

MDS5

Carmen Beltran Chably

2025-12-02

Ejercicio 5.1: Realiza el análisis anterior y comenta los resultados obtenidos

Kinshipdelta solución 2D

```
library(smacof)

data(kinshipdelta)
res = smacofSym(kinshipdelta)
res

## 
## Call:
## smacofSym(delta = kinshipdelta)
##
## Model: Symmetric SMACOF
## Number of objects: 15
## Stress-1 value: 0.264
## Number of iterations: 91

summary(res)

## 
## Configurations:
##          D1      D2
## Aunt     -0.3081  0.6436
## Brother   0.4239 -0.5302
## Cousin    0.2182  0.8434
## Daughter  -0.3914 -0.3705
## Father     0.1565 -0.6829
## Granddaughter -0.5192  0.1394
## Grandfather   0.7060 -0.1227
## Grandmother   -0.6987  0.1592
## Grandson     0.5207 -0.0935
## Mother      -0.4364 -0.5686
## Nephew       0.3822  0.4104
## Niece        -0.2024  0.5361
## Sister       -0.6011 -0.3124
## Son          0.2027 -0.4987
## Uncle        0.5473  0.4474
```

```

## 
## 
## Stress per point (in %):
##          Aunt      Brother     Cousin    Daughter    Father
##        6.24       7.49      6.20      4.04      4.88
## Granddaughter Grandfather Grandmother Grandson   Mother
##        8.50      11.17     11.25      8.55      4.86
##      Nephew      Niece     Sister      Son      Uncle
##        4.48       4.26      7.44      4.26      6.40

```

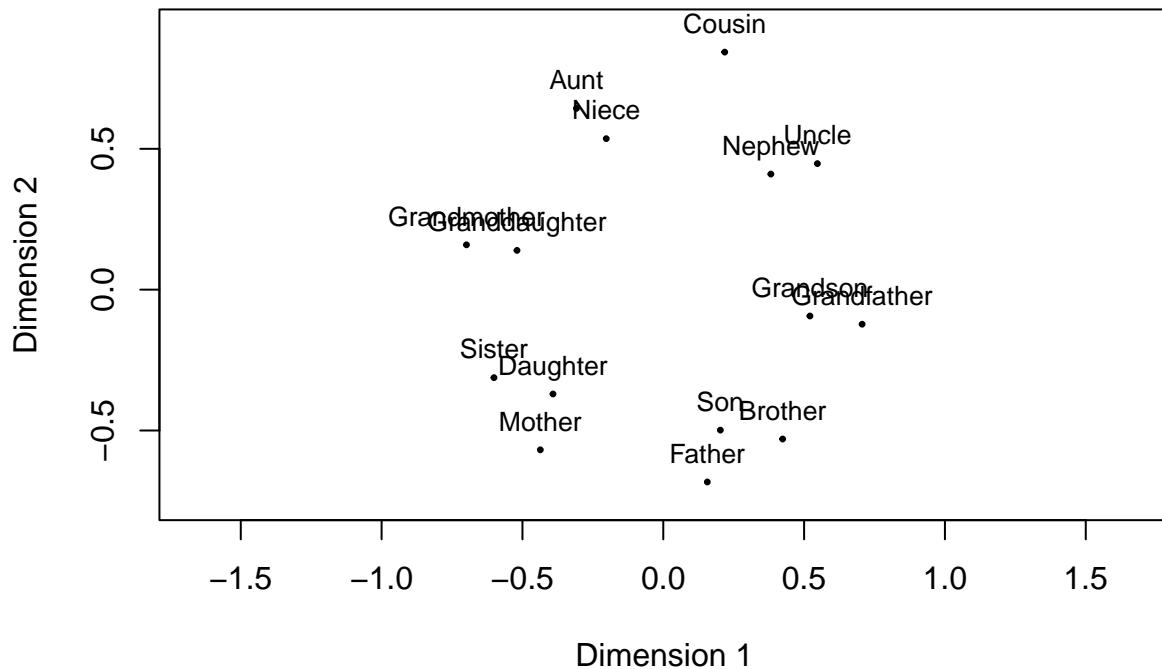
En el dataset kinshipdelta hay 15 relaciones de familiares que están representadas con disimilaridades. El modelo MDS se ejecutó sobre los datos de kinshipdelta utilizando la función smacofSym() en su configuración por defecto (MDS Métrico o Ratio, 2 Dimensiones). Vemos que el Stress-1 es de 0.264 por lo que podemos decir que según la escala de Kruskal es un ajuste pobre. Por lo que podemos decir que no encaja bien en un espacio 2D métrico.

Por otro lado si vemos la tabla de configuraciones vemos las dos dimensiones. La primera dimensión podríamos decir como hipótesis que representa la línea generacional, por tanto, diríamos que los valores positivos se refieren a los familiares más lejanos del ego en el árbol genealógico como por ejemplo los abuelos y nietos, mientras que los valores negativos representan generaciones cercanas como los padres, los hermanos o los hijos. La dimensión dos podríamos decir que se refiere a la distancia familiar, es decir, los valores positivo se refieren a los algo más alejados, como por ejemplo un tío o un primo y los valores negativos a los familiares más cercanos como un parente o un hijo.

Si miramos la tabla del Stress per point podemos ver cuales son las relaciones familiares que están peor representadas, las cuales serán el que mayor stress tengan. En este caso vemos que la abuela tiene un 11.25% el abuelo también con un 11.17% o el nieto con un 8.5%. Por tanto podemos ver que los extremos generacionales son difíciles de analizar en un modelo en dos dimensiones y creemos que lo mejor es utilizar una solución en 3D.

```
plot(res, plot.type = "conf", main = "Configuración Kinship (D1: Generación, D2: Distancia)")
```

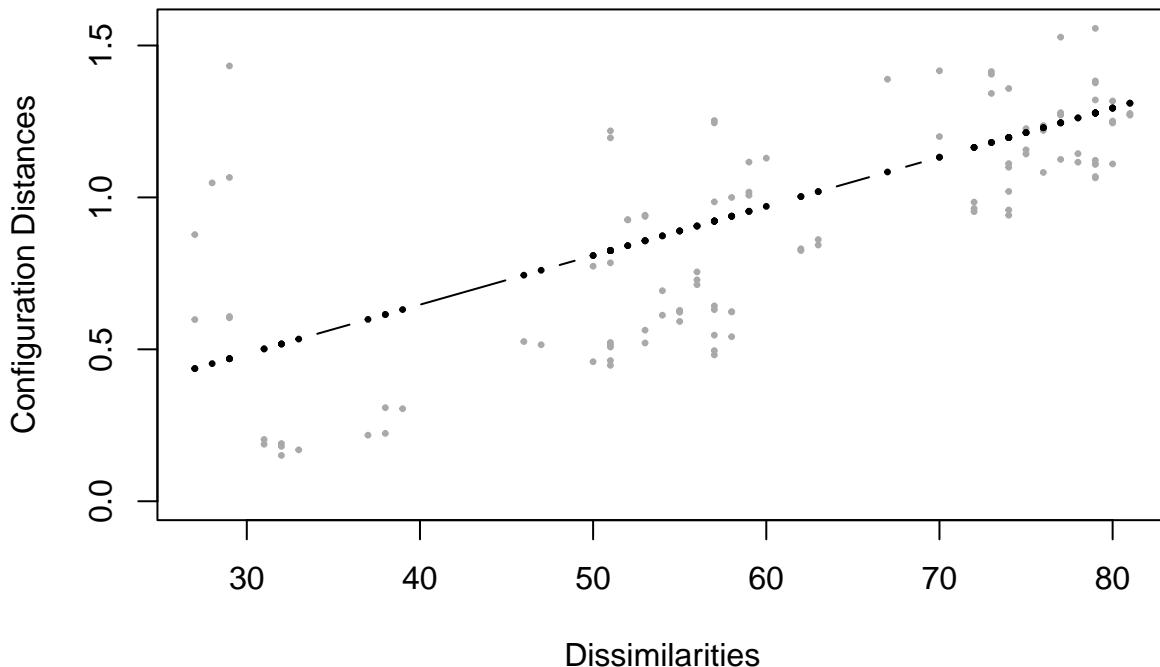
Configuración Kinship (D1: Generación, D2: Distancia)



En este gráfico podemos ver una nube de puntos la cual se desvía considerablemente de la curva monótona ideal. Podemos ver que hay una dispersión vertical la cual nos indica un error de ajuste alto, lo cual tiene sentido ya que el valor de STRESS del modelo es de 0.264. Por tanto, comolos datos no se ajustan bien al modelo métrico estricto, creemos que la idea de una solución con mayor dimensionalidad o un MDS no métrico sería más apropiado para representar la estructura de parentesco.

```
plot(res, plot.type = "Shepard", main = "Shepard Plot Kinship")
```

Shepard Plot Kinship



En este gráfico Shepard Plot del modelo Kinship vemos una nube de puntos que se desvía considerablemente de la curva monótona ideal. Esta dispersión vertical indica que hay un error del ajuste alto, lo cual tiene sentido ya que el valor del STRESS del modelo es de 0.264. Además al haber una variación vertical tan considerable podemos decir que las distancias del mapa 2D no reproducen bien las disimilaridades originales. Con todo esto podemos decir que los datos no se ajustan bien a un modelo métrico estricto, y creemos que la idea de una solución con mayor dimensionalidad o un MDS no métrico sería más apropiado para representarlo.

Solución en tres dimensiones.

```
res_kinship_3D <- smacofSym(kinshipdelta, ndim = 3, type="ordinal", ties="secondary")  
res_kinship_3D
```

```
##  
## Call:  
## smacofSym(delta = kinshipdelta, ndim = 3, type = "ordinal", ties = "secondary")  
##  
## Model: Symmetric SMACOF  
## Number of objects: 15  
## Stress-1 value: 0.136  
## Number of iterations: 89
```

```
summary(res_kinship_3D)
```

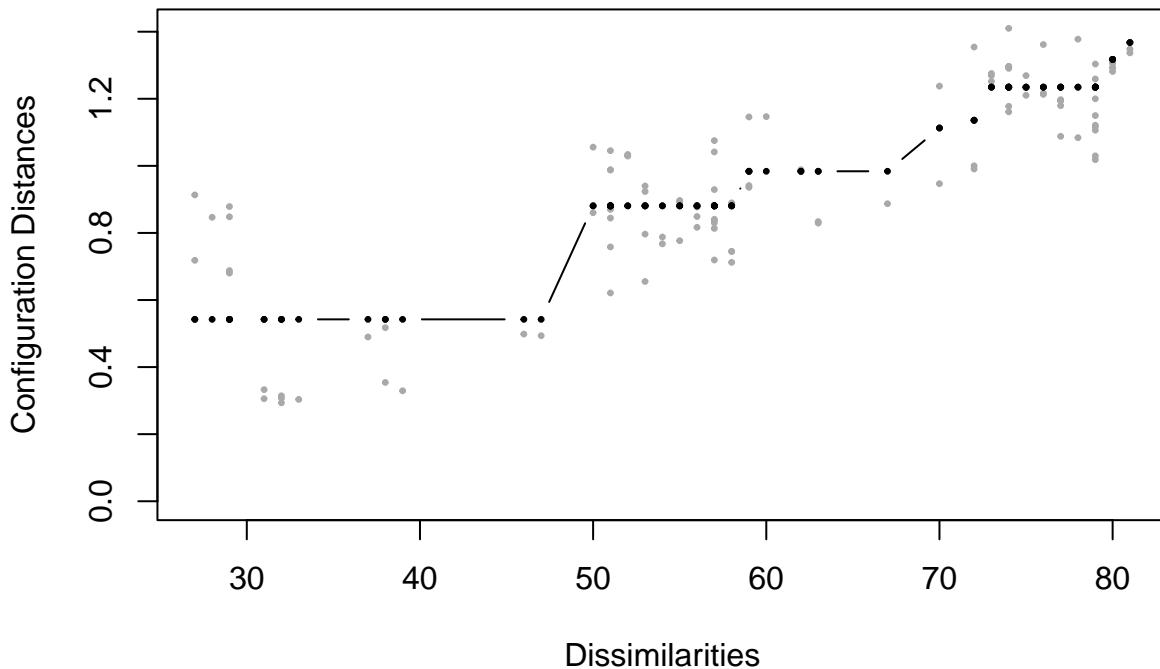
```
##  
## Configurations:  
##          D1      D2      D3  
## Aunt      -0.3338  0.5739  0.0776  
## Brother    0.3766 -0.1959  0.4645  
## Cousin     0.1822  0.6696  0.4700  
## Daughter   -0.5083 -0.3511  0.0664  
## Father     0.2303 -0.5047  0.3708  
## Granddaughter -0.4938 -0.2016 -0.4088  
## Grandfather   0.3212 -0.1502 -0.6084  
## Grandmother   -0.3520 -0.0099 -0.6030  
## Grandson     0.3705 -0.3590 -0.3906  
## Mother      -0.4335 -0.3574  0.3906  
## Nephew       0.4588  0.4597 -0.2041  
## Niece        -0.2442  0.6073 -0.1999  
## Sister       -0.4533 -0.0314  0.4334  
## Son          0.3212 -0.5274  0.0708  
## Uncle        0.5583  0.3780  0.0709  
##  
##  
## Stress per point (in %):  
##          Aunt      Brother     Cousin      Daughter      Father  
##          8.09      6.29       7.78       6.03       6.28  
## Granddaughter Grandfather Grandmother Grandson      Mother  
##          6.41      6.72       7.07       7.03       5.78  
##          Nephew     Niece      Sister      Son        Uncle  
##          5.96      5.72       6.71       6.05       8.09
```

Al incrementar una dimensión podemos ver que el modelo a mejorado de tal forma que ahora el stress-1 es 0.136 por lo que vemos que mejora en relación con el modelo 2D. Por tanto esto refuerza la idea de que el modelo 2D no era suficiente para representarlo. La solución ordinal es más flexible que la métrica ya que respeta solo el orden de disimilaridades, no las distancias exactas.

Mientras que la solución métrica de dos dimensiones tuvimos un STRESS algo alto (0.264), en este modelo no métrico en tres dimensiones vemos como ha disminuido considerablemente el error de ajuste. Esto indica que la estructura de parentesco no puede representarse adecuadamente en un plano, y que una tercera dimensión permite capturar relaciones que el espacio 2D distorsionaba. Además, el uso de un modelo ordinal es más apropiado conceptualmente, ya que las relaciones de parentesco son cualitativas y se basan en niveles de cercanía más que en distancias métricas estrictas. Por tanto, el modelo 3D proporciona una representación más fiel y coherente del sistema de parentesco.

```
plot(res_kinship_3D, plot.type = "Shepard", main = "Shepard Plot - Kinship 3D No Métrico")
```

Shepard Plot – Kinship 3D No Métrico



El Shepard Plot correspondiente al modelo MDS no métrico en tres dimensiones para los datos de parentesco muestra una mejora clara en el ajuste respecto al modelo métrico en dos dimensiones. En este gráfico, los puntos que representan las disimilaridades originales se distribuyen mucho más cerca de la curva monótona de ajuste, con una dispersión vertical considerablemente menor que en el Shepard del modelo 2D. Podemos ver que la línea de ajuste no es recta sino una curva monótona creciente. Esto se debe a que el MDS no métrico no impone una relación lineal entre las disimilitudes y las distancias, sino que ajusta una función que solo preserva el orden de las disimilitudes.

Trading solución no métrica 3D

```
data(trading)
res_trading = smacofSym(trading, ndim = 3, type="ordinal", ties = "secondary")
print(paste("STRESS-1 (Trading, 3D):", res_trading$stress))
```

```
## [1] "STRESS-1 (Trading, 3D): 0.106298812485902"
```

En primer lugar este dataset contiene una matriz de disimilitudes entre países basada en sus relaciones comerciales, Mide cuán distintos son los países entre sí en términos de comercio.

En primer lugar, podemos ver que le stress es 0.106 lo cual se traduce a que es un buen ajuste.

```
round(res_trading$conf, 3)
```

	D1	D2	D3
--	----	----	----

```

## Arge  0.455  0.193  0.286
## Aust  0.470  0.179 -0.112
## Braz  0.532  0.048  0.084
## Cana  0.535  0.292 -0.366
## Chin -0.044 -0.111 -0.821
## Czec -0.855  0.047 -0.279
## Egyp  0.069  0.375  0.303
## E.Ge -0.728  0.192 -0.012
## Fran  0.266  0.276  0.403
## Hung -0.632  0.424  0.067
## Indi  0.226  0.344 -0.063
## Ital  -0.027 -0.291  0.736
## Japa  0.214 -0.744 -0.252
## N.Ze  0.459  0.103 -0.434
## Pola -0.632  0.424  0.067
## Swed  0.338  0.454  0.095
## USA   0.169 -0.604  0.060
## USSR -0.781 -0.389 -0.089
## U.K.  0.140 -0.609  0.249
## W.Ge -0.175 -0.604  0.077

```

La matriz de coordenadas tridimensionales proporciona la posición de cada país en el espacio perceptual generado por el modelo. Podemos darle a cada dimensión un valor para que se entienda mejor. Por ejemplo, la dimensión uno puede hacer referencia al nivel de desarrollo socioeconómico, por lo que los valores positivos harán referencia a economías más desarrolladas como por ejemplo Francia o Suecia y los valores negativos a antiguas economías planificadas o emergentes como podría ser Hungría o Checoslovaquia.

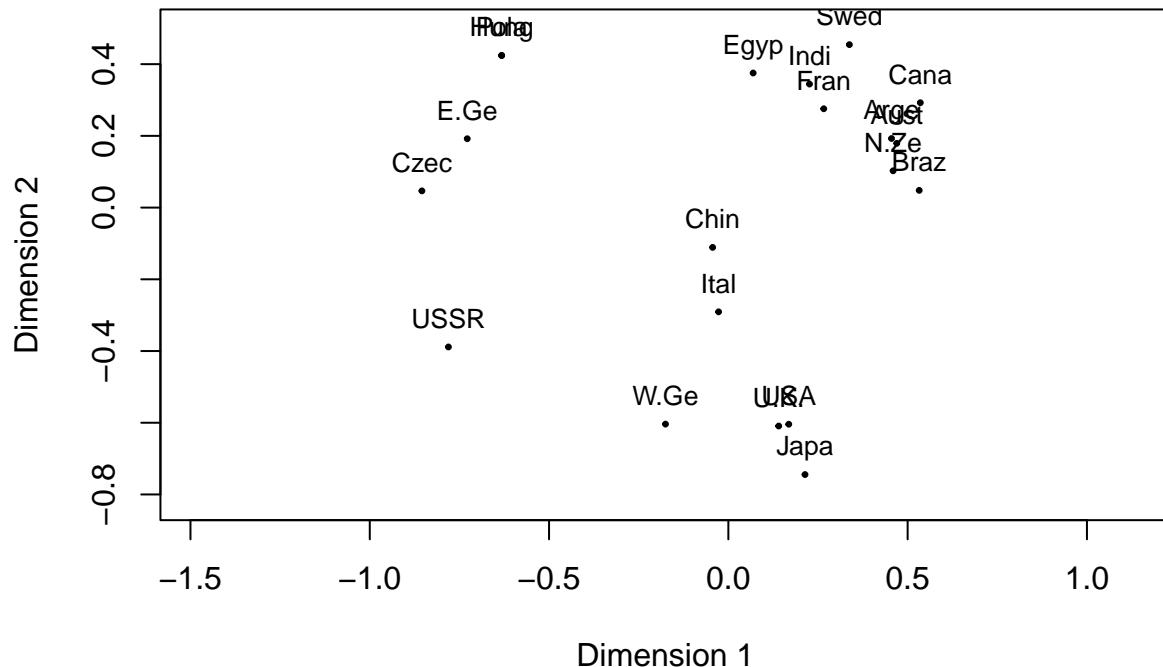
La dimensión 2 parece separar países según bloques geopolíticos o zonas de intercambio predominante, pero la solución 2D es una proyección de un espacio 3D, por lo que la interpretación debe hacerse con cautela.

La dimensión 3 puede representar el tipo de economía de tal forma que los valores positivos serán los países con industrias avanzadas como podría ser Italia o Francia y los valores negativos las economías más centradas en materias primas, en las que tengan un sector primario fuerte como Egipto o Nueva Zelanda.

El análisis MDS no métrico en tres dimensiones proporciona un ajuste muy bueno a los datos de comercio internacional ya que tiene un STRESS bastante bajo, de 0.106. Esta solución revela agrupamientos significativos como el bloque occidental, las economías socialistas y las naciones asiáticas, demostrando que la estructura comercial global puede representarse de forma clara y estable en un espacio de tres dimensiones.

```
plot(res_trading, plot.type = "conf", main = "Configuración Trading 3D (D1 vs D2)")
```

Configuración Trading 3D (D1 vs D2)



En este gráfico podemos ver que los países con más desarrollo económico como Canadá, Australia, Suecia o Francia están en la parte derecha de la dimensión 1 y en la parte izquierda están los países con economías planificadas como URSS, Polonia o Hungría.

En la dimensión dos vemos que los valores positivos, están en la parte superior del gráfico como Hungría, Polonia o Egipto y en la parte bajo los negativos como USA, UK o Nueva Zelanda siendo la zona de Asia-Pacífico. Aunque la solución es en 3D creemos que es pertinente usar el gráfico en 2D porque las dos primeras dimensiones suelen capturar la mayor parte de la variabilidad estructural.

Ejercicio 5.2: Realiza el análisis anterior y comenta los resultados obtenidos.

A lo largo de este ejercicio vamos a estimar tres modelos derivados del análisis INDSCAL aplicados a los datos perception: El primero será el modelo INDSCAL estándar el cual no tiene restricciones adicionales. El modelo Identity que tiene restricción ya que parte de la base de que todos los sujetos comparten los mismos pesos en todas las dimensiones. Y por último el modelo Idioscal que no tiene restricción ya que este modelo parte de la base de que cada sujeto puede tener pesos completamente libres y diferentes.

```
library(smacof)
data(perception)
# 1. INDSCAL (Estándar/Default) - El resultado se guarda en 'res'
res <- smacofIndDiff(perception)
res

##
## Call: smacofIndDiff(delta = perception)
##
```

```

## Model: Three-way SMACOF
## Number of objects: 16
## Stress-1 value: 0.166
## Number of iterations: 114

summary(res) # Muestra el STRESS y el resumen del ajuste

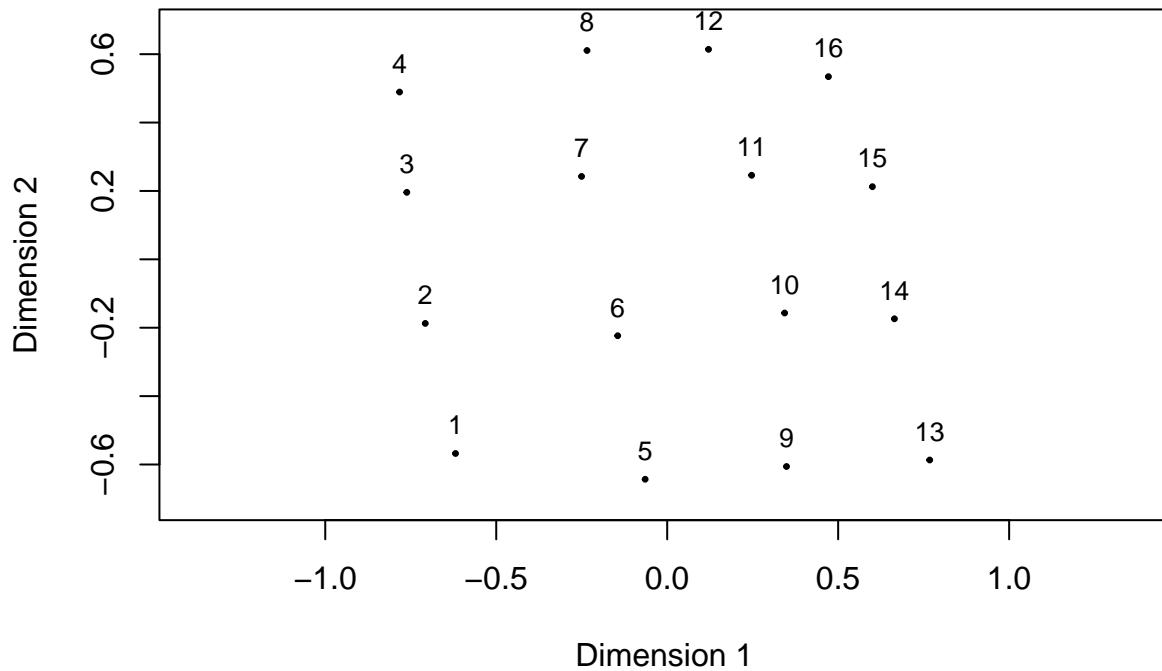
## 
## Group Stimulus Space (Joint Configurations):
##      D1      D2
## 1 -0.6189 -0.5678
## 2 -0.7075 -0.1874
## 3 -0.7614  0.1960
## 4 -0.7826  0.4892
## 5 -0.0644 -0.6431
## 6 -0.1447 -0.2234
## 7 -0.2503  0.2425
## 8 -0.2344  0.6107
## 9  0.3490 -0.6058
## 10 0.3435 -0.1570
## 11 0.2470  0.2459
## 12 0.1209  0.6141
## 13 0.7678 -0.5868
## 14 0.6645 -0.1738
## 15 0.6002  0.2125
## 16 0.4716  0.5343
##
##
## Stress per point:
##    1    2    3    4    5    6    7    8    9    10   11   12   13   14   15   16
## 7.03 5.03 5.55 8.24 6.89 6.43 7.32 7.40 5.49 4.87 4.84 6.15 9.31 5.08 5.52 4.86

```

Este es el modelo principal y representa un espacio común de percepción entre los estímulos y calcula diferencias individuales mediante pesos. El stress-1 obtenido es de 0.166 por lo que podemos decir que es un modelo aceptable y reproduce bien las diferencias percentuales. No es perfecto, pero sí lo suficientemente bueno para interpretar la estructura del espacio común.

```
plot(res, plot.type = "confplot", main = "Espacio Común INDSCAL (D1 vs D2) - PERCEPCIÓN")
```

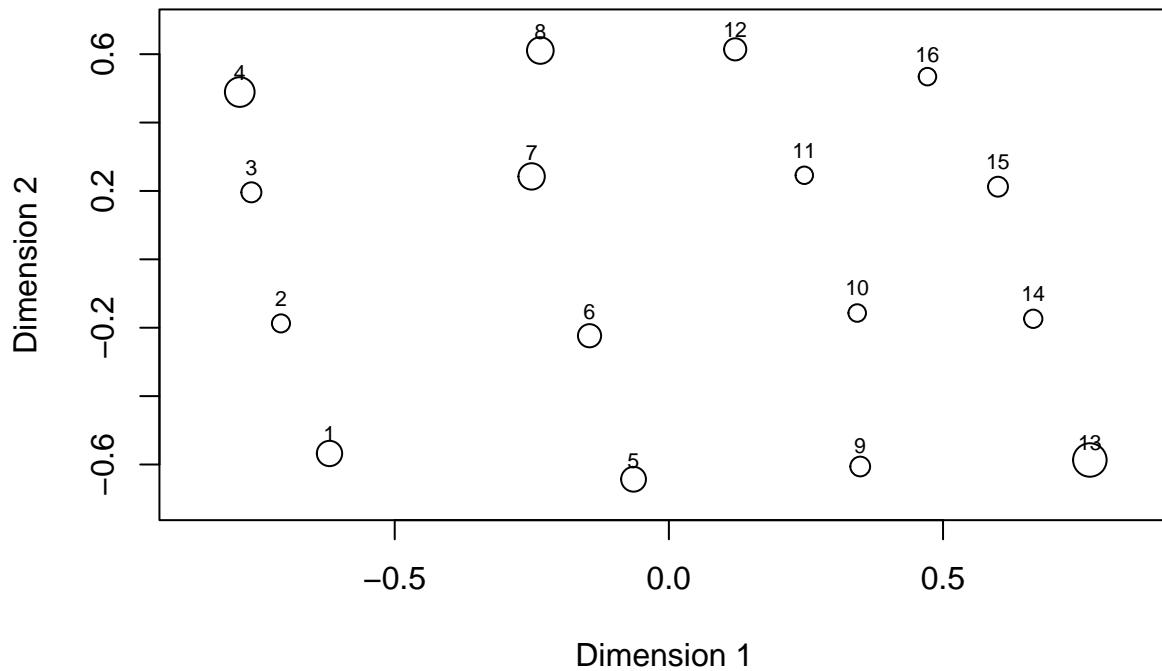
Espacio Común INDSCAL (D1 vs D2) – PERCEPCIÓN



Las distancias entre puntos representan cuán distintos los perciben los sujetos. De la primera dimensión podemos decir que representa diferencias sobre la intensidad o la complejidad donde los valores negativos son los estímulos más fáciles o menos cargados y los valores positivos los que más. En la dimensión dos podemos hablar de lo agradables o no que sean por ejemplo los valores positivos serán los estímulos más agradables en cambio los valores negativos son los estímulos menos agradables.

```
plot(res, plot.type = "bubbleplot", main = "Pesos Ponderados (D1 vs D2) – PERCEPCIÓN")
```

Pesos Ponderados (D1 vs D2) – PERCEPCIÓN



Este gráfico podemos ver cómo cada sujeto pondera las dos dimensiones, es decir, cómo usa cada dimensión para percibir los estímulos. El bubbleplot del modelo INDSCAL muestra diferencias individuales claras en la forma en que los sujetos utilizan las dos dimensiones perceptuales. Los vectores horizontales son sujetos que dependen más de la dimensión 1 en cambio los vectores más verticales son sujetos que dependen más de la dimensión dos. Los vectores 45° son sujetos que utilizan ambas dimensiones por igual. La dirección de los vectores revela tres perfiles: sujetos dominados por la dimensión de complejidad, sujetos que dependen de la agradabilidad y sujetos equilibrados que combinan ambas. La longitud de los vectores indica consistencia individual, siendo más largos para sujetos con juicios más estables. Esta estructura confirma que los participantes no utilizan las dimensiones de la misma manera, lo que justifica la mejora en el ajuste del modelo INDSCAL frente al modelo con restricción Identity, y demuestra la utilidad del modelo para capturar diferencias individuales reales.

```
# 2. Identity (Máx. Restricción)
res.id <- smacofIndDiff(perception, constraint = "identity")
```

El modelo Identity parte de la idea de que todas las personas usen exactamente las mismas dimensiones con la misma escala sin permitir que haya diferencias individuales en los pesos.

El stress es de 0.206, siendo más alto que el modelo anterior, por lo que podemos decir que es peor ajuste. Podríamos decir por tanto que los sujetos no están usando las dimensiones de la misma forma y por tanto por tanto no es adecuado asumir que todos los sujetos usan los mismos pesos dimensionales para explicar mejor el modelo ya que la restricción de identity es demasiado rígida y no representa bien las diferencias individuales

```
# 3. Idioscal (Mín. Restricción)
res.idio <- smacofIndDiff(perception, constraint = "idioscal")
```

Este modelo es completamente opuesto al identity ya que cada individuo puede deformar las dimensiones libremente, ya que se permiten diferencias individuales por tanto es el modelo menos restringido.

El stress que hemos obtenido es de 0.166. Es verdad que tiene un mejor ajuste que el modelo de identidad pero la mejora con INDSCAL es casi nula. Por tanto creemos que aunque es el más flexible, no mejora realmente el ajuste, así que no compensa usarlo frente al modelo INDSCAL normal.

```
# Comparación de STRESS
print("Valores de STRESS por Modelo de Restricción:")

## [1] "Valores de STRESS por Modelo de Restricción:"

print(paste("STRESS Identity (Máx. Restricción):", res.id$stress))

## [1] "STRESS Identity (Máx. Restricción): 0.206054290001244"

print(paste("STRESS INDSCAL (Estándar):", res$stress))

## [1] "STRESS INDSCAL (Estándar): 0.166304352792473"

print(paste("STRESS Idioscal (Mín. Restricción):", res.idio$stress))

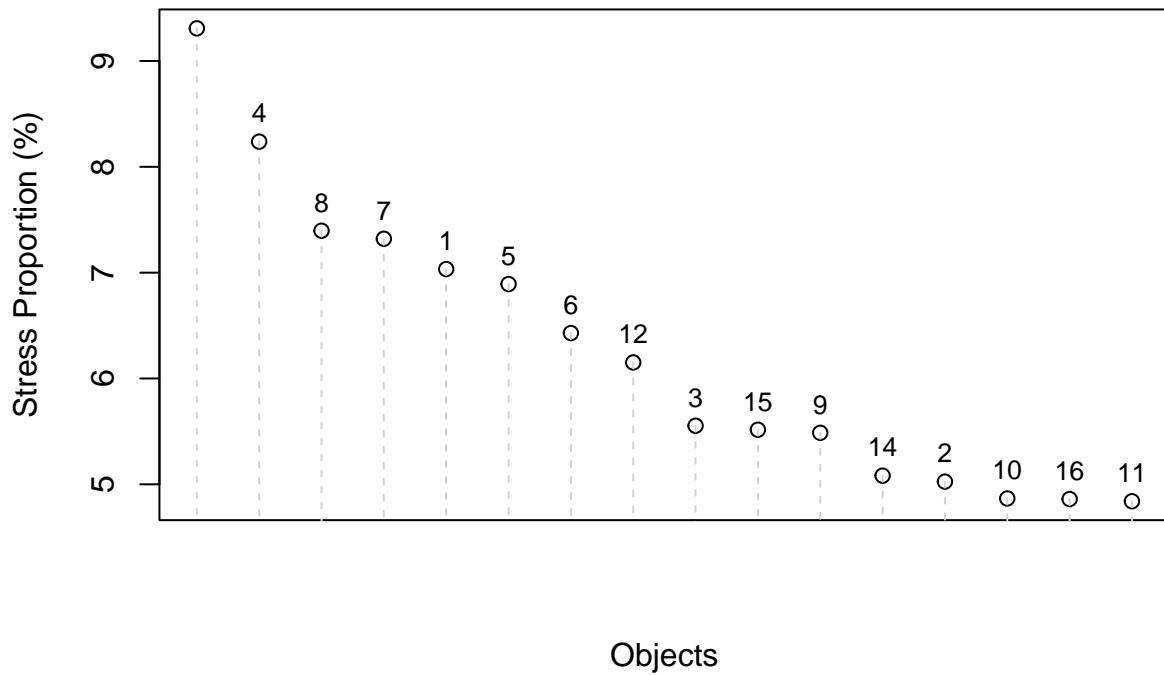
## [1] "STRESS Idioscal (Mín. Restricción): 0.166304285318029"
```

Podemos observar por tanto que el STRESS disminuye claramente cuando se permite que los sujetos ponderen las dimensiones de manera diferente. El modelo Identity, al imponer la misma estructura perceptual para todos, ofrece el peor ajuste. El modelo INDSCAL estándar mejora significativamente el STRESS al incorporar diferencias individuales mediante pesos.

El modelo Idioscal, que es el menos restrictivo, no muestra mejoras adicionales, por lo que el modelo óptimo es el INDSCAL estándar: equilibrado, interpretable y con el mejor compromiso entre ajuste y complejidad. Dado que el modelo INDSCAL y el modelo Idioscal tienen un STRESS virtualmente idéntico, se prefiere INDSCAL por ser más parsimonioso. Ofrece un excelente ajuste y, además, proporciona los pesos individuales, que son interpretables.

```
plot(res, plot.type = "stressplot", main = "Curva de Convergencia - INDSCAL")
```

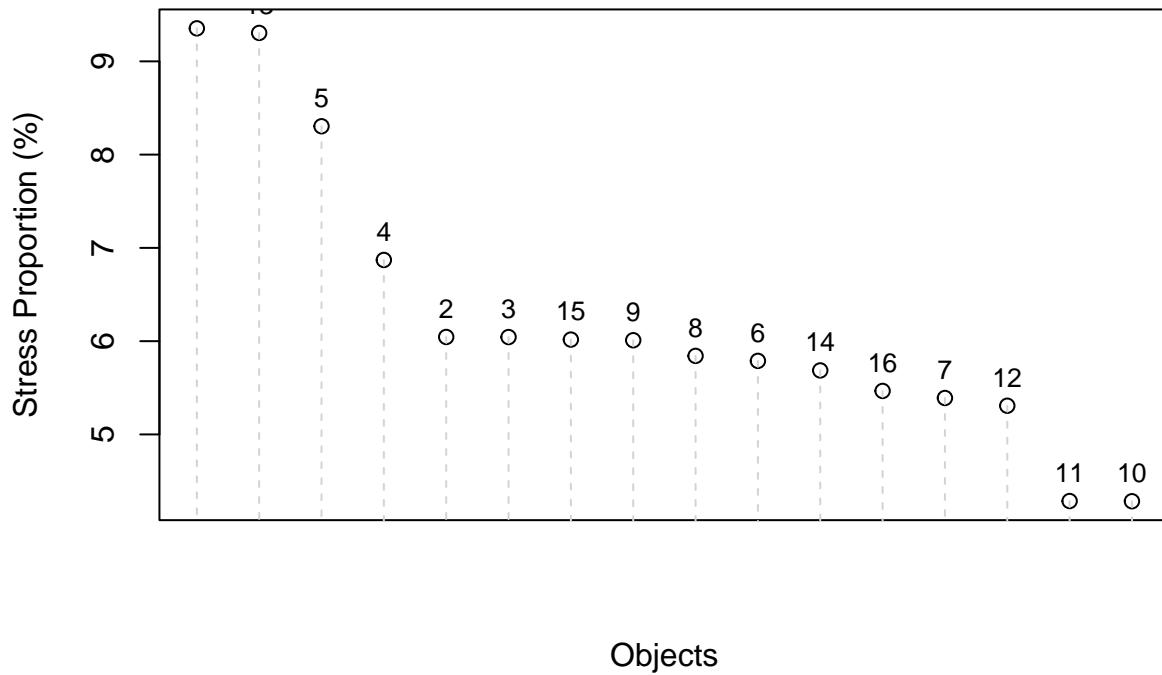
Curva de Convergencia – INDSCAL



La curva del modelo INDSCAL muestra una convergencia rápida y estable. Podemos ver como el stress disminuye de forma pronunciada durante las primeras iteraciones y se estabiliza en un mínimo bajo 0.166. La forma suave de la curva indica que el algoritmo no encuentra dificultades para ajustar el modelo, y que el mínimo alcanzado es fiable.

```
plot(res.id, plot.type = "stressplot", main = "Curva de Convergencia - Identity")
```

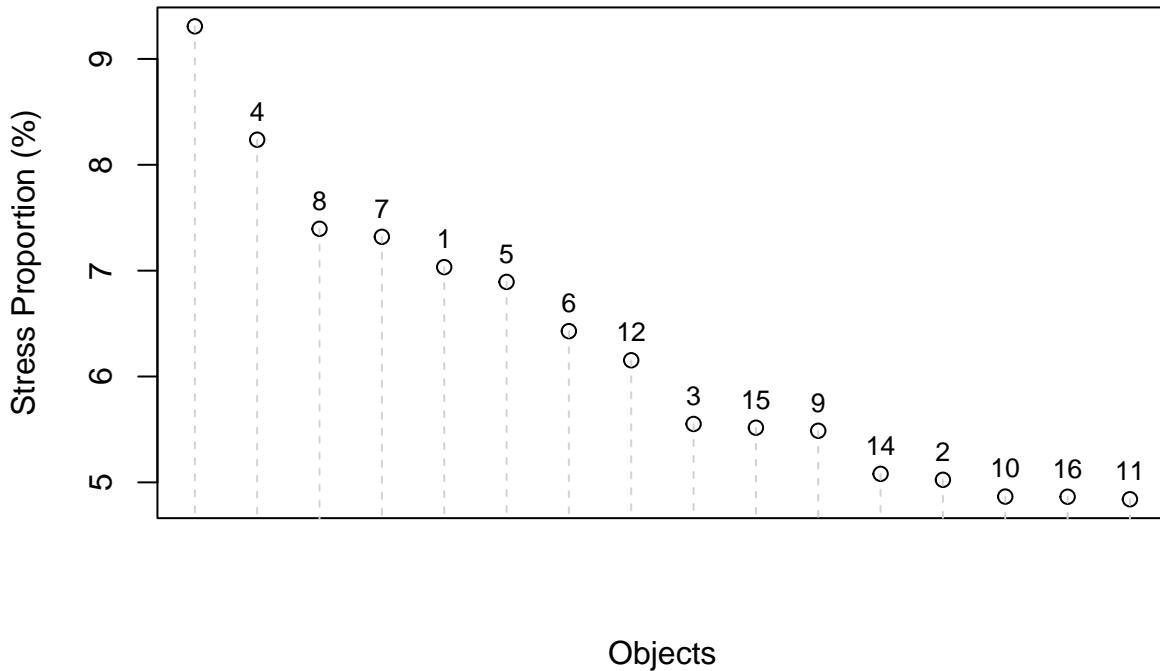
Curva de Convergencia – Identity



Si vemos el gráfico del modelo Identity vemos que se muestra una convergencia correcta, pero el nivel final de STRESS es mucho más alto 0.206 que en los demás modelos. La curva alcanza un mínimo estable, pero el mínimo está claramente por encima del obtenido con INDSCAL. Esto refuerza la idea de que los sujetos no comparten exactamente las mismas prioridades perceptuales, y que la restricción de identity resulta demasiado rígida para los datos.

```
plot(res.idio, plot.type = "stressplot", main = "Curva de Convergencia - Idioscal")
```

Curva de Convergencia – Idioscal



En cuanto al modelo Idioscal presenta una curva muy similar a la de INDSCAL. La convergencia es suave, y el STRESS final 0.166 coincide prácticamente con el del modelo estándar. La ausencia de mejora significativa sugiere que el modelo estándar ya captura adecuadamente las diferencias individuales sin necesidad de introducir mayor complejidad.

Los tres stressplots nos permite visualizar claramente cómo se comporta el algoritmo SMACOF en cada una de las variantes del modelo. El modelo Identity presenta el peor ajuste, con el valor más alto de STRESS. También hemos visto que INDSCAL y Idioscal alcanzan niveles de STRESS muy similares, aunque Idioscal no aporta mejora adicional respecto a INDSCAL. En conjunto, la comparación confirma que INDSCAL es el modelo óptimo, ya que combina buen ajuste, estabilidad de convergencia y parsimonia teórica. Por tanto podemos decir que la percepción de los estímulos no es homogénea. El modelo INDSCAL es la solución más robusta porque consigue un buen ajuste y, fundamentalmente, explica las desviaciones individuales mediante la tabla de pesos dimensionales.

Carmen Beltrán Chably
Ejercicio de análisis 5.
04/12/2025

Ejercicio 5.3. Ejercicio 5.3. Usando los datos de la Tabla 4.1 de colors de Helm, (1959): Leer los datos con SPSS. Realizar el análisis de los datos usando ALSCAL SPSS para el modelo identidad. Realizar el análisis de los datos con ALSCAL de SPSS para el modelo de diferencias individuales. Compara los resultados con los obtenidos mediante PROXSCAL.

Debemos tener en cuenta en primer lugar que el modelo de identidad de ALSCAL asume que todos los sujetos son iguales, es decir, que parte de la idea de que todos tienen la misma forma de ver el mundo, en cambio, el modelo INDSCAL asume que los sujetos tienen diferencias y, por tanto, le da a cada sujeto pesos diferentes. De tal forma que tiene en cuenta o pondera de forma diferente atributos diferentes, por ejemplo, como dijimos en el trabajo anterior puede ser la intensidad o el brillo del color en este caso. Para realizar en siguiente ejercicio hemos utilizado SPSS ALSCAL, mediante sintaxis. El primer modelo que usaremos es el modelo sin diferencias individuales.

En este ejercicio utilizamos **LEVEL = ORDINAL** porque los datos de colors de Helm no son distancias métricas reales, sino juicios de proximidad recogidos mediante comparaciones subjetivas entre colores. Es decir, los valores representan rangos de similitud, no magnitudes numéricas exactas. Además, no podemos asumir que la diferencia entre 1 y 2 tiene el mismo significado que entre 4 y 5. Los participantes solo expresan qué colores son más parecidos o diferentes, pero no cuánto. Es por esto por lo que el modelo debe respetar el orden de las disimilitudes, pero no sus valores absolutos, lo que encaja exactamente con un MDS no métrico u ordinal.

- **Ejercicio identidad.**

Alscal

Notas

Salida creada	03-DEC-2025 12:13:13
Sintaxis	<pre>ALSCAL VARIABLES=A C E G I K M O Q S /SHAPE=SYMMETRIC /LEVEL=ORDINAL /CONDITION=MATRIX /MODEL=EUCLID /CRITERIA=CONVERGE(0.00 1) STRESSMIN(0.005) ITER(30) CUTOFF(0) DIMENS(2,3)</pre>

- Comparación en tres dimensiones.

Iteration history for the 3-dimensional solution (in squared distances)

Young's S-stress formula 1 is used.

Iteration	S-stress	Improvement
1	,12671	
2	,12132	,00540
3	,11953	,00179
4	,11855	,00097

Este modelo converge en cuatro interacciones.

Matrix	Stress	RSQ	Matrix	Stress	RSQ
1	,047	,979	2	,060	,966
3	,045	,981	4	,053	,974
5	,065	,960	6	,072	,951
7	,053	,974	8	,079	,942
9	,112	,883	10	,099	,909
11	,097	,912			

Averaged (rms) over matrices

Stress = ,07430 RSQ = ,9483

Aquí podemos ver el Stress individuales que van de 0,045 a 0,112. Esto nos mide la falta de ajuste entre la configuración espacial del modelo y las disimilitudes observadas. Podemos observar que todos los valores del stress son bajos, ya que se encuentran por debajo de 0.20, por lo que podría ser un ajuste bueno.

Aunque nos gustaría destacar que hay variabilidades en la matriz número 4 el stress es muy pequeño, pero en la matriz 9 vemos que tiene el peor stress de todos. Los valores de RSQ nos da los datos de la proporción de la varianza que se explicada. Podemos ver que en todos los casos tienen un valor de RSQ bastante altos por lo que podríamos decir que este modelo explica entre el 88.3% y el 98.1% Si hablamos de variabilidad también vemos que la matriz 9 es la menos explicada ya que si nos fijamos en el Stress de esta matriz es de 0.112, por lo que podemos decir que es un ajuste aceptable, pero no es excesivamente bueno, por tanto, aunque es aceptable, este modelo de identidad al no tener en cuenta las diferencias entre sujetos puede perder información importante que no está teniendo en cuenta. En cuanto al RSQ global vemos que se explica casi el 90% de la varianza de las matrices.

Configuration derived in 3 dimensions

Stimulus Coordinates

Stimulus	Stimulus	1	2	3
Number	Name			
1	A	-1,5949	-,7059	,2367
2	C	-,3880	-1,7498	-,8995
3	E	,9797	-1,7221	-,1980
4	G	1,5264	-,3910	,6582
5	I	1,5819	,2912	,5213
6	K	1,2442	,9900	-,1749
7	M	,3980	1,3839	-,7462
8	O	-,7020	1,3836	-,3619
9	Q	-1,4749	,5535	,4677
10	S	-1,5702	-,0331	,4967

En la primera dimensión podemos ver que los colores I y G tienen unos altos positivos, en cambio los colores A, Q y S tienen un valor negativo -1,5949 y -1,4749 respectivamente. En la segunda dimensión vemos que los colores C y E en este caso son los colores negativos más altos, por tanto, los que más diferencian con M y O cuyos valores son altos y positivos. Esto puede ser con que M y O sean colores oscuros y C y E claros, por ejemplo.

En la tercera dimensión vemos que nos hay una dispersión de los valores tan grande como en las otras dimensiones, esto podría ser porque la diferencia es menos relevante que las demás dimensiones. El color con mas valor es el G y el menos la C en esta última dimensión.

• Comparación en dos dimensiones.

Iteration history for the 2 dimensional solution (in squared distances)

Young's S-stress formula 1 is used.

Iteration	S-stress	Improvement
1	,13736	
2	,13657	,00079

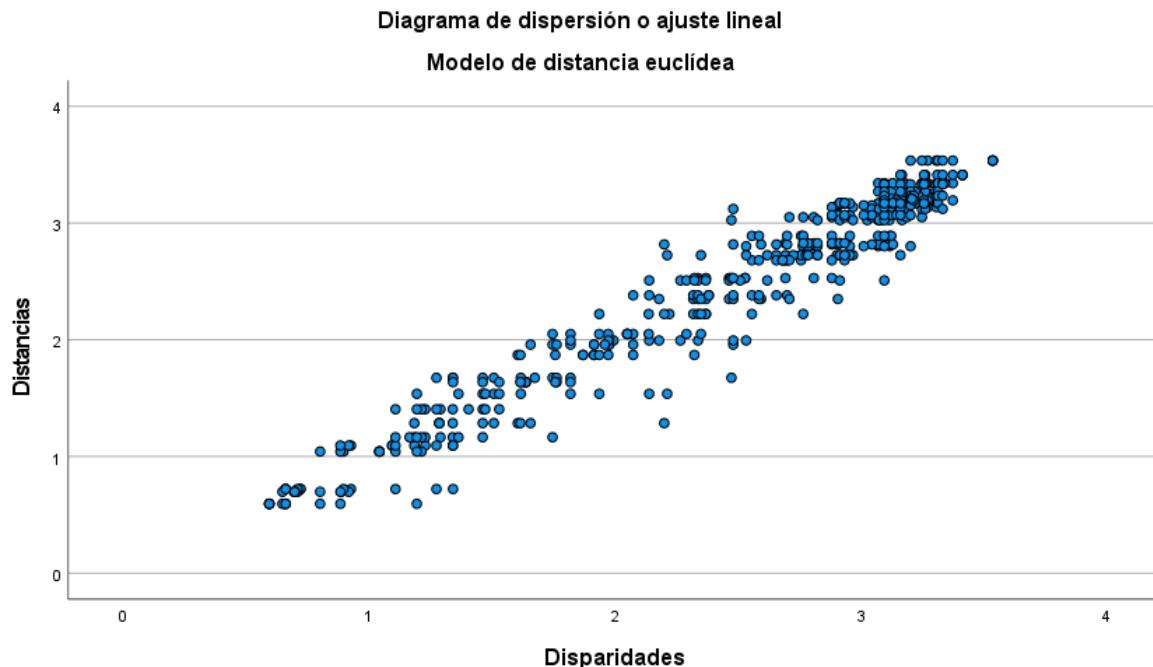
	Matrix	Stress	RSQ	Matrix	Stress	RSQ
1	,055	,976		2	,063	,969
3	,067	,964		4	,057	,975
5	,064	,968		6	,081	,949
7	,063	,969		8	,087	,941
9	,109	,908		10	,104	,916
11		,117	,894			

Averaged (rms) over matrices

Stress = ,08158 RSQ = ,94810

Podemos decir que en dos dimensiones también hemos obtenido un ajuste global bastante bueno también con un Stress y un RSQ bastante parecidos a los que hemos obtenido con tres dimensiones, tanto el stress como RSQ es más alto en 3D. Bajo esta premisa creemos que es mejor trabajar con tres dimensiones puesto que tanto el stress como el RSQ es mejor en 3D.

- **Análisis de los gráficos.**



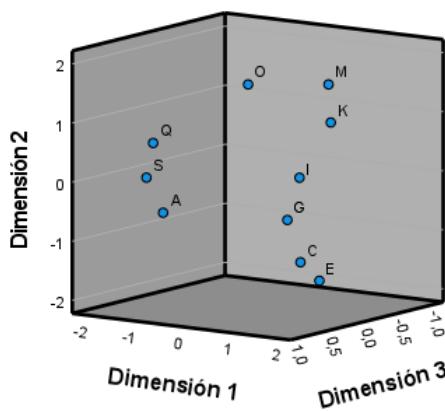
Como se ha justificado con los criterios estadísticos Stress = 0.08158\$ y RSQ = 0.94810\$), el modelo de tres dimensiones presenta un buen ajuste global a los datos de disimilitud. Como hemos dicho es un buen ajuste por lo que podemos ver que la nube de puntos es cercana a la línea de ajuste, esto nos confirma visualmente el buen ajuste global del modelo y la calidad de la representación espacial. Si los puntos estuvieran muy dispersos, el ajuste sería deficiente.

Podemos observar de igual forma una dispersión mayor en el lado de la izquierda, lo cual podría ser debido a que este ajuste no es capaz de posicionar correctamente a los pares de colores que han sido percibido como parecidos, lo cual nos da a entender que hay datos, por ejemplo, la saturación del color o la luminosidad que no se están teniendo en cuenta y por tanto el stress se acumula ahí.

Por tanto, creemos que el modelo de identidad no tiene la suficiente información para realizar el análisis de forma eficiente. Es por esto por lo que creemos que es necesario realizar un análisis del modelo mediante INSDCAL para así ver si esto se debe a que cada sujeto percibe los colores con criterios diferentes.

Configuración de estímulo derivado

Modelo de distancia euclídea



Los gráficos en 3D son más difíciles de analizar porque no se ven de forma tan clara como los de 2D. A veces parece que los puntos están juntos o separados solo por el ángulo desde el que se mira, así que para poder verlos bien deberíamos ir girando el gráfico. Si nos fijamos los colores en función de la dimensión podemos ver que según la dimensión 1 los colores con valores más positivos como A, S y Q están posicionados más a la derecha en cambio los colores con los valores más negativos están a la izquierda como por ejemplo G, I y K. Podrían deberse a la intensidad del color los colores más intensos a la derecha y los menos intensos a la izquierda. En cambio, si nos fijamos en el eje de la dimensión 2 vemos que los colores C y E están posicionados más arriba y los colores M y O están en la parte inferior del gráfico. Esto puede estar ligado con el brillo de los colores, es decir que los colores que se han entendido con más brillo estén arriba y los que menos abajo. Con todo esto entendemos que los colores más cercanos, como A y S son colores que se ven como muy parecidos y los que están más alejados entre sí como los más diferentes.

• EJERCICIO INDSCAL

El modelo con el que vamos a trabajar en el modelo INDSCAL, que es una técnica de escalamiento multidimensional a tres vías, cuyo propósito es analizar datos de proximidad pero que vienen de diferentes sujetos o condiciones. La diferencia de este modelo con el modelo de identidad es que este modelo asume heterogeneidad y en el de identidad se asumía

Este modelo pondera el espacio común, es decir, primero calcula una configuración de estímulos única que es el consenso o el promedio de las percepciones de todo el grupo. Pero además para cada sujeto y para cada dimensión calcula un peso dimensional. Este valor nos indica la importancia que le da cada sujeto a cada dimensión.

El modelo utiliza estos pesos para modificar la fórmula de la distancia Euclídea clásica. En lugar de calcular la distancia en el espacio común, calcula una distancia ponderada o modificada para cada sujeto

Alscal

Sintaxis	ALSCAL VARIABLES=A C E G I K M O Q S /SHAPE=SYMMETRIC /LEVEL=ORDINAL /CONDITION=MATRIX /MODEL=INDSCAL /CRITERIA=CONVERG E(0.001) STRESSMIN(0.005) ITER(30) CUTOFF(0) DIMENS(2,3) /PLOT=DEFAULT ALL /PRINT=DATA HEADER.
----------	--

• Solución en tres dimensiones.

Iteration history for the 3 dimensional solution (in squared distances)

Young's S-stress formula 1 is used.

Iteration	S-stress	Improvement
0	,12042	
1	,11700	
2	,11116	,00585
3	,10971	,00144
4	,10886	,00086

Iterations stopped because
S-stress improvement is less than ,001000

Iterations stopped because S-stress improvement is less than ,000010

Aquí podemos ver que spss paró en la interacción número 4 porque ya no puede seguir mejorando el modelo de manera significativa.

Matrix	Stress	RSQ	Matrix	Stress	RSQ
1	,043	,983	2	,054	,975
3	,041	,984	4	,057	,971
5	,063	,965	6	,063	,964
7	,042	,984	8	,076	,946
9	,103	,913	10	,097	,915
11	,089	,924			

Averaged (rms) over matrices

Stress = ,06942 RSQ = ,95674

En primer lugar, podemos ver que el stress es de 0.6942 y la varianza explicada del 95.7%. Por tanto, podemos decir que es un muy buen ajuste ya que se explica un 91% de la varianza y el stress ha disminuido por lo que podemos decir que representa mejor los datos teniendo cuando se tiene en cuenta los datos de forma ponderados por preferencias de los sujetos.

Podemos ver que el sujeto número diez junto con el nueve son los que tienen peor ajuste ya que tienen el stress más alto, pero eso no invalida el modelo. Podemos ver que el rango va desde el sujeto tres que tienen el stress más bajo a el sujeto nueve con un 0.103 que tiene el stress más alto por lo que podemos ver que hay variabilidad.

Podemos ver que el sujeto número uno ha mejorado con respecto al modelo anterior, ha pasado de un 0.055 a 0.043. El sujeto número nueve ha mejorado, pero no de forma significativa pasando de 0.109 a 0.103. El sujeto número 11 mejora también pasando de 0.117 a 0.089.

Stimulus	Stimulus	1	2	3
Number	Name			
1	A	1,3733	-,4471	,7905
2	C	1,2811	,9359	-1,4876
3	E	,3563	1,6757	-,4867
4	G	-,8003	1,0960	1,2691
5	I	-1,1732	,7199	1,0025
6	K	-1,3421	,0831	-,2968
7	M	-1,0012	-,6586	-1,5418
8	O	-,3238	-1,2694	-,7998
9	Q	,6004	-1,2472	,8098
10	S	1,0295	-,8883	,7409

En la primera dimensión vamos que los colores con valores positivos más altos son los colores I, K y M y los colores con valores negativos altos en esta dimensión son los colores A y C. Podríamos decir que los colores positivos en esta dimensión podrían ser los percibidos como colores cálidos, y los negativos como colores fríos.

En la segunda dimensión, los colores que tienen los valores positivos más alto son los colores O, y Q y los colores con los valores más altos pero negativos son E, y G. Por tanto, los sujetos entienden que estos colores son diferentes entre sí, esta dimensión podría ser la que se trae de la saturación de los colores, es decir, los más intensos podrían ser los colores con los valores positivos y los colores con valores negativos podrían corresponder a los colores más apagados.

Por último, en la tercera dimensión los colores con los valores positivos más altos son los colores G y I y los colores con valores negativos son C y M. Aquí podríamos darle el valor a la dimensión de los colores más claros, que serán representados por los colores positivos y los más oscuros que los pondremos con los colores cuyos valores son negativos.

Subject Weights

Dimension

Subject	Weird-	1	2	3
Number	ness			
1	,0642	,6548	,7350	,1196
2	,3198	,7766	,6072	,0587
3	,2172	,7078	,6694	,1859
4	,0360	,7092	,6744	,1141
5	,1315	,6417	,7372	,0975
6	,1120	,7713	,5980	,1085
7	,1046	,7838	,5947	,1261
8	,1562	,6633	,6924	,1623
9	,4495	,6611	,6893	,0321
10	,0245	,6998	,6417	,1181
11	,3177	,5524	,7587	,2076

A lo largo de esta tabla vamos a poder ver las diferencias individuales de los once sujetos interpretando el valor del peso que tienen en cada dimensión. Por tanto, un peso mayor en una de las dimensiones nos indicará que ese sujeto le pone más importancia a esa dimensión que a las demás.

Si hablamos en términos generales podemos ver que todos les dan más importancia a las dimensiones uno y dos, vemos que la dimensión tres pasa casi desapercibidas ya que tiene unos pesos muy bajos comparados con las dimensiones uno y dos.

Si nos fijamos un poco más en los datos podemos ver que el sujeto número siete es el sujeto que más peso tiene en la dimensión número uno, por lo tanto, según el ejemplo que hemos usado para nombrar a las dimensiones, podríamos decir que este sujeto es el más sensible al cambio del tono.

El sujeto once le da mucha más importancia a la dimensión dos que a la uno. Además, el sujeto que más peso le da a la dimensión tres en comparación con los demás sujetos. Pero la dimensión con mayor peso para el once es la segunda. Aun siendo el que más peso tiene en la dimensión tres, sigue siendo menor que el peso de las otras dimensiones. podríamos decir que este sujeto es el que más peso le da a la luminosidad de todo el grupo, por tanto, esto influye en su forma de ver los colores en este caso.

El sujeto 2,6 y 7 le da más peso a la dimensión uno, aunque la dos sigue siendo bastante importante y la dimensión tres casi irrelevante para ellos.

Si nos fijamos en los datos de weirdness podemos ver que el sujeto nueve, por tanto, sus patrones de ponderación son los esperados, en cambio, los sujetos diez y cuatro tienen los valores más altos, por ello se desvían un poco más de la media del grupo.

- **Solución en dos dimensiones.**

Averaged (rms) over matrices

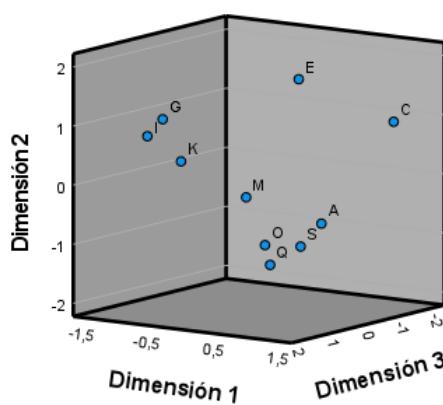
Stress = ,07585 RSQ = ,95556

Aquí podemos ver que el stress promedio en dos dimensiones es de 0.07585 y el RSQ de 0.95556. Vemos que la bondad del ajuste es muy buena porque es un valor alto. La solución en 2D ya explica más del 95% de la varianza. Hemos obtenido un error residual pequeño de 0.7585, pero aun siendo pequeño es un valor mayor que el que hemos obtenido en el análisis de tres dimensiones. Esa disminución del stress creemos que justifica el uso de la tercera dimensión ya que la caída del error en tres dimensiones es considerable, pasando de 0.07585(2D) a 0,06942 (3D).

- **Análisis de los gráficos**

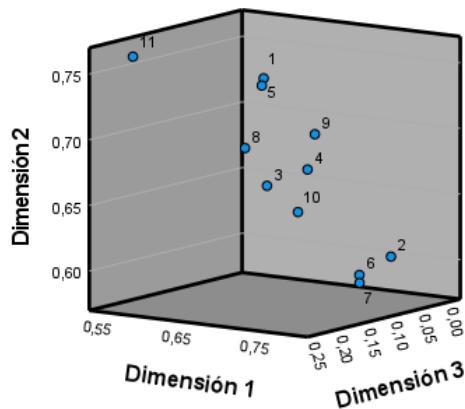
Los gráficos en 3D son más difíciles de analizar porque no se ven de forma tan clara como los de 2D. A veces parece que los puntos están juntos o separados solo por el ángulo desde el que se mira, así que para poder verlos bien deberíamos ir girando el gráfico. Por eso, aunque la solución 3D es más precisa, interpretar visualmente lo que pasa en el espacio tridimensional cuesta un poco más. Aun así, lo haremos mediante los datos obtenidos.

Configuración de estímulo derivado
Modelo de distancia euclídea de diferencias individuales (ponderado)



En este gráfico podemos ver los colores, cuyo posicionamiento es una configuración ponderada según los pesos de los sujetos. En la primera dimensión vamos que los colores positivos son I, K y M en contra posición de los colores negativos en esta dimensión que son los A, C y S. La segunda dimensión, el eje Y tiene positivos los colores O, Q y S en contra de los negativos que son los colores E, G y C. Como hemos dicho estos colores pueden estar diferenciados por los sujetos por la saturación de los colores. Y la última dimensión los colores positivos son los K, C y M y los negativos los Q, O y S que podríamos decir que son diferenciados como claros u oscuros.

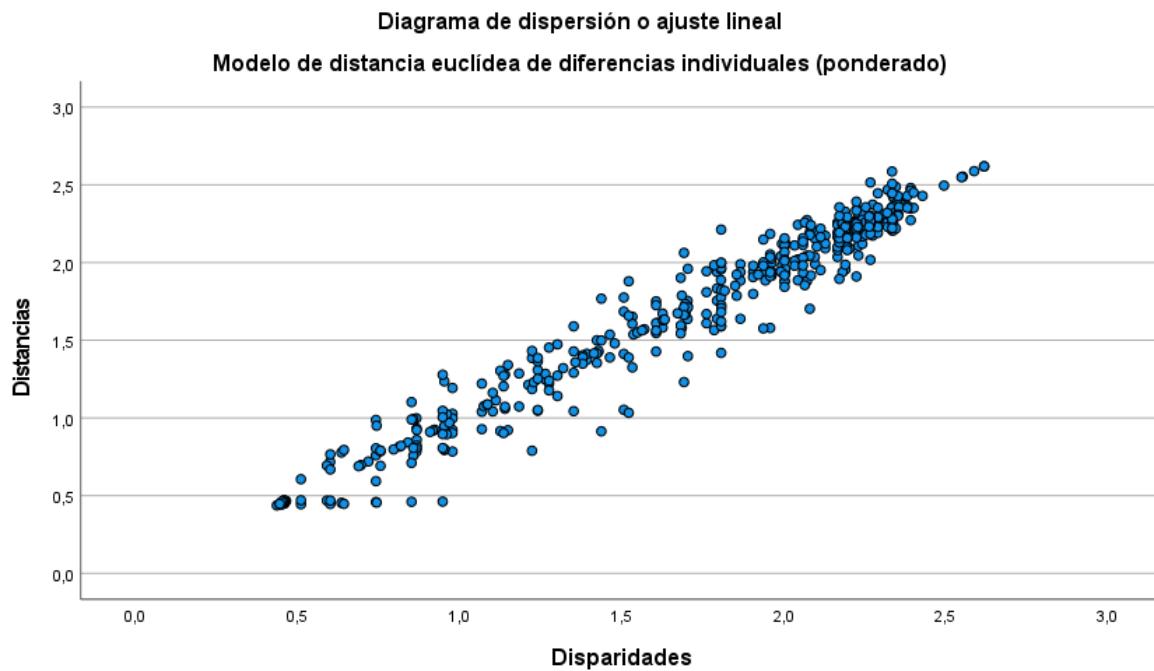
Ponderaciones de sujeto derivadas
Modelo de distancia euclídea de diferencias individuales (ponderado)



Este gráfico en tres dimensiones representa a los sujetos y su posición dentro del cubo nos dice que dimensión utiliza más. Todos los sujetos se concentran en el plano formado por la Dimensión 1 y la Dimensión 2. Esto se debe a que todos tienen pesos altos en D1 y D2, y bajos en D3. Los sujetos 4 y 10, es decir, los menos atípicos se ubican muy cerca entre sí y cerca del centro del grupo principal de puntos.

Además, el sujeto 9: Se aleja del grupo debido a su peso extremadamente bajo en D3 lo que significa que su punto estará proyectado casi exactamente sobre el plano D1-D2. Por último, el sujeto 11: Se aleja en la dirección del eje D3 debido a su peso alto y se ubicará más lejos a lo largo del eje D2, confirmando que tiene la percepción más equilibrada del grupo.

Lo que si se podemos ver es que algunos puntos se alejan un poco más, lo que refleja que ciertas parejas de colores son más difíciles de ajustar. Esto es normal en INDSCAL porque cada sujeto usa las dimensiones de forma distinta y el modelo tiene que encontrar un equilibrio entre todas esas percepciones. Aun así, la mayoría de los puntos siguen el patrón esperado y mantiene una relación lineal clara.



La dispersión de los puntos nos indica que existen diferentes estrategias cognitivas para juzgar las diferencias de color. La mayoría de los puntos se concentran en una región, lo que muestra el consenso de que las dimensiones de Tono (D1 y D2) son las más importantes. La dispersión de ciertos puntos lejos del centro de la nube identifica a los sujetos que tienen percepciones inusuales. La dispersión observada en este diagrama es la causa directa de la reducción del Stress en el modelo INDSCAL.

- **Comparación de los resultados obtenidos.**

En el modelo de identidad pudimos ver que el stress en la solución de tres dimensiones era de 0,06942 y tenía un RSQ de 0,94832 y aunque vimos que en este modelo era aceptable también vimos mucha variabilidad entre los sujetos, por ejemplo, los sujetos nueve y once tenían un stress alto comparado con los demás sujetos, por lo que pensábamos que no representaba bien a todos los participantes. Es por esto por lo que creímos oportuno realizar el modelo INDSCAL. También vimos que en el modelo de 2D el stress empeoraba un poco a 0,07430 aunque elegimos el de tres dimensiones porque el stress es mejor en dos dimensiones.

En el modelo INDSCAL vimos que el ajuste mejoró, el stress bajo a 0,06942 y el RSQ aumentó a 0,95674 en la solución en tres dimensiones. Hemos visto que sigue habiendo diferencias entre sujetos, el modelo sigue explicando mejor la variabilidad al haber introducido los pesos. De esta forma hemos podido ver que todos usan más las dos primeras dimensiones, pero por ejemplo el sujeto once usa más que los demás la dimensión 3 cosa que no pudimos ver con el modelo de identidad.

En resumen, podemos decir que el modelo INDSCAL nos da un ajuste mejor, ya que incorpora las diferencias individuales y con ello hemos conseguido reducir el error global, por tanto, este modelo en 3D es mejor que el ofrecido por el modelo de identidad.

- **Comparación con PROXSCAL.**

Si comparamos los datos que hemos obtenido con ALSCAL con los que obtuvimos con PROXSCAL vemos que son algo parecidos, pero encontramos diferencias que vamos a detallar a continuación. Si comparamos los modelos de identidad podemos ver que en el modelo de ALSCAL el stress en 2D era de 0.07430 y en 3D de ,06942 en PROXSCAL el Stress-1: 0.08120 y S-stress: 0.99341 En el primer caso decidimos usar el modelo en dos dimensiones y en ALSCAL en 3D. Pudimos ver que funciona bien, pero que no era perfecto el modelo de identidad la estructura en 2D es prácticamente igual tanto en ALSCAL como en PROXSCAL. En ambos vimos que los sujetos 2 y 11 eran los que peor se ajustaban al modelo.

Si hablamos del modelo ponderado hemos visto que en ALSCAL el stress en 3D era de ,06942 y en el modelo de identidad el stress es de de 0.07430 y que el stress por tanto mejoraba en INDSCAL y por eso elegimos usar la tercera dimensión porque mejoraba nuestro modelo. En PROXSCAL vimos que el s-stress 1 era de 0.12126 y el s-stress de 0.03815. En este modelo decidimos usar 2 dimensiones porque el stress apenas mejora en 3D. En cuando a las coordenadas son casi iguales en los dos modelos.

En general, ALSCAL y PROXSCAL obtienen configuraciones espaciales muy similares ya que los dos coinciden en que el modelo de identidad funciona bien en dos dimensiones y que añadir una tercera dimensión no mejora lo suficiente. Pero si hablamos del modelo INDSCAL en ALSCAL sí que vemos una mejora clara al pasar de dos a tres dimensiones, mientras que PROXSCAL hemos considerado que la mejora es mínima y prefiere mantenerse en dos dimensiones.

Aun así, las configuraciones de estímulos y los pesos de las dimensiones son muy coherentes entre ambos métodos. Esto indica que ambos procedimientos están captando la misma estructura general, aunque PROXSCAL es más conservador a la hora de aumentar la dimensionalidad También en ambos métodos los sujetos presentan patrones de ajuste distintos, especialmente algunos participantes concretos como los sujetos 2 y 11, los cuales sistemáticamente se han ajustado peor que los demás sujetos, lo que indica que existen diferencias individuales reales en la percepción del conjunto de estímulos.

Sin embargo, los dos programas difieren en la recomendación respecto a cuántas dimensiones utilizar. En el modelo de identidad, tanto ALSCAL como PROXSCAL muestran que la solución en dos dimensiones explica prácticamente lo mismo que la de tres, de forma que añadir una dimensión extra no aporta una mejora suficiente. Por eso, en ambos ejercicios la opción más razonable es quedarse con dos dimensiones en el modelo de identidad.

En cambio, en los modelos ponderados hemos visto que hay una diferencia importante entre los métodos. Con ALSCAL, la solución en tres dimensiones reduce de forma clara el stress y aporta información adicional, especialmente porque algunos sujetos utilizan más intensamente la tercera dimensión, como por ejemplo el sujeto 11. Esto justifica escoger la solución 3D. En cambio, con PROXSCAL la mejora al pasar de dos a tres dimensiones es mínima, y el modelo ya ajusta muy bien en 2D, por lo que PROXSCAL es más conservador y considera suficiente quedarse en dos dimensiones.

En resumen, aunque ALSCAL y PROXSCAL difieren en la recomendación sobre la dimensionalidad óptima en INDSCAL, ambos ofrecen configuraciones espaciales coherentes y captan la misma estructura perceptual general. La diferencia se debe principalmente a la forma en que cada método calcula y optimiza el stress.