



**Universidad de Jaén**

# Máster Universitario en Enfermería de los Cuidados Críticos, Urgencias y Emergencias (ECCUE)

Metodología y Diseños de Investigación Cuantitativa en Enfermería de Cuidados Críticos, Urgencias y Emergencias

**M<sup>a</sup> Virtudes Alba Fernández**  
**Rosa María Fernández Alcalá**  
**Dpto. Estadística e I.O.**



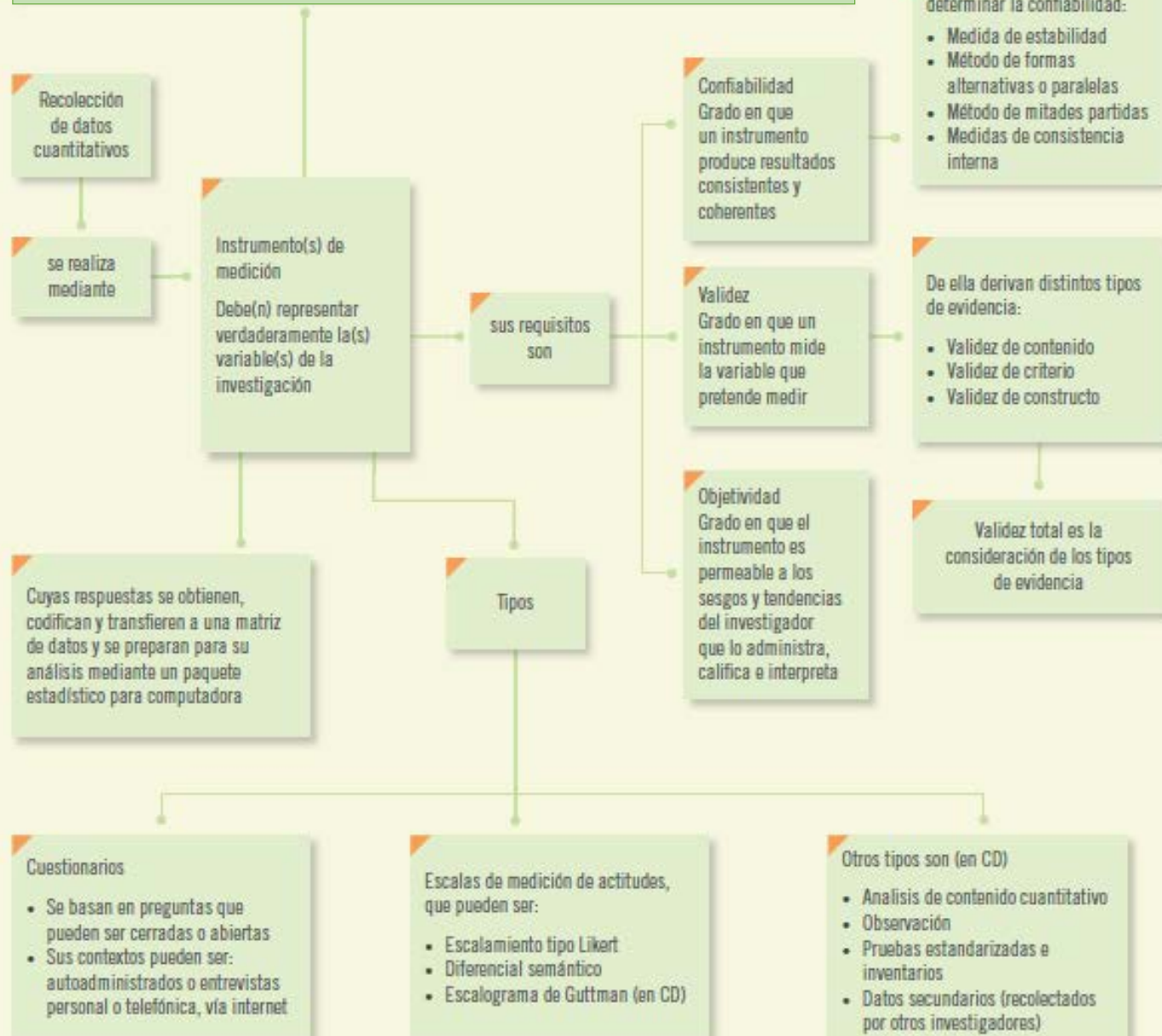
Universidad de Jaén

## **Tema 4: Diseño y análisis de la fiabilidad y validez de los instrumentos de medida.**

# Objetivo

Analizar los requisitos que un instrumento debe cubrir para recolectar apropiadamente datos cuantitativos: fiabilidad y validez.

# MAPA RESUMEN



# Motivación

¿Qué implica la etapa de recolección de datos?

**Recolectar los datos** implica elaborar un plan detallado de procedimientos que nos conduzcan a reunir datos con un propósito específico. Este plan incluye determinar:

- 1) ¿Cuáles son las fuentes de donde se obtendrán los datos? ¿Dónde se localizan tales fuentes?
- 2) ¿A través de qué medio o método vamos a recolectar los datos?
- 3) Una vez recolectados, ¿de qué forma vamos a prepararlos para que puedan analizarse y respondamos al planteamiento del problema?

El plan se nutre de diversos elementos:

- 1) Las variables, conceptos o atributos a medir
- 2) La muestra
- 3) Los recursos disponibles (de tiempo, apoyo institucional, económicos, etcétera).

# Motivación

## El proceso de medición

- Mediante el proceso de medición vinculamos conceptos abstractos con indicadores empíricos, clasificando los datos disponibles en términos del concepto que el investigador tiene en mente.
- En este proceso, el **instrumento de medición** o de recolección de datos tiene un papel central.
- Un **instrumento de medición adecuado** es aquel que registra datos observables que representan verdaderamente los conceptos o las variables que el investigador tiene en mente (Grinnell, Williams y Unrau, 2009).
- En toda investigación cuantitativa aplicamos un instrumento para medir las variables contenidas en las hipótesis (y cuando no hay hipótesis simplemente para medir las variables de interés).

# Motivación

¿Qué requisitos debe cubrir un instrumento de medición?

Toda medición o instrumento de recolección de datos debe reunir **dos requisitos esenciales: *fiabilidad* y *validez***.

## Fiabilidad

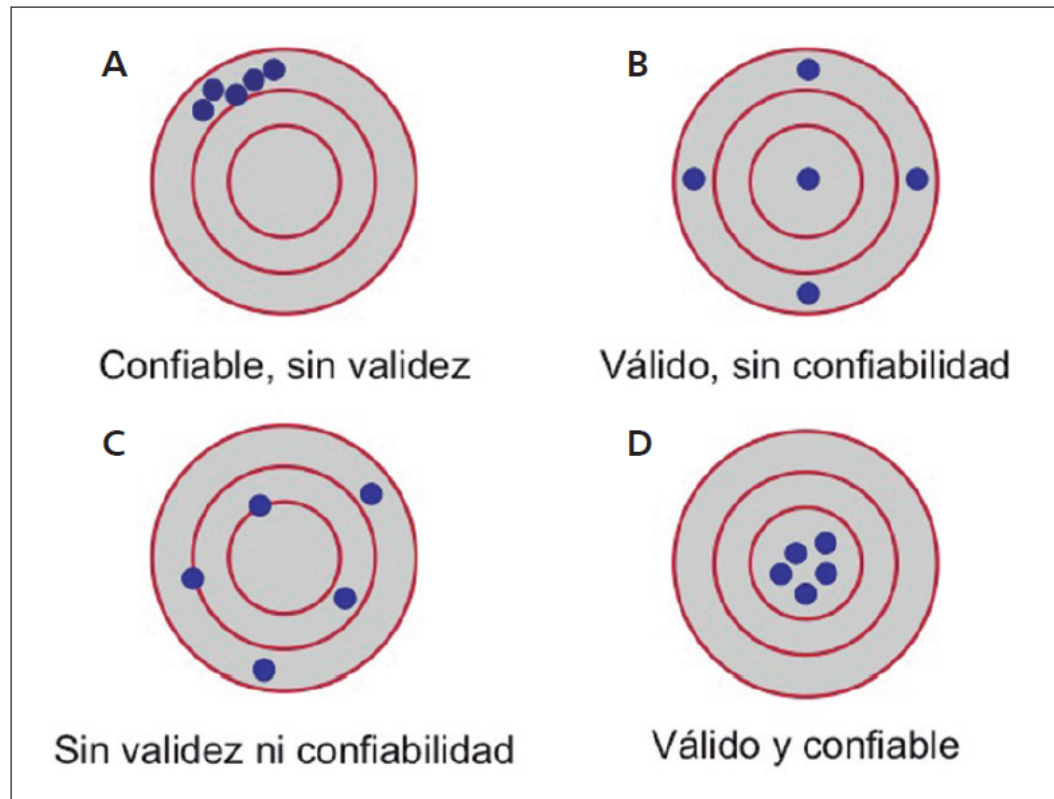
Mide el grado en el que las diferencias en los resultados son atribuibles al error aleatorio de medición o son atribuibles a diferencias en la característica/variable que se está midiendo.

## Validez

Grado en que un instrumento es capaz de medir aquello para lo que se ha diseñado

# Fiabilidad vs. Validez

El instrumento debe demostrar ser confiable y válido a la vez  
*Manterola et al. 2018*



**Figura 1.** Posibles escenarios de validez y confiabilidad. **A)** Todas las mediciones son parecidas, pero se alejan de la realidad. **B)** Las mediciones captan todo el espectro del fenómeno, pero son muy diferentes entre ellas. **C)** No se capta todo el fenómeno y las mediciones son muy disímiles entre ellas. **D)** Todas las mediciones son parecidas y se ajustan a la realidad de lo que se está midiendo.



## Factores que afectan a la Fiabilidad y a la validez

- **Improvisación** en la medición y selección del instrumento
- Utilización de **instrumentos extranjeros** sin previa adaptación y estandarización
- Instrumento **inadecuado** a las personas que se les aplica
- **Estilos personales** de los sujetos participantes
- **Condiciones** en las que se aplica el instrumento de medición
- **Falta de estandarización**
- **Aspectos mecánicos**

# ¿Cómo se sabe si un instrumento de medición es fiable y válido?

- **Fiabilidad**

- Estabilidad
- Equivalencia
- Consistencia interna

- **Validez**

- Validez de contenido
- Validez del constructo/concepto
- Validez de criterio

## Test-retest

- Un mismo instrumento de medición **se aplica dos (o más) veces a un mismo grupo de personas, después de cierto periodo.**
- Se mide la confiabilidad a través de la **correlación/adecuación** entre las puntuaciones de las distintas aplicaciones.
- Si la **correlación/adecuación** entre los resultados de las diferentes aplicaciones es **altamente positiva**, el instrumento se considera **confiable**.

### MEDIDA ESTADÍSTICA:

V. **Cuantitativa**-> Coeficiente de correlación de Pearson

$$r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

# Test-retest

**Ejemplo:** Sánchez  
Fernández et al. 2005

**TABLA 10. Medida de la presión arterial por 2 enfermeras diferentes**

Paciente	Enfermera A	Enfermera B	Diferencia (A-B)
1	140	132	8
2	150	124	26
3	150	144	6
4	113	112	1
5	128	124	4
6	182	161	21
7	143	138	5
8	180	170	10
9	134	128	6
10	161	152	9
Media	148,1	138,5	9,6
Desviación típica	21,77	18,22	7,85

## Formas paralelas o alternativas

- Se aplican dos (o más) versiones del mismo instrumento de medición a un mismo grupo de personas simultáneamente.
- Las versiones son similares en contenido, instrucciones, duración y otras características.
- El instrumento es **confiable** si la **correlación** entre los resultados de ambas versiones es **significativamente positiva**.

**MEDIDA ESTADÍSTICA:** Coeficiente de correlación de Pearson.

La **dificultad** radica en conseguir que dos instrumentos sean realmente “paralelos”, dada la dificultad que supone **realizar dos pruebas que midan exactamente lo mismo, pero con diferentes ítems..**

## Fiabilidad inter evaluador

Mide que tan similares son los puntajes asignados por diferentes evaluadores a un mismo fenómeno; para el caso en que los evaluadores entrevistan a las personas.

- Para **valores cualitativos**, Índice de Kappa y el porcentaje de acuerdo.
- Para **valores cuantitativos** se utiliza el coeficiente de correlación intra-clase

## Fiabilidad inter evaluador

- Para **valores cualitativos**, Índice de Kappa y el porcentaje de acuerdo.

		$\Gamma_1$	$\Gamma_2$	...	$\Gamma_k$	Total por filas
	$G_1$	$n_{11}$	$n_{12}$	...	$n_{1k}$	$n_{1+}$
	$G_2$	$n_{21}$	$n_{22}$	...	$n_{2k}$	$n_{2+}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	...
	$G_k$	$n_{k1}$	$n_{k2}$	...	$n_{kk}$	$n_{k+}$
Total por columnas		$n_{+1}$	$n_{+2}$		$n_{+k}$	$n$

$$PA = \frac{1}{n} \sum_{i=1}^k n_{ii}$$

$$k = \frac{\frac{1}{n} \sum_{i=1}^k n_{ii} - \frac{1}{n} \sum_{i=1}^k n_{+i} \frac{1}{n} \sum_{i=1}^k n_{i+}}{1 - \frac{1}{n} \sum_{i=1}^k n_{+i} \frac{1}{n} \sum_{i=1}^k n_{i+}}$$

**Ejemplo:** Sánchez  
Fernández et al. 2005

TABLA 4. Evaluación, por 2 enfermeras del tipo de piel de los pacientes

Enfermera B	Enfermera A			Total
	Normal	Seca	Grasa	
Normal	9	8	3	20
Seca	9	29	5	43
Grasa	0	3	14	17
Total	18	40	22	80

# Fiabilidad: Consistencia interna

## Coeficiente Alfa de Cronbach

- Mide la **consistencia interna de los ítems en cada escala y el instrumento en su conjunto**, indicando en que medida un conjunto de ítems se correlaciona entre sí.
- Aplicable en Escalas de Estimación, Escala de Lickert, cuestionario de preguntas cerradas con opciones policotómicas, test de aptitud verbal, test de aptitud no verbal , test psicológico.

$$\alpha = \frac{k}{k - 1} \left[ 1 - \frac{\sum_{i=1}^k S_i^2}{S_t^2} \right]$$

donde  $k$  es el número ítems  $S_i^2$  es la varianza del ítem  $i$  y  $S_t^2$  es la varianza total.

Interpretación:





# Fiabilidad: Consistencia interna

## Coeficiente Alfa de Cronbach

**Ejemplo:** Puntuaciones de 10 sujetos a un cuestionario de 8 items

i1	i2	i3	i4	i5	i6	i7	i8
1	1	1	2	3	1	1	1
5	4	5	4	4	3	2	1
3	3	4	4	3	4	3	2
5	5	5	4	4	4	5	3
2	1	2	2	3	1	1	1
4	3	4	4	3	5	5	3
1	2	1	2	1	2	2	1
5	5	5	4	5	5	4	2
3	3	1	2	3	3	1	1
5	5	5	5	4	4	3	1

# Fiabilidad: Consistencia interna

## Coeficiente de KUDER –RICHARDSON (KR-20)

- Mide la **consistencia interna del instrumento**, que proporciona la medida de todos los coeficientes de división por mitades para todas las posibles divisiones del instrumento en dos partes.
- Aplicable solo en instrumentos con ítems dicotómicos

$$r = \frac{k}{k-1} \left[ 1 - \frac{\sum_{i=1}^k p_i q_i}{S_t^2} \right]$$

donde  $p_i$  es la proporción de sujetos que cumple el ítem  $i$  y  $q=1-p$ .

Rangos	Magnitud
0,81 a 1,00	Muy Alta
0,61 a 0,80	Alta
0,41 a 0,60	Moderada
0,21 a 0,40	Baja
0,01 a 0,20	Muy Baja

Fuente: Tomado de Ruiz Bolívar (2002).

# Fiabilidad: Consistencia interna

## Split-halves o Mitades partidas

- Requiere una sola medición en un mismo grupo de sujetos.
- El conjunto total de ítems o reactivos **se divide en dos mitades equivalentes** y **se comparan las puntuaciones** o los resultados de ambas por medio de un **coeficiente de correlación**.
- Si el instrumento es confiable, las puntuaciones de las dos mitades deben estar muy correlacionadas.

### MEDIDAS ESTADÍSTICAS:

- Coeficiente de Spearman-Brown
- Coeficiente de Rulón
- Coeficiente de Guutman y Flanagan.

## Split-halves o Mitades partidas

- Coeficiente de **Spearman-Brown**,

$$r_{SB} = \frac{2r_{pi}}{1 + r_{pi}}$$

$r_{pi}$  coeficiente de correlación de Spearman entre las puntuaciones pares e impares.

- Coeficiente de **Rulón**,

$$r_R = 1 - \frac{S_d^2}{S_{p+i}^2}$$

$S_d^2$  varianza de la diferencia entre las puntuaciones pares e impares y  $S_{p+i}^2$  la varianza de la puntuación total.

- Coeficiente de **Gutman y Flanagan**,

$$r_{GF} = 2 \left[ 1 - \frac{S_p^2 + S_i^2}{S_{p+i}^2} \right]$$

$S_p^2, S_i^2$  varianzas de las puntuaciones pares e impares.

# Split-halves o Mitades partidas

Tabla 1. Resultados de los exámenes.

Alum	None	Pares	Totales	Calif.
1	23	25	48	96
2	22	23	45	90
3	24	24	48	96
4	21	22	43	86
5	25	23	48	96
6	25	24	49	98
7	21	18	39	78
8	24	25	49	98
9	25	24	49	98
10	21	20	41	82
11	24	22	46	92
12	25	23	48	96
13	20	21	41	82
14	22	23	45	90
15	17	18	35	70
16	22	21	43	86
17	17	13	30	60
18	22	20	42	84
19	24	23	47	94
20	25	23	48	96
21	25	23	48	96
22	24	24	48	96
23	25	25	50	100
			M in .	60
			M áx .	60
			P rom .	89.6
			S =Desv	10

Fuente: Tabla 1 de Datos Básicos del Apéndice

**Ejemplo:** Puntuaciones examen. Castillo et al.2006

# Fiabilidad

## Interpretación coeficientes

- No hay una regla que indique: a partir de este valor no hay fiabilidad del instrumento.
- Podemos decir —de manera más o menos general— que si obtengo 0.25 en la correlación o coeficiente, esto indica baja confiabilidad; si el resultado es 0.50, la fiabilidad es media o regular. En cambio, si supera el 0.75 es aceptable, y si es mayor a 0.90 es elevada, para tomar muy en cuenta.

## Validez de contenido

**Se refiere al grado en que un instrumento refleja un dominio específico de contenido de lo que se mide.**

- Un instrumento de medición requiere tener representados a todos o la mayoría de los componentes del dominio de contenido de las variables a medir.
- El dominio de contenido está definido o establecido por la literatura (teoría y estudios antecedentes)
- En cada estudio se debe probar que el instrumento utilizado es válido (revisión bibliográfica, meta-análisis, etc.).

**¿el instrumento mide adecuadamente las principales dimensiones de la variable en cuestión?**

## Evidencia de la validez de contenido: juicio de expertos

- Expertos en la materia
- Crean/mejoran el cuestionario para que se ajuste a los objetivos del estudio (claridad de preguntas, relevancia, número, etc)
- Fase piloto y definitiva



## Validez de expertos

**Se refiere al grado en que aparentemente un instrumento de medición mide la variable en cuestión, de acuerdo con expertos en el tema.**

- Vinculada a la validez de contenido.
- Se establece mediante la evaluación del instrumento ante expertos.

# Evidencia de la validez de contenido

## Técnica: Juicio de expertos

Ejemplo: Alonso Jiménez et al. 2013

**Tabla II. Resultados opinión expertos**

Cuestionario sobre satisfacción con los cuidados de enfermería del trabajo en los servicios de prevención

VALORACIÓN DEL CUESTIONARIO	Experto 1	Experto 2	Experto 3	Experto 4	Experto 5	Total	%
1 Es conveniente valorar la percepción que tiene el trabajador sobre los cuidados de la enfermería del trabajo	4	5	5	5	5	24	96
2 Las preguntas son sencillas y claras	5	4	4	4	5	22	88
3 Los datos figuran bien ordenados, está bien estructurada	4	4	4	4	5	21	84
4 Las preguntas abarcan todos los aspectos de la actividad de enfermería del trabajo	4	3	4	5	5	21	84
5 El trabajador podrá comprender las preguntas del cuestionario	5	4	4	5	5	23	92
6 El sistema de puntuación de las respuestas es adecuado	4	5	4	5	5	23	92
7 El cuestionario es fácil de cumplimentar	5	4	4	5	5	23	92
8 Puede ser un elemento más de ayuda a la enfermería del trabajo en la actividad diaria	3	5	5	4	5	22	88

## Validez de criterio

**Establece la validez de un instrumento de medición al compararlo con algún criterio externo que pretende medir lo mismo.**

- Se debe comparar con un estándar conocido o con él mismo. Si existe una alta correlación, la validez de criterio es alta
  - **Validez concurrente:** se compara la prueba con una medida establecida
  - **Validez predictiva:** se prueba durante un período de tiempo
- **Principio de la validez de criterio:** si diferentes instrumentos o criterios miden el mismo concepto/variable, deben arrojar resultados similares.

**¿en qué grado el instrumento comparado con otros criterios externos mide lo mismo?**

## Evidencia de la validez de criterio

### Técnica: Coeficiente de correlación

#### Coeficiente de Correlación entre X e Y:

Spearman: 0,83

Pearson: 0,87

La **PAA** es un instrumento  
válido en cuanto a criterio

Alumno	PAA (X)	L (Y)
Egledis	650	75
Ronny	710	87
Frank	682	85
Victor	700	83
Vanesa	691	80
Liévana	705	82
Ruben	600	81
Maybeth	690	90
Julio	709	90
Marbelis	715	78

## Validez de constructo

**¿el concepto teórico está realmente reflejado en el instrumento?**

**¿qué significan las puntuaciones del instrumento?**

**¿el instrumento mide el constructo y sus dimensiones?**

**¿por qué sí o por qué no?**

**¿cómo opera el instrumento?**

## Validez de constructo

**Se refiere a qué tan exitosamente un instrumento representa y mide un concepto teórico.**

- **Constructo:** variable medida que tiene lugar dentro de una hipótesis, teoría o un esquema teórico.
- La validez de constructo incluye tres etapas:
  1. Se establece y especifica la relación teórica entre los conceptos.
  2. Se correlacionan los conceptos y se analiza cuidadosamente la correlación.
  3. Se interpreta la evidencia empírica de acuerdo con el nivel en el que clarifica la validez de constructo de una medición en particular.
- El proceso de validación de un constructo está vinculado con la teoría.

## Evidencia de la validez de constructo

### Técnica estadística: Análisis factorial

Analiza las interrelaciones existentes entre las variables para explicarlas extrayendo los denominados “factores”. Dichos factores subyacentes explicarían diferentes dimensiones del constructo.

**Ejemplo:** Puntuaciones de 10 sujetos a un cuestionario de 8 items

i1	i2	i3	i4	i5	i6	i7	i8
1	1	1	2	3	1	1	1
5	4	5	4	4	3	2	1
3	3	4	4	3	4	3	2
5	5	5	4	4	4	5	3
2	1	2	2	3	1	1	1
4	3	4	4	3	5	5	3
1	2	1	2	1	2	2	1
5	5	5	4	5	5	4	2
3	3	1	2	3	3	1	1
5	5	5	5	4	4	3	1