecido)

☐ Galletas

☐ Cereales

☐ Pastelitos

☐ Yogurth

☐ Fruta

☐ Jugo preparado en casa

☐ Pan

☐ Otras cosas



Reliability (confiabilidad)

- La confiabilidad consiste en estimar la cantidad de variación en las puntuaciones de las pruebas que se debe a error en la medición.
- Dicho de otra forma, la confiabilidad de un instrumento de medición se refiere al **grado en su aplicación repetida al mismo sujeto u objeto produce resultados iguales**.
- Esta estimación se enfoca en los efectos de método de evaluar y en factores aleatorios.
- Las estimaciones de confiabilidad se llaman **coeficientes de confiabilidad**.



Validity (validez)

- La validez se refiere al grado en que un instrumento realmente mide la variable que pretende medir.
- Si una de mis variables es la edad, ¿cuál sería la mejor forma para obtener este dato en mi muestra?
 - ¿Cuándo naciste?
 - ¿Año de nacimiento?
 - Escribe día, mes y año de nacimiento
 - ¿Cuál es tu edad?



Reliability coefficients

- **Medidas de consistencia interna.** Tienen que ver básicamente con orígenes de error de dentro de la prueba y los procedimientos de puntuar.
- Medidas de estabilidad. Indican qué tan consistentes son las puntuaciones de las pruebas (testretest).
- Medidas de equivalencia. Dan una indicación de hasta qué punto las puntuaciones de formas alternativas de una misma prueba son equivalentes. Aquí se trata normalmente de dos versiones de una prueba aplicada al mismo grupo.



Principles

- Para hacer las pruebas de confiabilidad (y otras) es necesario hacer la descripción estadística de la muestra.
- Los conceptos de media, varianza y desviación estándar cobran gran valor.
- Hay diferentes algoritmos:
 - Medida Spearman-Brown
 - Kuder-Richardson
 - Coeficiente Alfa (o Alfa de Cronbach)

Evaluation

