

Máster en Ingeniería y Tecnología de Sistemas Software

Curso académico 2024/2025

Ingeniería del Software Experimental

Trabajo 1. Diseño de experimento Controlado

Silvia Abrahao

Departamento de Sistemas Informáticos y Computación

email: sabrahao@dsic.upv.es



UNIVERSIDAD
POLITECNICA
DE VALENCIA

DSIC

Sesión Práctica:

- Objetivo: cada grupo deberá elaborar un documento* con el **diseño de un experimento controlado** para comparar dos métodos/técnicas/herramientas en un dominio concreto.
- Introducid la información del grupo y tema elegido aquí:
https://docs.google.com/document/d/15YxaDxUgbjZtQh9fC6pBdKjS9k_5_bxMjK-Zsj39l4s/edit?usp=sharing

El documento a entregar deberá seguir las guías descritas en:

- A. Jedlischka, D. Pfajl, Reporting Guidelines for Controlled Experiments in Software Engineering, ISESE 2005: 95-104.
- **Experiment Process (en PoliformaT)**. El trabajo deberá incluir los siguientes apartados: *definition, implications and planning*



UNIVERSIDAD
POLITECNICA
DE VALENCIA

Proceso Experimental. Conceptos

Ejemplo: Queremos evaluar el efecto de un nuevo método de desarrollo en la productividad de los desarrolladores. Queremos introducir un **método dirigido por modelos** en lugar de un **método tradicional**

- **Variable dependiente:** productividad
- **Variable independiente:** el *método de desarrollo*, la *experiencia*, las *herramientas* y el entorno
- El **factor** es el método de desarrollo, ya que queremos evaluar el efecto cuando cambiamos el método
- Utilizamos **dos tratamientos** del factor: El método tradicional (viejo) y el método dirigido por modelos (nuevo)
- Los **objetos experimentales** en el ejemplo son los sistemas que deben ser desarrollados y los **sujetos** (desarrolladores)
- Un **test** puede ser que la persona N (sujeto) usa el método de desarrollo dirigido por modelos (tratamiento) para desarrollar el sistema A (objeto)



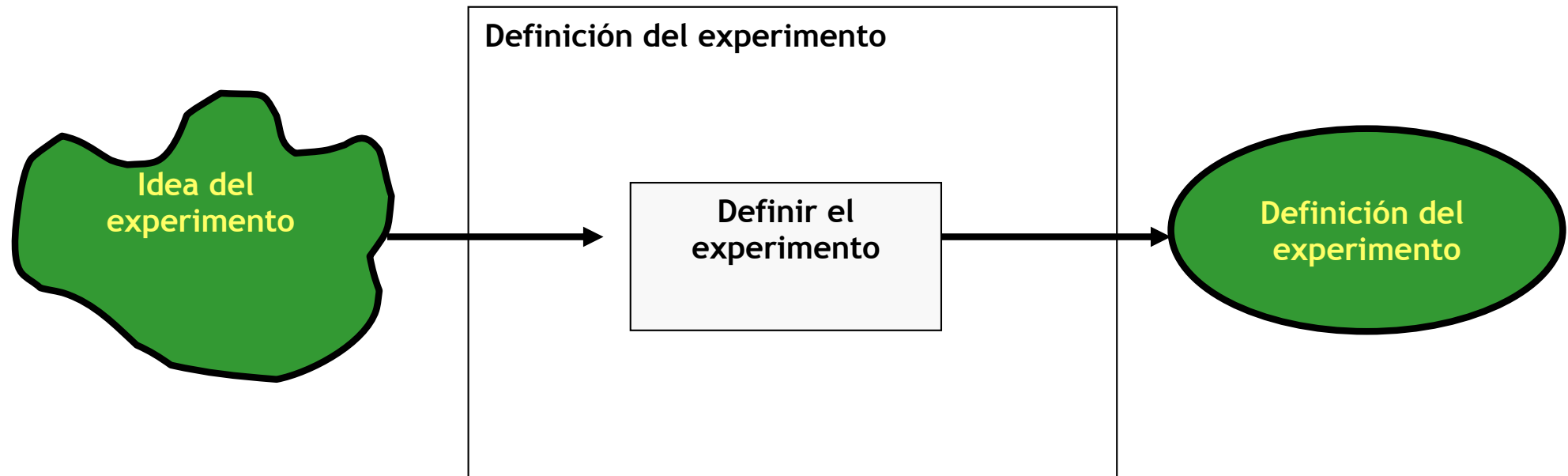
Diseño de un experimento controlado

PARTE 1: ETAPA DE DEFINICIÓN



UNIVERSIDAD
POLITECNICA
DE VALENCIA

Proceso Experimental. Etapa de Definición



- En esta etapa se establecen:
 - los fundamentos del experimento
 - los objetivos del experimento
 - “¿Por qué se realiza el experimento



Proceso Experimental. Etapa de Definición

Plantilla para la definición de objetivos

- El uso de una plantilla asegura que todos los aspectos importantes de un experimento son definidos antes que se realice la planificación y la ejecución del mismo.

- La plantilla para la definición de objetivos es la siguiente:

Analizar *<Objeto(s) de estudio>*

(¿qué es lo que se estudia?)

con el propósito de *<Propósito>*

(¿cuál es la intención?)

con respecto a *<Enfoque de calidad>*

(¿cuál es el efecto estudiado?)

desde el punto de vista de *<Perspectiva>*

(¿a la vista de quién?)

en el contexto de *<Contexto>*

(¿dónde se lleva a cabo el estudio?)



UNIVERSIDAD
POLITECNICA
DE VALENCIA

Proceso Experimental. Etapa de Definición

- **Objeto de estudio.** Es la **entidad** que es estudiada en el experimento. Pueden ser productos, procesos, recursos, modelos, métricas o teorías.
- **Propósito.** Define cuál es la **intención** del experimento. Puede ser evaluar el impacto de dos técnicas diferentes, o caracterizar la curva de aprendizaje de una organización.
- **Enfoque de calidad.** Es el **efecto** estudiado en el experimento. Puede ser la efectividad, el coste, la fiabilidad, etc.
- **Perspectiva.** Se refiere al **punto de vista** desde el cual los resultados del experimento son interpretados (*desarrollador, arquitecto de software, etc.*)
- **Contexto.** Es el **entorno** en el cual se ejecuta el experimento. Describe brevemente cuáles son los sujetos involucrados y que artefactos de software son utilizados en el experimento.



Proceso Experimental. Etapa de Definición

Este ejemplo define un Experimento de Inspección en el cual se investigan dos técnicas de inspección (Regnell et al., 1999):

- **Objeto de estudio.** Las técnicas *Perspective-Based Raeding (PBR)* y una técnica alternativa basada en *Checklist*
- **Propósito.** Evaluar las técnicas de inspección
- **Enfoque de calidad.** La *efectividad* y la *eficiencia* de las técnicas de inspección
- **Perspectiva.** Desde el punto de vista de los *investigadores*
- **Contexto.** El experimento se lleva a acabo con *alumnos de grado y de máster* de un curso de Análisis y Especificación de Requisitos de Software en la UPV y se utiliza un paquete de laboratorio con documentos de requisitos textuales.



Proceso Experimental. Etapa de Definición

El ejemplo se puede resumir de la siguiente manera:

Analizar las técnicas de inspección PBR y checklist

con el propósito de evaluarlas

con respecto a su efectividad y eficiencia

desde el punto de vista del investigador

en el contexto de alumnos de grado y máster

y un paquete de laboratorio

con documentos de requisitos textuales



UNIVERSIDAD
POLITECNICA
DE VALENCIA

Diseño de un experimento controlado

PARTE 1: ETAPA DE PLANIFICACION

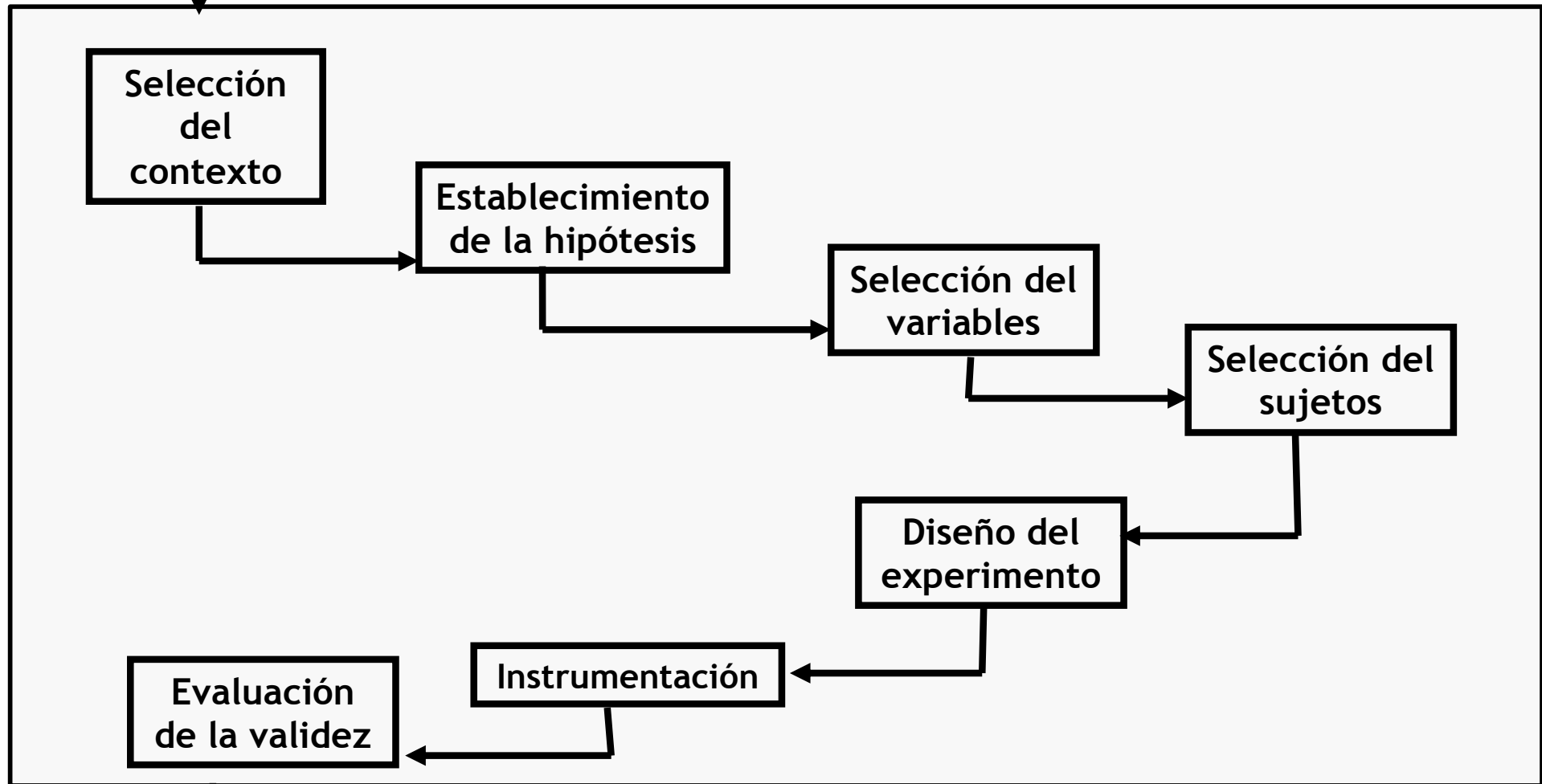


UNIVERSIDAD
POLITECNICA
DE VALENCIA

Proceso Experimental. Etapa de Planificación

Definición del experimento

“¿Cómo se realiza el experimento?”



Es un proceso “iterativo” y puede ser necesario volver atrás y refinar alguna etapa previa antes de continuar con las etapa posteriores.



Diseño del experimento

Proceso Experimental. Etapa de Planificación

Selección del contexto

- El contexto del experimento puede ser caracterizado de acuerdo a cuatro dimensiones:
 - ❖ Off-line vs. On-line
 - ❖ Estudiantes vs. Profesionales
 - ❖ Problemas de juguete vs. Problemas reales
 - ❖ Específico vs. General



Proceso Experimental. Etapa de Planificación

Establecimiento de la hipótesis

La base de cualquier análisis estadístico es “corroborar la hipótesis (**prueba de hipótesis**)

- Una hipótesis debe ser planteada formalmente y los datos recolectados durante el transcurso del experimento son utilizados para, si es posible, rechazar la hipótesis.
- Dos hipótesis deben ser planteadas (**por cada variable dependiente**):
 - ❖ Una hipótesis nula, H_0 : Es la hipótesis que se desea rechazar. Por ejemplo, que el método de inspección PBR permite encontrar en promedio la misma cantidad de errores que checklists.
$$H_0: \mu_{\text{checklist}} = \mu_{\text{PBR}}$$
 - ❖ Una hipótesis alternativa, H_1 : Es la hipótesis a favor de la cual la hipótesis nula es rechazada. Por ejemplo, que el método de inspección PBR en promedio permite encontrar más errores que Checklists.

$$H_1: \mu_{\text{checklist}} < \mu_{\text{PBR}}$$



Proceso Experimental. Etapa de Planificación

Selección de variables

- **Variables independientes:** variables que podemos controlar y modificar en el experimento.
 - Elegir las variables correctas no es fácil y requiere conocimiento del dominio.
 - La selección de variables independientes involucra también: elegir la escala de medición, el rango de las variables y a qué niveles específicos se realizarán los tests.
- **Variable(s) dependiente(s):** mide los efectos de los tratamientos.
 - Se deriva de las hipótesis (1 variable para cada hipótesis!)
 - Normalmente se miden a través de una medida indirecta.
 - Se debe elegir las **métricas** para medir dichas variables



Proceso Experimental. Etapa de Planificación

Selección de sujetos

- La selección de los sujetos tiene influencia directa sobre la posibilidad de generalizar los resultados.
- Para poder generalizar los resultados para una determinada población, es necesario que los sujetos considerados sean realmente **representativos de esa población**
- El **tamaño de la muestra** tiene un impacto considerable sobre los resultados, a la hora de querer generalizarlos. Cuanto **más grande sea la muestra menor posibilidad de error** al generalizar los resultados
- Para el trabajo (diseño del experimento) en cuestión indicad:
 - Perfil de los sujetos
 - Tamaño de la muestra
 - Cómo los sujetos serán seleccionados



Proceso Experimental. Etapa de Planificación

Diseño del experimento

Selección del diseño del experimento

- Para obtener resultados significativos de un experimento se aplican test estadísticos sobre los datos recolectados.
- Para que un experimento sea válido y sus resultados útiles, este debe ser planificado y diseñado muy cuidadosamente.
- La elección del **test** estadístico que podemos aplicar depende del **diseño previamente elegido** y de las **escalas de medición**.



Proceso Experimental. Etapa de Planificación

Diseño del experimento

Selección del diseño del experimento

- Un experimento consiste de una serie de **tests y tratamientos**. El diseño de un experimento describe cómo son organizados los tests!
- Durante el diseño se debe decidir **cuántos tests son necesarios** para estar seguros que el efecto del tratamiento sea visible.
- Además, un diseño adecuado servirá para replicar el experimento posteriormente.



Proceso Experimental. Etapa de Planificación

Diseño del experimento

Principios generales de diseño

- **Aleatorización.** Los test estadísticos para analizar los datos requieren que las observaciones sean sobre *variables aleatorias independientes*. La aleatorización se refiere a la *asignación de los objetos a los sujetos* y en que *orden los test* son realizados.
- **Bloqueo.** Algunas veces existe algún factor que probablemente tiene un efecto sobre la variable dependiente, pero no estamos interesados en tal efecto. Si el efecto del factor es conocido y controlable, lo que hacemos es “bloquearlo”. Esta técnica aumenta la precisión del experimento.
- **Balanceo.** Si asignamos los tratamientos de tal manera que cada tratamiento tenga el mismo número de sujetos, tenemos un diseño balanceado.



Proceso Experimental. Etapa de Planificación

Principios generales de diseño

Ejemplo: Una empresa desea realizar un experimento para investigar el efecto sobre la fiabilidad de un programa cuando se utiliza un diseño orientado objetos vs. un diseño tradicional.

El experimento usa un *programa A* como objeto experimental

El experimento es de tipo “multi-test within object design”

- **Aleatorización.** la selección de sujetos será representativa de los diseñadores de la empresa, seleccionando aleatoriamente los diseñadores. La asignación de cada tratamiento es asignada de forma aleatoria a cada sujeto.
- **Bloqueo.** Los sujetos tienen distintos niveles de *experiencia*. Algunos han utilizado orientación a objetos previamente y otros no. Para minimizar el efecto de la experiencia, las personas son agrupadas en dos grupos (bloques) uno con experiencia y otro sin experiencia.
- **Balanceo.** El experimento usa un diseño balanceado, lo que significa que hay la misma cantidad de personas en cada grupo (bloque)



Proceso Experimental. Etapa de Planificación

Tipos estándares de diseño (Montgomery, 1997)

- Presentamos los diseños que son adecuados para experimentos con:
 - Un factor con dos tratamientos
 - Un factor con más de dos tratamientos
 - Dos factores con dos tratamientos
 - Más de dos factores cada uno con dos tratamientos



Proceso Experimental. Etapa de Planificación

Tipos estándares de diseño (Montgomery, 1997)

■ Un factor con dos tratamientos

Ejemplo: El experimento investiga si un nuevo método de diseño produce software de más calidad que el método anterior.

El factor es el *método de diseño* y los tratamientos los dos métodos.

La variable dependiente puede ser la densidad de defectos (*número de defectos encontrados en el desarrollo/tamaño del proyecto*).

■ Hay dos posibilidades de diseño:

- **Diseño completamente aleatorio** (*completely randomized design*) - Between-subjects design
- **Diseño comparado por pares** (*paired comparison design*) - Within subjects design



Proceso Experimental. Etapa de Planificación

Tipos estándares de diseño (Montgomery, 1997)

- Un factor con dos tratamientos
- **Diseño completamente aleatorio** (*completely randomized design*)
 - Este es un ejemplo muy sencillo de experimento en el que se **comparan las “medias” de dos tratamientos**.
 - Este diseño utiliza **los mismos objetos para ambos tratamientos** y asigna los sujetos de forma aleatoria a cada tratamiento.
 - Cada **sujeto utiliza solo un tratamiento sobre cada objeto**.
 - Si tenemos el mismo numero de sujetos por tratamiento el diseño es balanceado.

Subjects	Treatment 1	Treatment 2
1	X	
2		X
3		X
4	X	
5		X
6	X	

Ejemplo: asignación de sujetos a tratamientos para el “diseño completamente aleatorio”



UNIVERSIDAD
POLITECNICA
DE VALENCIA

Proceso Experimental. Etapa de Planificación

Tipos estándares de diseño (Montgomery, 1997)

■ Diseño comparado por pares (*paired comparison design*)

- En este tipo de diseño cada sujeto usa **dos tratamientos** sobre el **mismo objeto experimental** (e.j., dos métodos sobre el mismo programa)
- Para minimizar el efecto del orden en el cual los sujetos aplican los tratamientos, el orden es asignado de manera aleatoria a los sujetos.
- Este tipo de diseño no siempre se utiliza ya que cada sujeto aprende mucho del tratamiento anterior.

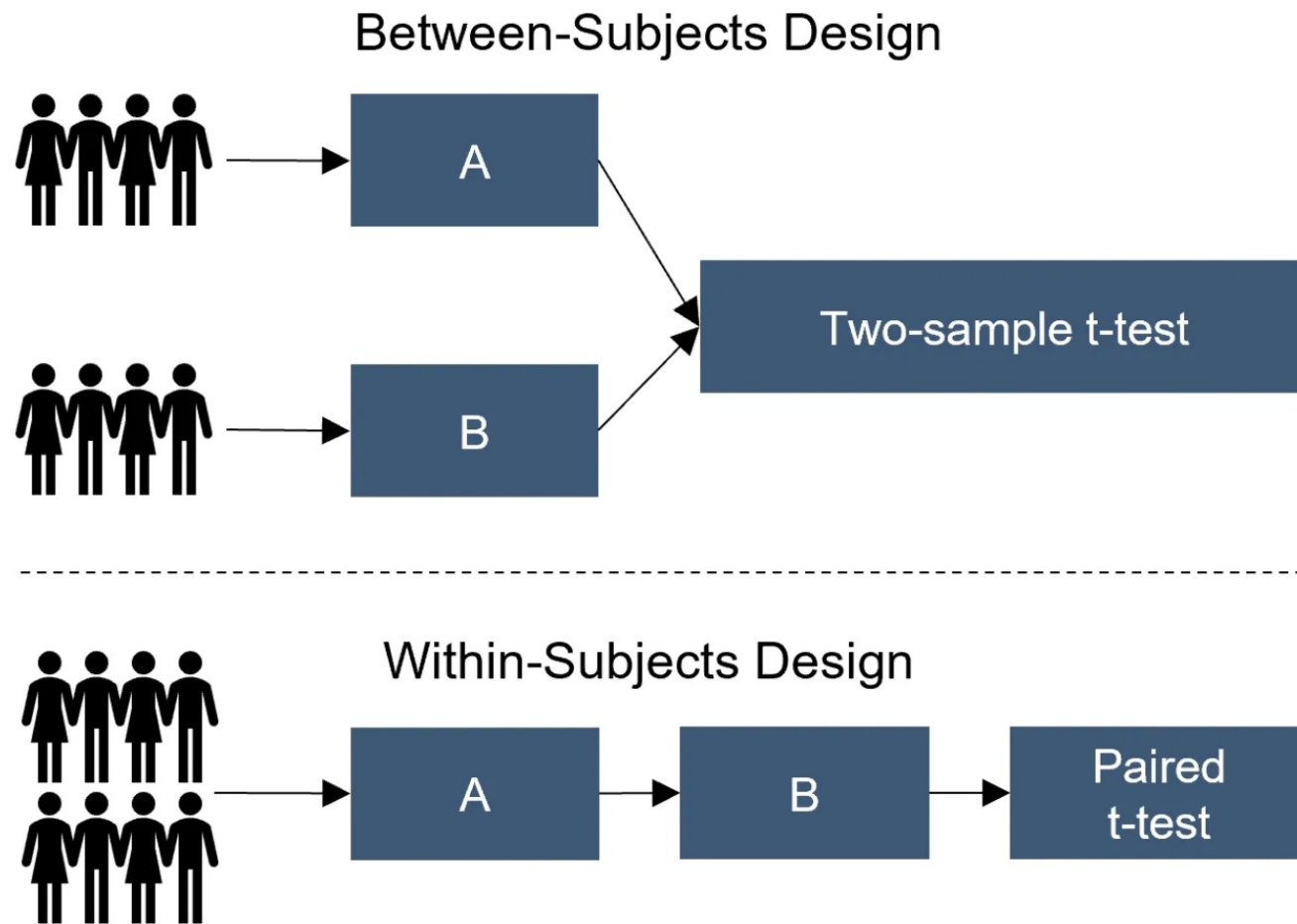
Subjects	Treatment 1	Treatment 2
1	2	1
2	1	2
3	2	1
4	2	1
5	1	2
6	1	2

Orden en el que el sujeto aplica los tratamientos a un mismo objeto experimental

Ejemplo: asignación de sujetos a tratamientos para el “diseño comparado por pares”

Proceso Experimental. Etapa de Planificación

Resumen



Proceso Experimental. Etapa de Planificación

Tipos estándares de diseño (Montgomery, 1997)

■ Diseño comparado por pares (*paired comparison design*)

- En este tipo de diseño cada sujeto usa **dos tratamientos** sobre el **mismo objeto experimental** (e.j., dos métodos sobre el mismo programa)
- Para minimizar el efecto del orden en el cual los sujetos aplican los tratamientos, el orden es asignado de manera aleatoria a los sujetos.
- Este tipo de **diseño no siempre se utiliza** ya que cada sujeto aprende mucho del tratamiento anterior!

Subjects	Treatment 1	Treatment 2
1	2	1
2	1	2
3	2	1
4	2	1
5	1	2
6	1	2

Orden en el que el sujeto aplica los tratamientos a un mismo objeto experimental

Ejemplo: asignación de sujetos a tratamientos para el “diseño comparado por pares”

Proceso Experimental. Etapa de Planificación

Tipos estándares de diseño (Montgomery, 1997)

- Un factor con más de dos tratamientos

Ejemplo: El experimento investiga la calidad del software obtenida al utilizar **diferentes lenguajes de programación**.

El **factor** es el *lenguaje de programación* y los **tratamientos** pueden ser C, C++ y java.

- **Diseño completamente aleatorio (*completely randomized design*)**
 - Este tipo de diseño requiere que el experimento sea realizado en un **orden aleatorio** por ello los tratamientos deben ser usados en un entorno tan uniforme como sea posible.
 - Este diseño usa **un objeto experimental para todos los tratamientos** y los sujetos son asignados de manera aleatoria a los tratamientos.



Proceso Experimental. Etapa de Planificación

Tipos estándares de diseño (Montgomery, 1997)

■ Dos factores

El experimento es más complejo cuando consideramos más de un factor

■ Se consideran tres hipótesis

1. El efecto del tratamiento i sobre el factor A
2. El efecto del tratamiento j sobre el factor B
3. El efecto de la interacción de las hipótesis anteriores

■ Diseño factorial 2x2

Este diseño tiene dos factores, cada uno con dos tratamientos

Ejemplo: El experimento investiga la comprensibilidad de un documento de diseño cuando se utiliza diseño estructurado o diseño orientado a objetos, basados en documento de requisitos “bueno” y otro” malo.

- El primer factor A, es el método de diseño.
- El segundo factor B, es el documento de diseño.
- Este diseño se llama 2x2 ya que ambos factores tienen dos tratamientos y todas las combinaciones de tratamientos son posibles.



Proceso Experimental. Etapa de Planificación

Tipos estándares de diseño (Montgomery, 1997)

■ Ejemplos de Diseño factorial 2x2

		Factor A	
		Treatment A1	Treatment A2
Factor B	Treatment B1	Subject 4, 6	Subject 1, 7
	Treatment B2	Subject 2, 3	Subject 5, 8

■ Ejemplo de balanced factorial design

Table 1. Experiment rounds

Round 1		Diagram Type	
		S	N
Domain	A	Group 1	Group 2
	B	Group 4	Group 3

Round 2		Diagram Type	
		S	N
Domain	A	Group 3	Group 4
	B	Group 2	Group 1

Independent and dependent variables. There are two independent variables in the experiment, the Diagram Type, with values: S (stereotyped) and N (nonstereotyped), and the Diagram Domain (A and B). By combining each level of the independent variables we obtain four treatments, as reflected by the four diagrams which are the objects of the experiment.

Genero et al, *Assessing the Influence of Stereotypes on the Comprehension of UML Sequence Diagrams: A Controlled Experiment*, MODELS 2008.

Proceso Experimental. Etapa de Planificación

Tipos estándares de diseño (Montgomery, 1997)

- Más de dos factores
- **Diseño factorial 2^k (2^k factorial design):** el caso del diseño 2×2 es un caso especial del 2^k , cuando $k=2$.

Este diseño **tiene k factores donde cada factor tiene dos tratamientos**. Esto significa que hay 2^k combinaciones posibles de tratamientos.

Para evaluar los efectos de los k factores, todas las combinaciones deben ser evaluadas.

Los sujetos son asignados de forma aleatoria a cada combinación

- **Diseño factorial fraccional 2^k (2^k fractional factorial design)**

Cuando el número de factores crece en un diseño factorial 2^k , el número de combinaciones de los factores crece, por ejemplo hay 8 combinaciones para un diseño factorial 2^3 , y 16 para un diseño factorial 2^4 .

A veces se considera que los efectos de ciertas interacciones no es importante, y entonces se puede realizar un diseño factorial fraccionado.



Proceso Experimental. Etapa de Planificación

Instrumentación

- Indicar los materiales que serán necesarios definir para llevar a cabo el experimento:
 - Objetos experimentales (cuántos y su estructura)
 - Guías e instrumentos de medición (e.j, cuestionarios)

Amenazas a la validez (*threats to validity*)

- Justificar las amenazas a la validez del experimento en función del diseño seleccionado:
 - Validez de constructo
 - Validez interna
 - Validez externa
 - Validez de conclusión



Diseño de un experimento controlado

PARTE 2: ANÁLISIS DE IMPLICACIONES



UNIVERSIDAD
POLITECNICA
DE VALENCIA

Análisis de implicaciones/impacto

El trabajo incluirá un apartado específico sobre las implicaciones/impacto de estas tecnologías en aspectos técnicos, jurídicos, comerciales, industriales, económicos, y/o sociales.

Ejemplo:

Cuestiones éticas: los procedimientos de investigación experimental en el ámbito de la Ingeniería del Software tienen un impacto en aspectos relativos a la **privacidad del sujeto, su contexto laboral y profesional** e incluso en aspectos individuales como la autoestima.

Existen diversos **códigos éticos** que pueden aplicarse, como pueden ser aquellos que hablan sobre el **consentimiento y voluntad de los sujetos**, el **anonimato de datos y participantes**, los riesgos y los daños que pudieran sufrir durante el experimento, etc.



Proceso Experimental. Etapa de Definición

Ejemplo: Experimento que compara dos técnicas de pruebas de software

Cuestiones económicas: La utilización de tecnologías y herramientas para realizar pruebas software conlleva un coste económico extra para los proyectos, en ocasiones elevado.

No obstante, los errores que en muchas ocasiones se dan en el software por dejar a un lado ciertos aspectos de la fase de pruebas y verificación de un proyecto, pueden llevar a pérdidas económicas aún mayores.

- En el artículo en **Wikipedia** [1] que recoge algunos de los bugs de software más famosos de la historia, muchos de los cuales tuvieron un **impacto económico importante para las empresas desarrolladoras y los stakeholders** de los proyectos.
- En otro artículo de la **NASA** [2], se hace un estudio profundo del coste (tanto económico como temporal) de corregir un fallo en el software, dependiendo del momento en el que se detecta.

[1] https://en.wikipedia.org/wiki/List_of_software_bugs

[2] <https://ntrs.nasa.gov/citations/20100036670>



UNIVERSIDAD
POLITECNICA
DE VALENCIA

Planificación del trabajo

- Hay varios ejemplos de diseños de experimentos en PoliformaT
- **Revisión del trabajo: 08 de octubre de 2024**
- Los grupos deberán trabajar de forma autónoma y solicitar tutorías si fuera necesario
- **Fecha de entrega: 21 de octubre de 2024**
- Entregar el trabajo mediante una tarea de PoliformaT

Entrega Diseño de Experimento Controlado



UNIVERSIDAD
POLITECNICA
DE VALENCIA