

Universidad de Valencia.

Facultad de Psicología.



**MEDICION DE CONSTRUCTOS AFECTIVOS BIPOLARES
MEDIANTE ESCALAS TIPO LIKERT:
EL DESAJUSTE ENTRE EL MODELO DE ESCALAMIENTO, EL
METODO DE ANALISIS Y EL OBJETO DE MEDIDA.**

TESIS DOCTORAL

**PRESENTADA POR:
SUSANA LLORET SEGURA**

**DIRIGIDA POR:
DR VICENTE GONZALEZ ROMA
DR JOSE M^a PEIRO SILLA**



[1995]

UMI Number: U607383

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U607383

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

UNIVERSIDAD DE VALENCIA
FACULTAD DE PSICOLOGÍA
BIBLIOTECA
Reg. de Entrada nº 8616
Fecha: 8-5-96
Signatura P-7446

BID. T 1560

D. 468072
L. 468093

A mis padres, Alejandro y Bárbara,
y a mi abuela, Bárbara.
Siempre estais...

| | |
|--|-----|
| PARTE I: EL ENFOQUE TRADICIONAL | 1 |
| CAPITULO I. INTRODUCCION | 3 |
| 1. LA ESTRUCTURA DEL AFECTO. PERSPECTIVA HISTORICA..... | 5 |
| 1.1. La primera etapa..... | 5 |
| 1.2. La segunda etapa..... | 9 |
| 1.3. La tercera etapa..... | 31 |
| CAPITULO 2. EL REPLANTEAMIENTO DEL PROBLEMA | 48 |
| 1. INTRODUCCION..... | 50 |
| 2. ESCALAS EXPLICITAMENTE BIPOLARES Y ESCALAS IMPLICITAMENTE BIPOLARES..... | 52 |
| CAPITULO 3. METODO | 60 |
| 1. PROCEDIMIENTO DE RECOGIDA DE DATOS Y MUESTRA..... | 62 |
| 2. MEDIDAS..... | 65 |
| 3. HIPOTESIS Y ANALISIS..... | 66 |
| CAPITULO 4. RESULTADOS Y CONCLUSIONES | 70 |
| 1. RESULTADOS..... | 72 |
| 2. CONCLUSIONES..... | 75 |
| PARTE II: ACERCAMIENTOS ALTERNATIVOS AL PROBLEMA DE LA MEDIDA DE DIMENSIONES BIPOLARES AFECTIVAS MEDIANTE ESCALAS TIPO LIKERT CONTRABALANCEADAS: DESDOBLAMIENTO DE ESCALAS | 80 |
| CAPITULO 5. INTRODUCCION | 82 |
| 1. INTRODUCCION..... | 84 |
| 2. EL MODELO DE MEDIDA TRADICIONAL FRENTE AL MODELO DE "DOS MITADES"..... | 89 |
| CAPITULO 6. METODO | 94 |
| 1. PROCEDIMIENTO DE RECOGIDA DE DATOS Y MUESTRA..... | 96 |
| 2. MEDIDAS..... | 96 |
| 3. HIPOTESIS Y ANALISIS..... | 97 |
| CAPITULO 7. RESULTADOS Y CONCLUSIONES | 100 |
| 1. RESULTADOS..... | 102 |
| 2. CONCLUSIONES..... | 121 |
| PARTE III: ACERCAMIENTOS ALTERNATIVOS AL PROBLEMA DE LA MEDIDA DE DIMENSIONES BIPOLARES AFECTIVAS MEDIANTE ESCALAS TIPO LIKERT CONTRABALANCEADAS: ESCALAMIENTO DE SUJETOS Y ESTIMULOS | 128 |
| CAPITULO 8. INTRODUCCION | 130 |
| 1. INTRODUCCION..... | 132 |

| | |
|---|------------|
| 2. EL MODELO DE ESCALAMIENTO DESPLEGADO DE COOMBS (1950)..... | 134 |
| 3. APROXIMACION PROBABILISTICA AL MODELO DE ESCALAMIENTO DESPLEGADO DE COOMBS (1950). | 139 |
| 4. EL MODELO DE ESCALAMIENTO DESPLEGADO Y EL DESDOBLAMIENTO DE FACTORES PRETENDIDAMENTE BIPOLARES..... | 146 |
| CAPITULO 9. METODO..... | 152 |
| 1. PROCEDIMIENTO DE RECOGIDA DE DATOS Y MUESTRA..... | 154 |
| 2. MEDIDAS..... | 154 |
| 3. HIPOTESIS..... | 154 |
| 4. ANALISIS..... | 155 |
| CAPITULO 10. RESULTADOS Y DISCUSION | 168 |
| 1. ESCALAMIENTO BAJO EL MODELO DE DESPLEGAMIENTO PROBABILISTICO..... | 170 |
| 1. ESCALA TENSO-CALMADO..... | 172 |
| 2. ESCALA DEPRIMIDO-ENTUSIASMADO..... | 180 |
| 3. ESCALA FATIGA-VIGOR..... | 185 |
| 4. ANALISIS CORRELACIONAL..... | 189 |
| 5. CONCLUSIONES..... | 193 |
| CAPITULO 11. ESCALAMIENTO ACUMULATIVO DE GUTTMANN BAJO EL MODELO DE DESPLEGAMIENTO..... | 196 |
| 1. INTRODUCCION..... | 198 |
| CAPITULO 12. METODO..... | 212 |
| 1. PROCEDIMIENTO DE RECOGIDA DE DATOS Y MUESTRA..... | 214 |
| 2. MEDIDAS..... | 214 |
| 3. HIPOTESIS..... | 214 |
| 4. ANALISIS..... | 215 |
| CAPITULO 13. RESULTADOS Y DISCUSION | 220 |
| 1. ESCALAMIENTO BAJO EL MODELO DE DESPLEGAMIENTO DE DOS ESCALAS ACUMULATIVAS. | 222 |
| 2. ESCALA TENSO-CALMADO..... | 222 |
| 3. ESCALA DEPRIMIDO-ENTUSIASMADO..... | 226 |
| 4. ESCALA FATIGA-VIGOR..... | 228 |
| 5. ANALISIS CORRELACIONAL..... | 229 |
| 6. CONCLUSIONES..... | 233 |
| PARTE IV: CONCLUSIONES | 240 |
| ANEXO | 254 |
| BIBLIOGRAFIA..... | 258 |

PARTE I

EL ENFOQUE TRADICIONAL.

CAPITULO I

**INTRODUCCION:
LA POLEMICA**

1. LA ESTRUCTURA DEL AFECTO: PERSPECTIVA HISTORICA.

1.1 La primera etapa

El estudio de la estructura o dimensionalidad del afecto, en el contexto de la psicología como disciplina científica, puede situarse a principios de la década de los cuarenta. Inicialmente, el estudio de la dimensionalidad del afecto parte de una premisa clara: el afecto es un constructo multidimensional y bipolar. Así lo muestra el indicador más claro del afecto, que es lenguaje que empleamos para manifestarlo. En este sentido, Guilford (1954) establece un precedente que continua vigente en la actualidad: "una escala para medir afecto es una escala bipolar" (Guilford, 1954, p. 264). En consecuencia, diferentes investigadores abordaron el estudio de la dimensionalidad del afecto bajo estas premisas. Tras identificar y definir teóricamente cuáles eran las dimensiones relevantes de la estructura del afecto, se seleccionaban un determinado número de adjetivos descriptivos de las mismas. Con el fin de incluir una muestra representativa de los adjetivos descriptivos de cada dimensión, por una lado, y para evitar determinados sesgos de respuesta, por otro, la selección de adjetivos se realizaba de modo que el número de adjetivos representativos de los dos polos de cada dimensión fuera igual o similar. La tarea del sujeto consistía generalmente en indicar la intensidad con que experimentaba cada uno de los estados afectivos descritos mediante los items, para lo que se le ofrecía una escala de tipo Likert de 4 o 5 puntos que oscilaba generalmente desde "nada en absoluto" a "extremadamente". De este modo quedaban constituidas las escalas. A continuación las intercorrelaciones entre tales items eran sometidas a un análisis factorial, esperando encontrar evidencia empírica en favor de la estructura factorial bipolar inicialmente hipotetizada.

Por ejemplo, Nowlis y Nowlis (1956) hipotetizaron que la estructura del afecto quedaría descrita a partir de cuatro dimensiones bipolares: placer-displacer, activación-desactivación, orientación social positiva-negativa y control-falta de control. A continuación seleccionaron una serie de adjetivos representativos de cada dimensión, y a través de sucesivos estudios, las respuestas obtenidas en cada uno

de ellos fueron sometidas a diversos análisis factoriales. No obstante, la evidencia empírica encontrada indicó que la estructura empírica no era bipolar. En lugar de obtener los cuatro factores bipolares inicialmente hipotetizados, las soluciones factoriales obtenidas identificaban ocho factores monopolares. Los polos opuestos de las dimensiones bipolares hipotetizadas habían dado lugar a factores monopolares e independientes.

Este resultado contraintuitivo despertó el interés de otros autores por el estudio de la estructura del afecto. Los sucesivos trabajos que se realizaron en esta "primera etapa" del estudio del afecto resultaron consistentes entre sí, pero incongruentes con las hipótesis de partida. En la mayor parte de ellos, los adjetivos que formaban los polos opuestos de las dimensiones bipolares hipotetizadas se desdoblaban, dando lugar a dos factores monopolares por cada una de esas dimensiones supuestamente bipolares (Nowlis, 1965; Borgatta, 1961; Clyde, 1963; McNair y Lorr, 1964; Tayer, 1967).

En consecuencia, la evidencia empírica observada llevó a gran parte de los investigadores del afecto a replantear la estructura del afecto como multidimensional pero monopolar (Bradburn, 1969; Hall, 1977; Izard, 1972; McLachlan, 1976; Westbrook, 1976). Ello dió lugar a la sustitución de las escalas bipolares inicialmente elaboradas por otras de carácter monopolar, acordes con este nuevo modelo.

De forma paralela se desarrollaron otros estudios para determinar el significado de las palabras. La aportación de mayor relevancia en este ámbito fue la realizada por Osgood, Suci y Tannenbaum (1957). El objetivo que guió la investigación de estos autores era la medida del significado del lenguaje. Para ello era necesario identificar si los sujetos empleaban alguna o algunas dimensiones comunes en los juicios de atribución de significado a un concepto dado (estímulo), y si esas dimensiones eran generalizables a cualquier concepto dado. Se trataba de determinar si existía un espacio semántico común a sujetos y conceptos, y se trataba de identificar qué dimensiones lo definían. En consecuencia, los autores realizaron una extensa serie de estudios encaminados a identificar la dimensionalidad del espacio semántico.

Estos estudios se caracterizaron por la riqueza de los métodos de obtención de datos y análisis empleados, en correspondencia con la variedad de perspectivas con que se abordó el problema de la dimensionalidad del espacio semántico. No obstante, la mayor parte de los datos empleados constituyeron juicios de atribución de significado a diferentes conceptos registrados sobre un tipo de escalas particulares. Cada ítem se componía de un concepto a enjuiciar, y de una escala que emparejaba dos adjetivos opuestos, separados por un continuo bipolar, dividido en 7 o 9 espacios, en el que el sujeto daba su respuesta. Finalmente concluyeron que el espacio semántico podía quedar explicado por tres dimensiones bipolares: evaluación, actividad y potencia, y desarrollaron un método para su medida. La medida de cada una de estas dimensiones para un concepto dado era posible a partir de la selección de un subconjunto de estas escalas máximamente cargado en cada una de las dimensiones identificadas. Este método de medida se denominó "diferencial semántico", y el formato de las escalas empleadas, generalizado a otros contextos, fue denominado "formato de diferencial semántico". De forma característica, este formato de escala de respuesta presenta como anclajes pares de ítems opuestos, que delimitaban un continuo bipolar dividido en 7 o 9 espacios, en el que el sujeto da su respuesta al estímulo objeto de medida.

La ausencia de evidencia empírica de bipolaridad en la estructura del afecto constatada en los estudios anteriormente mencionados, propició el cuestionamiento de la bipolaridad del significado afectivo del lenguaje. En este sentido Green y Goldfried (1965) desarrollaron una explicación alternativa para los resultados a favor de la bipolaridad encontrados por Osgood et al. (1957). Las escalas de Osgood et al. (1957) daban por supuesta la bipolaridad de las dimensiones subyacentes, y por el formato de respuesta empleado no solo forzaban la aparición de dimensiones bipolares, sino que impedían la contrastación empírica de la bipolaridad. Con este fin, Green y Goldfried (1965) construyeron una nueva versión de las escalas de Osgood et al. (1957), en la que los sujetos contestaban por separado a los adjetivos que en las escalas de Osgood constituían los extremos opuestos del continuo de respuesta. Si las dimensiones identificadas por Osgood et al. (1957) eran bipolares, entonces las correlaciones entre adjetivos opuestos que midieran la

misma dimensión deberían ser elevadas y negativas. Los resultados obtenidos mostraron en general que las correlaciones observadas no se ajustaban a ese patrón. La bipolaridad del espacio semántico parecía pues un artefacto del método de medida empleado, tal como estos autores habían hipotetizado.

Este resultado, de nuevo completamente contraintuitivo, estimuló la aparición de nuevos estudios diseñados para aclarar la confusión creada acerca de la bipolaridad. Una de las explicaciones de mayor relevancia para la falta de apoyo empírico de los datos respecto de la bipolaridad fue propuesta por Bentler (1969). Este autor explicaba la ausencia de correlaciones negativas y elevadas entre adjetivos opuestos como consecuencia del sesgo de aquiescencia. Este sesgo recoge la tendencia diferencial de los sujetos a asentir o disentir frente a los items, independientemente de su contenido. El efecto de la aquiescencia produciría relaciones positivas entre items opuestos, que disminuirían la magnitud de las correlaciones negativas existentes. Tras construir una nueva versión "disociada" de las escalas de diferencial semántico, análoga a la desarrollada por Green y Goldfried (1965), Bentler calculó una medida de aquiescencia y la empleó para obtener la correlación parcial entre cada par de escalas opuestas. La medida de aquiescencia empleada fue construida a partir de la suma de todos los items (sin invertir ninguno de ellos) excepto aquellos incluidos en el par de escalas opuestas para el que se iba a emplear. Este proceso se repitió para cada par de escalas opuestas. Las correlaciones parciales mostraron el signo esperado y un aumento en su magnitud en algunos casos espectacular: la correlación original entre items positivos y negativos de la dimensión evaluación fue de 0.03, tras eliminar el efecto de la aquiescencia en la forma expuesta, esta correlación pasó a ser de -0.76.

La bipolaridad del espacio semántico parecía finalmente fuera de duda, pero la del espacio afectivo permanecía cuestionada. En este sentido, Meddis (1972) llamó la atención sobre el tipo de escalas que empleaban dos de los trabajos en los que se había fundamentado la monopolaridad de las dimensiones del afecto. Tanto Nowlis (1965) como Thayer (1967) emplearon la misma escala de respuesta: 1. no; 2. indeciso; 3. un poco; y 4. completamente. De acuerdo con Meddis, la

asimetría de la escala de respuesta, que cuenta con una categoría de rechazo y dos de aceptación, junto con una opción de indeciso que puede interpretarse no sólo como un sentimiento cuya intensidad se encuentra entre las categorías adyacentes, podían ser responsables de la falta de bipolaridad observada. Con el fin de comprobar esta hipótesis, Meddis comparó las soluciones factoriales subyacentes a un conjunto de adjetivos similares a los empleados en los trabajos originales cuando eran contestados con la escala de respuesta original y cuando se contestaban en una escala de respuesta simétrica y exenta de la categoría "indeciso". Los resultados indicaron que el tipo de escala de respuesta determinaba en gran medida la solución factorial obtenida: predominantemente monopolar para la escala de respuesta original, y predominantemente bipolar para la escala simétrica.

1.2. La segunda etapa

Estos resultados, aunque esperanzadores, no alcanzaban a explicar la falta de bipolaridad observada en otros estudios que sí habían empleado escalas simétricas y sin la categoría 'indeciso'. De modo que los investigadores partidarios de la estructura bipolar del afecto, animados por la posibilidad de explicar la evidencia contraria a la bipolaridad en términos de deficiencias metodológicas, sacaron a la luz nuevas argumentaciones en su favor. Uno de los investigadores más activos en este sentido fue Russell. En un trabajo publicado en 1979 señalaba algunas de estas posibles deficiencias. Junto con la posibilidad de rotaciones alternativas a las realizadas y a la desigualdad de las distribuciones de respuesta para las variables incluidas en los análisis factoriales realizados, este autor enumeró una serie de sesgos específicamente dirigidos a oscurecer la aparición de la bipolaridad. En primer lugar, en la medida en que el muestreo de adjetivos no incluya elementos representativos de ambos extremos de un continuo bipolar, disminuye la probabilidad de que emerjan factores bipolares. En segundo lugar, cuanto mayor sea el referente temporal en el que se pide a los sujetos que evalúen su afecto, mayor es la posibilidad de que contesten con afectos opuestos. En tercer lugar, la proximidad entre items, ya sea temporal o espacial, puede aumentar espúriamente la correlación positiva entre items, independientemente de su contenido.

Ello podría atenuar las correlaciones entre items opuestos. Y en cuarto lugar, el sesgo de aquiescencia tendría sobre las correlaciones observadas un efecto similar al de la proximidad anteriormente mencionado, tal como mostró Bentler (1969) en el ámbito del estudio del significado.

A continuación, Russell (1979) realizó un estudio en el que se controlaron algunos de los sesgos identificados. El muestreo representativo de los items incluidos en el estudio quedó garantizado, según el autor, al seleccionar los items que mejor representaran los extremos opuestos de 3 continuos bipolares que la evidencia empírica acumulada indicaba como de mayor relevancia. Los continuos bipolares estudiados fueron los siguientes: 1. placer-displacer; 2. activación-letargo y 3. dominancia-sumisión. Junto a estas escalas también se incluyeron las desarrolladas por Thayer (1967) para medir la dimensión activación-letargo con mayor detalle. Finalmente, fueron incluidos una serie de adjetivos para medir depresión, pues según constató el autor, éste podría ser el opuesto de la escala de Activación General de Thayer (1967). En total se incluyeron 58 items.

Por otra parte, el referente temporal que se le daba a los sujetos era "hoy", con el fin de evitar el sesgo que pueden introducir periodos de tiempo más extensos. Es decir, se les pidió a los sujetos que respondieran a los items formulados bajo el enunciado "hoy me siento....." seguido de los diferentes adjetivos.

Finalmente, se elaboraron cuatro versiones del cuestionario, en las que se variaba la escala de respuesta. Las escalas de respuesta incluidas en el estudio fueron: 1) verdadero-falso; 2) el formato asimétrico empleado por Nowlis y Thayer y descrito en el estudio de Meddis (1972) arriba expuesto; 3) el formato empleado por McNair y Lorr: 1. nada de nada, 2. un poco, 3. bastante, 4. en extremo y 4) el formato empleado por Meddis (1972): 1. no me siento en absoluto(ej. triste), 2. no me siento(triste), 3. me siento un poco(triste), y 4. me siento absolutamente(triste).

Los cuatro cuestionarios fueron administrados a todos los sujetos, en días diferentes y espaciados un promedio de dos semanas entre pase y pase. La muestra estuvo compuesta por 150 estudiantes de ambos sexos. A continuación señalaremos las conclusiones más relevantes enunciadas por Russell (1979). En primer lugar, se constató que el formato de la escala de respuesta tenía efectos sobre las correlaciones entre los items. Comparativamente, los tres formatos alternativos al de Meddis (1972) introducían relaciones positivas entre los items, dificultando la aparición de evidencia en favor de la bipolaridad. No obstante, el tipo de formato de respuesta no presentó efectos significativos sobre las correlaciones obtenidas entre escalas opuestas, y aunque las correlaciones obtenidas con el formato de Meddis (1972) eran mayores que el resto, esta diferencia no resultó estadísticamente significativa.

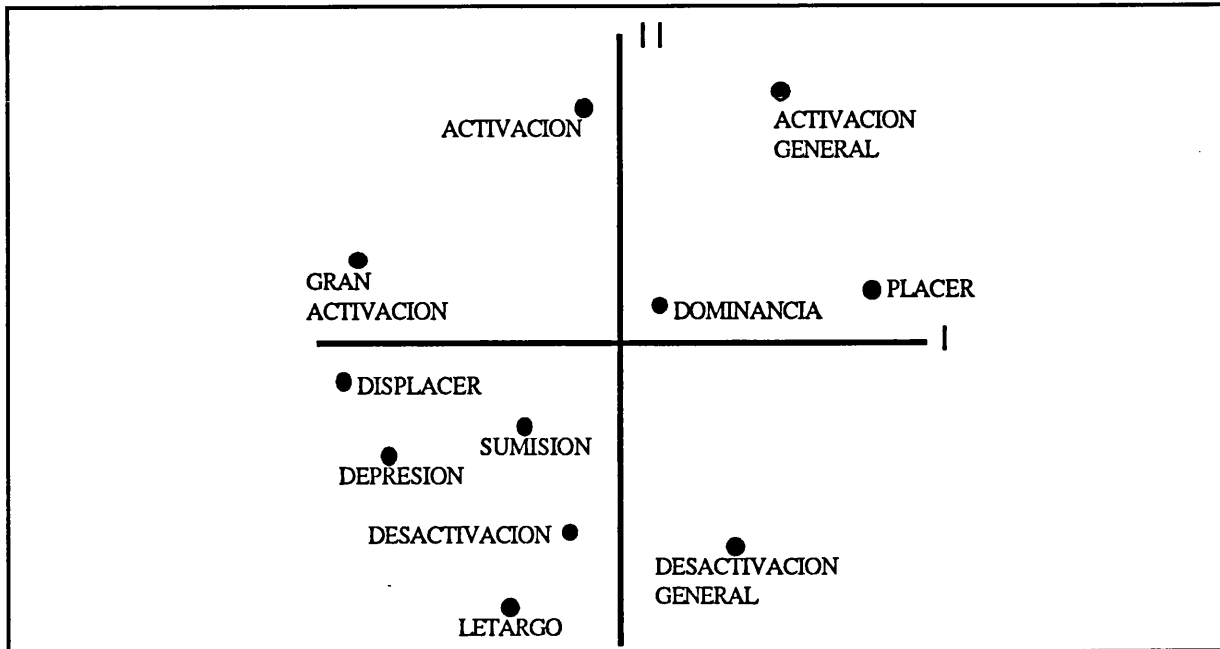
Por otra parte, no todas las escalas mostraron el mismo grado de relación. Consistentemente a través de los distintos formatos de respuesta, las escalas opuestas con las correlaciones negativas más altas fueron las correspondientes al continuo placer-displacer (entre $-.43$ y $-.71$, según el formato) y activación-letargo (entre $-.45$ y $-.62$). Las más bajas aparecieron en la dimensión dominancia-sumisión (entre $.27$ y $-.05$).

Para determinar si el sesgo de aquiescencia introducía variabilidad en contra de la bipolaridad esperada, se obtuvo una medida del efecto de este sesgo sobre cada par de escalas opuestas por el procedimiento empleado por Bentler (1969), y se calculó la correlación parcial consiguiente. Las correlaciones correspondientes a las escalas de placer-displacer y de activación-reposo mostraron un incremento en su magnitud, independientemente del formato de respuesta empleado, alcanzando valores indicativos de la bipolaridad de las mismas (entre $-.71$ y $-.80$ para el primer par y entre $-.66$ y $-.76$ para el segundo). Por el contrario, la eliminación del efecto de la aquiescencia no aportó evidencia en favor de la bipolaridad de las escalas de dominancia-sumisión. Las correlaciones para estas escalas oscilaron entre $.06$ para el formato de respuesta verdadero-falso y $-.19$ para el formato de respuesta de McNair y Lorr (1964). En cuanto a las escalas desarrolladas

por Thayer (1967), si bien la eliminación del sesgo de aquiescencia incrementó la magnitud de las correlaciones observadas entre escalas opuestas, estas no resultaron suficientemente elevadas como para ser consideradas evidencia de la bipolaridad de los continuos medidos. En concreto, las correlaciones más elevadas se obtuvieron para las subescalas de activación general-desactivación (-.66), y de activación general-depresión (-.60), mientras que las más bajas aparecieron entre las escalas de gran activación-desactivación general (-.25). En este sentido, las escalas de Thayer no aparecieron como extremos opuestos de continuos bipolares, tal como Meddis (1972) había indicado.

Sin dar por terminado su estudio en este punto, Russell cambió el enfoque analítico que había seguido hasta este punto por otro más holístico. La siguiente pregunta que se realizó fue ¿cuál es la estructura conjunta que presentan las once escalas? Consecuentemente, sometió las once escalas a un análisis de componentes principales, descubriendo una estructura bifactorial y bipolar. De acuerdo con la configuración espacial obtenida (ver figura 1.1), nombró a los dos componentes principales "placer-displacer" (I) y "grado de activación" (II). La ubicación de las escalas alrededor del espacio definido por estos dos factores ilustraba claramente, de acuerdo con Russell, por qué se había encontrado evidencia a favor de la bipolaridad para las escalas de activación-letargo y placer-displacer, en tanto que no había sido así para el resto: las primeras estarían "cargadas" casi exclusivamente de uno de los dos componentes identificados, en tanto que el resto se ubicaba en el espacio definido en función de la combinación en que incluían estos dos componentes, es decir, en función de la saturación que presentaban en cada componente. La falta de bipolaridad quedaría explicada por el desequilibrio entre la aportación de uno y otro componente a la escala, producida por un inadecuado muestreo de los adjetivos incluidos en las mismas. Así por ejemplo, no se encontraba evidencia empírica de bipolaridad entre la escala de activación general y ninguna de sus dos opuestas, depresión y desactivación, porque la primera de ellas se encontraba desequilibrada en exceso respecto del componente displacer, y la segunda desequilibrada por defecto respecto del mismo componente.

Figura 1.1. Estructura de las escalas analizadas por Russell (1979).



En resumen, la evidencia empírica acumulada en este trabajo sugería que el espacio afectivo era bipolar, y quedaba adecuadamente explicado a partir de la combinación de las dimensiones de placer y activación (aunque el autor no indica el porcentaje de varianza explicada por estos dos componentes). Las distintas dimensiones bipolares del afecto podrían describirse a partir de su ubicación respecto de estos dos ejes. La estructura del afecto así definida se revelaba como carente de "estructura simple", por cuanto que los distintos afectos no se agrupaban junto a los dos ejes mencionados, sino que se ubicaban en torno a su perímetro, en función de la combinación que los definiera. De forma complementaria, esta estructura del afecto daba una explicación a la ausencia de bipolaridad hallada en éste y otros estudios: en la medida en que el conjunto de adjetivos seleccionados para formar escalas de afecto ignorara la dimensión activación, se imposibilitaba la aparición de la bipolaridad existente.

Los prometedores resultados obtenidos en este trabajo llevaron a Russell a realizar un estudio más exhaustivo acerca de la estructura del afecto. La investigación anterior había limitado su alcance al estudio de

la estructura del afecto a través de medidas de autoinforme. Sin embargo, esta no era la única línea de investigación existente respecto del afecto y su estructura. Otros estudios anteriores habían tratado de identificar cuál era la representación mental del afecto, la forma en que categorizamos o representamos cognitivamente los estados afectivos. Además de conocer si existía alguna estructura común a través de los sujetos, en la que estos pudieran ser caracterizados en función de sus afectos -escalamiento de sujetos-, también interesaba conocer si estos emplean alguna estructura o representación mental para categorizar su conocimiento acerca de los distintos estados afectivos -escalamiento de estímulos-. En última instancia, lo que Russell (1980) esperaba es que el estudio desde ambas perspectivas hiciera emerger una única estructura del afecto, común a todos los individuos, tanto para la introspección como para la categorización de la experiencia externa.

En este sentido su trabajo se dividió en dos partes, la primera tenía como objeto de estudio la representación cognitiva del afecto, y la segunda la descripción del afecto propio. Ambas perspectivas contaban con evidencia empírica en favor de una estructura de dos factores bipolares: placer y activación. En este punto recogeremos únicamente la evidencia en favor de la representación cognitiva del afecto que sirvió de punto de partida para Russell, puesto que la referida a las medidas de autoinforme se presentó anteriormente.

El estudio de la representación o categorización de los distintos estados afectivos se había abordado tanto a partir de estímulos verbales como de estímulos no verbales. En los primeros se presentaba al sujeto una serie de caras con distintas expresiones faciales. Así, Scholsberg (1952) pidió a sus sujetos que categorizaran una serie de caras en un conjunto de categorías determinadas por él (sorpresa, disgusto etc.). Los resultados de sus análisis indicaron que los sujetos empleaban una representación circular del afecto alrededor de las dimensiones placer-displacer, atención-rechazo y, posteriormente (1954), letargo-tensión.

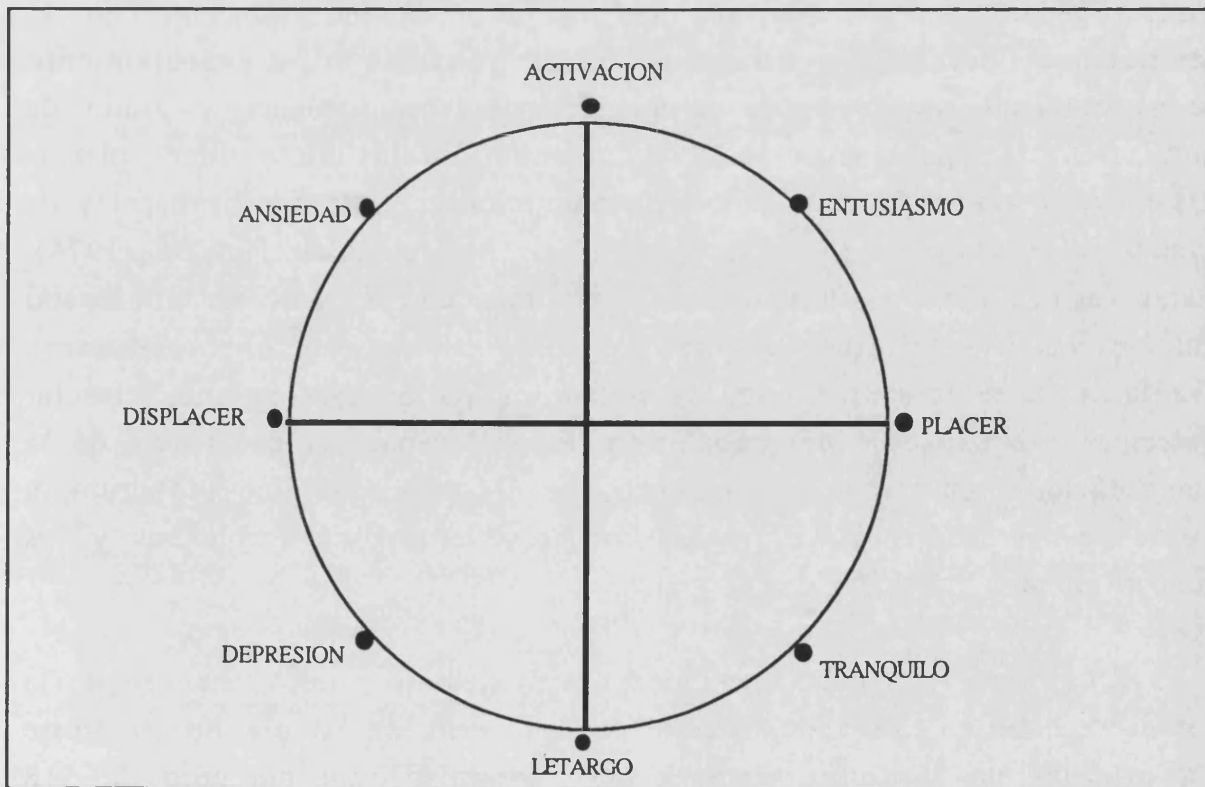
Experimentos posteriores (Abelson y Sermat, 1962; Shepard, 1962; Cliff y Young, 1968) redujeron la tarea experimental a juicios de similitud-disimilitud entre pares de expresiones, evitando la necesidad

de etiquetar los diferentes afectos. Los juicios de los sujetos fueron sometidos a escalamiento multidimensional, obteniéndose una estructura espacial bidimensional, cuyo primer eje era "placer-displacer" y el segundo una combinación de las dimensiones "atención-rechazo" y "tensión-letargo" "básicamente indiferenciables" (Russell, 1980, p. 1162).

Los trabajos que emplearon estímulos verbales también emplearon el escalamiento multidimensional aplicado a distintos adjetivos descriptivos del afecto. Estos estudios apoyaban la existencia de una representación cognitiva del afecto circular y bidimensional, siendo las dimensiones placer y activación las que explicaban la mayor parte de la varianza en las tareas de juicio de similitudes entre adjetivos (Bush, 1973; Neufeld, 1975, 1976; Russell, 1978).

En consecuencia, Russell (1980) desarrolló su modelo circunplejo acerca de la estructura del afecto que aparece en la figura 1.2, bajo la hipótesis de que éste era el modelo empleado tanto en la categorización de la experiencia como en la introspección del propio afecto.

Figura 1.2. Modelo circunplejo del afecto (Russell, 1980).



En base a su modelo, Russell seleccionó un conjunto de 28 adjetivos representativos del dominio del afecto, y los sometió a diversos análisis empleando distintas muestras. En primer lugar, sometió los 28 adjetivos a diferentes técnicas de escalamiento, con el fin de determinar si la representación cognitiva de los sujetos empleada para ordenarlos se ajustaba al modelo propuesto.

La primera técnica empleada fue el escalamiento circular de Ross (Ross, 1938). Los sujetos ($n=36$) tenían que ordenar los 18 adjetivos en 8 categorías: placer, entusiasmo, activación, ansiedad, displacer, depresión, letargo y contento. A continuación se les pidió que ordenaran las ocho categorías en un orden circular. Se evaluó el ajuste de las coordenadas polares teóricamente asignadas a cada uno de los 28 adjetivos con las obtenidas a partir de la ordenación de la muestra. Sobre un índice que oscilaba de 0 a 1, la correspondencia entre las coordenadas observadas y teóricas presentó valores entre .71 y .97, indicando un elevado grado de acuerdo entre sujetos en la ubicación de los adjetivos alrededor del círculo.

A continuación sometió los 28 adjetivos a escalamiento multidimensional ($n=34$) siguiendo un procedimiento basado en la similaridad percibida entre los 28 adjetivos, y a escalamiento unidimensional ($n=51$) en las dimensiones placer-displacer y grado de activación. Las puntuaciones de los adjetivos en las dimensiones placer-displacer y grado de activación se obtuvieron a partir de las escalas de placer-displacer y grado de activación de Mehrabian y Russell (1974). Estas escalas constan cada una de seis items con formato de diferencial semántico. Los resultados obtenidos bajo ambos modelos resultaron similares, y congruentes con los obtenidos en el escalamiento circular anteriormente expuesto. En conclusión, Russell señaló la constancia de la representación observada, y su ajuste con el modelo teórico propuesto, a pesar de las diferencias entre los modelos de medida empleados y los procedimientos de recogida de datos.

Una vez establecida la validez del modelo propuesto respecto de la representación cognitiva del afecto, Russell pasó al estudio de su ajuste en medidas de autoinforme. Para ello, administró un conjunto de 518

adjetivos descriptivos del afecto -entre los que se incluyeron los 28 adjetivos empleados en las tareas de escalamiento- y dos escalas para medir las dimensiones de activación-letargo y placer-displacer (Meherabian y Russell, 1974) a una muestra formada por 343 sujetos.

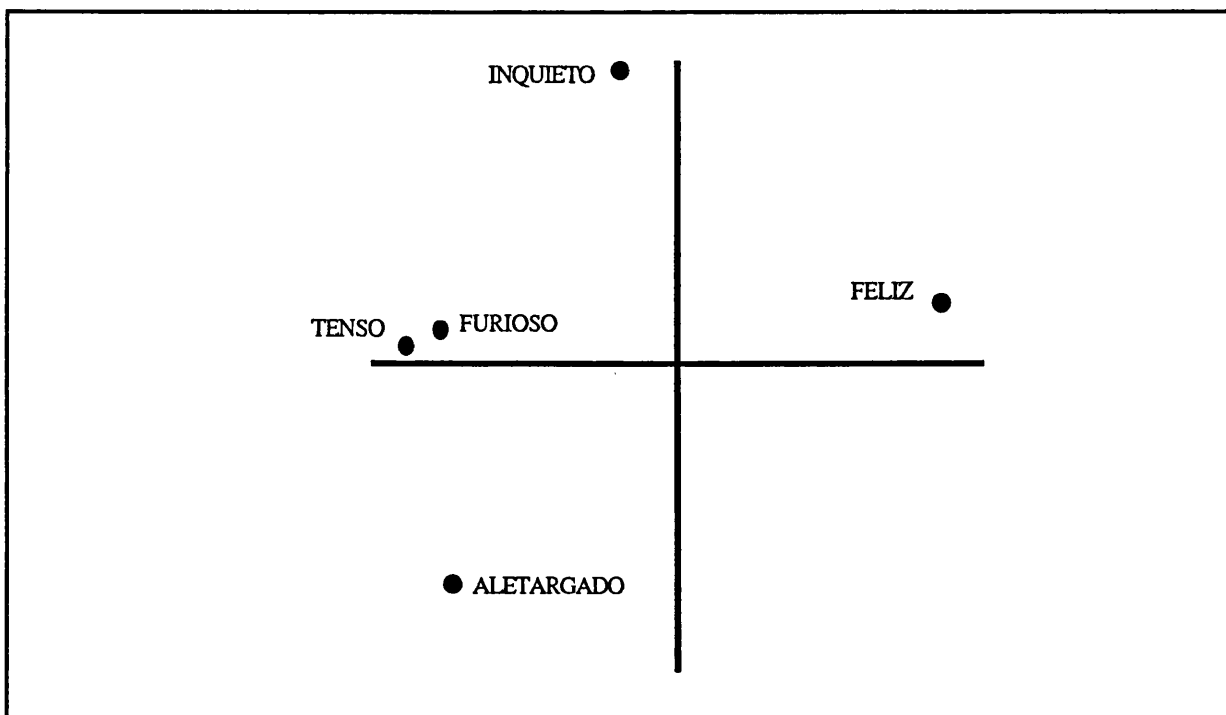
Como en el apartado anterior, Russell sometió los datos obtenidos a distintos tipos de análisis. En primer lugar, regresó cada uno de los 28 items empleados en los análisis de escalamiento sobre las escalas bipolares de activación-letargo y placer-displacer. Tomando los coeficientes de regresión estandarizados como coordenadas, representó gráficamente la posición de cada adjetivo en el espacio bidimensional bipolar formado por las dos escalas. La representación resultante era parecida a la hipotetizada por el modelo, aunque la mayor parte de los adjetivos se agruparon entorno a los extremos del eje placer-displacer. También calculó los coeficientes de correlación múltiple entre las dos dimensiones y cada uno de los adjetivos. Estas correlaciones, aunque estadísticamente significativas ($p < 0.01$, dos colas), no resultaron demasiado elevadas (entre .22 y .62).

A continuación evaluó el ajuste de los datos a un modelo de medida circunplejo. El ajuste de este modelo de medida a un conjunto de datos se evalúa a partir de los resultados de un análisis de componentes principales. Este análisis indica cuáles son los dos componentes de mayor peso subyacentes a los datos. A continuación, se toman estos dos componentes como ejes, y se representa gráficamente la ubicación de cada variable en función de la saturación que presentaba en cada eje. Finalmente, el ajuste se evalúa mediante la inspección visual del diagrama de puntos resultante, en la medida en que el espacio obtenido indique efectivamente una ordenación circular de las variables alrededor de los componentes identificados (Wiggins, 1979). Esta representación resultó similar a la anterior. Tal como indicaba la magnitud moderada de las correlaciones múltiples obtenidas en los análisis previos, ambos componentes sólo explicaron el 45.8% de la varianza total.

Considerando todos los componentes identificados en el análisis de componentes principales anterior, se retuvieron los cinco componentes

con valores propios superiores a 1. No obstante, la adición de los tres últimos componentes no incrementó sustancialmente la proporción de varianza total explicada. En conjunto, los cinco componentes explicaban el 58.9% de la varianza total. La solución de cinco componentes fue entonces sometida a rotación Varimax, obteniéndose finalmente los siguientes factores: feliz-triste, relajado-tenso, aletargado (monopolar), furioso (monopolar), e inquieto (monopolar). Estos cinco componentes fueron sometidos a un análisis de componentes principales de segundo orden. Los dos componentes principales de este análisis explicaron el 70 % de la varianza total. La representación gráfica de la solución obtenida se presenta en la figura 1.3.

Figura 1.3. Resultados del análisis de componentes principales de segundo orden (Russell, 1980).



Los resultados obtenidos muestran que ninguno de los dos acercamientos a los datos, -el circunplejo y el análisis de componentes principales de primer y segundo orden- representan adecuadamente el modelo del afecto hipotetizado. En primer lugar, tres de los cinco factores obtenidos en el análisis de componentes principales resultaron

monopolares. En segundo lugar, aunque los dos componentes principales de segundo orden obtenidos explican una gran parte de la varianza total (70%), hay que considerar que este porcentaje está obtenido sobre la varianza explicada por los cinco componentes identificados inicialmente, es decir, estos dos factores de segundo orden explican el 70% del 58% de la varianza total original (aproximadamente el 40.6% de la varianza total original). Por otra parte, la ubicación de los componentes en este espacio bipolar presenta una estructura simple, y no circular. En ella el primer componente puede ser interpretado como la dimensión activación-letargo, pero el segundo no parece corresponder tan claramente a la dimensión placer-displacer hipotetizada. Esto es así porque los componentes que saturan en él son tenso y furioso, ambos más cercanos a una combinación de displacer y activación, que a displacer únicamente, como indica su posición respecto a los dos ejes. En cuanto al análisis bajo el modelo de medida circunplejo, los dos componentes identificados sólo explicaron el 45.8% de la varianza total.

En definitiva, Russell (1980) concluyó que su modelo circunplejo dejaba una parte sustancial de la varianza observada sin explicar. En este sentido, además de señalar como posibles explicaciones cuestiones metodológicas de carácter secundario, como el error de medida o la presencia de sesgos en las respuestas, cuestionaba por primera vez el propio modelo de medida empleado. El método del análisis de componentes principales asume que las puntuaciones en las variables observables o items son combinaciones *lineales* de las puntuaciones de los sujetos en las dimensiones latentes. Lo mismo ocurre con las técnicas de regresión. Sin embargo, el autor sugiere que ésta no es la única combinación que puede explicar las puntuaciones observadas. En un trabajo anterior (Russell y Mehrabian, 1977), se encontró evidencia de que las puntuaciones observadas de algunos sujetos sí respondían a un patrón de combinación aditiva, en tanto que otros parecían reflejar un patrón de combinación multiplicativa. El método del análisis de componentes principales sólo explicaría la variabilidad producida por el patrón de combinación aditiva, quedando sin explicar cualquier otro patrón de combinación existente.

Volviendo a la cuestión que dió origen a este trabajo cabe preguntarse si se puede afirmar, en vista de los resultados obtenidos, que existe un modelo circunplejo del afecto que organiza tanto nuestra experiencia externa -el modo en que categorizamos e interpretamos la información del entorno- como interna -la introspección e interpretación del propio afecto-. En cuanto a la primera cuestión, el modelo circunplejo propuesto como representación cognitiva del afecto, la evidencia muestra un alto grado de congruencia entre los resultados de los distintos análisis en las distintas muestras. Además, los resultados obtenidos apoyan el modelo propuesto. No obstante hay que señalar que el tamaño de las muestras empleadas es claramente insuficiente (34 y 36 sujetos), y cuestiona la generalizabilidad de los resultados obtenidos. En cuanto a la validez del modelo propuesto para explicar la ubicación del sujeto en el espacio afectivo mediante introspección, los resultados no son más esperanzadores. Si bien la muestra sí alcanza un tamaño suficiente, la estructura propuesta no logra explicar el 50% de la variabilidad observada.

En cualquier caso el trabajo de Russell en el campo de la estructura del afecto ha supuesto una importante aportación por su carácter integrador. Su investigación deja abiertas cuestiones sumamente interesantes: ¿existe realmente una representación cognitiva del afecto?, ¿cuál es su estructura?, ¿cuál es la estructura del afecto tal como surge de las medidas de autoinforme?, y finalmente ¿qué relación existe entre la estructura del afecto obtenida mediante métodos de escalamiento de estímulos y la obtenida mediante métodos de escalamiento de sujetos?

Las investigaciones en torno a la estructura del afecto continuaron (Lorr, McNair y Fisher, 1982, Purcell, 1982). Mientras aumentaba el consenso acerca de la representación cognitiva del afecto, eminentemente bipolar y bidimensional, también se incrementaba la polémica acerca del número y tipo de factores necesarios para explicar la estructura del afecto tal como la reflejaban las medidas de autoinforme.

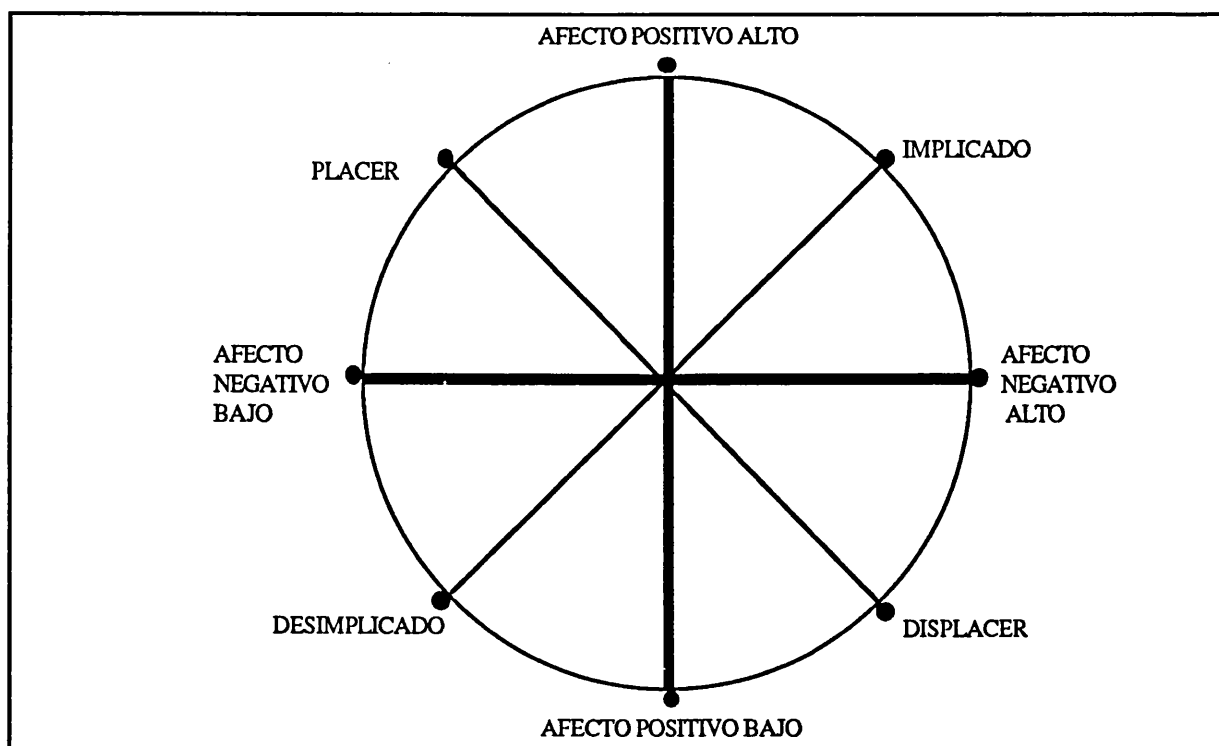
Uno de los trabajos centrados en este último punto es el realizado por Diener (1984). En una recopilación y síntesis acerca de los estudios de bienestar subjetivo, este autor ofrecía una revisión del problema de la conceptualización de la estructura del afecto. El autor situaba su origen en las argumentaciones de Bradburn (1969), quien a partir del desarrollo de la "Affect Balance Scale" concluía que el afecto general tenía dos componentes independientes: el afecto positivo, y el afecto negativo. Esta conceptualización había recibido apoyo empírico por cuanto que las correlaciones entre ambos componentes resultaban insignificantes, a la vez que cada componente presentaba un patrón de correlaciones con otras variables de interés claramente diferenciado. No obstante, sucesivos trabajos en que se replicaron los resultados de Bradburn ofrecieron resultados contradictorios. Tras la revisión mencionada, el autor propuso una nueva explicación, diferenciando intensidad y frecuencia del afecto. De acuerdo con Diener, mientras que ambos afectos tendían a suprimirse, y por lo tanto a relacionarse inversamente en lo referido a la frecuencia con que se experimentaban, también se relacionaban de forma positiva en lo referente a la intensidad con que eran experimentados. En apoyo a esta hipótesis el autor señalaba los resultados obtenidos por Warr, Barter y Brownbridge (1983), quienes tras la reformulación de la escala de Bradburn en términos de frecuencia habían observado una fuerte correlación inversa entre afecto positivo y negativo.

En este estado de cosas surgió un segundo modelo acerca del afecto que, según sus autores, presentaba una estructura capaz de integrar los resultados dispares de la investigación anterior en lo referente a las medidas de autoinforme. El modelo, representado en la figura 4, es el desarrollado por Watson y Tellegen (1985). El carácter innovador de este modelo le llevó enfrentarse tanto a los modelos multidimensionales propuestos hasta entonces, como al modelo circunplejo de Russell. Veamos cuáles eran las argumentaciones que presentaban sus autores para refutar ambas líneas de investigación.

Según los autores, la polémica que sostenían los investigadores partidarios de modelos multidimensionales respecto al número de componentes necesarios para explicar la estructura del afecto (entre 5 y

12) se debía al método de análisis empleado. El análisis de componentes principales junto con el uso del criterio de Kaiser para la extracción de los componentes daba lugar a la extracción de un número demasiado grande de factores, especialmente cuando el número de variables observadas resultaba elevado. Ello explicaría la variabilidad del número de componentes a través de los distintos estudios. Por el contrario, su modelo no pretendía ofrecer una representación exhaustiva de los componentes del afecto, sino clarificar cuáles eran las dimensiones *dominantes* en su estructura. Y de acuerdo con su experiencia, las dos dimensiones que aparecían consistentemente a través de distintas muestras y conjuntos de items eran "Afecto Positivo" y "Afecto Negativo" -en la línea propuesta por Bradburn (1969)-, que explicaban entre la mitad y dos tercios de la varianza total.

Figura 1.4. Modelo del afecto (Watson y Tellegen, 1985).



Respecto del modelo circumplejo de Russell, Watson y Tellegen (1985) señalaron que constituía una rotación alternativa de los ejes identificados por ellos (tal como ilustran las líneas de menor grosor representadas en la figura 1.4). Frente a los ejes propuestos por Russell,

justificaban la superioridad de los propios en la medida en que producían una estructura simple, con la mayor parte de los items localizados en las proximidades de uno u otro eje, y no en su perímetro, como proponía el modelo circunplejo de Russell.

El modelo hipotetizado por los autores resulta innovador en varios aspectos. Las dos dimensiones principales o ejes de la estructura del afecto son identificados como afecto positivo y afecto negativo, pero como se aprecia en la figura, no son los extremos opuestos de un continuo bipolar, sino que son dos dimensiones independientes y monopulares. En este sentido, los autores consideran que *el afecto positivo no viene definido por la mera ausencia de afecto negativo*, y viceversa, por lo que los consideran independientes. Por otra parte, cada una de estas dos dimensiones oscila desde bajo afecto (positivo o negativo) hasta alto afecto (positivo o negativo), de modo que no son conceptualizadas como bipolares. En concreto, la dimensión afecto positivo representa el grado en que una persona "siente entusiasmo por la vida" (Watson y Tellegen, 1985, p. 221), en tanto que la dimensión afecto negativo indica el grado en que una persona experimenta "sentimientos de inquietud". Detallando algo más el contenido de estas dimensiones, los autores las definen como descriptivamente bipolares pero afectivamente monopulares. La explicación a esta paradoja es que los polos de estas dimensiones confrontan la ausencia de afecto (bajo afecto) a la presencia extrema de ese afecto (alto afecto), en lugar de oponer la presencia de un afecto como el entusiasmo a la presencia del afecto opuesto, la depresión. En este sentido son dimensiones afectivamente monopulares o conceptualmente monopulares. No obstante, los estados de bajo afecto o ausencia de afecto suelen describirse no sólo a partir de negaciones de adjetivos que expresan la presencia de ese afecto, sino también afirmando adjetivos que describen su ausencia. Por ejemplo, un sujeto con ausencia de afecto positivo puntuará bajo en items como activo, entusiasmado, y vigoroso, y puntuará alto en items como aletargado, perezoso y amodorrado. De acuerdo con los autores, los seis items medirían la misma dimensión -afecto positivo-, y tanto la negación del primer subconjunto de items como la afirmación del segundo indicarían lo mismo: ausencia de afecto positivo o afecto positivo bajo. En este sentido ambos factores son

descriptivamente bipolares, y cabe esperar la aparición de correlaciones negativas y saturaciones negativas en items que miden cada una de las dimensiones monopolares hipotetizadas, como ocurriría con los tres últimos adjetivos arriba indicados.

Hasta aquí se ha expuesto el modelo hipotetizado por Watson y Tellegen (1985), así como las argumentaciones que ofrecían para justificar su superioridad respecto de los modelos -multidimensionales y circunplejo- propuestos en ese momento. Restaba someter a contrastación empírica la validez del mismo y evaluar su superioridad. Los autores abordaron esta tarea desde la siguiente perspectiva: en primer lugar seleccionaron una serie de trabajos en los que se habían identificado entre 6 y 16 factores, y un trabajo en el que se había identificado la estructura circunpleja propuesta por Russell. Todos los trabajos seleccionados ofrecían los datos necesarios para reanalizar la matriz de correlaciones correspondiente. También incluyeron en el estudio tres trabajos desarrollados por ellos mismos. Si el reanálisis de los trabajos de otros autores de acuerdo con su método -análisis de factores principales en el que se retenían sólo aquellos componentes dominantes- y el análisis de los trabajos propios mostraba de forma consistente la misma estructura de factores principales, y ésta resultaba congruente con su modelo, entonces quedarían empíricamente confirmadas las hipótesis de validez y superioridad mencionadas.

En concreto, los estudios reanalizados fueron los siguientes:

1. Thayer (1967), que realizó un análisis de componentes principales sobre 49 adjetivos descriptivos en el que identificó 16 factores ortogonales;

2. Hendrix y Lilly (1970), que obtuvieron 10 factores ortogonales sobre 44 adjetivos;

3. Borgatta (1961), que realizó un análisis de componentes principales sobre 40 adjetivos descriptivos en el que identificó 8 factores ortogonales;

4. McNair, Lorr y Droppleman (1971), que obtuvieron 4 factores oblicuos sobre 66 adjetivos, y 6 factores oblicuos sobre 65 adjetivos;

5. Russell y Ridgeway (1983), que realizaron análisis de componentes principales sobre dos grupos de 47 y 55 adjetivos descriptivos en los que identificaron 2 factores oblicuos.

En los estudios realizados por los autores analizaron conjuntos de adjetivos descriptivos mediante las técnicas R (sujetos x tests) y P (ocasiones x tests).

Las matrices de correlaciones correspondientes a todos estos estudios fueron sometidas a análisis de factores principales, observándose en todos ellos dos factores que explicaban en conjunto entre un medio y dos tercios de la varianza total, apreciándose un salto importante y consistente en los distintos conjuntos de datos entre el segundo y el tercer factor.

A continuación, los dos factores principales fueron sometidos a rotación Varimax, en busca de una estructura simple. De acuerdo con los autores, en todos los conjuntos de datos se pudo identificar un factor de afecto positivo y otro de afecto negativo. En general, se observó que el contenido de los factores resultaba similar a través de los distintos estudios, aunque se apreciaron diferencias en el número de adjetivos que definían los polos representativos de la ausencia de afecto positivo y negativo (se retuvieron aquellos adjetivos con saturaciones iguales o superiores a 0.25, en valores absolutos). Por ejemplo, en los datos de Borgatta (1961), el factor afecto positivo quedaba representado por 10 items en su polo alto, y por 5 en el polo bajo. A su vez, el factor afecto negativo quedaba representado por 10 items en el polo alto y por sólo 2 en el bajo, algo que también sucedía en el factor afecto negativo del estudio de McNair et al. (1971).

Para evaluar de forma más objetiva la convergencia entre los factores definidos sobre cada conjunto de datos, se calcularon las correlaciones entre las saturaciones de los adjetivos comunes a los distintos estudios. Las correlaciones observadas entre las saturaciones obtenidas en los factores de afecto positivo resultaron elevadas -la mayor parte por encima de 0.90-, y sólo una de ellas resultó inferior a 0.80. La convergencia entre las saturaciones de los factores de afecto



negativo fue también satisfactoria -la mayor parte por encima de 0.80-, y sólo 4 correlaciones resultaron inferiores a 0.63. Finalmente, las correlaciones entre los factores de afecto positivo y negativo resultaron menores, y de signo negativo, aunque demasiado elevadas para ser calculadas sobre items que pertenecían a factores supuestamente independientes: de un total de 81 correlaciones calculadas, 50 resultaron mayores -en valores absolutos- que 0.60.

Finalmente, las soluciones factoriales originalmente obtenidas por McNair et al. (1971), en las que se identificaban 4 y 6 factores oblicuos respectivamente, fueron sometidas a un análisis factorial de segundo orden. En ambas soluciones aparecieron dos factores de segundo orden ortogonales identificables como afecto negativo y positivo. Utilizando un procedimiento desarrollado por Hendrixson y White (1966), se estimaron las saturaciones de los adjetivos en estos factores de segundo orden. Las correlaciones entre las saturaciones obtenidas en el factor de primer orden afecto positivo -obtenido por el reanálisis de la matriz de correlaciones original- y el de segundo orden -obtenido a partir del análisis de la solución factorial original- oscilaron entre 0.92 y 0.99. De forma similar, las correlaciones entre las saturaciones correspondientes a los factores de afecto negativo de primer y segundo orden oscilaron entre 0.94 y 0.99. De acuerdo con los autores, ello indicaba que las dimensiones afecto positivo y negativo no sólo eran las dimensiones dominantes en la estructura del afecto, sino las dimensiones genéricas del mismo.

En resumen, los autores señalaron la consistencia de la evidencia empírica obtenida en favor del modelo propuesto, a pesar incluso de haber empleado conjuntos de adjetivos en que no se controlaron algunos de los problemas metodológicos señalados por otros autores como la disparidad entre adjetivos de diferentes estudios, la asimetría de las escalas de respuesta y el sesgo de aquiescencia. De acuerdo con los autores, estos resultados ponían de manifiesto no sólo la validez del modelo, sino el grado en que era robusto frente a los diferentes problemas mencionados.

La obtención de una solución con dos factores dominantes sobre conjuntos de datos que habían dado lugar a soluciones multifactoriales dispares entre sí apoyaba la hipótesis de Watson y Tellegen (1985) según la cual: 1) la variedad de soluciones factoriales podía ser debida al método de análisis -componentes principales con el criterio de Kaiser para la extracción de componentes-, y 2) una vez desestimadas ciertas fuentes de variabilidad secundarias, los factores dominantes en la estructura del afecto eran sólo dos, afecto positivo y negativo, y como tales, explicaban la mayor parte de la varianza. En este sentido, el modelo propuesto lograba consensuar resultados inicialmente incongruentes. Por otra parte, también se identificaron estos dos factores en el análisis realizado sobre los datos de Russell et al. (1983). De modo que la hipótesis de los autores acerca de que ambos modelos eran dos enfoques alternativos de la misma estructura, en los que variaba únicamente la elección de los ejes principales también obtuvo respaldo empírico. Si ambas conceptualizaciones eran admisibles, quedaba por determinar cuál de ellas era la mejor. En consecuencia, los autores enumeraron una serie de argumentaciones en favor de su conceptualización. En primer lugar, su modelo producía una estructura simple, frente al modelo circunplejo de Russell. En segundo lugar, en estudios anteriores (Zevon y Tellegen, 1982) y en una de las muestras empleadas en su estudio de 1985 se encontró evidencia de que las distribuciones de las puntuaciones factoriales en las dos dimensiones no eran comparables: la distribución de las puntuaciones en el factor afecto positivo era aproximadamente normal, relativamente simétrica y con un amplio espectro de valores moderados; en contraste, las puntuaciones en el factor afecto negativo tendían a permanecer de modo constante en torno a valores bajos. Ello respaldaba la obtención de puntuaciones separadas en afecto positivo y negativo. En tercer lugar, investigaciones anteriores habían mostrado que ambos tipos de afecto, medidos de forma separada, presentaban patrones de correlaciones consistentes y diferenciados con terceras variables (Costa y McCrae, 1980; Tellegen, 1984; Watson y Clark, 1984). Así, mientras que el afecto negativo se relacionaba con el neuroticismo pero no con la extraversión, el afecto positivo presentaba el patrón opuesto. Finalmente, la experiencia clínica también parecía mostrar patrones de relaciones diferenciados,

relacionando el afecto negativo alto con la ansiedad y el afecto positivo bajo con la depresión (Hall, 1977).

El estudio de Watson y Tellegen (1985) presenta un modelo del afecto innovador, que logra, de acuerdo con sus autores, dar explicación a la diversidad de modelos existentes en el momento de su formulación. No obstante, el análisis detallado de su conceptualización revela algunas insuficiencias que ponen en duda las conclusiones arriba presentadas. Los autores consideran que su modelo y el propuesto por Russell constituyen rotaciones alternativas del mismo espacio bifactorial en que puede ser estructurado el afecto. A su vez, definen las dos dimensiones dominantes de este espacio como independientes, y conceptualmente monopulares, siendo éstas afecto positivo alto/bajo y afecto negativo alto/bajo. En este sentido, argumentan que el afecto positivo no es la ausencia de afecto negativo, y distinguen entre presencia (polo alto) y ausencia (polo bajo) de afecto positivo, por una parte, y presencia (polo alto) y ausencia (polo bajo) de afecto negativo, por la otra. No obstante, de ser cierto que ambas son rotaciones alternativas definidas en el mismo espacio, la conceptualización anterior de la dimensión afecto positivo correspondería a la dimensión del modelo de Russell denominada depresión-entusiasmo, y la dimensión afecto negativo se correspondería con la dimensión tensión-relax. Sin embargo, la correspondencia mencionada no es posible porque mientras afecto positivo y negativo son consideradas dimensiones monopulares, depresión-entusiasmo y tensión-relax son conceptualizadas como bipolares. ¿Cuál es la diferencia entre la conceptualización bipolar y la monopolar? La ausencia de depresión tampoco implica la presencia de entusiasmo. Y lo mismo ocurre con la ausencia de tensión. Sin embargo depresión-entusiasmo y tensión-relax no son conceptualizadas como dimensiones independientes, sino como polos opuestos de una misma dimensión bipolar. La diferencia fundamental entre las dos conceptualizaciones se encuentra en el significado del polo "bajo" en un modelo y el significado del polo "negativo" en el otro: mientras Watson y Tellegen conceptualizan el polo "bajo" como la ausencia del afecto que identifica el polo alto, Russell conceptualiza el polo "negativo" como la presencia del afecto opuesto, incompatible al que define el polo positivo.

Realmente existen diferencias de contenido en la conceptualización de las dimensiones pretendidamente equivalentes bajo ambos modelos. Diferencias tan significativas como las mencionadas deberían pues tener alguna proyección sobre la evidencia empírica acumulada bajo uno y otro modelo. Pero no es así. La comparación del significado y patrón de saturaciones de los items adscritos a estas dimensiones bajo ambos modelos es muy similar. Ese es el motivo que llevó a Watson y Tellegen a definir sus dimensiones como "descriptivamente bipolares". Los items descriptivos del polo bajo de cada dimensión presentaban 1) correlaciones negativas y altas con los items del polo alto, y 2) saturaciones negativas en el factor correspondiente. Empíricamente no había ninguna diferencia entre los dos modelos, a excepción de que bajo el modelo circunplejo ambas dimensiones estarían correlacionadas (siendo las dimensiones principales y ortogonales placer y activación), y bajo el modelo de Watson y Tellegen serían independientes.

Por otra parte, la argumentación de Watson y Tellegen tiene como objetivo mostrar la superioridad de los ejes afecto positivo y negativo sobre los ejes placer y activación definidos por Russell por las razones ya mencionadas -estructura simple, patrón de relaciones diferentes, distribuciones diferentes...-, es decir, con fines claramente pragmáticos. En este sentido, Russell reconoce que si bien placer y activación son los ejes principales de la estructura del afecto, es su combinación la que cobra significado, y la que tiene sentido medir, ya que resultaría difícil y poco práctico elaborar escalas puras para medir placer y activación. Es el propio lenguaje el que ya emplea términos para designar afectos que generalmente combinan ambas dimensiones.

En definitiva, parece que la conceptualización de Watson y Tellegen tiene la misma capacidad explicativa de la estructura del afecto que la propuesta por Russell, añadiendo la paradoja de la monopolaridad conceptual-bipolaridad descriptiva. Sin embargo, presenta una ventaja clara respecto al modelo de Russell: no necesita probar empíricamente la bipolaridad "descriptiva" a la que alude, ya que es algo secundario, que puede aparecer en mayor o menor grado, o no aparecer. En su estudio, por ejemplo, aparecen dos conjuntos de datos en los que el polo "bajo" del factor afecto negativo, y en menor medida, del factor afecto positivo,

quedan definidos por sólo dos items. Pero bajo este modelo, ello no plantea problemas de validez para la estructura factorial propuesta. En este sentido, es un modelo mucho más flexible que el modelo circunplejo.

El mayor mérito de este modelo es poner de manifiesto una de las cuestiones más interesantes en el ámbito del estudio del afecto: *el afecto positivo no es la ausencia de afecto negativo, y por lo tanto no cabe esperar una relación directa e inversa entre ellos*. Por consiguiente, los autores concluyen que no son extremos de un continuo bipolar. La extensión de esta premisa al modelo bipolar de Russell lleva a la conclusión de que el placer no es la mera ausencia del displacer, y que el entusiasmo -una dimensión fruto de la combinación de placer y activación- no es la mera ausencia de depresión, y sin embargo sí se espera una relación directa e inversa entre ellos, porque son considerados extremos de un continuo bipolar. La cuestión es: ¿cómo es posible que partiendo de la misma premisa, se llegue a conclusiones completamente contrarias? Más adelante volveremos sobre este punto clave en la conceptualización de la estructura del afecto.

Como se recordará, los primeros trabajos presentados en este estudio fueron agrupados bajo el epígrafe de la "primera etapa" del estudio de la estructura del afecto. Estos estudios se caracterizaron por proponer una estructura multifactorial del afecto, y por iniciar la polémica acerca de la monopolaridad o bipolaridad de esos factores. Posteriormente, hemos presentado los trabajos más relevantes en los que se fundamentaron los dos modelos acerca de la estructura del afecto que vienen dominando este ámbito de estudio desde su publicación. Ambos modelos caracterizarían la "segunda etapa" en el estudio de la estructura del afecto. En esta etapa se logra consensuar el número de factores que definen la estructura básica del afecto. Sin embargo, la naturaleza de las dimensiones propuestas en cada uno -bipolar versus monopolar- continuaba generando polémica. A continuación presentaremos algunos de los trabajos que ilustran cómo el problema de la naturaleza de las dimensiones del afecto continúa vigente en nuestros días. Con ello abordaremos la "tercera etapa" del estudio del afecto.

1.3 La tercera etapa

Esta etapa se caracteriza por: 1) el desarrollo de nuevas escalas relacionadas con el afecto, unas bajo el modelo del afecto de Watson y Tellegen (1985), y otras bajo el modelo de Russell (1980); 2) la aparición de nuevos estudios en los que se continuaba buscando evidencia empírica que validara definitivamente una de las dos estructuras del afecto en discusión; y 3) la proliferación de estudios acerca de la validez factorial de escalas de afecto desarrolladas con anterioridad, generalmente bajo el modelo de dimensiones bipolares.

En cuanto al primer punto -desarrollo de nuevas escalas-, la revisión que presentaremos no es exhaustiva, aunque sí representativa de la coexistencia de los modelos presentados en la segunda etapa. Comenzaremos por la escala de Afecto Positivo y Afecto Negativo elaborada por Watson, Clark y Tellegen (1988). En su estudio, los autores reseñan la multiplicidad de estudios que, partiendo de la estructura monopolar del afecto que ellos defienden, llegan a resultados poco concluyentes e incluso contrarios a la independencia hipotetizada entre el afecto positivo y negativo. En este sentido argumentan que buena parte de estos resultados pudieran deberse nuevamente a problemas metodológicos, y en concreto, a la escasa validez o fiabilidad de las escalas empleadas.

En consecuencia, los autores elaboraron la escala de afecto positivo y afecto negativo (PANAS). Tomando como punto de partida un conjunto de 60 items, y tras someterlos a las técnicas de análisis R (sujetos x tests) y P (ocasiones x tests) (n=267), seleccionaron aquellos adjetivos con saturaciones promedio iguales o mayores a .40 en un factor y menores de .25 en valores absolutos en el otro factor. De este modo se retuvieron 10 items para cada factor -afecto positivo y negativo-, que fueron los que finalmente definieron cada una de las escalas. Las respuestas de los sujetos fueron obtenidas con una escala tipo Likert de cinco puntos en la que los sujetos indicaban el grado -1. nada de nada; 2. un poco; 3. moderadamente; 4. bastante; y 5. mucho- en que habían experimentado esos estados en las últimas semanas. Una vez elaborada la escala, se obtuvieron medidas en seis muestras (n=660, 657, 1002,

586, 649, 663) variando el referente temporal de las escalas: en este momento, hoy, en los últimos días, en las últimas semanas, durante el año y en general.

Los 10 items que definieron la escala de afecto positivo describían alto afecto positivo (entusiasta, interesado, decidido, animado, inspirado, atento, activo, fuerte, orgulloso, y despierto), y los 10 items de la escala de afecto negativo describían todos alto efecto negativo (asustado, temeroso, preocupado, angustiado, nervioso, inquieto, avergonzado, culpable, irritable y hostil). Con esta composición ninguno de los dos factores incluía items con saturaciones y correlaciones negativas. Ambos factores resultaron descriptiva y conceptualmente monopulares. En cuanto a las características psicométricas de las escalas, todas las muestras obtuvieron puntuaciones promedio mayores en la escala de afecto positivo que en la de afecto negativo, independientemente del referente temporal al que contestaron. La consistencia interna de las escalas en las diferentes muestras, evaluada a través del coeficiente alfa de Cronbach, osciló entre .86 y .90 para la escala de afecto positivo, y entre .84 y .87 para la escala de afecto negativo. La correlación entre las subescalas osciló entre -.12 y -.23, indicando su independencia.

Estos resultados, observados a través de diferentes referentes temporales, se mantuvo a niveles estadísticamente no significativos. En consecuencia, la hipótesis de que la relación entre ambas se incrementaba a medida que el referente temporal empleado era mayor no recibió apoyo en este estudio. En cuanto a la validez factorial de las escalas, los autores sometieron en cada muestra los 60 items iniciales a un análisis de factores comunes, tomando como estimaciones de las comunalidades la correlación múltiple al cuadrado. En todas ellas aparecieron dos factores identificables como afecto positivo y negativo, que explicaban entre la mitad y dos tercios de la varianza común. Finalmente, ambas escalas mantuvieron patrones de correlaciones claramente diferenciados con medidas como el componente ansiedad-estado de la escala de ansiedad rasgo-estado (STAI), y el inventario para la depresión de Beck (BDI).

En definitiva, los autores concluyen que todas las pruebas realizadas indican que esta escala reúne unas características psicométricas óptimas para su uso, y responden completamente al modelo del afecto positivo/negativo. El hecho de obtener consistentemente la misma estructura factorial a través de las 6 muestras valida además el modelo hipotetizado.

Indiscutiblemente, la evidencia empírica apoya estas conclusiones, pero cabe señalar una limitación importante, que los autores sin embargo pasan por alto. De acuerdo con el modelo expuesto con anterioridad (Watson y Tellegen, 1985), y como los propios autores señalan en la introducción de este trabajo, los factores que ellos mismos definen presentan una configuración teórica cualitativamente distinta de la que recogen las escalas mencionadas. Teóricamente el factor afecto positivo indica en su polo alto entusiasmo y energía, y en su polo bajo estaría caracterizado por el letargo y la tristeza; por su parte, el factor afecto negativo en su polo alto indica estados aversivos, en tanto que en su polo bajo estaría caracterizado por la calma y la serenidad. Sin embargo, las escalas sólo contienen items representativos del polo alto de cada dimensión.

Desde una perspectiva crítica puede aducirse que el tipo de items que dan forma a cada factor es el mismo que se obtendría si los items se hubieran agrupado en función de su signo, a causa de algunos de los sesgos mencionados anteriormente. La única diferencia entre una estructura factorial con contenido teórico y una puramente definida por el signo de los items es precisamente que la primera aún mantiene items de ambos signos, bajo la hipótesis de que unos miden el polo alto y los otros el polo bajo. En tanto que estas escalas sólo incluyen items de un signo, puede argumentarse bien que sus factores son un artefacto, o bien que carecen de validez de contenido. Inicialmente, el modelo teórico del afecto positivo/negativo soslayó uno de los problemas más difíciles con que se enfrentaba el modelo de factores bipolares: la evidencia empírica de la propia bipolaridad. Con esta escala se soslaya otro de los problemas de la medida del afecto: la configuración de los factores mediante items de ambos signos. De este modo, la evidencia empírica aducida en favor del modelo de afecto positivo/negativo

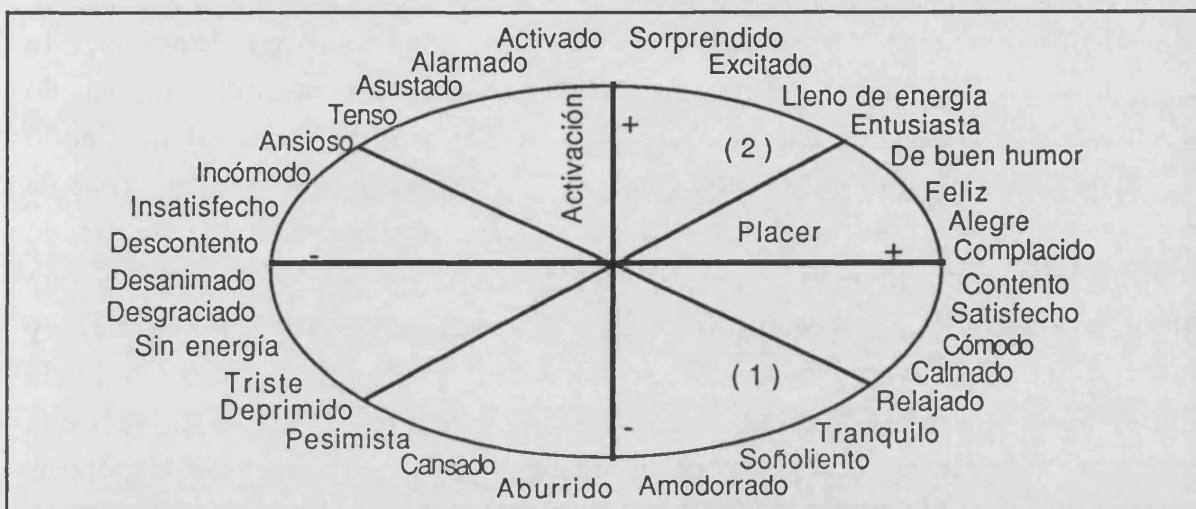
resulta cada vez más indiferenciable de la que se obtendría por efecto del sesgo de aquiescencia, entre otros artefactos. En cualquier caso queda patente que a pesar de lo que los propios autores mantienen, las dimensiones afecto positivo y negativo, tal como las miden estas escalas ya no pueden ser consideradas como análogas a las de depresión-entusiasmo y tensión-relax hipotetizadas bajo el modelo de Russell (1980).

Por lo que respecta a las escalas desarrolladas para medir constructos afectivos bajo el modelo de Russell (1980), algunas de las escalas más representativas fueron las desarrolladas por Warr (1990) para la medida del bienestar afectivo. Warr (1987) conceptualizó el bienestar afectivo como la combinación de las dos dimensiones ortogonales bipolares -placer y activación- identificadas por Russell (1980) en su modelo circunplejo. De este modo, el grado de bienestar afectivo experimentado por una persona dependería de su posición en ambas dimensiones.

Con el fin de obtener medidas del bienestar afectivo así definido, Warr (1987) retomó la idea planteada por Russell según la cual para obtener medidas directamente interpretables del afecto de los sujetos había que centrarse en la combinación de sus dimensiones básicas, y no en las propias dimensiones (ver figura 1.5). En lugar de elaborar escalas que midieran las dimensiones placer y activación de forma independiente, la medida debía centrarse en las dimensiones depresión-entusiasmo (2), y tensión-relax (1), que eran fruto de la combinación de ambas dimensiones básicas, y que representaban mejor el modo en que se experimenta el afecto. En consecuencia, este autor elaboró dos escalas para medir las dimensiones bipolares tenso vs. calmado, y deprimido vs. entusiasmado, cuya combinación podría interpretarse como la ubicación de los sujetos en el continuo bipolar bienestar afectivo negativo vs. bienestar afectivo positivo. Las escalas constaron de seis adjetivos cada una, cuyo referente temporal fue "las últimas semanas". La escala de respuesta, simétrica, oscilaba entre "1. nunca" y "6. siempre". Cada una de las escalas balanceaba el sentido de los items, de modo que cada

dimensión estaba representada por tres adjetivos típicos de un polo, y otros tres típicos del polo opuesto (negativo). La selección de los adjetivos incluidos en las escalas se realizó a partir de una serie de pre-tests de los que el autor no da más información.

Figura 1.5. El enfoque bidimensional del bienestar afectivo (Warr, 1987).



La evaluación de la dimensionalidad de las escalas (n=1686) se realizó mediante el análisis de componentes principales de la matriz de correlaciones parciales -de acuerdo con el procedimiento empleado por Lorr et al. (1989), expuesto más adelante- para soslayar el efecto de sesgos de respuesta como el de aquiescencia. La solución inicial fue sometida a una rotación Varimax. La estructura de las escalas mostró los 2 componentes bipolares esperados. Además, también se realizaron análisis de componentes principales de los 6 ítems "negativos" por una parte, y de los seis ítems "positivos" por otra, como procedimiento alternativo para evaluar la dimensionalidad de los ítems controlando los posibles sesgos mencionados. En concreto, estos análisis se realizaron bajo la hipótesis siguiente: si las repuestas de los sujetos estuvieran contaminadas por el sesgo de aquiescencia, los seis ítems que definían los polos positivos de las dos escalas saturarían en un único componente, en lugar de saturar en las dos dimensiones hipotetizadas, y lo mismo ocurriría con los ítems de los polos negativos. De nuevo se obtuvieron

dos componentes en cada análisis, correspondientes a las dimensiones depresión y tensión, y a las dimensiones entusiasmo y relax.

No obstante, el autor no indica el porcentaje de varianza total explicado por los componentes identificados en ninguna de las soluciones analizadas. Puesto que más adelante realizaremos un análisis exhaustivo de estas escalas, reservaremos para ese momento el resto de comentarios pertinentes a estas escalas.

Como indica el desarrollo de las escalas mencionadas, la configuración de la estructura del afecto continuaba siendo objeto de polémica, y así lo muestran también los trabajos que continuaban siendo publicados en los que el objetivo no era la creación de escalas, sino la propia validación del modelo del afecto. Lorr, quien ya había realizado estudios acerca de la estructura del afecto (Lorr y col. 1982), en los que había encontrado evidencia de una estructura multidimensional y bipolar, realizó un nuevo estudio (Lorr, Shi y Youniss, 1989) bajo la hipótesis de que la estructura del afecto constaba de 5 o 6 factores bipolares de primer orden. Tras el estudio de los conjuntos de ítems empleados en estudios anteriores -tanto bajo el modelo de factores bipolares como monopolares- seleccionó un conjunto de 60 ítems para describir los siguientes estados afectivos bipolares: relajado-tenso, agradable-malhumorado, animado-deprimido, seguro-inseguro, vigoroso-cansado y activo-apático. De los 10 ítems que conformaron cada escala teórica, 5 eran representativos del polo positivo y 5 del polo negativo. El referente temporal de la escala era "justo ahora", y la escala de respuesta presentaba cuatro alternativas: 1. no me siento en absoluto; 2. no me siento; 3. me siento un poco; y 4. me siento completamente.

Para controlar el sesgo de aquiescencia, los autores calcularon una medida de aquiescencia que resultó de la suma de los 60 ítems en su formulación original -es decir, sin invertir ningún ítem-. A continuación se calculó la correlación parcial entre cada par de ítems eliminando el efecto de la medida de aquiescencia. La matriz de correlaciones parciales constituyó el input para un análisis de componentes principales en el que el "scree test" indicó la existencia de 6 componentes. Estos componentes fueron sometidos a un rotación oblícua, de la que resultó eliminada un componente. Los cinco restantes explicaron el 53% de la

varianza total. De forma independiente se procedió a realizar otro análisis de componentes principales en que sólo se retuvieron los dos primeros componentes, que fueron sometidos a una rotación varimax, tal como indican Watson y Tellegen (1985). El primer componente identificado explicó el 27.8% de la varianza total. y el segundo el 10.5%.

La composición de los 5 componentes identificados en el primer análisis resultó sustancialmente diferente de la hipotetizada: sin entrar directamente en la definición de cada uno de ellos, basta con indicar que sólo cuatro de ellos presentaron items con saturaciones de ambos signos -lo que condujo a su interpretación como factores bipolares- pero de distribución muy desequilibrada -en contra de las dimensiones teóricas-: el primero constaba de nueve adjetivos para el polo negativo y sólo tres para el polo positivo; el segundo estuvo compuesto por 10 items en su polo negativo frente a 4 en el positivo; el tercero constaba de 10 items en el polo positivo y de sólo tres en el negativo, y el cuarto presentaba 2 items en el polo positivo y 7 en el negativo. El componente restante estaba enteramente compuesto por items positivos.

La composición de los dos componentes ortogonales identificados en la segunda solución era eminentemente bipolar, y los autores los asimilaron a las dimensiones placer y activación propuestas por Russell (1980). En este sentido atribuyen la emergencia de la bipolaridad a la eliminación del sesgo de aquiescencia. No obstante, hubiera sido interesante contar con la estructura factorial no rotada, en el sentido de que es la aproximación defendida por Russell en su modelo circunplejo. Por otra parte, existe una cuestión relativa a este trabajo que merece alguna mención. Tras el estudio de Bentler (1969), quien desarrolló la idea de elaborar una medida del efecto de la aquiescencia para poder eliminar su efecto, otros autores, como por ejemplo Russell (1979), han empleado esta estrategia para depurar las relaciones observadas entre subescalas opuestas. Sin embargo todos ellos (al menos aquéllos de los que tenemos constancia) emplearon el mismo procedimiento que dió a conocer Bentler para calcular la medida de la aquiescencia y eliminar su efecto. Este consistía en sumar todos los items (sin invertir ninguno de ellos) excepto aquellos incluidos en el par de escalas opuestas para el que se iba a emplear, y parcializar esta medida de la correlación entre el

par de subescalas pertinente. Este proceso se repetía para cada par de escalas opuestas estudiadas. Sin embargo, tanto la medida del efecto de aquiescencia como el procedimiento para su eliminación empleados en este estudio fueron diferentes, tal como se describen más arriba. Más adelante se retomará este punto, y se mostrarán las implicaciones de estos procesos de depuración de las relaciones observadas sobre la estructura factorial resultante.

En definitiva, parecía que la hipótesis que atribuía la falta de bipolaridad en la estructura factorial de las medidas de afecto a problemas metodológicos como el sesgo de aquiescencia, la asimetría de la escala de respuesta y las respuestas centradas en la intensidad y no en la frecuencia con que se experimentaban los estados afectivos muestreados, resultaba paulatinamente reforzada por la evidencia empírica acumulada.

Con todo, el empleo de escalas desarrolladas para medir componentes afectivos bajo el modelo bipolar -las de utilización más frecuente- continuaba generando resultados incongruentes con el modelo teórico correspondiente, como correlaciones demasiado elevadas entre sí que amenazaban su validez discriminante, distribuciones para los items "negativos" sustancialmente distintas de las que presentaban los items "positivos", y estructuras factoriales inadecuadas. Junto a ello, se produjo la aparición y difusión del análisis factorial confirmatorio, que permitía dar un nuevo enfoque al estudio de la validez factorial de estas escalas. Todo ello dió lugar a numerosos estudios, algunos de los cuales presentaremos a continuación. Con ello finalizaremos la tercera etapa en el estudio de la estructura del afecto.

Uno de los estudios acerca de la validez de escalas de uso extendido para la medida del afecto fue el desarrollado por Gotlib y Meyer (1986) sobre la escala Multiple Affect Adjective Check List (MAACL; Zuckerman y Lubin, 1965). MAACL es una escala desarrollada para medir tres de los afectos negativos con mayor relevancia en la práctica clínica: la ansiedad, la depresión y la hostilidad. De acuerdo con su autor, cada uno de estos afectos constituye un factor bipolar independiente o moderadamente relacionado con el resto. La escala en

realidad consta pues de tres escalas sobre las que se obtienen tres puntuaciones independientes. Estas escalas están formadas por aquellos adjetivos que discriminaron máximamente entre pacientes psiquiátricos con ansiedad y sujetos normales, entre sujetos deprimidos y normales, y entre un grupo en el que se indujo por hipnosis un estado de hostilidad y un grupo normal. En total se retuvieron 21 adjetivos para la escala de ansiedad, 40 para la escala de depresión y 28 para la escala de hostilidad. La escala de respuesta empleada sólo tiene dos alternativas: 1) si el sujeto considera que un adjetivo le describe, contesta marcando con una cruz el mismo; 2) si considera que no le describe, entonces deja en blanco el espacio destinado a la cruz. La puntuación en cada escala se obtiene sumando los adjetivos negativos marcados por los sujetos y los adjetivos positivos en blanco. La descripción y naturaleza de los factores hipotetizados por su autor permite la clasificación de esta escala bajo el modelo propuesto por Russell (1980).

Gotlib y Meyer abordaron el estudio de la validez de esta escala tras constatar en diversos estudios que las correlaciones entre las tres puntuaciones eran tan elevadas que cuestionaban la validez discriminante de los factores propuestos. Aplicaron la escala a una muestra de 475 sujetos, y, tras confirmar la estrecha relación que existía entre las tres escalas (las intercorrelaciones oscilaron entre 0.72 y 0.85), sometieron los datos a un análisis de componentes principales con rotación Varimax. El examen de la solución obtenida mediante un "scree test" indicó que sólo dos de los componentes debían ser retenidos y rotados. Estos componentes explicaban el 30% de la varianza total. La solución final mostraba un primer componente completamente definido por items negativos, y un segundo componente completamente formado por items positivos (en ambos casos se consideraron los items con saturaciones iguales o mayores que .40). La adscripción original de los items a las escalas de ansiedad, depresión y hostilidad no introducía diferencias: en los dos factores identificados aparecían aproximadamente el mismo número de items adscritos a estas escalas. Las implicaciones que extraen los autores de estos resultados son claras: aunque la escala había sido diseñada para obtener tres puntuaciones -ansiedad, depresión y hostilidad-, de modo que puntuaciones bajas indicaran el polo positivo de cada uno de estos afectos, los sujetos no estaban

respondiendo bajo esta lógica, y no podían extraerse esas inferencias a partir de sus puntuaciones. Los resultados mostraban claramente que las respuestas de los sujetos a ítems del polo positivo de cada dimensión teórica no se relacionaban con sus respuestas a ítems del polo negativo de la misma dimensión, sino con el resto de ítems positivos, independientemente de la dimensión teórica que supuestamente midieran.

Gotlib y Meyer aducen dos explicaciones a estos resultados. La primera consiste en admitir que afecto positivo y negativo son dimensiones independientes, *en lugar de mutuamente excluyentes*, como parte de la evidencia empírica anterior mostraba. No obstante, los autores señalan que a diferencia de la mayor parte de los estudios en los que aparecían estos dos factores, el referente temporal empleado en este estudio era "actualmente", lo que incrementaba la reticencia a aceptar que los sujetos realmente pudieran estar experimentando afecto negativo y positivo simultáneamente. Por otra parte, los factores identificados aquí no presentan la misma configuración que los identificados en el estudio de Watson y Tellegen (1985). El factor afecto positivo y el factor afecto negativo descritos por estos autores están configurados en su polo bajo por ítems descriptivos de estados de melancolía, fatiga y tristeza, en el primer caso, y por estados placenteros y de relax en el segundo. En el presente trabajo, los ítems que saturan en uno y otro componente son enteramente descriptivos de afectos positivos, en un caso, y negativos en el otro, como si estos componentes sólo constituyeran el polo alto de los definidos por Watson y Tellegen. Sin embargo, esta explicación no se sostiene cuando se observa que algunos de los ítems representativos del polo bajo de los factores definidos por Watson y Tellegen se encuentran incluidos también en estas escalas, aunque con un patrón de saturaciones diferentes.

La segunda explicación que plantean los autores es que los adjetivos positivos y negativos estén afectados por algún sesgo de respuesta que actúe de forma diferencial sobre ellos. En este sentido los autores señalan un sesgo diferente al de la aquiescencia. En concreto, argumentan que si existieran diferencias individuales en el umbral a partir del cual se marcaran los adjetivos negativos pero no los positivos,

el resultado sería la atenuación de las correlaciones entre las subescalas positivas y negativas de cada dimensión original -depresión, ansiedad y hostilidad-. Este sería el caso para dos sujetos que con el mismo nivel de infelicidad presentaran la misma tendencia a dejar sin marcar el ítem "feliz", pero difirieran en su tendencia a marcar el ítem "triste".

Otros estudios realizados sobre la validez factorial de la escala de Orientación Vital (Life Orientation Test, Scheier y Carver, 1985), que es una medida de disposición al optimismo y otras dos escalas para medir ansiedad y depresión realizados por Mook y sus colaboradores (Mook, Kleing y Ploeg, 1991; Mook, Ploeg, y Kleing, 1992; Mook, Kleing y Ploeg, 1992) ofrecieron resultados análogos a los del estudio anteriormente comentado. Las escalas estudiadas habían sido desarrolladas para medir dimensiones bipolares, y sin embargo la evidencia empírica mostraba que: 1) tras invertir los ítems correspondientes, las distribuciones de las puntuaciones de los sujetos en la subescala negativa -calculada como la suma y promedio de los ítems representativos del polo negativo de cada dimensión- mostraban medias significativamente mayores y varianzas menores que las obtenidas en la subescala positiva; y 2) el análisis de componentes principales, sometidos a rotación tanto oblícua como varimax, mostraba un componente en el que saturaban exclusivamente los ítems del polo negativo, y otro configurado exclusivamente por ítems del polo positivo de las dimensiones teóricas. Además esto ocurría tanto para medidas desarrolladas para medir afectos positivos (optimismo), como negativos (ansiedad y depresión).

La interpretación de estos resultados que realizaron los autores, a diferencia de Gotlib y Meyer (1986), no aludía a problemas metodológicos o sesgos de respuesta, sino que argumentaban que los factores identificados eran conceptualmente interpretables. En concreto, los autores indicaron que el componente que agrupaba todos los ítems positivos en la escala Orientación Vital se interpretaría como "presencia de afecto positivo", y que el componente que agrupaba los ítems negativos como "ausencia de afecto negativo", en línea con el modelo teórico del afecto de Watson y Tellegen (1985). La cuestión que plantean los autores es ¿por qué cuando un sujeto responde que raras veces está

deprimido se interpreta que casi siempre está contento? Su respuesta es que tal como indican sus resultados, esta interpretación es inadecuada, porque la ausencia de afecto positivo no implica la presencia de afecto negativo, como asumen las escalas desarrolladas bajo el modelo de afecto bipolar. En consecuencia, la inversión de los items "negativos" para obtener una única puntuación correspondiente al conjunto de items positivos y negativos sería cuestionable.

El desarrollo del análisis factorial confirmatorio (Jöreskog, 1969, 1971) permitió una nueva aproximación al problema. La diferencia fundamental que existe entre el análisis factorial exploratorio y el análisis factorial confirmatorio consiste en que en este último se especifica *a priori* un modelo en el que se indica las relaciones que se hipotetizan entre las variables observadas y los factores latentes, y las relaciones que éstos mantienen entre sí. En consecuencia, el salto cualitativo es enorme, pues permite *contrastar* si un modelo de medida previamente hipotetizado representa adecuadamente las relaciones observadas entre los items, en lugar de *explorar* su estructura subyacente.

En este sentido, Sevastos, Smith and Cordery (1992) aplicaron este método al estudio de la estructura factorial de las escalas tenso vs. calmado y deprimido vs. entusiasmado desarrolladas por Warr (1990) anteriormente descritas. Los resultados del análisis factorial confirmatorio, a pesar de utilizar la matriz de correlaciones parciales como input, no ofrecieron apoyo a la estructura factorial bipolar hipotetizada. El modelo que hipotetizaba los factores "ansioso-contento" y "deprimido-entusiasta" no representó adecuadamente los datos observados ($\chi^2 = 13614.91$, $gl=53$, $p<.001$; AGFI= .773; RSMR=.067¹; $n=3044$). Aun así, este modelo presentó un ajuste mejor que el ofrecido por un modelo que hipotetizaba un único factor de "bienestar afectivo general" ($\chi^2 = 1074.65$, $g.l.=1$, $p<.001$).

¹ La matriz sobre la que se realizaron estos análisis factoriales confirmatorios no se describe en el artículo, de modo que no es posible interpretar los RSMR obtenidos.

Lloret y González-Romá (1994) realizaron un nuevo estudio de la estructura factorial de las escalas desarrolladas por Warr (1990), de nuevo empleando el análisis factorial confirmatorio, e incluyendo además la escala correspondiente a la dimensión fatiga vs. vigor, desarrollada igualmente por este autor. Considerando que ninguna de las estructuras factoriales previamente analizadas mediante técnicas de análisis factorial confirmatorio habían ofrecido un ajuste satisfactorio (Sevastos et al., 1992), este trabajo evaluaba explícitamente la contribución de los efectos de método (*positive-negative wording*) a las relaciones observadas entre los items, además de la contribución de los factores inicialmente hipotetizados.

Los resultados de la contrastación de los modelos de rasgo -tres factores bipolares- y de método -dos factores: sentido positivo y negativo- ofrecieron un pobre ajuste, e indicaron que el mal funcionamiento de la escala fatiga-vigor podía ser responsable de ello. En consecuencia, Lloret y González-Romá (1994) procedieron a la contrastación del modelo de rasgo -un factor bipolar- y de método -dos factores de sentido- sobre la escala que medía el factor fatiga-vigor. Los resultados que obtuvieron para la escala de fatiga-vigor llevaron a nuevas consideraciones (ver tabla 1.1). El modelo que hipotetizaba que las relaciones entre los items de la escala se explicaban por efectos de método -el sentido de los items- reprodujo adecuadamente las relaciones observadas entre los items de esta escala. No obstante, existía un solapamiento total entre el sentido de los items y el polo del factor que medían, que impedía concluir cuál era la interpretación adecuada de esos factores: sentido positivo y negativo o fatiga y vigor.

Tabla 1.1. Indices de bondad de ajuste para los modelos referidos a la escala fatiga-vigor

| MODELO | χ^2 | d.f. | AGFI | RSMR | $\chi^2/d.f.$ |
|--------------------|----------|------|------|--------|---------------|
| factor de rasgo | 347.49* | 9 | .264 | .227** | 38.61 |
| factores de método | 26 (NS) | 8 | .915 | .075** | 3.25 |

*p<0.001; ** variables observadas estandarizadas; NS= p>0.05, dos colas

Por otra parte, los resultados del contraste del modelo del factores de rasgo y de método sobre las escalas restantes -depresión-entusiasmo y tensión-relax- ofrecieron nuevamente un pobre ajuste a los datos, aunque el modelo de factores de método era comparativamente el que menor desajuste presentaba.

El ajuste del modelo de dos factores de método en la escala fatiga-vigor planteaba la posibilidad de que los denominados factores de método pudieran ser reinterpretados como dos factores conceptuales o de rasgo, separados pero relacionados: fatiga y vigor. Además, esta reconceptualización de la estructura subyacente de las escalas podría explicar también los resultados obtenidos sobre las escalas de ansiedad-contenido y deprimido-entusiasmado. En este caso, ninguno de los dos modelos contrastados representaba adecuadamente la estructura factorial subyacente, de modo que no podía descartarse que ello se debiera a que los factores que medían no eran factores de rasgo bipolares ni factores de método, sino factores de rasgo simples y relacionados.

Con el fin de contrastar esta hipótesis Lloret y González-Romá (1994) añadieron la escala de competencia "Reported Job Competence" (Warr, 1990). Esta escala también constaba de 6 items, tres de los cuales eran items invertidos. La escala de respuesta de estos items era también de tipo Likert, con cinco alternativas que oscilaban desde "1. completamente en desacuerdo" hasta "5. completamente de acuerdo". Puesto que los items de esta escala medían una dimensión monopolar, su inclusión permitía comparar el ajuste de un modelo de factores de método, sentido positivo y negativo, con el ajuste de un modelo de rasgos simples -monopolares-, y con el ajuste del modelo que hipotetizaba un factor bipolar. Para ello se contrastaron tres nuevas estructuras factoriales: a) un modelo que hipotetizaba sólo factores de método (sentido positivo vs. negativo); b) un modelo que hipotetizaba dos factores correlacionados: fatiga-vigor (bipolar) y competencia (monopolar); y c) un modelo que hipotetizaba tres factores simples de rasgo relacionados: fatiga, vigor y competencia.

Tabla 1.2. Índices de bondad de ajuste de los diferentes modelos referidos a la escala Fatiga-Vigor (incluyendo la escala de Competencia).

| MODELO | χ^2 | d.f. | AGFI | RSMR | $\chi^2/d.f.$ |
|-----------------------|----------|------|------|------|---------------|
| factores de método | 289.63* | 53 | .768 | .101 | 5.46 |
| factores de rasgo | 506.42* | 53 | .628 | .174 | 9.55 |
| factores de rasgo (b) | 161.86* | 51 | .862 | .077 | 3.17 |

* $p < 0.001$; (b)= tres factores simples oblicuos.

Los índices de bondad de ajuste (ver tabla 1.2) mostraron claramente que el modelo que hipotetizaba tres factores de rasgo simples era el único que mostraba un ajuste aceptable a los datos observados.

De forma análoga se procedió a la contrastación de los modelos referidos a las escalas de ansioso-contento y deprimido-entusiasmado. Los resultados obtenidos para estos modelos mostraron que el modelo que hipotetizaba factores de rasgo simples en lugar de bipolares era el que ofrecía un mejor ajuste a los datos observados.

Por otra parte, estos autores también constataron un hecho ya detectado en estudios anteriores (Watson y Tellegen, 1985; Watson, Clark y Tellegen, 1988): las distribuciones de los ítems representativos del polo negativo y positivo de cada dimensión presentaron diferencias consistentes entre sí. En concreto, los adjetivos que describían el polo negativo de las escalas presentaban consistentemente puntuaciones medias (una vez invertidos) más altas y varianzas menores que las que presentaban los adjetivos positivos en las tres escalas. Asimismo se observó que las correlaciones ítem-total fueron sistemáticamente mayores para los adjetivos positivos -probablemente debido a su mayor variabilidad-.

Los resultados obtenidos en este estudio son análogos a los obtenidos por Mook y colaboradores (1992). Sin embargo, Lloret y González-Romá (1994) no se limitan a interpretar estos resultados como

indicativos de la independencia del afecto positivo y negativo. En concreto, indican que aunque la bipolaridad de las dimensiones teóricas bajo estudio es claramente cuestionada por la evidencia empírica, ello no conduce *directamente* a la desconfirmación de la aproximación teórica subyacente. Por el contrario, señalan algunas argumentaciones alternativas, como el tipo de escala empleado y el efecto de la deseabilidad social. En primer lugar, aluden al uso de una escala tipo Likert, en lugar de emplear una escala de elección forzada (Girelli y Stake, 1993). En este sentido, los anclajes que proporcionan ambos tipos de escala son sustancialmente diferentes: mientras que las escalas de elección forzada hacen coincidir los anclajes o extremos de la escala con elementos representativos de los dos polos de la dimensión (ej. pensando en las últimas semanas me he sentido: 1. muy tenso, 2. tenso, 3. algo tenso, 4. algo relajado, 5. relajado, 6. muy relajado), las escalas tipo Likert empleadas proporcionan anclajes relativos a uno de los polos de la dimensión (pensando en las últimas semanas, con qué frecuencia se ha sentido tenso: 1. nunca, 2. raras veces, 3. algunas veces, 4. bastantes veces, 5. la mayor parte de las veces, 6. siempre), de modo que se responde de forma independiente a cada polo de la dimensión. En definitiva, ello podría dar lugar a sesgos sistemáticos en la evaluación del sujeto de su posición en el continuo bajo estudio. En segundo lugar, señalan que existe un solapamiento perfecto entre la deseabilidad social (socialmente deseable/indeseable) y la adscripción de los items a uno u otro polo (vigor-fatiga). Tal solapamiento implica que los items que miden el polo fatiga son socialmente indeseables, mientras que aquéllos que miden el polo vigor son socialmente deseables. De forma similar, existe solapamiento entre los polos a los que se adscriben los items y la dimensión afecto positivo/negativo.

En definitiva, este estudio muestra la necesidad de realizar nuevas investigaciones que aclaren las razones por las cuales las escalas estudiadas no miden las dimensiones bipolares pretendidas. ¿Son esas dimensiones simples en lugar de bipolares, o son las escalas inapropiadas para medir las dimensiones bipolares?

CAPITULO 2

**EL REPLANTEAMIENTO DEL PROBLEMA:
DEL CUESTIONAMIENTO DE LA ESTRUCTURA TEORICA AL DE
LA ESTRUCTURA EMPIRICA.**

1. INTRODUCCION

Pese a la multiplicidad de estudios presentados, todavía no se ha llegado a consolidar ninguno de los modelos acerca de la estructura del afecto presentados. Aun así, la aceptación de cada uno de ellos es diferente. El modelo circunplejo de Russell (1980) es el que goza de mayor aceptación. De hecho, si analizamos la evolución que ha seguido el estudio de la estructura del afecto mediante instrumentos de autoinforme, se constata que la pauta en el estudio del afecto la ha marcado el modelo bipolar. Sólo la falta de evidencia empírica a su favor, y la reinterpretación de los resultados obtenidos en los sucesivos trabajos en los que se ha intentado validar, ha dado paso a la aparición del modelo de factores monopolares, y a la independencia entre el afecto positivo y negativo.

Casi podría decirse que la conceptualización del afecto positivo como independiente del negativo se ha desarrollado a la sombra de la conceptualización bipolar, aprovechando las "debilidades" que mostraba para ganar terreno. Además, esto no ha resultado difícil: mientras que el modelo bipolar necesitaba indicadores empíricos muy concretos para apoyar la bipolaridad de sus dimensiones -correlaciones elevadas y negativas entre ítems opuestos y factores definidos por igual número de ítems de uno y otro polo aproximadamente-, el modelo monopolar era mucho más flexible: si las correlaciones entre ítems representativos de alto y bajo afecto positivo aunque negativas, no eran elevadas, y si el factor resultante presentaba casi exclusivamente ítems representativos de alto afecto positivo, ello no invalidaba el modelo, sólo que algunos ítems eran mejores indicadores que otros del factor subyacente. Si, además, los ítems de bajo afecto negativo presentaban saturaciones mayores en el factor afecto positivo, ello tampoco invalidaba el modelo, ya se sabía que existe cierto solapamiento entre los polos bajo y alto de cada factor.

Revisando los trabajos presentados, se observa que el desajuste entre la estructura factorial bipolar hipotetizada y la empíricamente observada ha generado básicamente dos reacciones a través de los

diferentes trabajos: 1) se ha atribuído a la falta de validez de la estructura hipotetizada, en favor de la estructura monopolar; y 2) se ha atribuído a diferentes problemas metodológicos que afectaban a las puntuaciones de la escala y a las relaciones entre los items. Estos problemas "secundarios" eran, según los autores, los que impedían que las relaciones existentes de los items entre sí y respecto de la dimensión medida aparecieran con claridad, eliminando la bipolaridad. Llamamos secundarios a estos problemas no por su menor importancia, sino porque no cuestionaban la bipolaridad, que se asumía, sino que impedían que emergiera.

En este sentido, unos autores atribuyeron este desajuste al tipo de escala de respuesta, que debía de ser simétrica, aunque en otros estudios en que se cumplía este requisito continuaban apareciendo factores monopulares "desdoblados".

Otros estudios, los que más, atribuyeron su desdoblamiento al efecto del sesgo de aquiescencia. No obstante, en la mayor parte de los casos este sesgo no podía explicar el desdoblamiento de los factores bipolares observado. Cuando se factorizaban escalas multidimensionales, generalmente la agrupación de items no se realizaba sólo sobre el sentido en que estaban formulados, sino que también se diferenciaban las dimensiones teóricas a las que pertenecían. No aparecían todos los items formulados para medir el polo positivo de las distintas dimensiones agrupados en un solo factor, sino que cada factor agrupaba un subconjunto de items que se caracterizaba por estar formulado para medir un polo de una dimensión concreta, y así para cada polo de cada dimensión. Evidentemente, un sesgo que actuaba independientemente del contenido de los items no podía producir semejante especificidad. Por otra parte, estudios en los que se había eliminado este sesgo persistían en no obtener la estructura factorial bipolar esperada. Incluso cuando se había empleado un procedimiento que, a parte de eliminar el efecto de este sesgo de las correlaciones entre items opuestos, además también lo eliminaba de las correlaciones entre items del mismo polo (Lorr et al., 1988; Sevastos et al., 1992).

También cabía argumentar la calidad de los items o adjetivos empleados en las escalas. Pero puesto que han sido muchos y muy diversos los adjetivos empleados en las escalas revisadas, y sin embargo han llevado de forma casi constante al desdoblamiento de los factores bipolares propuestos, consideraremos éste como un factor controlado, que afectaba por igual a todos los estudios analizados, y por lo tanto no introducía una fuente de variación de interés para explicar el fenómeno del desdoblamiento.

Finalmente, también se ha estudiado el efecto del sentido de los items sobre la configuración de los factores obtenidos, sin que ello consiguiera explicar el desdoblamiento observado.

A pesar de los sucesivos intentos de explicación del desajuste entre la estructura factorial bipolar hipotetizada y la observada en términos de problemas metodológicos, todavía quedaban algunos argumentos por explorar. Los trabajos más recientes acerca de la validación de distintas escalas de respuesta apuntaban en una nueva dirección: ¿por qué cuando un sujeto respondía que raras veces estaba contento se infería que a menudo estaba triste?, ¿era el modelo bipolar una mala representación de la realidad, o eran las escalas empleadas un mal instrumento para validarlo?

2. ESCALAS EXPLICITAMENTE BIPOLARES Y ESCALAS IMPLICITAMENTE BIPOLARES.

Inicialmente, la medida de dimensiones bipolares se había realizado con escalas de diferencial semántico. En estas escalas la escala sobre la que responde el sujeto constituye un "isomorfismo" respecto de la dimensión que se pretende medir. Al sujeto se le presenta el estímulo al que debe responder, seguido de una escala explícitamente bipolar, cuyos anclajes los constituye el par de adjetivos representativos del dominio de la dimensión, entre los cuales se presenta un continuo de 7 o 9 puntos en el que el sujeto se ubica en función de su posición respecto de ambos anclajes. Así por ejemplo para medir el estado de ánimo de un sujeto se le preguntaría:

DIFERENCIAL SEMANTICO

EN LAS ULTIMAS SEMANAS SE HA SENTIDO:

DEPRIMIDO :.....:.....:.....:.....:.....:.....:.....:.....: ENTUSIASMADO

EL SUJETO RESPONDE SOBRE UN CONTINUO EXPLICITAMENTE BIPOLAR

Estas escalas impedían pues la contrastación empírica de la bipolaridad de la dimensión bajo estudio, ya que estaba explícitamente asumida en el formato de respuesta, y no se podía falsar. En consecuencia, se necesitaba cambiar las escalas empleadas de modo que se pudiera contrastar empíricamente el supuesto de bipolaridad. Las escalas que vinieron a sustituir el diferencial semántico fueron escalas tipo Likert. Estas escalas no presentaban ya explícitamente el continuo bipolar que se asumía para la dimensión a medir. En lugar de ello, los adjetivos representativos del dominio de la dimensión bipolar eran presentados por separado, convirtiéndose en los estímulos a los que respondía el sujeto:

ESCALA TIPO LIKERT

EN LAS ULTIMAS SEMANAS,
INDIQUE LA FRECUENCIA CON QUE SE HA SENTIDO:

ENTUSIASMADO

1. NUNCA 2. 3. 4. 5. 6. SIEMPRE

DEPRIMIDO

1. NUNCA 2. 3. 4. 5. 6. SIEMPRE

EL SUJETO RESPONDE A UN ESTIMULO REPRESENTATIVO UNICAMENTE
DE UN POLO DE LA DIMENSION

La suma y promedio de las respuestas de los sujetos a escalas de diferencial semántico formadas por pares de adjetivos representativos del continuo bipolar bajo estudio se interpretaba como su ubicación en ese continuo bipolar. La suma y promedio de las respuestas de los sujetos a adjetivos representativos de un único polo de la dimensión a medir -en las escalas tipo Likert-, tras invertir las puntuaciones obtenidas en los adjetivos representativos de uno de los polos -generalmente el negativo-, también era interpretada como su ubicación en ese continuo bipolar. Sólo que es este tipo de escalas, los sujetos no emitían sus respuestas en un continuo bipolar. Entonces, ¿cómo pueden interpretarse sus respuestas de ese modo?

Los supuestos en los que se basa el escalamiento tipo Likert indican que cada uno de los items debe estar monotónicamente relacionado con la *dimensión* a medir, y que la suma de las respuestas a los distintos items también debe estar monotónicamente relacionada con la dimensión a medir. De modo que todos los items también deben estar monotónicamente relacionados entre sí.

Las escalas tipo Likert empleadas presentan estímulos relativos a uno de los polos de la dimensión, por lo que se responde de forma independiente a cada polo de la dimensión. De modo que cabe preguntarse bajo qué supuestos adicionales: 1) las respuestas de los sujetos ante un estímulo representativo de un polo de la dimensión pueden considerarse indicativas de su posición en el continuo bipolar en toda su extensión; y 2) las respuestas de los sujetos ante estímulos representativos de un polo pueden transformarse -invertirse-, para obtener la puntuación que el sujeto hubiera dado ante los estímulos correspondientes del opuesto.

Respecto al primer punto, sólo se puede interpretar la respuesta ante un estímulo representativo de un polo de la dimensión como indicativa de la ubicación en el continuo si se *infiere* que el sujeto, al responder, hace coincidir el extremo inferior de la escala con algún punto de un polo de la dimensión, y el extremo superior de la escala con algún punto del polo opuesto. De este modo la suma y promedio de las puntuaciones de los sujetos a lo largo del conjunto de items permite inferir la posición del sujeto en la dimensión bajo estudio, respecto de

los dos polos que la caracterizan, y en relación al resto de los sujetos. Puntuaciones bajas indicarían proximidad al polo negativo, y puntuaciones altas proximidad al polo positivo. Denominaremos a este supuesto particular de la medida de dimensiones bipolares mediante escalas tipo Likert como el supuesto de "bipolaridad de la escala de respuesta".

1. BIPOLARIDAD DE LA ESCALA DE RESPUESTA

ANIMADO

1. NUNCA 2. 3. 4. 5. 6. SIEMPRE

DEPRESION

ENTUSIASMO

AL RESPONDER, EL SUJETO INTERPRETA LA ESCALA DE RESPUESTA
ABARCANDO AMBOS POLOS

Partiendo de la base de que todos los items -positivos y negativos- son respondidos *en la misma escala de respuesta bipolar*, es cuando se puede considerar que además las respuestas dadas a los items positivos y negativos son *complementarias*. Este supuesto es el que permite invertir las puntuaciones obtenidas en los items correspondientes a uno de los polos y promediar después las puntuaciones en todos los items para obtener una puntuación total que indica la posición del sujeto en la dimensión estudiada.

2. COMPLEMENTARIEDAD DE LA ESCALA DE RESPUESTA

ANIMADO

| | | | | | |
|----------------|-------|-------|-------|-------|------------|
| 1. NUNCA | 2. | 3. | 4. | 5. | 6. SIEMPRE |
| DEPRESION..... | | | | | ENTUSIASMO |
| 6. SIEMPRE | 5. | 4. | 3. | 2. | 1. NUNCA |

DESANIMADO

ANTE ITEMS OPUESTOS, EL SUJETO EMPLEA LA MISMA ESCALA CON LOS POLOS INVERTIDOS.

El desarrollo e interpretación de las escalas desarrolladas por Warr para medir las tres dimensiones mencionadas del bienestar afectivo ilustran este aspecto fundamental. Este autor sometió a análisis un gran conjunto de adjetivos característicos de cada uno de los polos de las dimensiones que pretendía medir, a continuación seleccionó un número igual de adjetivos de uno y otro polo, hasta llegar a un total de seis por escala. Los tres adjetivos representativos del polo positivo no eran los opuestos de los tres adjetivos del polo negativo. En lugar de introducir seis ítems referidos al mismo polo, el autor sustituyó tres de los que designaban el polo positivo por sus opuestos. La lógica subyacente es que a efectos de medida, es indistinto utilizar uno u otro polo para medir la posición del sujeto en la dimensión. La ubicación del sujeto en la dimensión bipolar resulta *invariable* al tipo de estímulo, positivo o negativo, que elicitó la respuesta. Se asume que ambos estímulos son intercambiables, complementarios. En definitiva, si preguntamos a un sujeto con qué frecuencia se siente triste, en una escala de 1. "nunca" a 6. "siempre", y contesta 2. "raras veces", asumimos que de haberle planteado la pregunta opuesta, con qué frecuencia se siente alegre, hubiera respondido 4. "la mayor parte de las veces". En este sentido, la ausencia de afecto negativo sí implica la presencia de afecto positivo, ya que se interpreta una puntuación baja en un ítem representativo de

afecto negativo como una puntuación alta en el ítem opuesto, representativo de afecto positivo.

En definitiva, las escalas de diferencial semántico fueron sustituidas por las escalas tipo Likert porque asumían que la bipolaridad existía, e impedían su contrastación empírica. Por su parte, las escalas tipo Likert fueron empleadas bajo la lógica de que:

A. si la dimensión subyacente era bipolar,

B. entonces aparecerían relaciones lineales inversas entre los ítems representativos de polos opuestos, que darían lugar a:

1) correlaciones elevadas y negativas entre ítems opuestos, y

2) factores bipolares, en los que los ítems representativos de polos opuestos presentarían saturaciones con signos opuestos.

Sin embargo, todo ello sólo era posible si se asumía implícitamente que la escala de respuesta era bipolar, y por lo tanto complementaria. Era la bipolaridad de la escala de respuesta, y consecuentemente la complementariedad, lo que permitía hipotetizar relaciones lineales e inversas entre ítems de polos opuestos. La constatación de esta peculiaridad referida únicamente a escalas tipo Likert contrabalanceadas elaboradas para medir dimensiones bipolares lleva a la conclusión de que, cuando se mide un continuo efectivamente bipolar:

A'. sólo si la escala de respuesta es también percibida como bipolar y por lo tanto complementaria

B'. aparecerán las relaciones lineales e inversas entre ítems opuestos que darán lugar a:

1) correlaciones elevadas y negativas entre ítems opuestos, y

2) factores bipolares, en que los ítems representativos de polos opuestos presentarían saturaciones con signos opuestos,

C. que a su vez respaldarán empíricamente las hipotetizadas dimensiones bipolares latentes.

Desde esta perspectiva se observa que para obtener la evidencia empírica que tradicionalmente se ha considerado a la hora de validar una dimensión bipolar, no basta con que la dimensión sea efectivamente bipolar, sino que además la escala debe reunir las condiciones expuestas.

Sólo bajo el supuesto de que la escala de respuesta presentada, aunque explícitamente monopolar, es interpretada por los sujetos como bipolar, y por lo tanto complementaria, pueden esperarse relaciones lineales entre subescalas opuestas. Invirtiendo el razonamiento se deduce que la no aparición de la evidencia empírica esperada en favor de la bipolaridad de la dimensión medida puede deberse a que cualquiera de las dos premisas se incumpla. Sin embargo, ninguno de los estudios anteriormente presentados han considerado la existencia de la premisas relativas a la escala, de modo que la falta de evidencia empírica en favor de la bipolaridad de las dimensiones se atribuía al modelo teórico, considerándose que era una representación inadecuada de la realidad. No obstante, es posible que aún cuando las dimensiones objeto de estudio sean bipolares, si las condiciones mencionadas para la escala no se cumplen, no aparezcan relaciones inversas y lineales entre items opuestos, y por lo tanto no se obtenga evidencia empírica que apoye la bipolaridad de las dimensiones. Pero en este caso la validez del modelo en sí no podría ser cuestionada.

Más aún, la falta de cumplimiento de estos supuestos puede ofrecer una explicación razonable para la estructura de factores monopulares generalmente hallada. Si la escala de respuesta no es complementaria para items de uno y otro polo, entonces cada sujeto tendrá dos puntuaciones distintas para cada dimensión, una sobre el polo positivo, y la otra sobre el polo negativo. Ello implicaría que la escala de respuesta tampoco es bipolar, y por lo tanto, inadecuada para medir factores que sí lo son. No habría fundamento para esperar las relaciones inversas y lineales entre los items opuestos que darían lugar a los esperados factores bipolares. Finalmente, esta doble ubicación de los sujetos sobre cada continuo bipolar explicaría el desdoblamiento de los factores bipolares hipotetizados. Todo ello sin comprometer la validez del modelo teórico objeto de estudio.

En consecuencia, el primer objetivo de esta tesis consiste en contrastar empíricamente el cumplimiento de los supuestos necesarios acerca de la escala de respuesta para la medida de factores bipolares mediante escalas balanceadas tipo Likert.

CAPITULO 3

METODO.

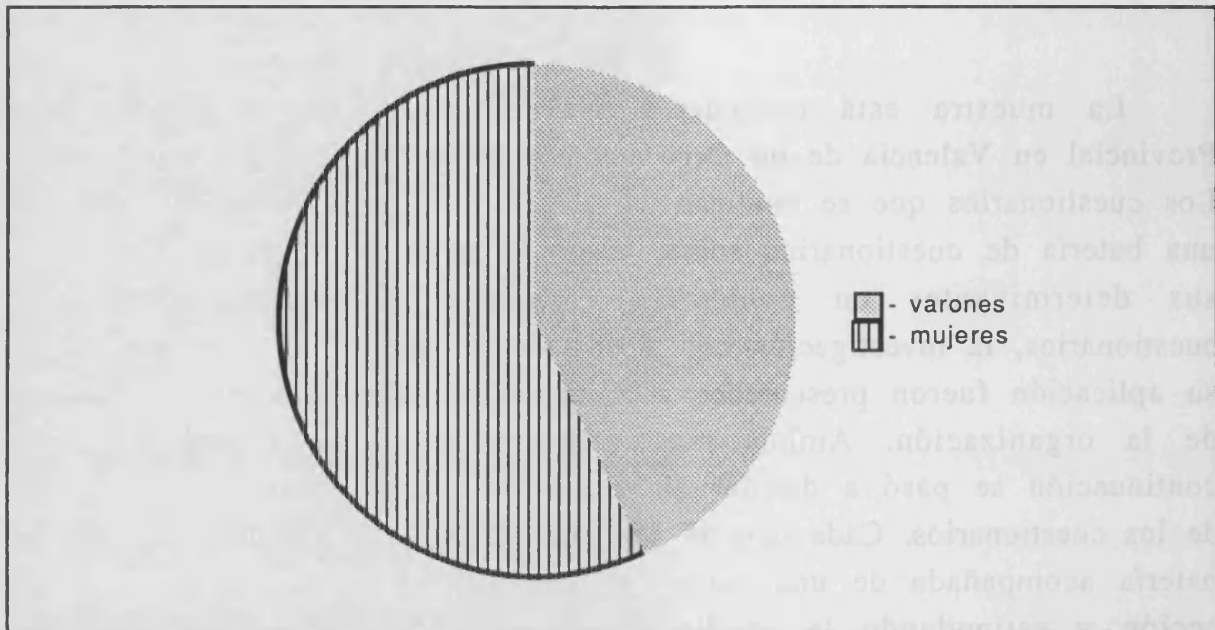


1. PROCEDIMIENTO DE RECOGIDA DE DATOS Y MUESTRA.

La muestra está compuesta por empleados de la Delegación Provincial en Valencia de un Organismo de la Administración del Estado. Los cuestionarios que se analizan en la presente obra formaban parte de una batería de cuestionarios sobre bienestar psicológico, salud mental, y sus determinantes en contextos organizacionales. La batería de cuestionarios, la investigación en la cual se enmarcaba, y el propósito de su aplicación fueron presentados a la dirección y al comité de empresa de la organización. Ambos mostraron su apoyo a la acción, y a continuación se pasó a diseñar el proceso de administración y recogida de los cuestionarios. Cada uno de los empleados recibió un ejemplar de la batería acompañada de una carta del director mostrando su apoyo a la acción y estimulando la colaboración, una carta de presentación del equipo de investigación, y un sobre. Una vez cumplimentada la batería de cuestionarios, ésta fue remitida a la Unidad Técnica de la organización dentro del sobre cerrado. Se entregaron 675 cuestionarios, de los que 298 fueron debidamente cumplimentados y devueltos. Ello implica que la tasa de respuesta fue el 42.57%.

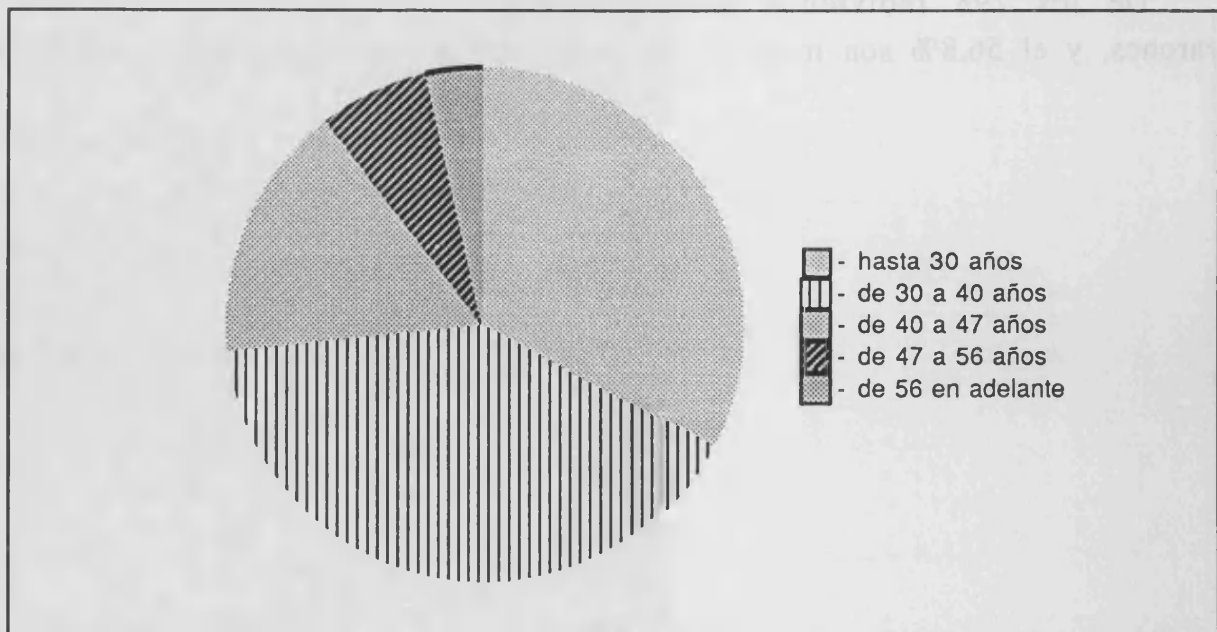
De los 298 individuos que componen la muestra, el 43.1% son varones, y el 56.8% son mujeres, tal como aparece en la tabla 3.1.

Tabla 3.1. Distribución de la muestra por sexo.



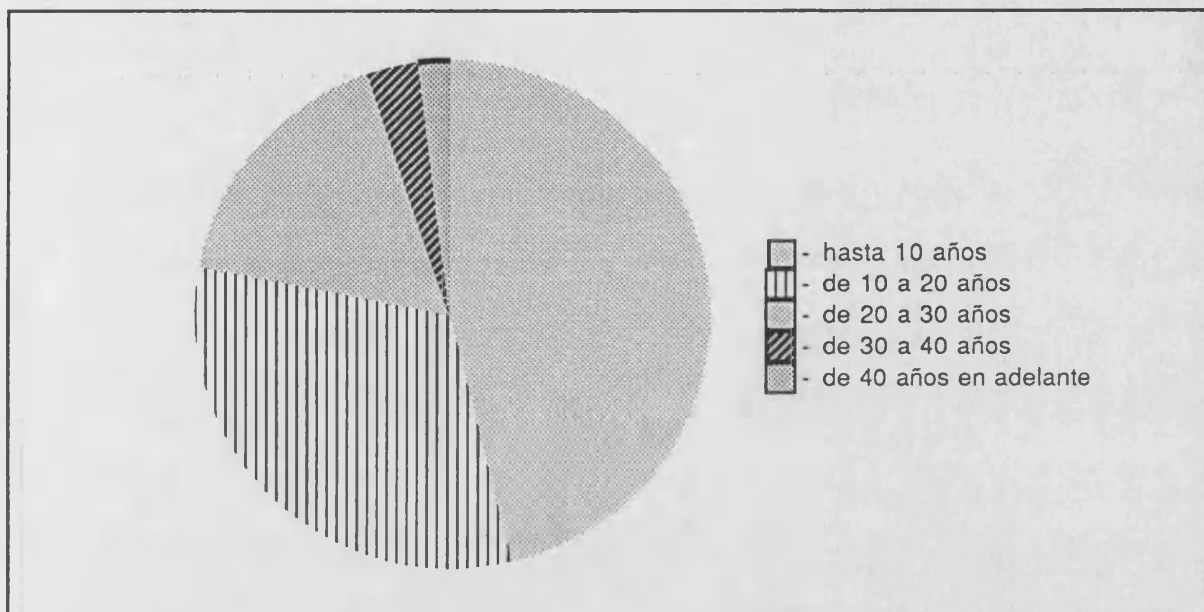
En cuanto a la edad, la media de los sujetos que componen la muestra es de 35.56 años, con una desviación típica de 8.55. La tabla 3.2 muestra la distribución de la muestra por edades.

Tabla 3.2. Distribución de la muestra por edad.



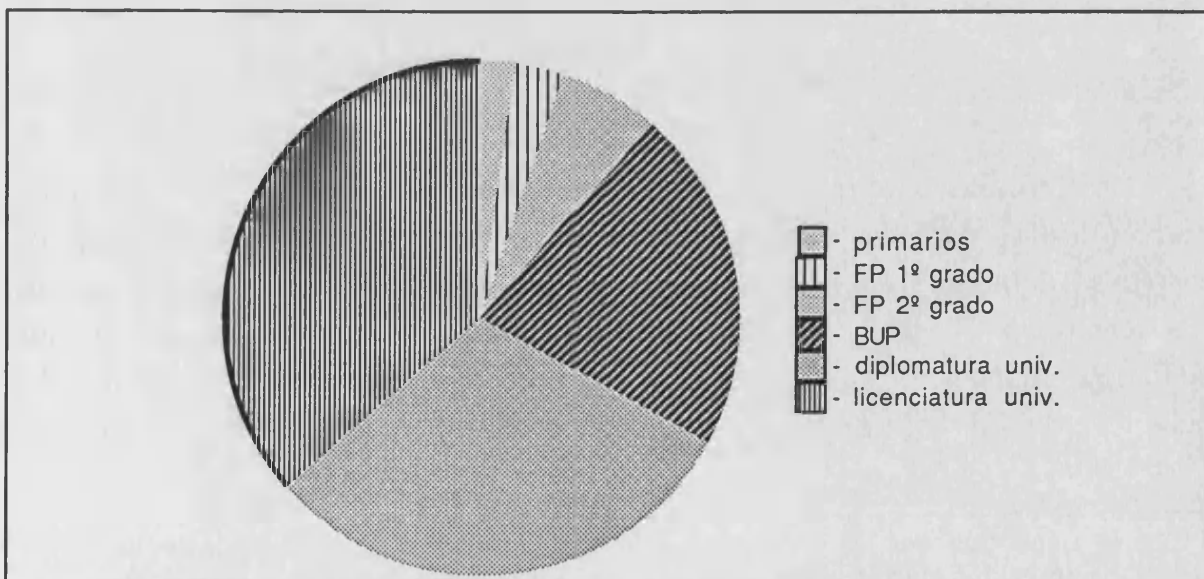
La experiencia laboral de los sujetos que componen la muestra es, como promedio 13.97 años, con una desviación típica de 9.59. La distribución de esta variable en la muestra aparece en la tabla 3.3.

Tabla 3.3. Distribución de la muestra en función de la experiencia laboral.



En cuanto al nivel de estudios, la distribución de esta variable en la muestra aparece en la tabla 3.4.

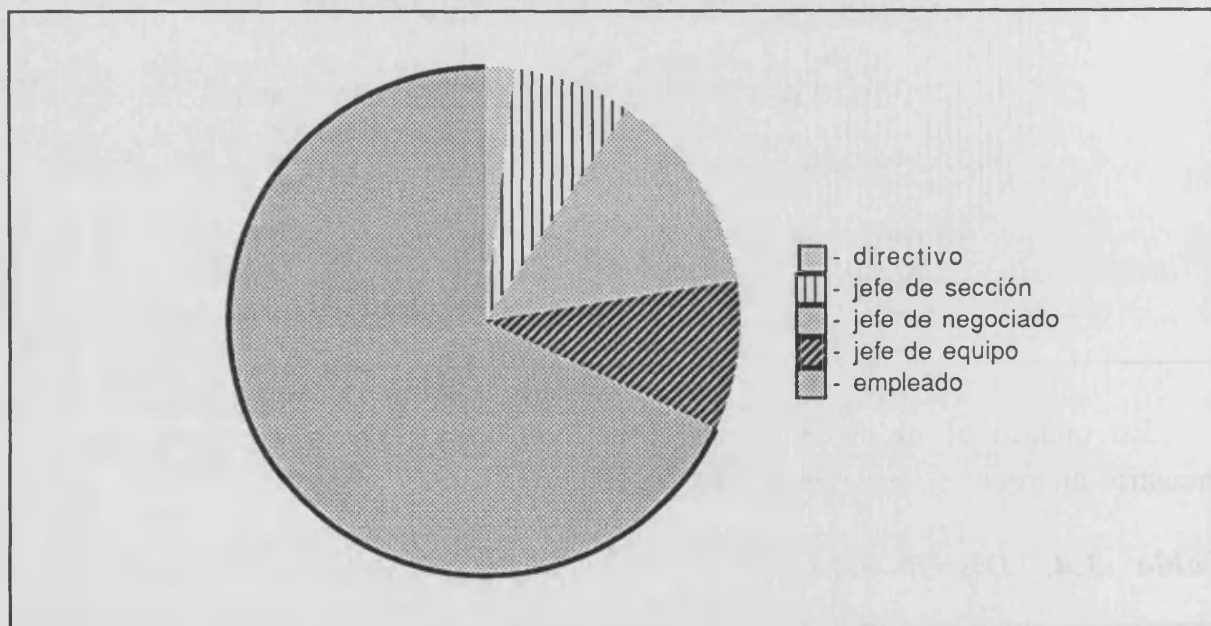
Tabla 3.4. Distribución de la muestra por nivel de estudios.



El 42.5% de los sujetos cursaron estudios hasta el bachillerato o BUP incluido, y el 24.1% cursaron alguna diplomatura universitaria o ingeniería técnica.

Finalmente, el nivel jerárquico de los sujetos que componen la muestra se distribuye del modo siguiente: el 2.06% son directivos; el 7.61% son jefes de sección o agencia; el 13.14% son jefes de negociado; el 9.34% son jefes de equipo; y el 67.82% son empleados. Esta distribución se muestra en la tabla 3.5.

Tabla 3.5. Distribución de la muestra por nivel jerárquico.



2. MEDIDAS.

Las escalas empleadas en este estudio son las desarrolladas por Warr (1990)² para la medida de las dimensiones del bienestar afectivo tenso vs. calmado, deprimido vs. entusiasmado y fatigado vs. animado (ver anexo 1). Todas las escalas constan de seis items cada una, y su escala de respuesta, simétrica, oscila entre "1. nunca" y "6. siempre". Cada una de las escalas consta de tres de adjetivos típicos de uno de los

² Para la traducción de las escalas al castellano se utilizó el procedimiento de Back-translation, al que amablemente se prestaron dos profesores británicos.

polos del eje (positivo), y otros tres típicos del polo opuesto (negativo). No obstante, las escalas no están estructuradas como pares de adjetivos opuestos, tal como se configurarían si se hubieran desdoblado escalas tipo diferencial semántico. El autor indica que los items que configuran las escalas fueron seleccionados en función de pre-tests realizados en distintos contextos, aunque no indica en que consistieron esos pre-tests. En consecuencia, los análisis realizados en este estudio no implican la comparación directa entre puntuaciones obtenidas en pares de items representativos de polos opuestos de una misma dimensión, sino entre subescalas opuestas.

3. HIPOTESIS Y ANALISIS.

Partiendo de la tesis de que las escalas de respuesta empleadas son efectivamente bipolares y complementarias, entonces podemos formular una serie de hipótesis empíricamente contrastables, que confirmarán o no la tesis enunciada. Puesto que sólo puede cumplirse el supuesto de complementariedad si también se cumple el supuesto de bipolaridad, y de él depende la aparición de relaciones lineales e inversas entre items opuestos, inicialmente nos centraremos en el contraste de este último supuesto.

Como ya se ha indicado, las subescalas que miden uno y otro polo no están compuestas por tres adjetivos representativos de un polo y sus tres opuestos, de modo que las comparaciones no pueden realizarse entre pares de items, sino entre pares de subescalas. Para mayor claridad de exposición, denominaremos "puntuación sobre el polo positivo" al sumatorio de las puntuaciones de cada sujeto sobre los items representativos del polo positivo, y denominaremos "puntuación sobre el polo negativo" al sumatorio de las puntuaciones de cada sujeto sobre los items representativos del polo negativo.

Las hipótesis son las siguientes:

1) Si los pares de subescalas opuestas presentan puntuaciones promedio complementarias, entonces la suma de ambas será aproximadamente igual a la puntuación máxima de la escala mas uno

(siempre que el límite inferior de la escala sea uno, y sin invertir las puntuaciones del polo negativo) (Askevis-Leherpeux y Matalon, 1990). En consecuencia, las distribuciones de estas variables "suma" fueron agrupadas en tres categorías: menor que 6.5 (NCD: no complementaria por defecto), entre 6.5 y 7.5 (C: complementaria), y mayor que 7.5 (NCE: no complementaria por exceso). En este análisis se utilizaron las puntuaciones promediadas en las subescalas, con el fin de que su interpretación resultara más sencilla. La posibilidad de que las respuestas obtenidas se distribuyeran al azar entre estas tres categorías, o entre las dos categorías de no complementariedad fue contrastada mediante el análisis de las tablas de contingencia correspondientes y el estadístico ji-cuadrado asociado.

2) En segundo lugar, de ser complementarias, las distribuciones correspondientes a cada par de subescalas serán similares. En consecuencia, se calcularon los estadísticos descriptivos de las distribuciones correspondientes a las subescalas positivas, y, tras invertir los items pertinentes, los de las subescalas negativas y se compararon las medias y varianzas observadas mediante pruebas t y pruebas F, respectivamente.

3) Las puntuaciones obtenidas sobre los polos opuestos de una misma dimensión presentarán correlaciones altas y negativas, y mayores, en valores absolutos, que las que puedan presentar respecto de otras subescalas del mismo u otro polo de distintas dimensiones; y, finalmente,

4) Los sujetos con puntuaciones bajas sobre un polo presentarán puntuaciones altas sobre el polo opuesto, mientras que los sujetos con puntuaciones medias sobre un polo presentarán puntuaciones medias sobre el polo opuesto. En consecuencia, las puntuaciones originales en cada subescala fueron agrupadas en tres categorías: baja -puntuaciones entre 3 y 7-; media -puntuaciones entre 8 y 13-; y alta -puntuaciones entre 14 y 18, y se obtuvieron las tablas de contingencia para cada par de subescalas. Con ello determinamos si ambas variables presentaban algún grado de asociación, y el porcentaje de sujetos bien clasificados de acuerdo con el supuesto de complementariedad.

CAPITULO 4

RESULTADOS Y CONCLUSIONES.

1. RESULTADOS.

Para contrastar la hipótesis 1 se calcularon los porcentajes de casos en los que la suma de las puntuaciones promediadas sobre ambos polos resultó menor que 6.5 (NCD: no complementaria por defecto), entre 6.5 y 7.5 (C: complementaria, con un margen de medio punto por exceso y medio punto por defecto), y mayor que 7.5 (NCE: no complementaria por exceso). En la tabla 4.1 se ofrecen los resultados obtenidos. Estos resultados indican que la combinación más frecuente entre las puntuaciones de subescalas opuestas no es complementaria, sino de tipo no complementario por defecto. Por ejemplo, la tabla muestra que sólo el 44.89% de las respuestas de los sujetos a las subescalas tenso y calmado se combinaron de forma complementaria, el 46.93% de los sujetos presentaron puntuaciones con la combinación no complementaria por defecto, y el 8.16% presentaron respuestas con la combinación no complementaria por exceso.

Sólo en el caso de las subescalas fatigado-animado aparece una distribución equiprobable entre las tres categorías consideradas. En las subescalas tenso-calmado y deprimido-entusiasmado las distribuciones obtenidas no son equiprobables, siendo la más frecuente la combinación no complementaria por defecto. La frecuencia de aparición de este tipo de no complementariedad respecto de la no complementariedad por exceso fue significativamente superior.

Tabla 4.1. Análisis acerca de la complementariedad de las subescalas.

| ESCALA | NCD | C | NCE | χ^2_a | χ^2_b |
|------------------------|--------|--------|--------|------------|------------|
| TENSO/CALMADO | 46.93% | 44.89% | 8.16% | 84.78** | 80.22** |
| DEPRIMIDO/ENTUSIASMADO | 87.06% | 12.58% | 0.34% | 392.25** | 253.01** |
| FATIGADO/ANIMADO | 37.41% | 38.77% | 23.80% | 12.15 | --- |

Nota: χ_a corresponde al supuesto de que las tres categorías (NCD, C, y NCE) son equiprobables, χ^2_b corresponde al supuesto de que NCD y NCE son equiprobables. ** $p < .001$.

Para contrastar la hipótesis 2, se obtuvieron las distribuciones obtenidas para las puntuaciones sobre el polo positivo, y las correspondientes al polo negativo, una vez invertidos los items pertinentes, y se procedió a su comparación. Los resultados se presentan en la tabla 4.2. Los resultados indican que la distribución de las puntuaciones obtenidas sobre el polo negativo presentan medias significativamente mayores y varianzas significativamente menores que las correspondientes al polo positivo, en cada una de las tres escalas estudiadas.

Tabla 4.2 Descripción y comparación de las distribuciones observadas.

| ESCALA | MEDIA | | | VARIANZA | | |
|---------------------|-------------|-------------|-----------------|-------------|-------------|--------|
| | Subescala + | Subescala - | Prueba <i>t</i> | Subescala + | Subescala - | F |
| TENSO/CALMADO | 11.73 | 13.12 | 10.01** | 7.23 | 5.92 | 1.24* |
| DEPRIMIDO/ENTUSIASM | 10.34 | 15.25 | 28.84** | 9.39 | 5.20 | 1.80** |
| FATIGADO/ANIMADO | 12.06 | 12.88 | 3.99** | 10.01 | 6.23 | 1.60** |

* $p < 0.05$, dos colas; ** $p < 0.001$ dos colas.

La matriz de correlaciones entre subescalas aparece en la tabla 4.3. Como se esperaba de acuerdo con la tercera hipótesis, se observa que las correlaciones entre subescalas opuestas (en cursiva) son significativas y del signo esperado. Sin embargo, las subescalas presentan correlaciones mayores (en valores absolutos) o muy similares con otras subescalas diferentes (en negrita) que las que presentan con sus opuestas.

Tabla 4.3 Correlaciones entre subescalas.

| ESCALA | CORRELACIONES | | | | |
|--------------|---------------|--------------|--------------|---------------|--------------|
| | TENSO | CALMADO | DEPRIMIDO | ENTUSIAS | FATIGADO |
| TENSO | 1 | | | | |
| CALMADO | <i>-.524</i> | 1 | | | |
| DEPRIMIDO | .354 | <i>-.425</i> | 1 | | |
| ENTUSIASMADO | <i>-.153*</i> | .565 | <i>-.435</i> | 1 | |
| FATIGADO | .458 | <i>-.331</i> | .325 | <i>-.197*</i> | 1 |
| ANIMADO | <i>-.197*</i> | .465 | <i>-.422</i> | .682 | <i>-.286</i> |

Nota: todas las correlaciones son significativas $p < 0.01$, * $p > 0.05$, contraste bilateral.

Así, la subescala calmado presenta una correlación de *-.524* con su opuesta (tenso), y una correlación de **.565** con la subescala entusiasmado ($t=1.55$, $p > 0.05$, dos colas); la subescala deprimido muestra una correlación de *-.435* con su opuesta (entusiasmado), y de *-.422* con la

subescala animado ($t=.550$, $p>0.05$, dos colas); y la subescala animado muestra una correlación de $-.286$ con su opuesta (fatigado), y de $.682$ con la subescala entusiasmado ($t=3.55$, $p<0.01$, dos colas).

Las tablas de contingencia construidas a partir de las variables "puntuación alta/ media/ baja en la subescala negativa" y "puntuación alta/ media/ baja en la subescala negativa" revelaron que la ubicación de los sujetos en una de las subescalas estaba asociada a su ubicación en la escala opuesta. No obstante, el porcentaje de sujetos para los que la ubicación en una y otra subescala concordaba con la hipotetizada por el supuesto de la complementariedad resultó ser, como máximo, del 53%. La tabla 4.4 muestra el porcentaje de sujetos que presentan puntuaciones congruentes con el supuesto de complementariedad. En concreto, las columnas alto-bajo, medio-medio y bajo-alto representan las tres ubicaciones esperadas según el mencionado supuesto. Estas columnas contienen el porcentaje de sujetos clasificados de ese modo en las subescalas. Por ejemplo, la tabla 4.4 muestra que sólo el 1.20% de los sujetos presentan puntuaciones altas en la subescala tenso y puntuaciones bajas en la subescala calmado, mientras que el 20.40% de los sujetos presentaron puntuaciones medias en ambas subescalas, y el 23.22% puntuaciones bajas en la subescala tenso y puntuaciones altas en la subescala calmado. Finalmente, se observa que el porcentaje total de sujetos con las combinaciones esperadas es del 44.72%. Por su parte, el estadístico ji cuadrado resulta estadísticamente distinto de cero ($\chi^2 = 42.89$, $p<0.01$, dos colas), indicando que la combinación de las puntuaciones en ambas subescalas no se debe al azar.

Tabla 4.4. Ubicación de los sujetos en subescalas opuestas.

| ESCALA | ALTO-BAJO | MEDIO-MEDIO | BAJO-ALTO | % TOTAL | χ^2 |
|------------------------|-----------|-------------|-----------|---------|----------|
| TENSO/CALMADO | 1.20% | 20.40% | 23.12% | 44.72% | 42.89* |
| DEPRIMIDO/ENTUSIASMADO | 0.68% | 4.42% | 17.00% | 22.10% | 22.16* |
| FATIGADO/ANIMADO | 0.68% | 21.76% | 30.27% | 52.71% | 23.40* |

* $p<0.01$, dos colas. Los porcentajes indican los casos con la combinación de puntuaciones que aparece en la columna en las subescalas positiva y negativa de la fila correspondiente.

2. CONCLUSIONES.

A la vista de estos resultados cabe plantearse algunas cuestiones. En primer lugar, ninguna de la hipótesis formuladas acerca de la complementariedad de las puntuaciones en ambas subescalas resulta confirmada. Las puntuaciones promedio de los sujetos sobre subescalas opuestas difieren significativamente. Ello nos lleva a admitir alguna de las siguientes opciones: 1) mantener el supuesto de la bipolaridad de la escala de respuesta, asumiendo que cada sujeto puede presentar dos puntuaciones distintas sobre la misma dimensión; y 2) rechazar el supuesto de bipolaridad de la escala de respuesta.

Respecto de la primera opción, cabe admitir que las puntuaciones en las subescalas no son perfectas, están afectadas de error de medida, y posiblemente de diferentes sesgos de respuesta que distorsionan la complementariedad esperada. En contra de esta interpretación se puede argumentar que el error de medida se distribuiría al azar, por lo que se esperaría igual número de combinaciones no complementarias por exceso como por defecto. Además se podría añadir que la probabilidad de obtener combinaciones entre subescalas no complementarias aún cuando afectas de error disminuiría por el hecho de que la combinación de una puntuación en una subescala con error por defecto con su opuesta con error por exceso daría por resultado una combinación complementaria.

En cuanto al efecto de posibles sesgos de respuesta, como el de aquiescencia, esperaríamos que este sesgo se manifestara por igual en sus dos tendencias posibles: contestando con puntuaciones altas en las dos subescalas, o contestando con puntuaciones bajas en las dos subescalas. Ello llevaría nuevamente a esperar igual número de combinaciones no complementarias por exceso y defecto.

La segunda opción mantiene que las escalas de *respuesta* no son bipolares, y por lo tanto no puede esperarse que las respuestas a subescalas opuestas sean complementarias. Bajo este supuesto pueden suceder dos cosas 1) que la falta de bipolaridad se deba a que cada subescala mide una dimensión distinta e independiente (en línea con lo

hipotetizado por el modelo de afecto positivo/negativo), y 2) que la falta de bipolaridad se deba a que los sujetos no reconstruyen ante cada estímulo, independientemente del polo que represente, el continuo bipolar en toda su extensión, sino que responden sólo en función de uno de los polos de la misma.

De estas dos posibilidades, de la primera se concluye que las dimensiones pretendidamente bipolares son, en cambio, monopolares e independientes. Sin embargo, los resultados obtenidos indican que aunque las puntuaciones en subescalas opuestas no son complementarias, tampoco son independientes. Lo que queda patente es que la relación entre subescalas no es complementaria, y por lo tanto no es lineal e inversa, como se esperaba, pero sí existe una relación entre ambas.

La segunda opción no cuestiona directamente la bipolaridad de las dimensiones que se pretende medir, sino la bipolaridad de la escala de respuesta, que como hemos visto, es el supuesto añadido a las escalas tipo Likert contrabalanceadas cuando se emplean para medir dimensiones bipolares. Es en base a este supuesto que se pueden esperar relaciones lineales inversas -complementariedad- entre subescalas opuestas. El hecho de que no se cumpla este supuesto impide que aparezcan relaciones lineales e inversas entre subescalas. La pregunta es ¿es ello un indicador directo de la falta de bipolaridad de las dimensiones subyacentes?

Desde esta perspectiva se puede constatar que los estudios presentados en la introducción han atribuido la no complementariedad de las puntuaciones, y por lo tanto la ausencia de relaciones lineales e inversas entre subescalas opuestas -aún sin formular el problema desde esta perspectiva- únicamente a dos de las explicaciones señaladas. Los investigadores partidarios del modelo bipolar mantenían el supuesto de la bipolaridad de la escala de respuesta, y atribuían la falta de complementariedad al error de medida y al efecto de sesgos de respuesta como el de aquiescencia. Por su parte, los investigadores partidarios del modelo monopolar atribuían la falta de

complementariedad a la independencia entre las dimensiones medidas por subescalas opuestas.

Queda pues por explorar la explicación que atribuye la falta de complementariedad a la falta de bipolaridad de la escala de respuesta, o, en términos generales al modelo de medida o incluso al método de escalamiento empleado. Si los sujetos, al responder a ítems representativos de un polo de un continuo bipolar no interpretan los extremos de la escala de respuesta como los extremos del continuo bipolar, entonces su ubicación en el continuo no tiene que ser invariante respecto del polo del que son representativos los estímulos, y por lo tanto sus respuestas a estímulos de ambos polos no estarán lineal e inversamente relacionadas. En este caso, la inadecuación del modelo de medida empleado explicaría los resultados en contra de la bipolaridad de las dimensiones.

Por otra parte, si los sujetos ni siquiera perciben los diferentes adjetivos representativos de un mismo polo como réplicas unos de otros, sino que parte de la variabilidad sistemática observada se debiera a diferencias entre adjetivos del *mismo* polo, entonces lo que resultaría inadecuado sería el propio método de escalamiento empleado, en favor de otros métodos que consideraran no sólo los sujetos, sino también los ítems como fuente de variabilidad sistemática.

Estas dos posibles explicaciones constituyen el objeto de estudio de los siguientes capítulos. A continuación se analizará la explicación centrada en la inadecuación del modelo de medida. En la tercera parte nos ocuparemos de la explicación centrada en la inadecuación del método de escalamiento.

PARTE II

**ACERCAMIENTOS ALTERNATIVOS AL PROBLEMA DE LA
MEDIDA DE DIMENSIONES BIPOLARES AFECTIVAS MEDIANTE
ESCALAS TIPO LIKERT CONTRABALANCEADAS:
DESDOBLAMIENTO DE LAS ESCALAS.**

CAPITULO 5

INTRODUCCION:

LA BIPOLARIDAD DESDE UNA PERSPECTIVA INTERDEPENDIENTE.

1. INTRODUCCION.

A la luz del replanteamiento del problema de medida que nos ocupa, se observa que la evidencia empírica que tradicionalmente se ha interpretado como prueba de la validez de una dimensión bipolar - relaciones negativas e inversas entre subescalas opuestas- es consecuencia directa del cumplimiento del supuesto de bipolaridad de la escala de respuesta, y por consiguiente, del cumplimiento del supuesto de complementariedad. De otro modo, sólo si el supuesto de bipolaridad de la escala de respuesta se cumple, puede aparecer la complementariedad entre subescalas opuestas; y sólo si aparece la complementariedad, se obtendrán las relaciones lineales e inversas entre subescalas que darán lugar a las correlaciones elevadas y negativas, y que propiciarán, a su vez, la aparición de factores bipolares. Finalmente, la aparición de estos factores validaría el modelo de dimensiones bipolares, objetivo de todo el proceso.

Al contrario, si la evidencia empírica no ofrece relaciones inversas y lineales entre subescalas opuestas, entonces no se da la complementariedad, y por lo tanto la bipolaridad de la escala de respuesta. Lo que no implica *necesariamente* que las dimensiones latentes no sean bipolares, sino que puede ocurrir que el modelo de medida empleado sea inadecuado.

El modelo de medida empleado implícitamente en los estudios anteriores relacionados con la estructura del afecto -al que denominaremos modelo tradicional- supone que la puntuación que se atribuye a un sujeto refleja su posición en un continuo bipolar. Aunque al sujeto sólo se le presenta un estímulo que describe uno de los polos de la dimensión, se asume que éste reconstruye el continuo bipolar, y responde en consecuencia. Los resultados aquí obtenidos muestran que esta no es la lógica bajo la que responde el sujeto.

Aunque en el ámbito del estudio del afecto -que es el marco conceptual en el que se desarrolla este estudio- no se ha cuestionado el fenómeno de la bipolaridad en sí mismo, sino su adecuación para

representar la estructura del afecto (al menos no tenemos constancia de ello), en otras áreas de conocimiento sí se ha hecho. A continuación expondremos algunos trabajos que presentan interesantes reflexiones acerca del significado de la bipolaridad y cómo se manifiesta empíricamente, a la vez que muestran la importancia y vigencia de este problema desde el punto de vista conceptual, y por extensión, metodológico. Esta revisión no pretende ser exhaustiva, pero sí representativa del estado actual de la cuestión.

Para ubicar adecuadamente las reflexiones acerca de la bipolaridad como fenómeno, y como objeto de medida, hay que señalar que prácticamente todos los trabajos revisados parten de la premisa de que la bipolaridad es determinista, en el sentido de que cuando se emplean escalas en que los estímulos -adjetivos o sentencias- son representativos de *un polo* de la dimensión bipolar a medir, la posición de un sujeto cuando se ubica en un polo debe determinar su posición en el polo opuesto. Cuando una dimensión es conceptualizada como bipolar, se considera que lo es porque ambos polos de la dimensión son mutuamente excluyentes. Con un ejemplo, un sujeto no puede ser muy introvertido y muy extrovertido a la vez. La traducción metodológica de esta conceptualización ha llevado a esperar fuertes relaciones lineales e inversas entre las puntuaciones de los sujetos en ambos polos, cuando estos eran medidos de forma independiente -con escalas tipo Likert contrabalanceadas-. Esta premisa es la que ha prevalecido en áreas de conocimiento como el estudio del afecto (Warr, 1992), o de dimensiones de personalidad como la masculinidad-feminidad (Bem, 1974, 1981), o dimensiones actitudinales como la satisfacción laboral (Herzberg, Mausner y Snyderman, 1959) y el liberalismo-conservadurismo (Kerlinger, 1984). En esta última área de conocimiento es donde más se ha reflexionado e investigado acerca del fenómeno de la bipolaridad en sí. El estudio y medida de dimensiones actitudinales generalmente se ha realizado mediante el uso de escalas tipo Likert. Mediante su empleo diversos autores han constatado que la posición de los sujetos en un polo no determinaba su posición en el polo opuesto, en contra de lo que se esperaba. Las reacciones suscitadas por este comportamiento contraintuitivo también se han centrado básicamente en dos cuestiones: 1) limitar el uso de escalas explícitamente bipolares, a menos que se

compruebe que el objeto de medida es efectivamente una dimensión bipolar; y 2) elaborar nuevas teorías que expliquen por qué las dimensiones bipolares aparecen, en cambio, como empíricamente monopolares.

El estudio de Askevis y Matalon (1990) es ilustrativo del primer punto. En él se hace referencia al uso extensivo de escalas explícitamente bipolares en el estudio de creencias, juicios y actitudes (p.e. "la longevidad es producto de" 1. ambiente..... 6. herencia). El uso de estas escalas se basa en que se considera que lo que no es atribuible al ambiente, lo es a la herencia, y por lo tanto que ambos forman un continuo bipolar en que ambos factores -herencia y ambiente- se combinan de forma *complementaria*. Estos autores pusieron en duda el mencionado supuesto, y para comprobar su adecuación, emplearon escalas tipo Likert en que los extremos del continuo bipolar anterior se desdoblan. Finalmente, constató que sus resultados no apoyaban la complementariedad supuesta entre los factores estudiados, y que la falta de complementariedad no era atribuible a efectos del error de medida o del método de medida. Su conclusión fue que esos factores no formaban un continuo bipolar, y que por lo tanto no era adecuada su medida en términos de escalas bipolares. Más aún concluye que *"nuestros resultados muestran que esta hipótesis (la de la bipolaridad) no es evidente.....los resultados críticos que hemos citado, poco numerosos ciertamente, fueron publicados, fueron probablemente leídos, y sin embargo, aparentemente no han sido tenidos en cuenta. Como en otros casos, se constata que una técnica cómoda, a fuerza de ser utilizada, adquiere cierta legitimidad y una especie de inmunidad a la crítica, incluso si sus fundamentos no han sido jamás justificados, o incluso han sido puestos en duda. Que nosotros sepamos, ninguna revista ha rechazado un artículo porque el autor no haya considerado estas críticas"* (Askevis y Matalon, 1990, p.248).

Respecto del segundo punto, referido a la elaboración de nuevas teorías que expliquen por qué las dimensiones bipolares aparecen como empíricamente monopolares, cabe citar el trabajo realizado por Kerlinger (1984) en su libro "Liberalismo y Conservadurismo". Tras constatar la gran cantidad de estudios en los que escalas tipo Likert

empleadas para medir este continuo bipolar habían dado lugar a factores independientes de contenido liberal, por un lado, y conservador, por otro, este autor concluye que no hay bipolaridad porque los contenidos que son relevantes para los conservadores no son relevantes para los liberales y viceversa. En otros términos, "*la posición de un sujeto en un polo no determina su posición en el otro*" (p.2). Frente a un conjunto de referentes relevantes para una actitud, las personas pueden mostrarse más o menos implicadas o indiferentes, de modo que lo que comparten personas con la misma actitud es su implicación o reacción positiva frente al conjunto de referentes que les describen, de lo que no se infiere una reacción negativa sobre referentes representativos de otra actitud -la supuestamente opuesta-. En consecuencia, no se puede predecir la reacción de las personas ante un conjunto de referentes no pertinentes para una persona actitudinalmente definida (liberal o conservadora). Sin embargo, el autor reconoce que en ocasiones sí aparece la bipolaridad esperada, de modo que sujetos que aceptan los referentes liberales, también rechazan los referentes conservadores. En este sentido, Kerlinger argumenta que ello sólo ocurre bajo ciertas condiciones, y no como norma. En concreto, las condiciones son:

- 1) la muestra está compuesta por un gran número de sujetos máximamente implicados;
- 2) cuando existe una coyuntura interpretada como amenazante (que a su vez incrementa la implicación de los sujetos); y
- 3) por efecto de aspectos técnicos como la técnica de puntuación empleada en las escalas.

A continuación, pone a prueba sus hipótesis mediante la factorización de las respuestas dadas a escalas tipo Likert en diferentes muestras y escalas, y empleando análisis factorial exploratorio de primer y segundo orden, y análisis factorial confirmatorio. En general, los análisis realizados revelaron la existencia de dos factores independientes, aunque no totalmente puros, que fueron identificados como "liberalismo" y "conservadurismo". En definitiva, la ausencia de las relaciones lineales e inversas, consideradas como prueba empírica de la bipolaridad de este continuo, llevó al autor a la confirmación de su

modelo, en el que liberalismo y conservadurismo son dos actitudes independientes.

Estos dos trabajos cuestionan distintos aspectos del fenómeno de la bipolaridad. En primer lugar, Askevis y Matalon (1990) ponen de manifiesto que la bipolaridad no debe darse por asumida, y explicitan por primera vez uno de los supuestos fundamentales para la interpretación de las puntuaciones obtenidas con escalas tipo Likert como indicativas de la ubicación de un sujeto en un continuo bipolar: la complementariedad de las puntuaciones obtenidas ante estímulos de polos opuestos.

Kerlinger, por su parte, pone de manifiesto que:

a) cuando se estudia un fenómeno bipolar, se asume una relación lineal e inversa entre los polos opuestos, de modo que

b) si la relación entre polos "opuestos" no es necesariamente lineal e inversa

c) entonces la dimensión subyacente no es bipolar.

En este sentido, Kerlinger (1984) expone la asociación generalmente asumida entre bipolaridad y relación inversa y lineal entre opuestos, siendo la bipolaridad el fenómeno teórico, y la relación lineal e inversa su manifestación empírica. A continuación, argumenta que las relaciones entre los llamados "polos opuestos" bajo el modelo bipolar no son necesariamente lineales e inversas. La explicación que aduce puede ilustrarse partiendo de un sujeto que se identifica con uno de los "polos" de la dimensión supuestamente bipolar. Si entendemos que un sujeto se identifica con un polo cuando presenta puntuaciones altas a los items representativos de ese polo, entonces el modelo tradicional predice que ese sujeto presentará puntuaciones bajas en los items representativos del polo opuesto, dando lugar a las relaciones lineales e inversas esperadas. En cambio, Kerlinger interpreta que si un sujeto se identifica con un polo, entonces el otro ya no le es pertinente, no le describe, y, por consiguiente, no cabe esperar relaciones sistemáticas entre las puntuaciones que se atribuya en uno y otro polo. Finalmente, concluye que puesto que los polos "opuestos" no presentan

relaciones lineales e inversas, la dimensión subyacente no puede ser bipolar.

Sin embargo, la conclusión final a la que llegan ambos estudios todavía parte de la premisa de que una dimensión bipolar debe producir necesariamente relaciones lineales e inversas entre los ítems que miden sus polos opuestos. Si consideramos que esto sólo ocurre cuando además las escalas empleadas para medirla son implícitamente bipolares, de modo que el sujeto responde sobre el mismo continuo bipolar independientemente del polo del que sea representativo el estímulo, entonces los resultados que obtienen pueden ser interpretados de otra forma. Askevis y Matalon (1990) no encontraron evidencia a favor de la complementariedad, precisamente debido a que las escalas empleadas no cumplían previamente el supuesto de bipolaridad de la escala de respuesta; y Kerlinger obtuvo los factores desdoblados por la misma razón. La observación de que la ubicación de un sujeto en un polo no predice directamente su ubicación en el polo opuesto sería correcta. Esto sólo ocurriría si el sujeto se ubicara ante los estímulos de ambos polos en el mismo continuo bipolar. La ausencia relaciones lineales e inversas entre los ítems representativos de los polos de conservadurismo y liberalismo no desconfirmaría directamente la bipolaridad de esta dimensión actitudinal.

2. EL MODELO DE MEDIDA TRADICIONAL FRENTE AL MODELO DE "DOS MITADES".

A diferencia de lo que supone el modelo de medida tradicional, puede ser que los sujetos, al reaccionar ante un estímulo representativo de un sólo polo de la dimensión bajo estudio, evalúen su afecto sólo con respecto a ese polo de la dimensión. En ese caso, la escala de respuesta sería interpretada por el sujeto haciendo coincidir el extremo inferior de la escala de respuesta con la no pertinencia de ese polo para describir su afecto, y por lo tanto se le haría corresponder el valor de "neutro". En tanto que el extremo superior de la escala se haría corresponder con el polo positivo o negativo de la dimensión, en función de los ítems empleados. En otros términos, un sujeto que contestara con "1. nunca" al ítem contento, estaría indicando que ese ítem no describe su afecto, y en

la medida en que le resultara más pertinente o descriptivo, así iría dando como respuesta alternativas más cercanas al extremo superior. De este modo los sujetos se ubicarían sobre la parte del continuo bipolar "neutro-positivo" cuando contestaran a items del polo positivo, y sobre la parte del continuo "neutro-negativo" cuando contestaran a items del polo negativo.

¿Qué consecuencias tiene esta reinterpretación sobre la complementariedad de los items de uno y otro polo? Si la escala de respuesta no es percibida por los sujetos como bipolar, entonces la invarianza en la ubicación del sujeto respecto de la dimensión medida a través de los distintos items (positivos y negativos) no se sostiene, y por lo tanto no puede esperarse que las puntuaciones observadas en uno y otro subconjunto de items sean complementarias. ¿Qué argumento justifica entonces la inferencia de que un sujeto que presenta sobre el polo positivo una puntuación baja, presentará sobre el polo negativo una puntuación alta? Ambos polos se encuentran en los extremos de un continuo, en cuyo centro existe un punto medio o neutral, y puesto que la dimensión es bipolar, cabe la posibilidad de que un sujeto presente puntuaciones bajas en ambos polos, en cuyo caso se ubicaría en ese punto central de la dimensión. Una persona puede estar raras veces triste, y raras veces eufórica, si su estado de ánimo habitual no se caracteriza especialmente por ninguno de los extremos de la dimensión deprimido-entusiasmado, por ejemplo. Lo que resulta poco plausible es presuponer que una persona que contesta estar raras veces triste está necesariamente la mayor parte del tiempo contenta.

En otras palabras, en un continuo bipolar, la ubicación de un sujeto en uno de los polos no nos da de forma directa su ubicación en el otro. En primer lugar, un sujeto sólo puede tener una ubicación en todo el continuo bipolar. Si sobre los items del polo positivo se da una puntuación baja, lo que eso indica desde esta perspectiva es que ese polo de la dimensión no le describe, no caracteriza su grado o tipo de afecto, o no es relevante. Y el hecho de que un polo no describa el afecto de un sujeto no implica necesariamente que sí lo haga el polo contrario. En definitiva, puntuaciones bajas en un polo de una dimensión bipolar indican ausencia del afecto que caracteriza ese polo, pero ello no implica

la presencia del afecto contrario. Así podría suceder que un sujeto caracterizado por puntuaciones bajas en el polo de la depresión, por ejemplo, presentara cualquier combinación de puntuaciones en el polo del entusiasmo. A una persona que no le caracteriza el ánimo deprimido puede no caracterizarle tampoco el ánimo entusiasta (punto neutro de la dimensión), o caracterizarle un ánimo medianamente entusiasta, o sentirse decididamente entusiasta. Por el contrario, una persona decididamente entusiasta (con puntuaciones altas en este polo de la dimensión), no puede presentar también puntuaciones altas en el polo contrario de la dimensión, así como difícilmente se puede entender que se atribuya puntuaciones medias. En resumen, si consideramos que los sujetos evalúan su afecto en escalas "unipolares", en el sentido de que sólo consideran una mitad de la dimensión bipolar bajo estudio en la interpretación y uso de la escala de respuesta, entonces su ubicación en ambos polos no es completamente complementaria, aunque tampoco es independiente. No es complementaria por cuanto que la ausencia del afecto que caracteriza a uno de los polos de la dimensión (denotada por puntuaciones bajas) no informa del grado en que se experimenta el afecto opuesto. No es independiente en la medida en que la presencia del afecto que caracteriza uno de los polos determina la ausencia del afecto opuesto.

Operativizando las premisas expuestas, lo que se hipotetiza es que, si categorizamos las respuestas obtenidas en uno y otro polo de una dimensión bipolar en las categorías bajo, medio y alto, sin invertir previamente los ítems correspondientes al polo negativo, encontraremos que los sujetos con puntuaciones bajas en un polo presentan tanto puntuaciones bajas, como medias o altas en el otro, que los sujetos con puntuaciones medias en un polo presentan puntuaciones bajas en el otro, y que los sujetos con puntuaciones altas en un polo presentan puntuaciones bajas en el otro. Frente a esta hipótesis, la interpretación tradicional de las escalas hipotetiza que los sujetos con puntuaciones bajas en un polo presentarían puntuaciones altas en el opuesto, y que los sujetos con puntuaciones medias en un polo presentarían puntuaciones medias en el otro. De modo que mientras el modelo de medida tradicional sólo admite como evidencia de la bipolaridad de la dimensión bajo estudio una relación lineal e inversa -supuesta la

complementariedad y bipolaridad de la escala de respuesta-, el modelo de medida que se propone no asume ni la bipolaridad ni la complementariedad de la escala de respuesta, sino que considera este tipo de relación como uno de los diferentes subtipos posibles: el que presentarán los sujetos con puntuaciones altas en uno de los polos. El conjunto de relaciones que cabe esperar bajo este modelo podría denominarse "de interdependencia". Estas relaciones podrían explicar la variedad de resultados obtenidos en el proceso de factorización expuestos en la introducción de este trabajo. En la medida en que la mayor parte de los sujetos se ubicara en el polo bajo de una dimensión, sus puntuaciones variarían libremente en el polo opuesto, dando lugar a factores monopolares independientes; y en la medida en que se ubicaran en el polo alto, sus puntuaciones bajas en el polo opuesto darían lugar a relaciones inversas y factores bipolares. Finalmente, este modelo también explicaría las relaciones observadas en el capítulo 4 de este trabajo, donde no se encuentra la complementariedad esperada entre subescalas opuestas, pero tampoco resultan ser independientes. Lo que se concluye es que no son linealmente dependientes.

De este modo, el problema de investigación que vamos a abordar es el siguiente: bajo el modelo de medida tradicional, la evidencia empírica esperada como indicador de la bipolaridad de las dimensiones latentes puede no aparecer por causas debidas al modelo de medida y al método de análisis. En consecuencia, se ha presentado un modelo de medida alternativo, que no asume algunos de los supuestos correspondientes al modelo tradicional. Bajo este modelo se pueden formular una serie de hipótesis que se cumplirán si las dimensiones latentes son efectivamente bipolares. En la medida en que estas hipótesis resulten confirmadas, se podrá concluir que los resultados obtenidos bajo el modelo tradicional en contra de la bipolaridad de las dimensiones pueden atribuirse al modelo de medida y al método de análisis empleados, en lugar de al modelo teórico hipotetizado.

CAPITULO 6

METODO.

1. PROCEDIMIENTO DE RECOGIDA DE DATOS Y MUESTRA.

Esta parte del estudio ha sido realizada sobre la misma muestra que se empleó en la primera parte. Allí se describe el procedimiento de recogida de datos, y la composición y descripción de la muestra.

2. MEDIDAS.

Las escalas empleadas en esta parte del estudio son las escalas desarrolladas por Warr (1990) para la medida de las dimensiones del bienestar afectivo tenso vs. calmado, deprimido vs. entusiasmado y fatigado vs. animado, que ya quedaron descritas en la primera parte del estudio. Recordaremos que estas escalas constan de seis ítems cada una, y su escala de respuesta, simétrica, oscila entre "1. nunca" y "6. siempre". Cada una de las escalas consta de tres de adjetivos típicos de uno de los polos del eje (positivo), y otros tres típicos del polo opuesto (negativo). Como argumentan Lloret et al. (1995), los análisis realizados en este estudio no implican la comparación directa entre puntuaciones obtenidas en pares de ítems representativos de polos opuestos de una misma dimensión, sino entre subescalas opuestas. Además se incluyó un ítem en formato de diferencial semántico por cada una de las escalas a evaluar. Estos ítems pedían los sujetos que indicaran su estado de ánimo en tres escalas de 7 puntos como la que aparece en la tabla 6.1. Cada escala contaba en sus extremos con el par de adjetivos tenso-calmado; deprimido-entusiasmado y cansado-leno de energía, respectivamente.

Tabla 6.1. Escala de respuesta de la escala criterio correspondiente a la dimensión tenso-calmado.

| | | |
|-------|---------------------------|---------|
| TENSO | _ : _ : _ : _ : _ : _ : _ | CALMADO |
|-------|---------------------------|---------|

Entre cada par de ítems opuestos que se dispuso una línea continua dividida en 7 espacios, de entre los que el sujeto marcaba el

que mejor definiera su estado de ánimo en el mismo referente temporal que en las escalas tipo Likert.

3. HIPOTESIS Y ANALISIS.

Partiendo de la tesis de que las escalas de respuesta empleadas pueden ser separadas en una subescala positiva y una negativa, y que las respuestas a las subescalas positivas ubican al sujeto en la mitad positiva de la dimensión bipolar correspondiente, y que las respuestas a las subescalas negativas ubican al sujeto en la mitad negativa de la dimensión bipolar correspondiente, entonces podemos formular una serie de hipótesis empíricamente contrastables, que confirmarán o no la tesis enunciada.

Estas hipótesis son las siguientes:

1) Los sujetos con puntuaciones altas en una subescala presentarán puntuaciones bajas en la subescala opuesta, los sujetos con puntuaciones medias en una subescala presentarán puntuaciones bajas en la subescala opuesta, y los sujetos con puntuaciones bajas en una subescala podrán presentar puntuaciones altas, medias o bajas en la subescala opuesta. Categorizaremos las puntuaciones en tres categorías, en función de la escala de respuesta. Puesto que las puntuaciones en cada subescala se han calculado como el sumatorio de las puntuaciones en cada ítem, y los ítems se contestan mediante escalas de respuesta que oscilan de 1 a 6, el rango de puntuaciones en las subescalas oscila de 3 a 18. En consecuencia, las categorías quedaron definidas por los siguientes puntos de corte: bajo -puntuaciones entre 3 y 7-; medio -puntuaciones entre 8 y 13-; y alto -puntuaciones entre 14 y 18. La contrastación de esta hipótesis se efectuará mediante análisis de tablas de contingencia.

2) Utilizando la misma categorización de las puntuaciones en las subescalas -baja, media y alta-, se observarán diferencias en las puntuaciones obtenidas en los ítems con formato de diferencial semántico en función de la ubicación de los sujetos en ambas subescalas. Los sujetos con las puntuaciones más bajas en los ítems de diferencial

semántico serán aquellos con la combinación alto (subescala negativa)-bajo (subescala positiva), los sujetos con puntuaciones medias serán aquellos con puntuaciones bajas en ambas subescalas, y los sujetos con las puntuaciones más altas serán aquellos con la combinación bajo (subescala negativa)-alto (subescala positiva). La contrastación de esta hipótesis se realizará mediante análisis de varianza de dos factores (el par de subescalas opuestas) con tres niveles cada uno (bajo, medio y alto), tomando como variable dependiente el ítem con formato de diferencial semántico correspondiente.

3) Las puntuaciones obtenidas en cada par de subescalas, tras ser apropiadamente transformadas en una única puntuación final, presentarán correlaciones mayores con las escalas criterio que las presentadas por la puntuación final calculada bajo el modelo tradicional.

CAPITULO 7

RESULTADOS Y CONCLUSIONES.

1. RESULTADOS.

En primer lugar, se obtuvo la representación gráfica de las distribuciones (ver figuras 7.1, 7.2 y 7.3) conjuntas de las puntuaciones en cada par de subescalas. De acuerdo con la primera de las hipótesis formuladas, se esperaba encontrar que los sujetos con puntuaciones bajas en una subescala presentaran puntuaciones altas en la opuesta, y que lo sujetos con puntuaciones medias en una subescala presentaran puntuaciones medias en la opuesta. Sin embargo, ninguna de las representaciones gráficas obtenidas mostró el perfil típico de las variables lineal e inversamente relacionadas. Sin embargo, tampoco parecía el perfil típico de variables no relacionadas. Todas ellas mostraban sistemáticamente una diferencia fundamental respecto de la representación de variables independientes: no había casos en que se presentara la combinación alto-alto an ambas subescalas. Sin embargo, sí había sujetos con la combinación bajo-bajo en ambas subescalas, así como sujetos que con una puntuación baja en una de las subescalas, presentaban cualquier combinación posible en la subescala opuesta.

Figura 7.1 Representación gráfica de las puntuaciones obtenidas en las subescalas positiva (BPANIM) y negativa (BPDEP). Dimensión fatigado-animado.

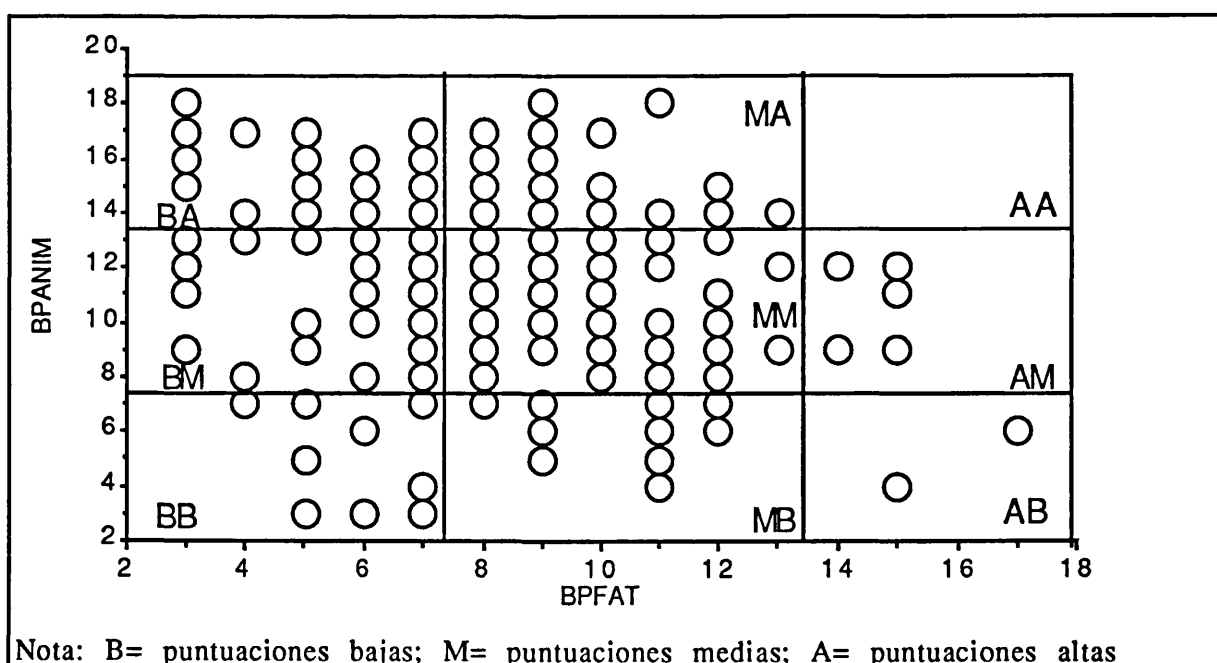


Figura 7.2. Representación gráfica de las puntuaciones obtenidas en las subescalas positiva (BPCALM) y negativa (BPTEN). Dimensión tenso-calmado.

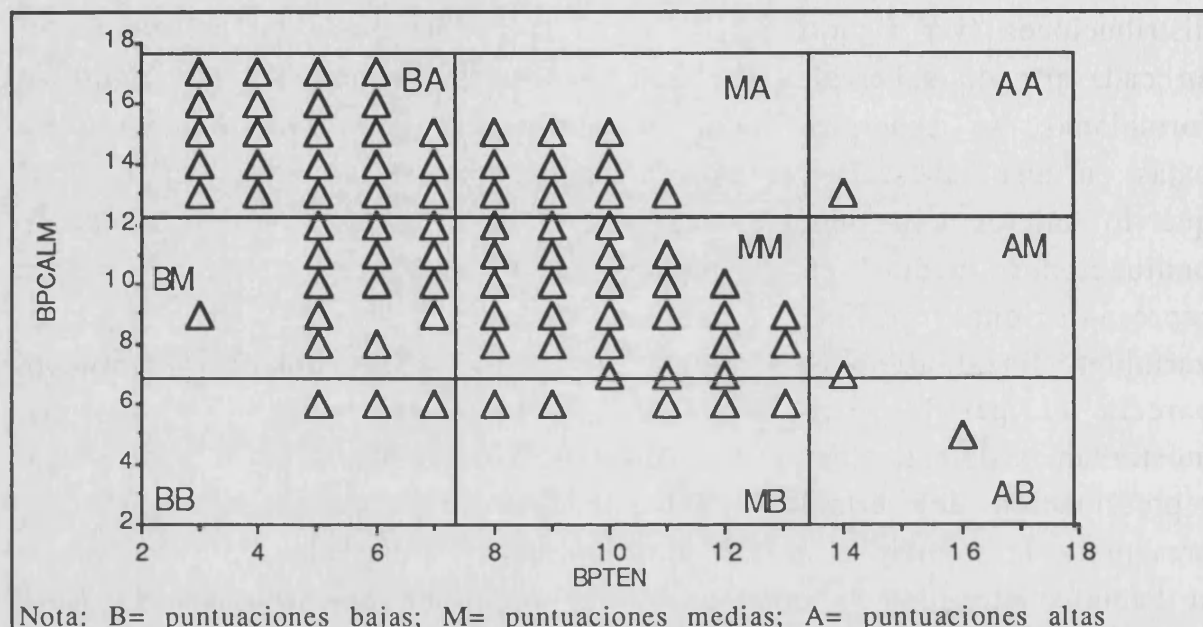
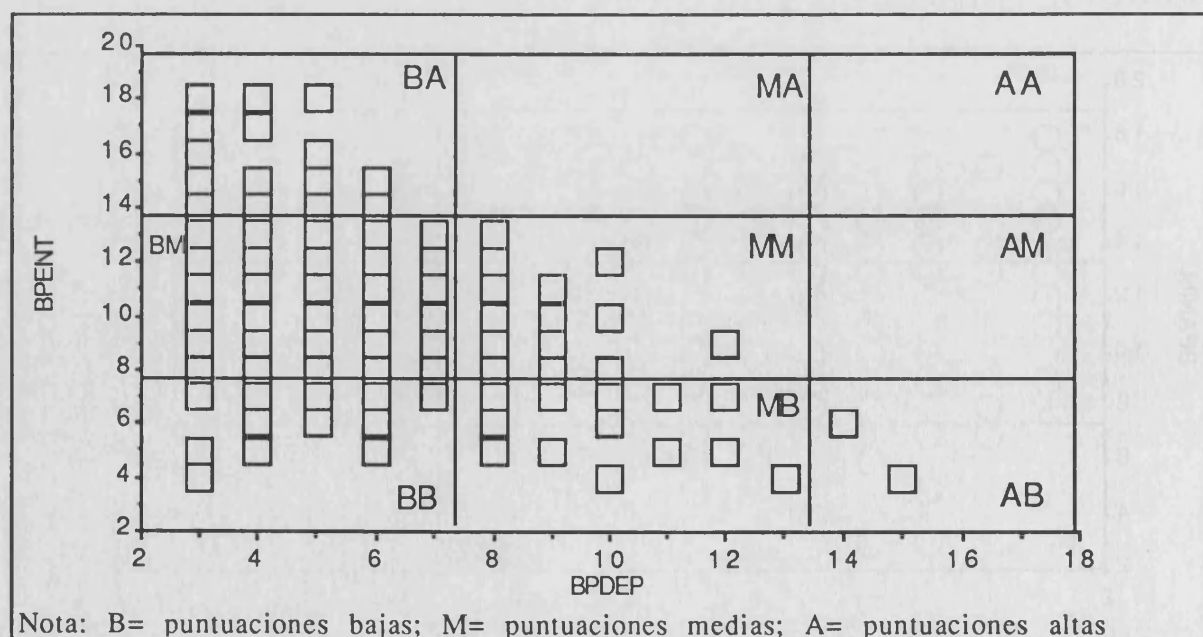


Figura 7.3. Representación gráfica de las puntuaciones obtenidas en las subescalas positiva (BPENT) y negativa (BPDEP). Dimensión deprimido-entusiasmaso.



Las tablas de contingencia indicaron que en todos los casos, las puntuaciones observadas en subescalas opuestas estaban significativamente relacionadas, aunque esa relación no era lineal. Para identificar la frecuencia con que se observaban combinaciones entre las puntuaciones de subescalas opuestas congruentes con el supuesto de que en cada una, los sujetos se ubicaban únicamente respecto de la mitad del continuo bipolar "activado" por los items de la subescala, se contabilizaron los casos en que se presentaron las siguientes combinaciones (en primer lugar indicaremos la ubicación respecto del polo negativo, y en segundo respecto del polo positivo): bajo-bajo; bajo-medio; bajo-alto; medio-bajo; y alto-bajo, señaladas en negrita en la tabla. Con fines comparativos, también se calculó el porcentaje de casos en que aparecían las combinaciones esperadas bajo el supuesto tradicional, estas frecuencias se señalan en la tabla en cursiva (las frecuencias contabilizadas bajo ambos modelos se señalan en cursiva y negrita).

Tabla 7.1 Tabla de contingencia de las subescalas vigor x fatiga.

| Observed Frequency Table | | | | |
|--------------------------|-----------|-----------|------|---------|
| | BAJO | MEDIO | ALTO | Totals: |
| BAJO | 15 | 18 | 2 | 35 |
| MEDIO | 73 | <i>64</i> | 5 | 142 |
| ALTO | 89 | <i>28</i> | 0 | 117 |
| Totals: | 177 | 110 | 7 | 294 |

Nota: filas= vigor, columnas=fatiga. Las frecuencias en negrita corresponden a las combinaciones esperadas bajo el supuesto de que cada subescala indica la posición del sujeto en una mitad de la dimensión bipolar. Las frecuencias en cursiva indican las combinaciones esperadas bajo el modelo tradicional.

En concreto, en las subescalas de la dimensión fatiga-vigor (ver tabla 7.1), el 67% de los sujetos presentaron las combinaciones predichas bajo el supuesto de que ofrezcan la ubicación de los sujetos en cada una de las mitades de un continuo bipolar. Bajo el supuesto tradicional -

relaciones lineales e inversas entre subescalas opuestas- este porcentaje fue del 52%. Ambas subescalas (fatiga y vigor) estaban significativamente relacionadas ($\chi^2= 23.40$, 4 g.l., $p<.0001$).

En el caso de las subescalas correspondientes a la dimensión tensión-calma (ver tabla 7.2), el porcentaje de casos en que aparecieron las combinaciones esperadas bajo el supuesto de las "dos mitades" fue del 74%, en tanto que bajo el modelo tradicional, este porcentaje fue del 44%. Ambas subescalas estaban significativamente relacionadas ($\chi^2= 42.89$, 4 g.l., $p<.0001$).

Tabla 7.2 *Tabla de contingencia de las subescalas calma x tensión.*

| | | Observed Frequency Table | | | |
|---------|--------------|--------------------------|----------|------|---------|
| | | BAJO | MEDIO | ALTO | Totals: |
| BAJO | 1 0 | 2 4 | 3 | 37 | |
| MEDIO | 1 1 3 | 6 0 | 1 | 174 | |
| ALTO | 6 8 | 1 5 | 0 | 83 | |
| Totals: | 191 | 99 | 4 | 294 | |

Nota: filas= calma, columnas= tensión. Las frecuencias en negrita corresponden a las combinaciones esperadas bajo el supuesto de que cada subescala indica la posición del sujeto en una mitad de la dimensión bipolar. Las frecuencias en cursiva indican las combinaciones esperadas bajo el modelo tradicional.

En el caso de las subescalas correspondientes a la dimensión depresión-entusiasmo (ver tabla 7.3), el porcentaje de casos en que aparecieron las combinaciones esperadas bajo el supuesto de las "dos mitades" fue del 95%, en tanto que bajo el modelo tradicional, este porcentaje fue del 22%. Ambas subescalas estaban significativamente relacionadas ($\chi^2= 22.16$, 4 g.l., $p<.0002$).

Tabla 7.3 Tabla de contingencia de las subescalas entusiasmo x depresión.

| Observed Frequency Table | | | | |
|--------------------------|------------|-----------|----------|---------|
| | BAJO | MEDIO | ALTO | Totals: |
| BAJO | 63 | <i>17</i> | <i>2</i> | 82 |
| MEDIO | <i>149</i> | <i>13</i> | 0 | 162 |
| ALTO | <i>50</i> | 0 | 0 | 50 |
| Totals: | 262 | 30 | 2 | 294 |

Nota: filas= entusiasmo, columnas= depresión. Las frecuencias en **negrita** corresponden a las combinaciones esperadas bajo el supuesto de que cada subescala indica la posición del sujeto en una mitad de la dimensión bipolar. Las frecuencias en *cursiva* indican las combinaciones esperadas bajo el modelo tradicional.

En las tres escalas estudiadas se confirma pues que la ausencia -operativizada como puntuaciones bajas- de afecto negativo no informa acerca del grado de afecto positivo del sujeto, y viceversa. En cambio, la presencia -operativizada como puntuaciones altas- de afecto positivo si determina la ausencia de afecto negativo, y viceversa.

Generalmente, el debate acerca de la polaridad de las dimensiones del afecto -bipolares versus monopolares- se ha centrado en la oposición de las siguientes argumentaciones: mientras el modelo bipolar ha defendido que ambos afectos no pueden ser experimentados de forma simultánea, el modelo monopolar ha defendido lo contrario, es decir, que ambos afectos sí pueden experimentarse de forma simultánea. Por consiguiente, se puede decir que el debate se ha centrado sobre la presencia de ambos afectos. Operativizando estas argumentaciones, bajo el modelo bipolar se han esperado relaciones lineales e inversas entre las escalas que miden afectos opuestos, mientras que bajo el modelo monopolar se ha esperado la ausencia de

relación entre escalas que miden afectos "independientes" ¿Qué ocurre si en lugar de centrarnos en las implicaciones que la presencia de un afecto presenta para el afecto opuesto, nos centramos en su ausencia? Bajo el modelo monopolar, que asume independencia entre ambos afectos, puede esperarse la ausencia simultánea de ambos afectos (Watson y Tellegen, 1985, Watson et al. 1989). Por otra parte, bajo el modelo bipolar la bipolaridad del afecto se ha conceptualizado de modo que ambos afectos no se pueden dar simultáneamente, pero no se ha planteado si pueden *no* darse simultáneamente: la presencia de uno difícilmente no excluye la presencia del otro, pero la ausencia de uno no implica la presencia del otro, son mutuamente excluyentes, pero no mutuamente "incluyentes". En este sentido, sólo cabría esperar relaciones inversas y lineales bajo el modelo bipolar en aquéllos sujetos caracterizados por presentar uno de los dos afectos. Los sujetos que no se caracterizaran por experimentar ninguno de los dos afectos tenderían a presentar puntuaciones bajas en ambas subescalas, dado lugar a relaciones lineales y positivas entre subescalas opuestas. Y la combinación de relaciones lineales inversas en un subgrupo de sujetos con relaciones lineales y positivas en otro pudiera dar lugar a las correlaciones bajas que vienen dando apoyo al modelo monopolar del afecto. En consecuencia, pudiera ser que la independencia entre ambos afectos surja como artefacto porque pueda no darse ninguno de los dos afectos representativos de los extremos del continuo bipolar, que es lo que muestran nuestros datos. Frente a esta explicación, la aparición de relaciones positivas ha sido generalmente atribuida al efecto del sesgo de aquiescencia, bajo el supuesto de que sujetos con puntuaciones bajas en subescalas opuestas respondían en función de este sesgo.

La segunda de las hipótesis extendía las repercusiones de este nuevo enfoque sobre las relaciones que cabría esperar respecto de un criterio externo que pudiera indicar con alguna seguridad la posición de los sujetos en el continuo bipolar en toda su extensión. En la medida en que las diferentes combinaciones entre las puntuaciones en subescalas opuestas determinaran diferentes ubicaciones en un continuo explícitamente bipolar, y en la medida en que esas diferentes ubicaciones presentaran la ordenación esperada, con los sujetos bajo-bajo aproximadamente en el centro de la escala bipolar, el enfoque de

"las dos mitades" podría quedar validado. En este sentido cabe recordar que en muchas ocasiones se ha atribuido la ausencia de relaciones lineales e inversas bajo el modelo tradicional a sesgos de respuesta como el de aquiescencia. Si esto fuera así, hubieran aparecido sujetos con puntuaciones altas en subescalas opuestas, cosa que no ha sucedido, pero aún queda por determinar si los sujetos con puntuaciones bajas en subescalas opuestas están respondiendo de acuerdo con este sesgo. En caso de que fuera así, los sujetos con esta combinación también habrían respondido con puntuaciones bajas en las escalas de diferencial semántico, o al menos no se observaría una correspondencia interpretable entre su ubicación en las subescalas opuestas y las escalas de diferencial semántico.

Tabla 7.4. Analisis de varianza de los factores "subescala fatiga" y "subescala vigor" sobre una escala de diferencial semántico.

| VIGOR | | BAJO | MEDIO | ALTO | Totales |
|---------|-------|-------------|--------------|--------------|--------------|
| FATIGA | BAJO | 15 3.583 | 73 4.285 | 89 4.936 | 177 4.553 |
| | MEDIO | 18 2.843 | 64 3.594 | 28 4.369 | 110 3.668 |
| | ALTO | 2 1.667 | 5 2.833 | 0 . | 7 2.5 |
| Totals: | | 35 3.093 | 142 3.922 | 117 4.801 | 294 4.173 |

Nota: la escala criterio oscila de 1. (fatiga) a 7. (vigor). En la parte superior de las celdillas se muestra la frecuencia observada para la combinación correspondiente en ambas subescalas; en la parte inferior se muestra la puntuación media en la escala criterio que presentan los sujetos con esa combinación; en negrita se muestran las combinaciones esperadas bajo el modelo de dos mitades.

Tabla 7.5 Resultados de las fuentes de variación para la escala fatiga-vigor.

| FACTORES | G.L. | F | P (2 colas) |
|--------------|------|--------|-------------|
| ANIMADO | 2 | 39.570 | 0.0001 |
| FATIGADO | 1 | 29.652 | 0.0001 |
| FAT. X ANIM. | 2 | 3.048 | 0.049 |

Nota: en este análisis se eliminó el nivel "alto" de la subescala fatigado. $R^2=0.346$

Los resultados obtenidos para las subescalas fatiga-vigor (ver tablas 7.4 y 7.5) indican que todas las fuentes de variación consideradas presentan efectos sobre las puntuaciones observadas en la escala "criterio" de diferencial semántico empleada. Pero lo interesante para corroborar la hipótesis de partida no es únicamente que las fuentes de variación consideradas, y especialmente la interacción entre las subescalas, introduzcan diferencias entre las puntuaciones de los sujetos en la escala criterio, sino que además la ordenación de las medias en la escala criterio en función de la combinación de las subescalas presente un orden determinado. Y sucede así: tal como muestra la tabla 7.4 los sujetos con la combinación alto-bajo (el primer miembro del par hace referencia a la subescala negativa y el segundo a la positiva) son los que presentan las puntuaciones más bajas en la escala bipolar (1.667), los sujetos con la combinación medio-bajo presentan una puntuación algo mayor y así sucesivamente (siguiendo las medias señaladas en negrita) hasta llegar a los sujetos con la combinación bajo-alto, que son los que presentan la mayor puntuación (4.936). Además, los sujetos con la combinación bajo-bajo presentan puntuaciones en el punto medio de la escala criterio (4), con lo que la interpretación de esta combinación por efectos de la aquiescencia queda descartada. Frente a ello, cabe señalar lo que ocurre con los sujetos que no presentan combinaciones esperadas bajo el supuesto de las "dos mitades" (33%). Los sujetos con puntuaciones medias en ambas subescalas también presentan puntuaciones medias en la escala criterio. Ello puede deberse a que: 1) sus respuestas están afectadas por el sesgo de tendencia central, o 2) no tienen clara su posición respecto a la dimensión evaluada, por lo que se quedan en el término medio. En cualquier caso ello no desconfirma la hipótesis de que las subescalas opuestas sitúan a los sujetos en las "mitades" correspondientes del continuo bipolar latente. En cuanto a los sujetos con las dos combinaciones restantes -medio-alto y alto-medio-, no constituyen un grupo numeroso, y su ubicación en la escala criterio tiene sentido, en cuanto que presentan puntuaciones similares a las de los sujetos con las combinaciones bajo-medio y medio-bajo, respectivamente.

Tabla 7.6. Analisis de varianza de los factores "subescala tensión" y "subescala calma" sobre una escala de diferencial semántico.

| | | CALMA | BAJO | MEDIO | ALTO | Totals: |
|---------|-------|-------|-------------|--------------|-------------|--------------|
| TENSION | BAJO | | 10 3.9 | 110 4.882 | 66 5.667 | 186 5.108 |
| | MEDIO | | 23 2.522 | 60 4 | 15 5.4 | 98 3.867 |
| | ALTO | | 3 1.333 | 1 7 | 0 . | 4 2.75 |
| Totals: | | | 36 2.806 | 171 4.585 | 81 5.617 | 288 4.653 |

Nota: la escala criterio oscila de 1. (tensión) a 7. (calma). En la parte superior de las celdillas se muestra la frecuencia observada para la combinación correspondiente en ambas subescalas; en la parte inferior se muestra la puntuación media en la escala criterio que presentan los sujetos con esa combinación; en negrita se muestran las combinaciones esperadas bajo el modelo de dos mitades.

Tabla 7.7. Resultados de las fuentes de variación para la escala tenso-calmado.

| FACTORES | G.L. | F | P (2 colas) |
|-------------|------|-------|-------------|
| TENSO | 2 | 25.14 | 0.0013 |
| CALMADO | 1 | 22.58 | 0.0001 |
| TEN X CALM. | 2 | 3.09 | 0.09 |

Nota: en este análisis se eliminó el nivel "alto" de la subescala fatigado. $R^2=0.277$

En lo que respecta a los resultados obtenidos acerca de las subescalas tensión-calma y la escala bipolar criterio (ver tablas 7.6 y 7.7), son análogos a los comentados para las subescalas fatiga-vigor. De nuevo se producen los efectos esperados y se observa que las medias en la escala criterio para cada combinación en las subescalas se ordenan de acuerdo con lo hipotetizado. Sólo señalar que en este caso el número de sujetos con combinaciones no esperadas es menor que en el caso anterior (26.38%). Los resultados obtenidos para las subescalas depresión-entusiasmo (ver tablas 7.8 y 7.9) también presentan una interpretación análoga.

Tabla 7.8. Analisis de varianza de los factores "subescala depresión" y "subescala entusiasmo" sobre una escala de diferencial semántico.

| ENTUSIAMO | | BAJO | MEDIO | ALTO | Totals: |
|-----------|-------|--------------------|---------------------|-------------------|--------------|
| DEPRESION | BAJO | 63 3.741 | 149 4.402 | 50 5.31 | 262 4.416 |
| | MEDIO | 17 2.765 | 13 3.538 | 0 . | 30 3.1 |
| Totals: | | 80 3.533 | 162 4.332 | 50 5.31 | 292 4.281 |

Nota: la escala criterio oscila de 1. (depresión) a 7. (entusiasmo). En la parte superior de las celdillas se muestra la frecuencia observada para la combinación correspondiente en ambas subescalas; en la parte inferior se muestra la puntuación media en la escala criterio que presentan los sujetos con esa combinación; en negrita se muestran las combinaciones esperadas bajo el modelo de dos mitades.

Tabla 7.9. Resultados de las fuentes de variación para la escala depresión-entusiasmo.

| FACTORES | G.L. | F | P (2 colas) |
|-------------|------|-------|-------------|
| DEPRESION | 2 | 39.57 | 0.001 |
| ENTUSIASMO | 1 | 29.65 | 0.001 |
| DEP X ENTUS | 2 | 3.04 | 0.049 |

Nota: en este análisis se eliminó el nivel "alto" de la subescala fatigado. $R^2=0.346$

La tercera de la hipótesis formuladas indicaba que una transformación aproximadamente lineal de las puntuaciones obtenidas en ambas subescalas presentaría correlaciones mayores con la escala criterio que las observadas bajo el modelo tradicional. De acuerdo con la correspondencia entre subescalas observada en las figuras 7.1, 7.2, y 7.3, que también resulta corroborada por los análisis de varianza comentados, se ha procedido a realizar la siguiente transformación de las puntuaciones en las subescalas opuestas:

1. hemos categorizado las puntuaciones promediadas en las subescalas de la siguiente forma: [1, 2[: muy bajo; [2, 2.5[: bajo; [2.5, 3.5[: medio-bajo; [3.5, 4.5[: medio-alto; [4.5, 6]: alto. El motivo por el sólo hay una categoría para las puntuaciones altas, mientras que las categorías

baja y media han sido desdobladas en la forma expuesta responde a la distribución que presentan las puntuaciones. Como se aprecia en las figuras 7.1, 7.2, y 7.3, la mayor parte de sujetos presentan puntuaciones bajas y medias en una u otra subescala. En orden a representar adecuadamente la variabilidad observada, no parecía aconsejable dividir las puntuaciones en tres categorías de igual amplitud. El hecho de considerar dos categorías para las puntuaciones bajas permite diferenciar a los sujetos con la combinación "muy bajo en la subescala positiva-bajo en la subescala negativa" de los sujetos con la combinación "bajo en la subescala positiva-muy bajo en la subescala negativa". En este caso consideramos que aunque el sujeto no se identifica con ninguno de los afectos que representan las subescalas, comparativamente el afecto que menos le describe es el que representa la escala en la que obtiene puntuaciones muy bajas. En concreto, la transformación que proponemos hará corresponder al primer caso puntuaciones bajas pero correspondientes al polo negativo, y, en el segundo, puntuaciones bajas pero correspondientes al polo positivo. De forma análoga se proponen dos categorías de puntuaciones medias, de modo que sea posible distinguir los sujetos con puntuaciones "medio altas en la subescala positiva-medio bajas en la subescala negativa" de los sujetos con puntuaciones "medio bajas en la subescala positiva-medio altas en la subescala negativa". Por otra parte, los sujetos que presentan puntuaciones altas en una subescala generalmente presentan también puntuaciones bajas en la subescala opuesta, de modo que no es necesario establecer mayores matizaciones.

2. La transformación de las puntuaciones se ha realizado considerando la interdependencia hipotetizada y observada entre las puntuaciones de ambas subescalas, de modo que diferentes niveles de una u otra subescala llevaran aparejadas diferentes combinaciones de sus puntuaciones. El objetivo era conseguir una serie de combinaciones tales que los sujetos con puntuaciones alto-bajo en las subescalas negativa y positiva presentaran las puntuaciones transformadas más bajas. A continuación debían aparecer los sujetos con puntuaciones medio-bajo, bajo-bajo, bajo-medio, y bajo-alto. Además era necesario ubicar a los sujetos con combinaciones no esperadas como medio-medio, alto-medio, y medio-alto.

Con este objetivo comenzamos a desarrollar la transformación para las combinaciones menos dificultosas. Así, a los sujetos con puntuaciones bajas en una subescala y puntuaciones medias o altas en la subescala opuesta se les ha asignado la suma de las puntuaciones en ambas subescalas, multiplicando esta puntuación por la constante -1 cuando la puntuación baja se daba en la subescala positiva. De este modo quedaron determinados los límites de las puntuaciones en la escala transformada entre -8.5 y 8.5. A su vez, esta transformación dividió la escala transformada en dos partes, una positiva y otra negativa, separadas por un punto central, cero, que se correspondían con los dos polos de la dimensión bipolar hipotetizada.

A continuación, se asignó la constante 0 a los sujetos con puntuaciones muy bajas en una subescala y puntuaciones muy bajas también en la subescala opuesta. De este modo se asignaba a estos sujetos una puntuación "neutral". La subdivisión de las puntuaciones bajas en dos categorías permitió asignar una puntuación baja pero representativa del polo negativo para los sujetos con la combinación "muy bajo en la subescala positiva-bajo en la subescala negativa" y una puntuación baja pero representativa del polo positivo para los sujetos con la combinación "bajo en la subescala positiva-muy bajo en la subescala negativa". En concreto, estas puntuaciones fueron -2 y 2, respectivamente. Finalmente, a los sujetos con la combinación bajo-bajo también se les asignó la constante 0. En definitiva, el propósito de la transformación era reflejar las diferentes ubicaciones de los sujetos en el continuo bipolar, de acuerdo con las hipótesis del modelo de "dos mitades". En los casos en que esto no era posible, se optó por asignar a los sujetos la puntuación neutral.

Continuando con la transformación, a los sujetos con puntuaciones medio bajas en ambas subescalas y a los sujetos con puntuaciones medio altas en ambas subescalas también se les asignó la constante 0. La consideración de dos categorías medias también permitió afinar en la asignación de puntuaciones para los sujetos con estas puntuaciones. De este modo, a los sujetos con puntuaciones "medio altas en la subescala positiva-medio bajas en la subescala negativa" se les asignó la constante 3, y a los sujetos con las puntuaciones "medio bajas en la subescala

positiva-medio altas en la subescala negativa" se les asignó la constante -3.

Únicamente restaba por dilucidar la transformación que necesitaban los sujetos con puntuaciones medio bajas o medio altas en un polo y altas en el otro. De acuerdo con el modelo de "dos mitades", estos sujetos se identificarían máximamente con el polo de la dimensión en el que presentarían puntuaciones altas. Pero, por otra parte, la identificación de los sujetos con ese polo no sería tan clara como la de los sujetos con puntuaciones bajas, en lugar de medias, en el polo opuesto. Finalmente se optó por reflejar esta diferencia asignando a estos sujetos la puntuación que directamente habían obtenido en la subescala con la puntuación alta. De este modo, mientras los sujetos con la combinación bajo-alto recibían en la escala transformada la suma de las puntuaciones en ambas subescalas, los sujetos con la combinación medio-alto recibirían sólo la puntuación correspondiente a la subescala alta, es decir, una puntuación generalmente menor. En el caso de que la puntuación alta se obtuviera en la subescala negativa, la puntuación se multiplicó por la constante -1.

Las instrucciones empleadas en estas transformaciones se ofrecen en la tabla 7.10.

Tabla 7.10. Instrucciones empleadas en la linealización de las puntuaciones observada en subescalas opuestas.

```
>if ((neg<2.5) and (pos>2.5 or pos=2.5)) then let tot= (neg+pos)
>if ((pos<2.5) and (neg>2.5 or neg=2.5)) then let tot= ((neg+pos)*(-1))
>if ((pos<2) and (neg<2)) then let tot=0
>if (pos=2 or (pos>2 and pos<2.5)) and (neg<2) then let tot=2
>if (pos=2 or (pos>2 and pos<2.5)) and (neg=2 or (neg>2 and neg<2.5)) then let tot=0
>if (pos=1 or (pos>1 and pos<2)) and (neg=2 or (neg>2 and neg<2.5)) then let tot=(-2)
>if (pos=2.5 or (pos>2.5 and pos<3.5)) and (neg=2.5 or (neg>2.5 and neg<3.5)) then let tot=0
>if (pos=3.5 or (pos>3.5 and pos<4.5)) and (neg=3.5 or (neg>3.5 and neg<4.5)) then let tot=0
>if (pos=2.5 or (pos>2.5 and pos<3.5)) and (neg=3.5 or (neg>3.5 and neg<4.5)) then let tot=(-3)
>if (pos=3.5 or (pos>3.5 and pos<4.5)) and (neg=2.5 or (neg>2.5 and neg<3.5)) then let tot=3
>if (pos=2.5 or (pos>2.5 and pos<4.5)) and (neg=4.5 or neg>4.5) then let tot=neg* (-1)
>if (pos=4.5 or pos>4.5) and (neg=2.5 or (neg>2.5 and neg<4.5)) then let tot=pos
```

Nota: neg=subescala negativa; pos=subescala positiva; tot=puntuación transformada

Una vez realizadas las transformaciones anteriores, se calcularon las correlaciones entre la puntuación asignada a cada sujeto en cada escala en base a tales transformaciones y la variable criterio correspondiente. Estas correlaciones fueron comparadas con las que mostraban las puntuaciones obtenidas bajo el modelo tradicional con los mismos criterios.

La tabla 7.11 ofrece los resultados correspondientes a la escala tenso-calmado. En ella se aprecia que la correlación entre la puntuación tradicional y la escala criterio es de 0.62, en tanto que la correlación entre ésta y la puntuación bajo el modelo de dos mitades es de 0.567. Aunque la correlación que presenta la puntuación del modelo tradicional es mayor, esta diferencia no es estadísticamente significativa. Tras aplicar el procedimiento de comparación de dos correlaciones con un criterio detallado por Cohen y Cohen (1983), se obtuvo un valor $t=1.01$, ($p>0.05$, dos colas). Además se observa que las puntuaciones bajo ambos modelos de medida presentan una correlación casi perfecta (0.908). Nuevamente se observa también que la correlación entre las subescalas opuestas no es suficiente para ser consideradas realmente "opuestas".

Tabla 7.11. Análisis correlacional. Escala tenso-calmado

| | ds | mt | md | ten | cal |
|-----|-------|-------|-------|-------|-----|
| ds | 1 | | | | |
| mt | .62 | 1 | | | |
| md | .567 | .908 | 1 | | |
| ten | -.484 | -.854 | -.789 | 1 | |
| cal | .594 | .896 | .802 | -.534 | 1 |

Nota: ds=puntuación en la escala criterio; mt=puntuación bajo el modelo tradicional; dm=puntuación bajo el modelo de "dos mitades"; ten=puntuación en la subescala tensión; cal=puntuación en la subescala calma. $r>0.194$, $p<0.05$; $r>0.253$, $p<0.01$

Los resultados del análisis correlacional para la escala deprimido-entusiasmado presentan un patrón similar al de la escala anterior. En la

tabla 7.12 se observa que la correlación entre la puntuación tradicional y la escala criterio es de 0.69, en tanto que la correlación entre ésta y la puntuación bajo el modelo de dos mitades es de 0.612. No obstante, la diferencia entre estas dos correlaciones sí resultó estadísticamente significativa ($t=3.25$; $p<0.05$, dos colas). También se observa que la correlación entre las puntuaciones bajo ambos modelos de medida es muy alta (0.85). Finalmente, la correlación entre las subescalas opuestas tampoco es suficiente para ser consideradas realmente "opuestas".

Tabla 7.12. Análisis correlacional. Escala deprimido-entusiasmado

| | ds | mt | dm | DEP | ENT |
|-----|-------|-------|-------|-------|-----|
| ds | 1 | | | | |
| mt | .69 | 1 | | | |
| dm | .612 | .85 | 1 | | |
| DEP | -.528 | -.785 | -.676 | 1 | |
| ENT | .627 | .895 | .755 | -.427 | 1 |

Nota: ds=puntuación en la escala criterio; mt=puntuación bajo el modelo tradicional; dm=puntuación bajo el modelo de "dos mitades"; DEP=puntuación en la subescala deprimido; ENT=puntuación en la subescala entusiasmado. $r>0.194$, $p<0.05$; $r>0.253$, $p<0.01$

Los resultados del análisis correlacional para la escala fatiga-vigor confirman la tendencia mostrada por las subescalas anteriores. En la tabla 7.13 se observa que la correlación entre la puntuación tradicional y la escala criterio es de 0.725, en tanto que la correlación entre ésta y la puntuación bajo el modelo de dos mitades es de 0.694. La correlación que presenta la puntuación del modelo tradicional es mayor, aunque nuevamente esta diferencia no es estadísticamente significativa ($t=1.74$; $p>0.05$, dos colas). También se observa que la correlación entre las puntuaciones bajo ambos modelos de medida es casi perfecta (0.906), y que la correlación entre las subescalas opuestas tampoco es suficiente para ser consideradas realmente "opuestas".

Tabla 7.13. Análisis correlacional. Escala fatigado-vigor

| | ds | mt | dm | fat | vig |
|-----|-------|-------|-------|-------|-----|
| ds | 1 | | | | |
| mt | .725 | 1 | | | |
| md | .694 | .906 | 1 | | |
| fat | -.541 | -.752 | -.685 | 1 | |
| vig | .616 | .846 | .764 | -.285 | 1 |

Nota: ds=puntuación en la escala criterio; mt=puntuación bajo el modelo tradicional; dm=puntuación bajo el modelo de "dos mitades"; fat=puntuación en la subescala fatiga; vig=puntuación en la subescala vigor. $r > 0.194$, $p < 0.05$; $r > 0.253$, $p < 0.01$

Como indican los resultados de los análisis correlacionales realizados, las correlaciones que presentaron las puntuaciones calculadas bajo el modelo de "dos mitades" con el criterio no fueron mayores que las correspondientes al modelo tradicional, en contra de lo esperado por la tercera de las hipótesis formuladas. Tal como estaba formulada esta hipótesis, quedaba implícito que las correlaciones del modelo tradicional con el criterio podían ser mejoradas. No obstante, no había evidencia de ello. En consecuencia, cabía la posibilidad de que estas correlaciones fueran tan altas que no pudieran más que ser igualadas. Y, en ese caso, la formulación de esta hipótesis no sería la adecuada.

Para determinar si la formulación de la tercera hipótesis era adecuada, era necesario obtener algún tipo de evidencia que mostrara que las correlaciones de las puntuaciones bajo el modelo tradicional y las escalas criterio eran superables ¿Cómo obtener esa evidencia? Comparando las correlaciones de otra escala explícitamente bipolar con el mismo criterio. Si contáramos con las puntuaciones de los sujetos de la muestra en otra escala explícitamente bipolar, cabría esperar que las correlaciones mayores aparecerían entre la nueva escala explícitamente bipolar y la escala criterio, seguidas de las correlaciones obtenidas en la escala tipo Likert -bajo el modelo tradicional o dos mitades- también con ese criterio.

obtenidas en la escala bipolar bajo el modelo tradicional, y entre las puntuaciones en la escala criterio y la nueva escala explícitamente bipolar indicó que ambas correlaciones no diferían significativamente ($t=1.069$, $p>0.05$, dos colas). Por su parte, la comparación de las correlaciones entre las puntuaciones en escala la criterio y las obtenidas en la escala bipolar bajo el modelo de "dos mitades", y entre las puntuaciones en la escala criterio y la nueva escala explícitamente bipolar también indicó que ambas correlaciones no diferían significativamente ($t=0.159$, $p>0.05$, dos colas).

Tabla 7.16. Análisis correlacional II. Escala fatigado-vigorouso (N=284)

| | DS | PL | MT | DM |
|----|-------|-------|-------|----|
| DS | 1 | | | |
| PL | 0.690 | 1 | | |
| MT | 0.728 | 0.578 | 1 | |
| DM | 0.696 | 0.555 | 0.906 | 1 |

Nota: Ds=puntuación en la escala criterio; PL=puntuación en la nueva escala explícitamente bipolar; MT=puntuación bajo el modelo tradicional; DM=puntuación bajo el modelo de "dos mitades" $r>0.194$, $p<0.05$; $r>0.253$, $p<0.01$

Los resultados de estos nuevos análisis indican que las correlaciones que presentan las puntuaciones calculadas bajo el modelo tradicional con el criterio son tan altas como las que se observan entre la escala criterio y otra escala explícitamente bipolar. En otras palabras, los resultados indican que las correlaciones entre las puntuaciones calculadas bajo el modelo tradicional y las escalas criterio *no pueden ser mejoradas*, pues son tan altas como las que presenta otra escala explícitamente bipolar. Por otra parte, estos resultados indican lo mismo respecto de las correlaciones que ofrecen las puntuaciones calculadas bajo el modelo de "dos mitades".

A la luz de estos análisis cabe concluir la inadecuación de la tercera hipótesis en su formulación original. Puesto que las correlaciones con el

Por su parte, la comparación de las correlaciones entre las puntuaciones en la escala criterio y las obtenidas en la escala bipolar bajo el modelo de "dos mitades", y entre las puntuaciones en la escala criterio y la nueva escala explícitamente bipolar también indicó que ambas correlaciones no diferían significativamente ($t=1.871$, $p>0.05$, dos colas).

Las correlaciones relativas a la escala deprimido-entusiasmado aparecen en la tabla 7.15. La comparación de las correlaciones entre las puntuaciones en la escala criterio y las obtenidas en la escala bipolar bajo el modelo tradicional, y entre las puntuaciones en la escala criterio y la nueva escala explícitamente bipolar indicó que ambas correlaciones no diferían significativamente ($t=0.610$, $p>0.05$, dos colas). Por su parte, la comparación de las correlaciones entre las puntuaciones en la escala criterio y las obtenidas en la escala bipolar bajo el modelo de "dos mitades", y entre las puntuaciones en la escala criterio y la nueva escala explícitamente bipolar también indicó que ambas correlaciones no diferían significativamente ($t=0.416$, $p>0.05$, dos colas).

Tabla 7.15. Análisis correlacional II. Escala deprimido-entusiasmado (N=285)

| | DS | PL | MT | DM |
|----|-------|-------|-------|----|
| DS | 1 | | | |
| PL | 0.640 | 1 | | |
| MT | 0.681 | 0.639 | 1 | |
| DM | 0.608 | 0.535 | 0.850 | 1 |

Nota: Ds=puntuación en la escala criterio; PL=puntuación en la nueva escala explícitamente bipolar; MT=puntuación bajo el modelo tradicional; DM=puntuación bajo el modelo de "dos mitades" $r>0.194$, $p<0.05$; $r>0.253$, $p<0.01$

Finalmente, las correlaciones relativas a la escala deprimido-entusiasmado aparecen en la tabla 7.16. Nuevamente la comparación de las correlaciones entre las puntuaciones en la escala criterio y las

Para comprobar esta nueva hipótesis se calcularon las correlaciones de las escalas bajo estudio y de las escalas criterio con tres nuevas escalas explícitamente bipolares. Cada escala constaba de un ítem con el encabezado "pensando en las últimas semanas, me he sentido:". La escala de respuesta ofrecía las siguientes alternativas 1. muy *tenso*; 2. bastante *tenso*; 3. algo *tenso*; 4. ni *tenso* ni *calmado*; 5. algo *calmado*; 6. bastante *calmado*; 7. muy *calmado* (cada escala presentaba el par de adjetivos en cursiva que correspondían a cada una de las tres dimensiones medidas). A continuación se calcularon las correlaciones de las puntuaciones obtenidas en esta nueva escala con las obtenidas: 1) en la escala criterio; 2) en la escala objeto de estudio bajo el modelo tradicional; y 3) en la escala objeto de estudio bajo el modelo de "dos mitades" en cada una de las dimensiones estudiadas.

Las correlaciones relativas a la escala tenso-calmado aparecen en la tabla 7.14. La comparación de las correlaciones entre las puntuaciones en la escala criterio y las obtenidas en la escala bipolar bajo el modelo tradicional, y entre las puntuaciones en la escala criterio y la nueva escala explícitamente bipolar indicó que ambas correlaciones no diferían significativamente ($t=0.797$, $p>0.05$, dos colas).

Tabla 7.14. Análisis correlacional II. Escala tenso-calmado (N=280)

| | DS | PL | MT | DM |
|----|-------|-------|-------|----|
| DS | 1 | | | |
| PL | 0.653 | 1 | | |
| MT | 0.618 | 0.481 | 1 | |
| DM | 0.566 | 0.456 | 0.908 | 1 |

Nota: Ds=puntuación en la escala criterio; PL=puntuación en la nueva escala explícitamente bipolar; MT=puntuación bajo el modelo tradicional; DM=puntuación bajo el modelo de "dos mitades" $r>0.194$, $p<0.05$; $r>0.253$, $p<0.01$

criterio bajo el modelo tradicional son tan altas como las que presenta otra escala explícitamente bipolar, no es pertinente esperar correlaciones mayores bajo ningún otro modelo, y sólo cabe que las obtenidas bajo el modelo de "dos mitades" las igualen. De modo que los resultados inicialmente obtenidos apoyan la validez del modelo de "dos mitades", por cuanto que consigue igualar las correlaciones que presentan las puntuaciones obtenidas bajo el modelo tradicional con las escalas criterio.

2. CONCLUSIONES.

En definitiva, los resultados obtenidos apoyan las tres hipótesis formuladas acerca de la bipolaridad de las dimensiones medidas bajo el modelo de las "dos mitades". La evidencia empírica muestra que la relación entre las puntuaciones en las subescalas opuestas no es lineal e inversa, tal como indicaba la primera hipótesis. Sin embargo, la relación que aparece confirma la bipolaridad de las dimensiones medidas, de acuerdo con la segunda hipótesis. Finalmente, las correlaciones de las puntuaciones calculadas bajo este modelo con la escala criterio no difieren significativamente de las que presenta el modelo tradicional con los mismos criterios (excepto en el caso de la escala deprimido-entusiasmado). Todo ello invalida claramente la inferencia de que los sujetos al contestar se ubican en un continuo bipolar, y que éste es invariante respecto del contenido positivo o negativo del ítem.

Desde esta perspectiva, puede explicarse tanto la aparición como la ausencia de relaciones lineales inversas entre las puntuaciones en subescalas opuestas, es decir, los resultados que han favorecido la polémica acerca de la estructura del afecto -y de otras dimensiones diferentes-. En principio, y de acuerdo con la observación de Kerlinger (1984), cabe esperar este tipo de relaciones inversas en la medida en que la muestra incluya gran número de sujetos extremos en cualquiera de los polos. Los sujetos que presentan puntuaciones altas en una de las subescalas -extremos- se identifican con ese extremo de la dimensión, y consecuentemente presentan puntuaciones bajas en la subescala opuesta. Ello da lugar a la aparición de relaciones lineales e inversas, incluso cuando la escala de respuesta no es percibida como bipolar ni

complementaria. Sin embargo, a medida que la muestra esté formada por sujetos para los que: 1) la dimensión medida no sea relevante, 2) se ubiquen en el centro de la dimensión, o 3) se identifiquen sólo moderadamente con uno de los polos de la misma, no aparecerán las relaciones lineales e inversas esperadas como evidencia de la bipolaridad.

Lo que también muestran con bastante claridad los resultados obtenidos es que a pesar de que las subescalas opuestas estudiadas no presenten el tipo de relaciones inversas esperadas bajo el modelo bipolar tradicional -cada par de subescalas comparte entre un 8% y un 27% de varianza-, su combinación en la forma tradicional -inversión y promedio de items- da lugar a una puntuación final válida, y en cualquier caso mejor que cualquiera de las puntuaciones de cada subescala por separado.

De este modo, se observa que la correlación entre las puntuaciones de los sujetos en las subescalas fatiga y vigor es de $-.285$ (ver tabla 7.13), lo que indica que ambas subescalas apenas comparten un 8% de su varianza. Ello indica claramente que ambas subescalas no pueden ser consideradas, desde el modelo tradicional, como parte de un continuo bipolar. No aparece una relación lineal inversa de la magnitud esperada. Sin embargo, la correlación de la puntuación total calculada bajo el modelo tradicional con la obtenida en la escala de diferencial semántico es de $.72$. Finalmente, las correlaciones de las subescalas por separado con la escala de diferencial semántico son inferiores a la que presentan combinadas ($-.541$ y $.616$).

En la tabla correspondiente a las subescalas tensión-calma (ver tabla 7.11) se observa que la correlación ambas subescalas es de $-.534$, lo que indica que ambas subescalas comparten un 27% de su varianza. Aunque la relación observada es mayor que en el caso anterior, aún no es suficiente como para considerarlas parte de un continuo bipolar. La correlación de la puntuación total calculada bajo el modelo tradicional es de $.62$, en tanto que las correlaciones de las subescalas por separado con la escala de diferencial semántico son inferiores a la que presentan combinadas ($-.48$ y $.57$).

Finalmente, en la tabla correspondiente a las subescalas depresión-entusiasmo (ver tabla 7.12) se observa que la correlación entre ambas subescalas es de $-.427$, lo que indica que ambas subescalas comparten un 18% de su varianza. La relación observada es menor que en el caso anterior, e insuficiente como para considerarlas parte de un continuo bipolar. La correlación de la puntuación total calculada bajo el modelo tradicional con la obtenida en la escala de diferencial semántico es de $.69$, y las correlaciones de las subescalas por separado con la escala de diferencial semántico son inferiores a la que presentan combinadas ($-.528$ y $.627$).

El hecho de que la correlación entre la escala total y la escala de diferencial semántico sea mayor que las que presentan sus subescalas por separado resulta congruente con el modelo de "las dos mitades", por cuanto que cada subescala indica la ubicación de los sujetos sólo respecto de un polo, en tanto que la escala total contiene información referida a la dimensión en toda su amplitud, aunque combinada en una forma inadecuada. Lo que realmente resulta sorprendente, después de la evidencia obtenida en los análisis anteriores, es que la puntuación final obtenida bajo el modelo tradicional, presente una correlación tan alta con la obtenida bajo el modelo de las dos mitades (entre $.85$ y $.906$). Sin embargo, ésta puede ser la explicación de porqué se ha encontrado reiteradamente la paradoja por la que escalas que presentaban una estructura factorial monopolar, sin embargo ofrecían puntuaciones en las subescalas monopulares que combinadas linealmente resultaban predictores válidos de puntuaciones en escalas bipolares, como es el caso aquí. Parece que estas puntuaciones, aunque calculadas bajo premisas erróneas, reproducen en gran medida la ordenación de los sujetos que se obtiene de aplicar el modelo de "las dos mitades".

Otro aspecto que puede ser reintrepretado a la luz del modelo de "las dos mitades" es el del efecto de la aquiescencia. En ocasiones, el control o eliminación del sesgo de aquiescencia ha conducido a la obtención de resultados en favor de la bipolaridad en sentido tradicional, por cuanto que tras su eliminación correlaciones no significativas o de escasa magnitud entre subescalas opuestas alcanzaban valores suficientes para

ser considerados evidencia en favor de la bipolaridad de la dimensión medida. En otros términos, la parcialización de la medida de aquiescencia incrementaba las relaciones lineales inversas hasta alcanzar los niveles esperados. Reflexionando sobre la composición de esa medida se comprende el efecto producido por su parcialización. Esta medida se ha calculado de dos formas diferentes, bien sumando todos los items excepto los del par de escalas en que se iba a aplicar, tras lo cual se calculaba la correlación parcial entre esas subescalas, o bien sumando todos los items, y recalculando la matriz de correlaciones parciales. En ambos casos, la medida de aquiescencia era tal que los sujetos con puntuaciones bajas en todos los items presentaban puntuaciones bajas, los sujetos con puntuaciones altas en todos los items presentaban puntuaciones altas, y entre ambos extremos se encontraban los sujetos con las restantes combinaciones. Por consiguiente, se consideraba que los sujetos con puntuaciones bajas o altas en items o subescalas de signo opuestos estaban contestando bajo el efecto del sesgo de aquiescencia. Al parcializar esta medida, se eliminaba la influencia de los sujetos con puntuaciones iguales en subescalas o items opuestos, lo que incrementaba el peso relativo de los sujetos con puntuaciones diferentes en items o subescalas opuestas, que a su vez incrementaba la relación inversa esperada. Bajo el modelo de las mitades, sólo los sujetos con puntuaciones altas en items opuestos estarían afectados por este sesgo. Los sujetos con puntuaciones bajas en ambas subescalas no estarían afectados por este sesgo, como muestran los resultados de los análisis de varianza anteriormente comentados. Su eliminación consigue efectivamente linealizar las relaciones observadas, pero no a costa de eliminar un sesgo de respuesta, sino oscureciendo las relaciones de interdependencia que parecen existir entre ambas subescalas.

En resumen, los resultados obtenidos indican que, tal como se hipotetizaba, es el uso de inferencias no contrastadas empíricamente - como las de la bipolaridad de la escala de respuesta y la complementariedad consiguiente- lo que explica la falta de evidencia empírica en favor de la bipolaridad de las dimensiones medidas. La doble implicación de que si la dimensión es bipolar, entonces deben aparecer relaciones lineales inversas, y si no aparecen estas relaciones,

es que la dimensión no es bipolar no es totalmente cierta. Sólo si la dimensión es bipolar, y las escalas cumplen determinados supuestos, podemos esperar esas relaciones, y si no aparecen, puede ser por efecto de que alguna de las dos condiciones no se cumpla. En este capítulo se ha mostrado evidencia que indica que es ésto lo que puede haber explicado la polémica en torno a la estructura del afecto, pues se demuestra cómo la falta de evidencia en favor de la bipolaridad de la dimensión se debe a los supuestos incorrectos que se realizan sobre la escala. Cuando estos supuestos se eliminan, la distribución de las puntuaciones indica que la dimensión medida es efectivamente bipolar.

Por otra parte, también se ha dado una explicación parcial a la paradoja de la falta de validez factorial, que sin embargo se acompaña de una buena validez de criterio. En este sentido, se ha observado cómo la ordenación de los sujetos en el continuo bipolar bajo el modelo tradicional es análoga a la que se obtiene bajo el modelo de "dos mitades", a pesar de emplear otras premisas como punto de partida. Esto podría explicar por qué se obtienen buenos indicadores de validez de criterio.

La reinterpretación de los resultados obtenidos bajo el modelo de "dos mitades" invalida pues el uso de métodos de análisis factorial exploratorio o confirmatorio para evaluar la estructura subyacente en este tipo de escalas. Estos métodos parten del supuesto de que las relaciones entre subescalas o ítems opuestos son lineales e inversas, y ésta es una premisa equivocada.

No obstante, el modelo de "dos mitades" no es el único que se ha desarrollado para explicar la paradoja de la bipolaridad anteriormente mencionada. A continuación pasamos a exponer y analizar otro enfoque alternativo cuyo objetivo es también explicar este fenómeno.

PARTE III

**ACERCAMIENTOS ALTERNATIVOS AL PROBLEMA DE LA
MEDIDA DE DIMENSIONES BIPOLARES AFECTIVAS MEDIANTE
ESCALAS TIPO LIKERT CONTRABALANCEADAS:
ESCALAMIENTO UNIDIMENSIONAL DESPLEGADO.**

CAPITULO 8

INTRODUCCION:

ESCALAMIENTO DE SUJETOS FRENTE A ESCALAMIENTO DE SUJETOS Y ESTIMULOS.

1. INTRODUCCION.

Como hemos visto, las escalas construidas bajo el modelo tradicional presentan, sin embargo, una estructura factorial en la que las pretendidas dimensiones bipolares se desdoblan en dos factores monopolares, en los que los items saturan en función del polo del continuo bipolar del que son representativos. Y también hemos visto que estas escalas ofrecen puntuaciones finales válidas, en mayor medida que las obtenidas en cada subescala por separado, pues predicen adecuadamente las puntuaciones observadas con otros indicadores. Esto llevaba a la paradoja de admitir que escalas empíricamente monopolares eran mejores indicadores del afecto cuando eran tratadas como escalas bipolares que cuando eran tratadas como las escalas monopolares que eran. Y esta paradoja ha llevado recientemente a la aplicación y desarrollo de otro enfoque alternativo al presentado anteriormente, que cuestiona no ya el modelo de medida, sino el método de escalamiento empleado: el enfoque del modelo de desplegamiento unidimensional (Coombs, 1950, 1952, 1964).

Las diferencias entre los enfoques anteriormente expuestos y el que se mostrará a continuación pueden ser ilustradas siguiendo la clasificación que Torgerson (1958) realiza de las diferentes aproximaciones al problema de la medida. En este sentido, Torgerson (1958) distingue entre la aproximación "centrada en el sujeto", la aproximación "centrada en el estímulo" y la aproximación "centrada en la respuesta". Cada una de estas aproximaciones se caracteriza por la fuente a la que se atribuye la variabilidad en las puntuaciones observadas. La variabilidad sistemática observada en las puntuaciones de una escala puede ser explicada por los sujetos que contestan, y/o por los estímulos a los que contestan. La aproximación centrada en los sujetos supone que los estímulos o items son réplicas unos de otros, de modo que toda la variabilidad sistemática queda atribuida a las diferencias individuales entre los sujetos. En consecuencia, se dice que el objetivo de esta aproximación es el escalamiento de sujetos. Por el contrario, en la aproximación centrada en el estímulo se parte del supuesto de que los sujetos son réplicas unos de otros, por lo que toda la

variabilidad sistemática observada en las puntuaciones es atribuida a los estímulos. Las técnicas basadas en esta aproximación tienen por objetivo el escalamiento de estímulos. Finalmente, la aproximación centrada en la respuesta considera ambas fuentes de variabilidad sistemática: sujetos y estímulos. A diferencia de las aproximaciones anteriores, ésta considera que la variabilidad en las puntuaciones observadas se debe a las diferentes ubicaciones que tanto sujetos como estímulos presentan en la dimensión bajo estudio. El objetivo de los métodos centrados en esta aproximación es el escalamiento tanto de sujetos como de estímulos.

Como se recordará, bajo el modelo tradicional bipolar la puntuación final, obtenida tras la inversión de los ítems pertinentes, se interpreta como la ubicación del sujeto en la dimensión bipolar bajo estudio. Esta interpretación presupone que los ítems son réplicas unos de otros, que la escala de respuesta es bipolar e invariante al signo del ítem, y que las relaciones entre los ítems, y entre éstos y la dimensión latente es monótona. Esta aproximación puede clasificarse entre las "aproximaciones centradas en el sujeto", tal como se definieron anteriormente. En definitiva, la variabilidad sistemática observada en las puntuaciones se atribuye a diferencias individuales en la dimensión objeto de medida, y se espera que la ubicación del sujeto en los ítems representativos de un polo determine su puntuación en los ítems correspondientes al polo opuesto.

Bajo el modelo de "dos mitades" anteriormente presentado, la puntuación obtenida en los ítems de un polo sólo determina la puntuación en los ítems correspondientes al polo opuesto si se trata de una puntuación alta o media. Se hipotetiza que la puntuación en cada subescala indica la posición del sujeto respecto del polo de la dimensión bipolar de la que los ítems son representativos. Los ítems del mismo polo son réplicas unos de otros, pero no los de polos opuestos. Por lo tanto las relaciones entre los ítems y la dimensión latente sólo es monótona con respecto al polo de la dimensión de la que éstos son representativos. Además, mientras las relaciones entre ítems de igual signo sí son lineales, la relación entre ítems o subescalas opuestas no es lineal e inversa sino interdependiente. Con ello se invalida el uso del

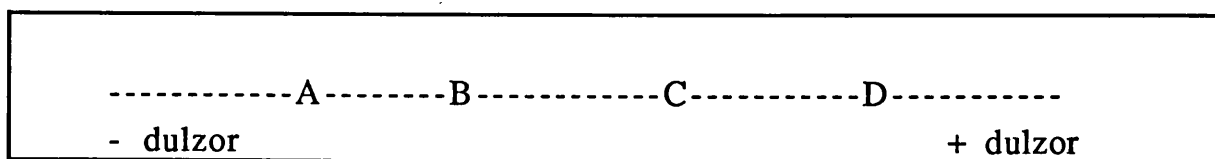
análisis factorial como método de evaluación de la dimensionalidad de la escala. A pesar de las diferencias señaladas, ésta también es una aproximación "centrada en el sujeto".

En cambio, bajo el enfoque del modelo de despliegamiento unidimensional, los items no son considerados réplicas unos de otros, no se trata de una aproximación "centrada en los sujetos". Por el contrario, este enfoque se clasifica entre las aproximaciones "centradas en la respuesta", que como se indicó anteriormente, atribuyen la variabilidad sistemática observada en las puntuaciones no sólo a las diferentes ubicaciones de los sujetos en el continuo latente, sino también a las diferentes ubicaciones de los items, que pueden ser ordenados en función del grado en que contienen el atributo medido. Bajo el modelo de despliegamiento unidimensional desarrollado por Coombs (1950, 1952, 1964), la respuesta del sujeto al ítem indicaría la distancia psicológica que el sujeto percibe entre su posición en el continuo, y la que le atribuye al ítem.

2. EL MODELO DE ESCALAMIENTO DESPLEGADO DE COOMBS (1950).

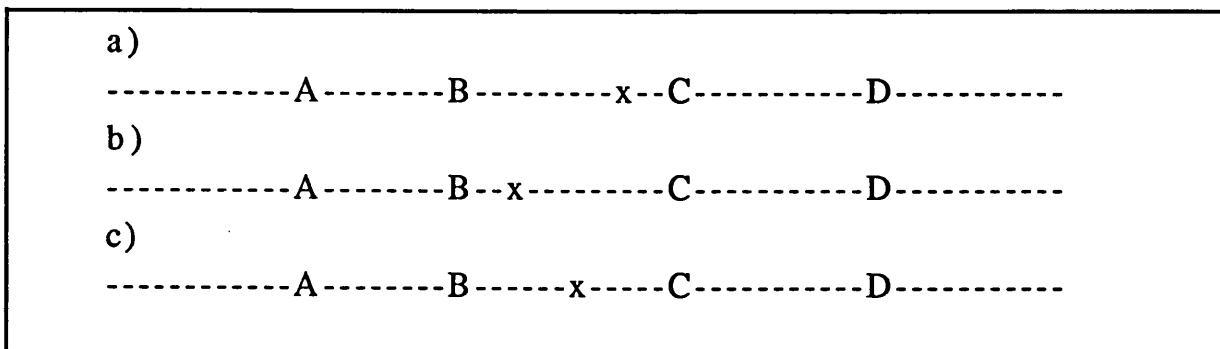
Para comprender la lógica que subyace a este modelo tomaremos un ejemplo sencillo (Torgerson, 1958). Imaginemos cuatro helados A, B, C y D, tales que su ordenación en la dimensión "dulzor" resulta conocida: $A < B < C < D$, si representamos gráficamente la posición de cada helado en la dimensión elegida obtendríamos la siguiente configuración:

Figura 8.1. Ubicación de los estímulos A B C y D en la dimensión grado de dulzor.



Ahora imaginemos que pedimos a un sujeto que deguste cada uno de los helados, y los ordene de acuerdo con sus preferencias en función del dulzor que presentan. El sujeto podría dar la siguiente ordenación: CBAD. Puesto que conocemos el orden de los helados en la escala, y sabemos que es invariable a las preferencias de un sujeto particular, podemos tomar la ordenación que ha realizado el sujeto para inferir su punto ideal en la dimensión dulzor, es decir, su punto de dulzor preferido. El sujeto ha ordenado los helados en función de sus preferencias particulares, colocando en primer lugar el que se encuentra más cercano a su punto ideal de dulzor, en segundo lugar al que le sigue, y así sucesivamente. En definitiva, nos ha indicado la distancia que existe entre cada estímulo y su punto ideal. Si sabemos el orden absoluto de los estímulos, y el orden relativo que toman para un sujeto dado -para un punto ideal determinado-, entonces podemos determinar cuál es el punto en la dimensión para el que los estímulos se ordenan en la forma indicada por el sujeto, es decir, el punto ideal del sujeto. En este caso, la ordenación CBAD correspondería a un punto ideal situado entre los puntos C y B, pero cualquier punto dentro de este intervalo no produciría la ordenación señalada. Tal como muestra la figura 8.2, sólo la representación gráfica "c" ofrece una representación adecuada.

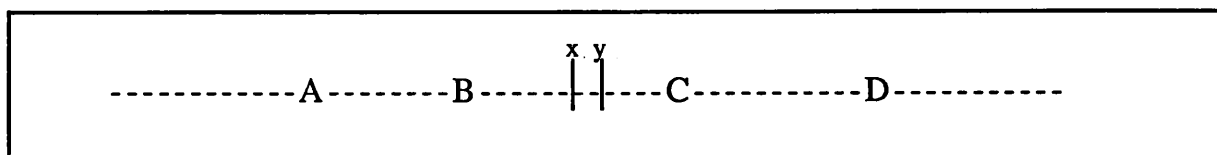
Figura 8.2. Representación de la ubicación de un sujeto que presenta la ordenación CBAD.



La representación del punto ideal que ofrecen las gráficas a y b producirían la ordenación CBDA y BCAD, respectivamente, ya que en el primer caso, la distancia desde el punto ideal (x) es menor respecto del estímulo D que respecto del estímulo A, y en el segundo caso, la distancia desde el punto ideal (x) es menor respecto del estímulo B que

respecto del estímulo C. El punto a partir del cual la representación adecuada es CBDA, BCAD o CBAD depende del punto medio entre los estímulos B y C, y los estímulos A y D.

Figura 8.3. Representación de la ubicación de un sujeto que presenta la ordenación CBAD.



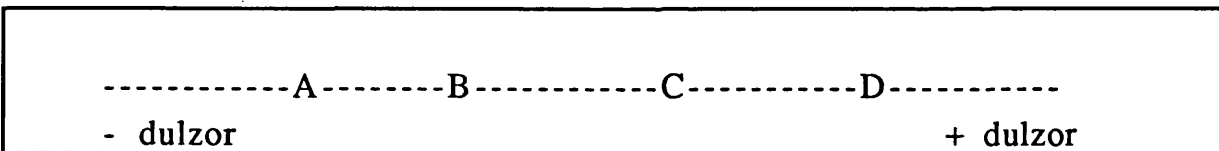
El punto x indica el punto medio entre los estímulos B y C, y el punto y indica el punto medio entre los estímulos A y D. La ordenación CBAD se obtendrá para cualquier punto ideal ubicado a la derecha del punto medio x , y a la izquierda del punto medio y , lo que indica respectivamente que el estímulo C está más próximo al punto ideal del sujeto que el estímulo B, y que el estímulo A está más cerca de ese punto que el estímulo D.

Este ejemplo ilustra cómo el orden de preferencia de un sujeto sobre un conjunto de estímulos puede ser empleado para inferir la ubicación del sujeto en el continuo subyacente, siempre que se conozca la ubicación de los estímulos en ese continuo. La ordenación del sujeto indica la distancia a la que se encuentra el punto ideal del sujeto de cada estímulo, de modo que cuanto menor es esa distancia, mayor es la identificación del sujeto con la posición atribuida al ítem. En este sentido las repuestas del sujeto son denominadas de "preferencia" (Coombs, 1952) o comparativas (Torgerson, 1958), pues indican el orden de preferencia que el sujeto atribuye a los estímulos, en función de la distancia a que se encuentran del punto ideal en el que se ubica el sujeto. En consecuencia, este modelo identifica primero la ordenación de los adjetivos a lo largo del continuo bipolar -escalamiento de estímulos-, y a continuación, determina la posición de los sujetos en ese continuo en función de la ordenación que realizan de los adjetivos -escalamiento de sujetos-. Puesto que el orden de los estímulos en la dimensión es independiente de la ubicación de los sujetos, y las respuestas de los

sujetos indican la distancia entre la ubicación que se atribuyen los sujetos y la que presenta el ítem, es posible determinar cuál es la ubicación de los sujetos: se trata de identificar en qué lugar del continuo las distancias relativas de los ítems se ordenan en la forma indicada por el sujeto.

No obstante, puede darse el caso de que un sujeto dé una ordenación de los estímulos que no pueda ser representada atendiendo a una única dimensión. Continuando con el ejemplo anterior, un sujeto que ofreciera la ordenación ACBD no podría ser representado porque no hay ningún punto en la dimensión objeto de estudio en que A sea el estímulo más próximo, seguido de C, y seguido de B y D (ver figura 8.1). Si A es el estímulo más proximo respecto del punto ideal de un sujeto dado, a continuación debe serlo B.

Figura 8.1. Ubicación de los estímulos A B C y D en la dimensión grado de dulzor.



La ordenación absoluta de los estímulos en la dimensión bajo estudio, entendiendo por absoluta el hecho de que es independiente de los juicios o preferencias particulares de los sujetos, es denominada por Coombs (1952) "escala J", e indica la posición de cada estímulo en la dimensión, o lo que es lo mismo, el grado en que cada estímulo posee el atributo que mide la dimensión bajo estudio. La ordenación particular que cada sujeto realiza del conjunto de estímulos que forman la escala J es denominada por Coombs (1952) escala I, e indica la ordenación de los estímulos respecto del punto ideal del sujeto. Como hemos visto con el ejemplo anterior, no todas las escalas I tienen representación en una escala J determinada. El número de ordenaciones posibles de un conjunto de estímulos viene dado por el factorial del número de estímulos (n!), en tanto que el número de representaciones posibles para una escala J dada viene dado por la expresión:

$$\binom{n}{2} + 1 = \left[\frac{n!}{2!(n-2)!} \right] + 1$$

De modo que para el caso de cuatro estímulos, se pueden obtener 24 ordenaciones posibles, de las cuales sólo 7 son representables o compatibles con la escala J correspondiente. El modo en que pueden identificarse las 7 ordenaciones posibles puede encontrarse en Coombs (1952), Torgerson (1958) y Arce (1994) entre otros.

Hasta aquí se ha ejemplificado la lógica de este modelo de escalamiento sobre el supuesto de que la escala J era conocida. No obstante, en la práctica, y especialmente en la medida de atributos psicológicos, esto no es así. De hecho, el primer paso en la aplicación de este método consiste en recoger las ordenaciones de los sujetos, escalas I. A partir de estas escalas se puede proceder a la identificación de la escala J. Para ello es necesario que las escalas I reúnan, al menos, las siguientes características (Arce, 1994; McIver y Carmines, 1986; Coombs 1952):

- 1) las escalas I han de acabar siempre en uno de dos estímulos; y
- 2) existirán dos escalas I tales que una es imagen en espejo de otra.

Si las escalas I satisfacen estos supuestos, entonces puede aplicarse el procedimiento ideado por Coombs (1952) para la identificación de la escala J subyacente. Este procedimiento se encuentra ilustrado en los textos de Torgerson y Arce arriba citados. No obstante, en la práctica es muy improbable que todas las escalas I observadas en una muestra de sujetos y para una muestra de estímulos se ajusten a los requisitos mencionados anteriormente. En este sentido cabe destacar el hecho de que el modelo de escalamiento desplegado tal como fue desarrollado por Coombs (1952) es un modelo completamente determinista, ya que "el modelo, en sí mismo, no preve ninguna fuente de variación en las respuestas más que la que se explica por la relación que mantienen los sujetos y los estímulos con la dimensión subyacente" (Torgerson, 1958, p. 404). Consecuentemente, el propio Coombs (Coombs, 1964) sugiere que si es posible ajustar el 50%+1 de las escalas I sobre una escala J

cuantitativa común, en la que se representa tanto el orden de los estímulos como las distancias relativas que los separan (para una explicación detallada de la diferencia entre escalas J cualitativas y cuantitativas remitimos al lector a los textos citados), se puede aceptar la representación conjunta de estímulos y sujetos en un único continuo latente en la forma en que especifica su modelo. En general, la literatura especializada (Torgerson, 1958; Arce, 1994; McIver y Carmines, 1986) considera insuficiente el grado de desarrollo alcanzado en la evaluación de la capacidad de reproductibilidad que este modelo puede presentar respecto de un conjunto de datos.

3. APROXIMACION PROBABILISTICA AL MODELO DE ESCALAMIENTO DESPLEGADO DE COOMBS (1950).

Una de las aportaciones más importantes en la aplicación práctica de esta aproximación al escalamiento simultáneo de sujetos y estímulos, basada en la idea de proximidad psicológica, es la realizada por Van Schuur (1988). Este autor ha desarrollado una aproximación probabilística al modelo de escalamiento desplegado propuesto por Coombs, así como un programa (Van Schuur y Post, 1990) mediante el cual es posible identificar y evaluar el ajuste de una determinada escala J a un conjunto de datos.

Al igual que en el modelo determinista, desde este modelo probabilístico las puntuaciones dadas a los ítems no indican directamente la posición del sujeto en el continuo latente, indican la proximidad que existe en ese continuo entre la posición del sujeto y la del ítem. El método de escalamiento es obviamente diferente al escalamiento tipo Likert, y el método de análisis para identificar la dimensionalidad de escalas que presentan esta estructura subyacente no es, por tanto, el análisis factorial, sino el método de desplegamiento unidimensional -en el caso de que los ítems se ordenen en función de una sola dimensión-.

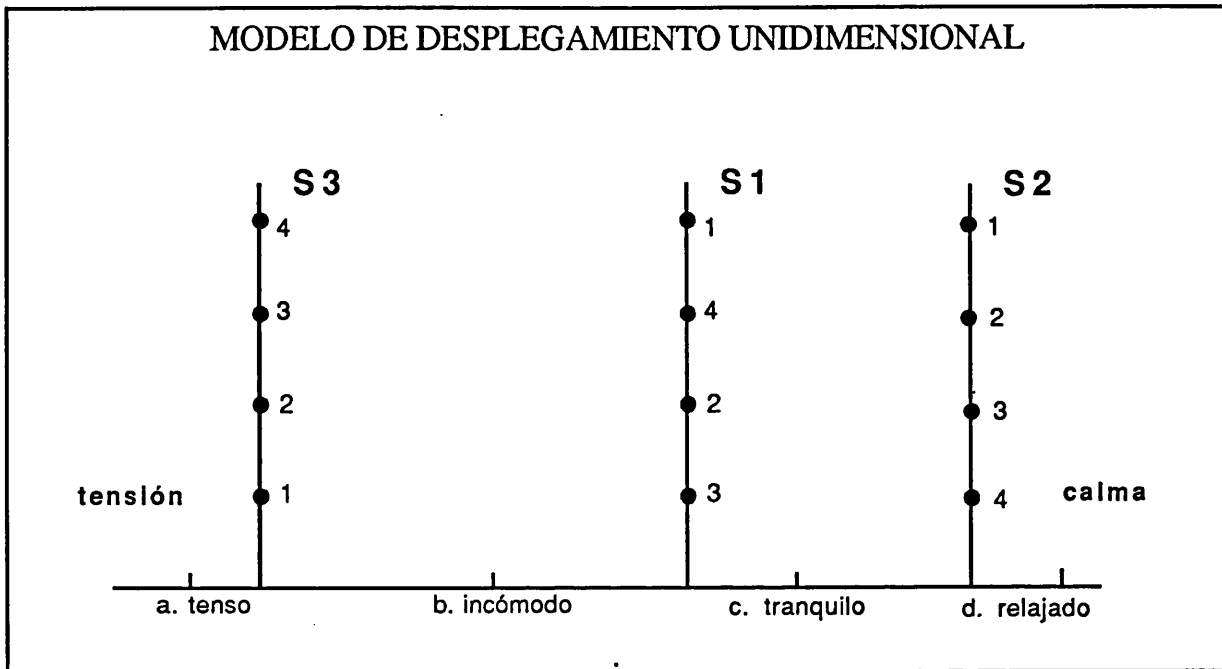
Para ejemplificar el modelo probabilístico de escalamiento unidimensional desplegado, imaginemos que se selecciona el siguiente

conjunto de adjetivos para la medida del continuo bipolar tensión-calma: a.tenso; b. incómodo, c. tranquilo, d. relajado. A continuación, este conjunto de adjetivos es presentado a una muestra de sujetos para que indiquen en qué orden prefieren o se sienten identificados con cada uno de los adjetivos. Esto nos daría el conjunto de escalas I que contiene la información necesaria para la identificación de la escala J.

A diferencia del modelo determinista, el modelo probabilístico desarrollado por Van Schuur y Post (1990) simplifica el problema de la métrica de la escala J, es decir, el problema de la cuantificación de las distancias que existen entre los estímulos representados en la dimensión. Bajo el modelo determinista, se asume que los sujetos muestran acuerdo respecto de las distancias que existen entre los estímulos, es decir, que todos los sujetos emplean la misma métrica en sus juicios. Este supuesto es el que permite identificar una escala J cuantitativa. No obstante, este supuesto es, en la práctica, altamente improbable (Coombs, 1964). De modo que bajo el modelo probabilístico que nos ocupa el objetivo es encontrar una escala J ordinal, que recoja el orden en que los estímulos se encuentran ubicados en la dimensión, independientemente de las distancias relativas que puedan existir entre ellos. Para ello, no es necesario contar con la ordenación completa que los sujetos realizan del conjunto de estímulos, sino que únicamente resulta relevante conocer cuál o cuáles son los estímulos preferidos o más próximos al punto ideal de cada sujeto. Con este fin, cada una de las escalas I es transformada en patrones de ceros y unos, codificando como 1 aquellas respuestas que indican máxima preferencia, y codificando como 0 el resto.

Continuando con el ejemplo anterior, imaginemos que tres sujetos, S1, S2 y S3, nos dan la siguiente ordenación de estímulos en función del grado con que se identifican con cada uno de ellos: el sujeto S1 nos da la ordenación c, b, d, a; el sujeto S2 nos da la ordenación c, b, d, a; y el sujeto S3 nos da esta otra a, b, c, d. Imaginemos que la ordenación correcta de los estímulos en la dimensión latente es la que aparece en la figura 8.4. La posición de cada uno de los sujetos podría representarse de la siguiente forma:

Figura 8.4. Representación de modelo de desplegamiento



Si redujéramos la información contenida en las escalas I de estos sujetos, dicotomizando las respuestas obtenidas en "estímulo preferido" y "estímulo no preferido", y codificáramos estas alternativas como 1 y 0 respectivamente, obtendríamos el siguiente patrón de respuestas para cada sujeto:

Tabla 8.1.

| SUJETOS | ESTIMULOS | | | |
|---------|-----------|----------|-----------|----------|
| | tenso | incómodo | tranquilo | relajado |
| S1 | 0 | 1 | 1 | 0 |
| S2 | 0 | 0 | 1 | 1 |
| S3 | 1 | 1 | 0 | 0 |

El procedimiento que proponen Van Schuur y Post (1990) para la identificación de la escala J correspondiente a un conjunto dado de escalas I transformadas de este modo, comienza por dividir el conjunto de estímulos en tantas ternas o combinaciones de tres elementos sin

repetición como sea posible. En este caso son posibles cuatro combinaciones:

| | | |
|--------------|-----------|-----------|
| 1. tenso | incómodo | tranquilo |
| 2. tenso | incómodo | relajado |
| 3. tenso | tranquilo | relajado |
| 4. tranquilo | incómodo | relajado |

Cada terna representa la unidad mínima de información necesaria para construir una escala J. A continuación, la lógica del procedimiento es la siguiente:

1) se evalúa la cantidad de error asociado a cada terna bajo el supuesto de que sea una escala J, es decir, se evalúa su grado de ajuste;

2) conocido el error asociado a cada posible escala J inicial, se identifica la que produce el mejor ajuste a los datos, y si este ajuste es suficiente, la escala se amplía añadiendo el estímulo que proporcione una nueva escala con el mejor ajuste posible; y

3) el paso segundo se repite hasta agotar los estímulos o hasta que la adición de un nuevo estímulo reduce el ajuste de la escala resultante.

Este procedimiento, brevemente esbozado aquí, se verá con mayor claridad ilustrado con el ejemplo que estamos siguiendo, y se operativizará con mayor detalle en apartados posteriores. No obstante, ya se puede apreciar la diferencia que existe entre este procedimiento y el que esbozamos para el modelo determinista de desplegamiento. En esta aproximación probabilística, el problema fundamental del modelo de desplegamiento, es decir, la identificación de la escala J, se resuelve precisamente a través de la consideración del error de medida que cada una de las escalas candidatas contiene.

Tomemos una de las ternas anteriormente identificadas:

| | | |
|-------|----------|-----------|
| tenso | incómodo | tranquilo |
|-------|----------|-----------|

Supongamos que esta terna forma la escala J inicial que estamos buscando. En ese caso, aunque estos sean los tres estímulos que forman la escala J inicial, no necesariamente han de presentar este orden en su ubicación en la dimensión latente. De hecho, es posible obtener 6 permutaciones diferentes con estos tres estímulos:

| | | | |
|---|-----------|-----------|-----------|
| 1 | tenso | incómodo | tranquilo |
| 2 | tranquilo | incómodo | tenso |
| 3 | tranquilo | tenso | incómodo |
| 4 | incómodo | tenso | tranquilo |
| 5 | incómodo | tranquilo | tenso |
| 6 | tenso | tranquilo | incómodo |

Bajo el supuesto de que estos tres estímulos en un orden o permutación determinada constituyen una escala J inicial, entonces hay un patrón de respuesta que no deberíamos obtener para esa permutación, el patrón de respuesta 101. Este patrón indica que el sujeto al que corresponde se siente identificado o prefiere los estímulos de los extremos de la permutación pero no el central. Puesto que en cada permutación los estímulos se encuentran ordenados según su disposición en el continuo latente, y las respuestas del sujeto indican su proximidad a los estímulos, no es posible que un sujeto se encuentre de forma simultánea cerca de los estímulos extremos y lejos del estímulo central, y que la ordenación de los estímulos que presenta esa permutación corresponda a la escala J que buscamos. No hay ningún punto representable en la dimensión subyacente que pueda representar tal ordenación. Si a una persona se le pide que indique en qué medida los siguientes tonos de cabellos representan adecuadamente el suyo: a. negro, b. castaño, c. rubio, no puede responder que los tonos que mejor le describen son negro y rubio, y que el castaño no le describe. Una persona puede tener el cabello decididamente negro, o rubio, o castaño, o bien entre negro y castaño -sin pasar por el rubio-, o entre rubio y castaño -sin pasar por el negro-, pero no entre negro y rubio -sin pasar por el castaño-. Una persona que constestara de este modo estaría incurriendo en algún tipo de error o confusión.

De modo que, una vez conocemos cómo identificar el término de error bajo este modelo, podemos cuantificar su incidencia en cada una de las permutaciones anteriormente identificadas. Para ello recordaremos que los patrones de respuesta de los tres sujetos del ejemplo eran los siguientes:

Tabla 8.1.

| SUJETOS | ESTIMULOS | | | |
|---------|-----------|----------|-----------|----------|
| | tenso | incómodo | tranquilo | relajado |
| S1 | 0 | 1 | 1 | 0 |
| S2 | 0 | 0 | 1 | 1 |
| S3 | 1 | 1 | 0 | 0 |

y las permutaciones que hemos identificado para la terna tenso, incómodo y relajado eran las siguientes:

| | Permutaciones | | | Pat. de resp. de los sujetos | | |
|---|---------------|-----------|-----------|------------------------------|-----|-----|
| | | | | S1 | S2 | S3 |
| 1 | tenso | incómodo | tranquilo | 011 | 001 | 110 |
| 2 | tranquilo | incómodo | tenso | 110 | 100 | 011 |
| 3 | tranquilo | tenso | incómodo | 101 | 100 | 011 |
| 4 | incómodo | tenso | tranquilo | 101 | 001 | 110 |
| 5 | incómodo | tranquilo | tenso | 110 | 010 | 101 |
| 6 | tenso | tranquilo | incómodo | 011 | 010 | 101 |

A la derecha se han representado los patrones de respuesta correspondientes a los tres sujetos para cada permutación, considerando únicamente los estímulos que forman la permutación de la fila, en ese orden preciso. En primer lugar, se observa que cada par de permutaciones en que se mantiene el estímulo central presentan el mismo patrón de respuesta en orden inverso. Así, el sujeto S2 en la primera permutación presenta el patrón de respuesta 001, en tanto que en la segunda permutación presenta el patrón 100. Esto es así porque ambas permutaciones comparten el mismo estímulo central, y se diferencian únicamente en que se intercambian los estímulos de los extremos. A efectos de cálculo de patrones de error, en que los estímulos de los extremos son codificados como 1 y el central como 0, es

irrelevante la diferencia entre ambas permutaciones. Eso es lo que ocurre, por ejemplo, con el patrón de respuesta del sujeto S3 en las permutaciones quinta y sexta. Por otra parte, se observa que aparecen patrones de error en cuatro de las permutaciones estudiadas: 3ª tranquilo-tenso-incómodo; 4ª incómodo-tenso-tranquilo; 5ª incómodo-tranquilo-tenso; y 6ª tenso-tranquilo-incómodo. Si reducimos el número de permutaciones a estudiar eliminando las permutaciones pares, que son la imagen en espejo de las permutaciones impares, observamos que de las tres permutaciones restantes (1, 3, y 5), sólo la primera permutación no presenta ningún patrón de error. Y ésta es precisamente la permutación que presenta la escala J cuando sólo se consideran los tres estímulos que estamos estudiando, tal como se muestra en la figura 8.4. Esto es precisamente lo que debe ocurrir para que una terna dada sea tomada como escala J inicial, siendo la permutación en que no hay patrones de error la que indica el orden adecuado de los estímulos en la escala. Cuando el análisis de los patrones observados en las tres permutaciones relevantes en que puede aparecer una terna dada (en nuestro caso las permutaciones 1, 3 y 5), muestra que dos de sus permutaciones producen patrones de error, y una tercera no, se considera que se ha encontrado una terna con ordenación "única", que presenta un único orden adecuado, que es precisamente lo que garantiza su calidad como escala J inicial.

En la práctica, es difícil encontrar ternas que produzcan permutaciones libres de error. De hecho, este modelo contempla la posibilidad de que incluso la permutación que corresponde a la escala J presente algún error de medida. Por ello en lugar de contabilizar directamente la frecuencia con que aparecen patrones de error en cada permutación, bajo este modelo se compara el error observado con el error que cabría esperar si las respuestas dadas a los diferentes estímulos fueran independientes, y si no existieran diferencias individuales entre los sujetos en la probabilidad de preferir uno u otro estímulo. En otros términos, se comparan los resultados observados en cada permutación con los que se esperaría obtener bajo el modelo nulo de independencia estadística. Esta comparación se realiza mediante un coeficiente denominado coeficiente de escalabilidad (Van Schuur y Post, 1990), que a su vez, es la base de la evaluación del ajuste con que una

determinada escala J reproduce las relaciones observadas entre sujetos y estímulos. Más adelante nos ocuparemos con detalle de la formulación y utilidad de este coeficiente.

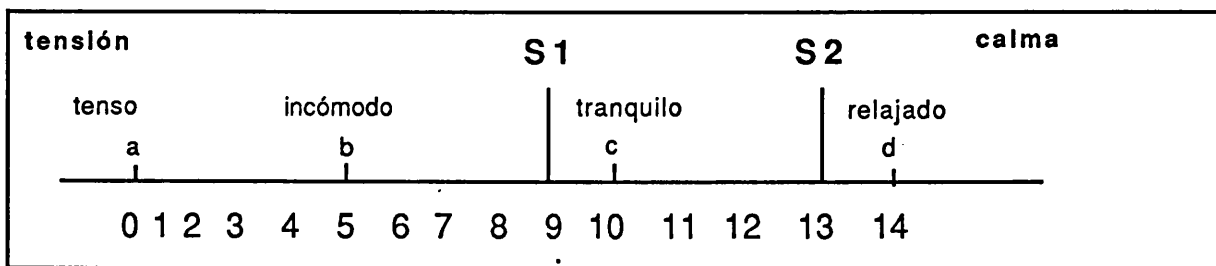
Una vez conocida la lógica con que opera el modelo de desplegamiento probabilístico propuesto por Van Schuur y Post (1990), podemos identificar la relación que existe entre este modelo de escalamiento y la problemática del desdoblamiento de factores pretendidamente bipolares en factores monopolares.

4. EL MODELO DE ESCALAMIENTO DESPLEGADO Y EL DESDOBLAMIENTO DE FACTORES PRETENDIDAMENTE BIPOLARES.

Retomando la ilustración presentada en la figura 8.4, se recordará que no se asignaron valores escalares ni a los sujetos ni a los items. La representación de ambos se realizó únicamente en base al orden absoluto de los estímulos, y al orden relativo que presentaban las respuestas de cada sujeto, del que dedujimos su ubicación.

Supongamos que conociéramos la métrica de la escala, y que ésta fuera tal como se representa en la figura 8.5.

Figura 8.5. Escala de desplegamiento con métrica conocida.



Entonces, podemos cuantificar las distancias a que se encuentra cada estímulo de la posición ideal de cada sujeto. Así, los sujetos S1 y S2 presentarían el siguiente conjunto de puntuaciones:

| | tenso | incómodo | tranquilo | relajado |
|----|-------|----------|-----------|----------|
| S1 | -9 | -4 | 1 | 5 |
| S2 | -13 | -8 | -3 | 1 |

que expresadas en distancias al cuadrado corresponderían a:

| | tenso | incómodo | tranquilo | relajado |
|----|-------|----------|-----------|----------|
| S1 | 81 | 16 | 1 | 25 |
| S2 | 169 | 64 | 9 | 1 |

Si a su vez suponemos que los sujetos S1 y S2 sólo disponen de una escala tipo Likert de 6 puntos para expresar su ubicación en la dimensión arriba descrita, de modo que una mayor puntuación expresara una mayor identificación con el ítem, entonces podríamos obtener las siguientes puntuaciones:

| | tenso | incómodo | tranquilo | relajado |
|----|-------|----------|-----------|----------|
| S1 | 2 | 4 | 6 | 3 |
| S2 | 1 | 3 | 5 | 6 |

Estas puntuaciones constituirían una de las posibles transformaciones monotónicas decrecientes de las distancias cuadráticas originales. El procedimiento de transformación sólo conserva el orden que los sujetos asignan a los estímulos en función de su punto ideal. Otra posible transformación posible hubiera sido:

| | tenso | incómodo | tranquilo | relajado |
|----|-------|----------|-----------|----------|
| S1 | 1 | 4 | 5 | 3 |
| S2 | 1 | 2 | 5 | 6 |

La aportación de Van Schuur y Kiers (en prensa) consiste en relacionar el modelo probabilístico de escalamiento desplegado con el estudio de la dimensionalidad de dimensiones bipolares, y se basa precisamente en esta reconceptualización de las puntuaciones obtenidas con escalas tipo Likert. Generalmente, el estudio de la dimensionalidad de datos obtenidos con escalas tipo Likert se ha realizado por medio de métodos de análisis factorial y de componentes principales. La

aplicación de estos métodos presupone que las relaciones entre las variables observadas y la dimensión latente o factor es lineal, o al menos monotónica. Sin embargo, cabe la posibilidad de que estas relaciones no sean monotónicas, sino cuadráticas -en términos de distancias-, en cuyo caso las puntuaciones observadas constituirían una transformación monotónica de esas distancias.

Bajo el modelo de medida que se asume en el análisis factorial, cuanto mayor es la puntuación de los sujetos en el factor, mayor es su puntuación en los ítems o indicadores empleados. Bajo el modelo de medida que subyace al método de escalamiento desplegado, cuanto mayor es la distancia entre la ubicación del sujeto y la del ítem en la dimensión subyacente, menor es la puntuación del sujeto en el ítem.

En consecuencia, el mismo conjunto de datos puede corresponder a uno u otro modelo de medida. Pero la aplicación de métodos de análisis factorial cuando los datos corresponden al modelo desplegado da como resultado la identificación de $r+1$ factores, en lugar de los r factores esperados (Coombs y Kao, 1960; Davison, 1977; Ross y Cliff, 1964). Además, el factor "extra" quedaría definido por las variables que reciben las puntuaciones más altas por parte de los sujetos (Ross y Cliff, 1964).

A pesar de que la relación entre el análisis factorial y el modelo de desplegamiento se conoce desde hace tres décadas, y de las implicaciones que puede tener el uso indiscriminado del análisis factorial, sólo recientemente se ha reconsiderado la posibilidad de que los resultados anómalos obtenidos por métodos de análisis factorial se deban al uso de este método sobre datos que no cumplen los supuestos necesarios para ello. En concreto, Van Schuur y Kiers (en prensa), retomando la discusión iniciada varias décadas atrás, consideran que la identificación de dos factores monopolares cuando lo que se espera es encontrar un factor bipolar puede ser un artefacto debido al uso del análisis factorial -y del modelo de medida subyacente- cuando los datos se corresponden al modelo de desplegamiento. En este sentido informan de resultados que apoyan esta hipótesis referidos al estudio de la androginia (Van Schuur, 1993b), donde dos factores independientes

podieron ser representados adecuadamente como las mitades de un continuo bipolar desplegado.

Estos autores presentan algunas pruebas que pueden ayudar a identificar si un determinado conjunto de datos corresponde al modelo de medida del análisis factorial o al modelo de desplegamiento. Entre estas pruebas destacan la forma de la matriz de correlaciones y el cálculo de las correlaciones parciales. Si se conoce el orden en que se ubican los estímulos en la dimensión latente, y se disponen los estímulos en ese orden, entonces la matriz de correlaciones presentará correlaciones altas entre estímulos contiguos, y decrecerán monotónicamente a partir de la diagonal hacia abajo y hacia la izquierda, tomando valores positivos y elevados en torno a la diagonal, próximos a cero a medida que nos alejamos de ella, y negativos y elevados en la parte inferior izquierda de la matriz (Davison, 1977), tal como se muestra en la tabla 8.2.

Tabla 8.2. Matriz de correlaciones entre estímulos esperada bajo el modelo de desplegamiento.

| | tenso | incómodo | tranquilo | relajado |
|-----------|-------|----------|-----------|----------|
| tenso | 1 | 0.9 | 0.0 | -0.5 |
| incómodo | 0.9 | 1 | 0.9 | -0.6 |
| tranquilo | 0.0 | 0.9 | 1 | 0.9 |
| relajado | -0.5 | -0.6 | 0.9 | 1 |

Si ordenamos la matriz de correlaciones de modo que los estímulos muestran el orden que ocupan en la dimensión latente, entonces los estímulos contiguos en la matriz son también contiguos en la dimensión latente. Las correlaciones que corresponden a estímulos contiguos en la dimensión latente se encontrarán entonces ubicadas en torno a la diagonal de la matriz. Además, a medida que nos alejamos de la diagonal de la matriz tanto en el sentido de las filas como en el de las columnas, las celdillas correspondientes contienen la relación entre estímulos más distantes entre sí en la dimensión latente. Esta representación de los datos facilita la interpretación de los resultados obtenidos, independientemente de los coeficientes concretos que contenga la matriz

que en este caso son coeficientes de correlación, ya que los estímulos contiguos se ubican adyacentes a la diagonal, y la distancia entre estímulos en la dimensión latente queda representada por la distancia en la matriz tomando como punto de referencia su diagonal.

No obstante, es posible encontrar matrices de correlaciones con todos sus elementos positivos que se ajusten al modelo de desplegamiento, y otras que cumplan estos criterios, y no se ajusten al modelo de desplegamiento (Davison, 1977). En cuanto a las correlaciones parciales, si los datos se ajustan al modelo de desplegamiento unidimensional, dados tres estímulos contiguos A, B y C, la correlación parcial entre A y B será positiva en tanto que C se encuentre a la derecha o izquierda de ambos (A y B), y negativa si C se encuentra entre ambos (Davison, 1977). No obstante, esta prueba sólo es útil si los datos se ajustan perfectamente al modelo de desplegamiento. Su uso en datos empíricos, como aquéllos que sólo son transformaciones monotónicas de las distancias cuadráticas, no siempre se acompaña de los resultados esperados (Davison, 1977).

En definitiva, actualmente la mejor forma de determinar si un conjunto de datos se corresponde con el modelo de desplegamiento consiste en someter los datos a un análisis de desplegamiento, y evaluar el ajuste de la escala obtenida.

Tomando todo ello en consideración, en esta parte procederemos a la contrastación empírica de la dimensionalidad de las escalas estudiadas bajo el modelo de desplegamiento unidimensional.

CAPITULO 9

METODO.

1. PROCEDIMIENTO DE RECOGIDA DE DATOS Y MUESTRA.

Esta parte del estudio ha sido realizada sobre la misma muestra que se empleó en la primera parte, en la que figura su composición y descripción.

2. MEDIDAS.

Las escalas empleadas en esta parte del estudio son las escalas desarrolladas por Warr (1990) para la medida de las dimensiones del bienestar afectivo: tenso vs. calmado, deprimido vs. entusiasmado y fatigado vs. animado, que ya quedaron descritas en la primera parte del estudio. Recordaremos que estas escalas constan de seis ítems cada una, y su escala de respuesta, simétrica, oscila entre "1. nunca" y "6. siempre". Cada una de las escalas consta de tres adjetivos típicos de uno de los polos del eje (positivo), y otros tres típicos del polo opuesto (negativo). También se incluyó un ítem en formato de diferencial semántico por cada una de las escalas a evaluar, contabilizando un total de tres ítems. Estos ítems presentaban en cada extremo el par de adjetivos opuestos que caracterizan cada una de las dimensiones medidas, entre los que se dispuso una línea continua dividida en 7 espacios. El sujeto respondía marcando una cruz sobre el espacio que mejor definiera su estado de ánimo en el mismo referente temporal que en las escalas tipo Likert.

3. HIPOTESIS.

Partiendo de la tesis de que las escalas estudiadas ofrecen datos interpretables bajo el modelo de medida de desplegamiento, es decir:

1) que cada escala mide un continuo bipolar desplegado, en el que los adjetivos presentan un orden determinado en función de su significado, común para todos los sujetos, y

2) que las puntuaciones de los sujetos son transformaciones monotónicas de las distancias cuadráticas que perciben entre la ubicación de cada ítem y su posición ideal,

entonces podemos formular una serie de hipótesis empíricamente contrastables, que confirmarán o no la tesis enunciada.

Estas hipótesis son las siguientes:

1) Para cada escala será posible encontrar una ordenación de los ítems correspondientes en un continuo bipolar, de modo que las puntuaciones de los sujetos indiquen su posición relativa respecto de los estímulos que la componen. La contrastación de esta hipótesis se efectuará mediante análisis de desplegamiento de cada escala, utilizando el programa MUDFOLD (Multiple UniDimensional unFOLDding, Van Schuur y Post, 1990).

2) Las puntuaciones obtenidas bajo este modelo de medida presentarán correlaciones iguales con los ítems de diferencial semántico correspondientes que las obtenidas bajo los modelos de medida tradicional o "dos mitades", que hemos visto son tan altas como las ofrecidas por otra escala explícitamente bipolar. Ello indicará que este modelo iguala la validez mostrada por los modelos anteriores.

4. ANALISIS.

El proceso de búsqueda y evaluación de las escalas bajo el modelo desplegado se ha llevado a cabo mediante el programa MUDFOLD. Este programa busca la representación del número máximo de estímulos de entre el total de ellos (n) y del número máximo de sujetos (N) en una dimensión, usando para ello sólo los juicios de preferencia -o identificación- de cada sujeto. El método que emplea este programa requiere el uso de puntuaciones dicotomizadas: las respuestas a los estímulos con los que se identifican los sujetos son codificadas como 1, y el resto se codifican como 0.

Se pueden distinguir dos aproximaciones básicas en la codificación de los datos. Si se opta por emplear los estímulos "preferidos" de los sujetos, entonces se indica al programa el valor o valores que definen un estímulo como preferido -en una escala tipo Likert de 6 puntos se emplearían las puntuaciones que expresan el mayor acuerdo o frecuencia, es decir, 5 y 6-, y el programa dicotomiza los datos de modo que la puntuación de un sujeto en un ítem será 1 si su respuesta ante el ítem fue 5 o 6, y 0 si dió cualquier otra respuesta. Así, cada sujeto estará representado por un patrón de ceros y unos que oscilará desde n ceros hasta n unos. Más adelante veremos que no todos los patrones de ceros y unos resultan incluidos en el proceso de búsqueda y evaluación de una escala desplegada. Este es el procedimiento a seguir bajo el modelo de desplegamiento probabilístico.

La segunda aproximación no se basa en las preferencias absolutas expresadas por los sujetos, en tanto que el criterio para considerar una respuesta como de preferencia viene dado por la amplitud de la escala de respuesta y el criterio del investigador. En lugar de ello, este procedimiento se basa en la ordenación que cada sujeto realiza de los estímulos. Se codifica como 1 las puntuaciones de los sujetos a los k estímulos a los que responde con la puntuación más alta, y como cero al resto. El investigador es quien determina en cada caso el valor de k , es decir, el número de estímulos preferidos por cada sujeto que se va a considerar.

Los autores diferencian ambos procedimientos denominando "toma cualquier ítem entre n " en el primer caso, y "toma k ítems entre n " en el segundo. En el primer caso se codifican como 1 las respuestas ante cualquier número de estímulos que el sujeto haya respondido con puntuaciones de máxima preferencia. En el segundo, sólo se codifican como 1 las respuestas a los k estímulos con la mayor puntuación para cada sujeto. En este último caso, los sujetos presentan patrones de k unos y $n-k$ ceros, en cualquier orden.

Una vez obtenidas las puntuaciones dicotómicas, el programa analiza todas las combinaciones de tres estímulos que se pueden formar, y para cada terna, considera tres permutaciones. Supuesta una terna

que se ajusta perfectamente a una escala desplegada, y una vez ordenados los estímulos en su posición en la escala, hay un patrón de ceros y unos que no puede ocurrir: el patrón en el que las respuestas a los estímulos de los extremos están codificadas como unos, y la respuesta del estímulo central como cero (101). Dado que los estímulos están dispuestos en el orden adecuado, un sujeto no puede preferir los estímulos de los extremos sin preferir también el estímulo central. Este patrón es el término de error en el modelo de desplegamiento, y únicamente asume información ordinal tal como se vió en la introducción. Para la terna ABC, si consideramos que el orden adecuado de los estímulos es ABC, el patrón de error correspondiente sería 101, que a su vez es el mismo que para CBA. Para la permutación BAC, si tomamos los estímulos en orden alfabético para designarlos, el patrón de error sería 011 ($A=0$, $B=1$, $C=1$), que también lo sería para la permutación CAB. Finalmente, el término de error para la permutación ACB sería, tomando los estímulos también en orden alfabético 110, que también sería el error de la permutación BCA. En consecuencia, para cada terna de estímulos, sólo se consideran tres permutaciones, en cada una de las cuales el lugar central está ocupado por uno de los tres estímulos. Por consiguiente, cada permutación de una terna dada tiene un error observado asociado, que se calcula como el número de patrones de error que aparecen en las respuestas de los sujetos, es decir, el número de veces que las ordenaciones proporcionadas por los sujetos violan el orden de la permutación. En este punto retomamos algo ya avanzado anteriormente: no todos los sujetos presentan patrones de respuesta falsables. Los sujetos que en la modalidad "toma cualquier ítem entre n " (a partir de ahora "toma cualquiera/ n ") eligen ninguno, uno o todos los estímulos no pueden ser contabilizados en el proceso de elaboración de la escala, y por lo tanto son eliminados del análisis. Los sujetos que en la modalidad "toma k ítems entre n " (a partir de ahora "toma k/n ") responden a un número de estímulos inferior a k , o bien presentan "empates" en el grado de preferencia de modo que no es posible identificar que k estímulos son los preferidos, también son eliminados del análisis. De modo que se calcula el número de patrones de error observados para cada permutación de cada terna en el subconjunto de sujetos que cumplen las condiciones para ello. A continuación, se calcula el número de patrones de error que cabría

esperar bajo el modelo nulo de independencia estadística. Este modelo asume que los sujetos no presentan diferencias en la probabilidad con que prefieren cada uno de los estímulos, y que las respuestas a los estímulos son dadas de forma independiente (Van Schuur y Post, 1990). El error esperado bajo este modelo se calcula como el producto de las frecuencias relativas con que cada estímulo de la terna es preferido, multiplicado por el número de sujetos, en el caso de la modalidad "toma cualquiera/n", y de forma algo más compleja en el caso "toma k/n" (para una explicación detallada de este procedimiento ver Van Schuur y Post, 1990).

Una vez calculado el error observado, y estimado el error esperado bajo el modelo nulo, se puede calcular el coeficiente de escalabilidad análogo al coeficiente H de Loevinger, como el cociente

$$H_{(i,j,k)} = 1 - \frac{O_{(i,j,k)}}{EO_{(i,j,k)}}$$

donde $H_{(i,j,k)}$ es el coeficiente de escalabilidad para la terna i, j, k en su permutación ijk, $O_{(i,j,k)}$ es la frecuencia observada del patrón de error, y $EO_{(i,j,k)}$ es la frecuencia esperada bajo el modelo nulo del patrón de error. Para cada terna se obtienen tres coeficientes, uno en cada una de las permutaciones estudiadas: $H_{(i,k,j)}$, $H_{(j,i,k)}$ y $H_{(i,j,k)}$.

La escalabilidad de un determinado estímulo en la escala también puede evaluarse. Para ello se contabiliza el número de errores observados y esperados en todas las permutaciones de ternas que contienen ese estímulo, procediéndose al cálculo del coeficiente de escalabilidad $H_{(i)}$ en forma análoga a la expuesta. En caso de escalabilidad perfecta este coeficiente tomará el valor $H=1$, lo que indica que no hay error observado, mientras que cuando $H=0$, el error observado es igual al error esperado bajo el modelo nulo. Finalmente, cuando H toma valores negativos, el error observado es mayor que el esperado. De modo análogo se puede obtener e interpretar el coeficiente de escalabilidad para la escala completa en su ordenación final.

Para permitir la inferencia estadística acerca de los coeficientes de escalabilidad mencionados, se han obtenido las distribuciones de estos coeficientes bajo el modelo nulo de independencia estadística. El programa ofrece directamente el valor del estadístico T, que presenta una distribución normal estándar.

Una vez conocidos los coeficientes de escalabilidad de las ternas en cada una de sus permutaciones, así como la significación del mismo, el programa selecciona la "mejor terna" como inicio de la escala desplegada. Esta terna se caracteriza por presentar coeficientes de escalabilidad negativos en dos de sus permutaciones, y positivo y mayor que 0.30 en la restante. Con ello se garantiza que se trata de una ordenación única, que sólo funciona en una de sus permutaciones. Si una terna presenta un error menor del esperado bajo el modelo nulo en una de sus permutaciones, y presenta un error mayor que el esperado en las dos restantes, entonces se puede decir que existe acuerdo en la ordenación de esa terna tal como indica la permutación con el coeficiente de escalabilidad positivo. Esto era lo que sucedía en el ejemplo que analizamos en la introducción relativo a la terna incómodo-tranquilo-tenso. Como se recordará, de las tres permutaciones no redundantes (1, 3, y 5), sólo la primera se caracterizaba por no mostrar ningún patrón de error, en tanto que el resto de permutaciones posibles daban como resultado la aparición de patrones de error. Esto era así porque esa permutación reproducía la ordenación adecuada en la dimensión o escala J. Si en lugar de contar con los patrones de respuesta de tres sujetos, hubiéramos manejado un número de sujetos mayor, probablemente esta permutación también hubiera producido patrones de error, atribuibles al error de medida. No obstante, el número de patrones de error asociado a esta permutación hubiera sido menor que el esperado bajo el modelo nulo de independencia estadística, ya que el número de patrones de error asociados a las permutaciones restantes hubiera sido mayor que el esperado bajo el modelo nulo de independencia estadística. Finalmente, ello hubiera conducido a la obtención de coeficientes de escalabilidad negativos en las permutaciones incorrectas, y positivo en la permutación correcta. En caso de que exista más de una "mejor terna", se consideran otros criterios como el nivel de significación de T.

Una vez identificada la terna inicial en su permutación adecuada, el programa añade los estímulos restantes en cuarto lugar, y recalcula todo el proceso con la nuevas ternas que se pueden formar con ese conjunto de cuatro estímulos. Finalmente, añade como cuarto estímulo aquél que presenta un coeficiente de escalabilidad mayor que el fijado por el investigador (por omisión es .30), y que no disminuye el coeficiente de escalabilidad de la escala por debajo de un valor dado (por omisión .30). En caso de que existan varios candidatos, se retiene aquel estímulo con posición única, es decir, que sólo puede ser añadido en una posición al subconjunto de tres estímulos ya existentes, o en función de la significación de los coeficientes de escalabilidad. El programa continua este procedimiento iterativo hasta que ya no queda ningún estímulo por añadir, o hasta que ninguno de los que quedan cumple los requisitos expuestos.

El coeficiente de escalabilidad descrito no indica de forma unívoca que los datos analizados se ajustan al modelo de despliegamiento. En el caso de "toma cualquiera/n" pueden obtenerse coeficientes altos al tiempo que los datos se ajusten al modelo determinista de despliegamiento, o al modelo determinista de escalamiento acumulativo de Guttman.

Los modelos de despliegamiento pueden distinguirse de los modelos acumulativos en base a las curvas características de los ítems (CCI) que asumen ambos modelos. Los modelos de despliegamiento se basan en relaciones de proximidad, y por lo tanto asumen CCI unimodales, es decir, asumen que la probabilidad de que un sujeto con una posición dada en el continuo elija un ítem es una función unimodal de la dimensión latente. Por el contrario, los modelos acumulativos se basan en relaciones de dominancia, que muestran CCI no decrecientes. Sin embargo, los modelos acumulativos pueden ser considerados como un caso limítrofe de despliegamiento (Van Schuur y Post, 1990). Más adelante volveremos sobre este punto.

En consecuencia, el programa ofrece algunas matrices para las que se conocen las particularidades que presentan en el caso de que los datos se ajusten al modelo determinista de desplegamiento.

La matriz de dominancia es una matriz cuadrada y asimétrica que contiene en sus celdillas el porcentaje de sujetos que elijen el estímulo de la fila pero no de la columna. Bajo el modelo determinista de desplegamiento se espera que estos porcentajes disminuyan desde la izquierda hasta la diagonal, y que aumenten desde la diagonal hasta la izquierda, tal como se indica en la tabla 9.1.

Tabla 9.1. Matriz de dominancia esperada bajo el modelo determinista de desplegamiento.

| | tenso | incómodo | tranquilo | relajado |
|-----------|-------|----------|-----------|----------|
| tenso | - | < | < | < |
| incómodo | > | - | < | < |
| tranquilo | > | > | - | < |
| relajado | > | > | > | - |

Como ya se indicó en la introducción, la disposición de los estímulos en las matrices en la forma en que se ordenan en la dimensión subyacente permite aplicar una serie de reglas que facilitan la interpretación de los resultados. En concreto, esta disposición de los datos ofrece una configuración de la matriz tal que los valores que ocupan las celdillas adyacentes a la diagonal corresponden a estímulos contiguos en la dimensión latente, y a medida que las celdillas se alejan de la diagonal, también se incrementa el número de estímulos ubicados entre el par de estímulos al que corresponde esa celdilla. En consecuencia, si la matriz de dominancia contiene en sus celdillas el porcentaje de sujetos que elijen el estímulo de la fila pero no el de la columna, cabe esperar que a mayor proximidad entre dos estímulos, menor sea este porcentaje. Puesto que los pares de estímulos contiguos en la dimensión se encuentran representados adyacentes a la diagonal, esperaremos que este porcentaje sea menor en las celdillas adyacentes a la diagonal, y se incremente a medida que nos alejamos de ella.

La matriz de adyacencia es una matriz triangular que contiene el porcentaje de casos en que se elige el estímulo de la fila y el de la columna. Bajo el modelo determinista de desplegamiento se espera que estos porcentajes aumenten de izquierda hacia la diagonal, y que disminuyan de la diagonal hacia abajo (ver figura 9.2). Como se recordará, en esta matriz también se disponen los estímulos en el orden en que aparecen en la dimensión. Esto hace que la proximidad entre dos estímulos aumente a medida que su ubicación en la matriz se acerque a la diagonal, tanto en el sentido de las filas como en el de las columnas. De modo que el porcentaje de casos en que se elige el estímulo de la fila y el de la columna, bajo el modelo de desplegamiento, aumentará a medida que nos acercamos a la diagonal, es decir, a medida que aumente la proximidad entre los estímulos considerados. Bajo el modelo determinista acumulativo se espera que estos porcentajes aumenten de izquierda hacia la diagonal, y que también aumenten de la diagonal hacia abajo.

Tabla 9.2. Matriz de adyacencia esperada bajo el modelo determinista de desplegamiento.

| | tenso | incómodo | tranquilo | relajado |
|-----------|-------|----------|-----------|----------|
| tenso | - | | | |
| incómodo | <<< | - | | |
| tranquilo | < | <<< | - | |
| relajado | < | < | <<< | - |

Nota: a mayor número de símbolos "<", mayor porcentaje de sujetos en esa celdilla.

No obstante, bajo el modelo probabilístico de desplegamiento no es necesario que aparezca este patrón de monotonía. Bajo el modelo probabilístico la monotonía no depende del ajuste de la escala al modelo desplegado, sino de la distribución de los sujetos. Por ello, es posible que no se cumplan estas condiciones, siendo que los datos no se ajustan al modelo determinista, pero sí al probabilístico. Para determinar si esto es así, el programa también ofrece un conjunto de matrices complementarias de las que se conoce su comportamiento cuando los datos se ajustan al modelo de desplegamiento probabilístico.

La matriz de adyacencia condicional es una matriz cuadrada y asimétrica que contiene la frecuencia relativa con que se elige el estímulo de la fila cuando se elige el estímulo de la columna ($p_{(i/j)}$). Bajo el modelo probabilístico de desplegamiento se espera que:

1) cada fila presente una distribución unimodal, y

2) el máximo de una fila esté a la izquierda o sobre el máximo de la fila siguiente.

La ilustración del cumplimiento de estas condiciones resulta algo compleja, sobre todo para un número pequeño de estímulos. No obstante, se puede emplear el argumento contrario, es decir, la ilustración de su falta de cumplimiento. Se puede comprender de forma sencilla lo que implican estas condiciones si se considera que su incumplimiento produciría una distribución tal que, como muestra la tabla 9.3 (la frecuencia relativa mayor se ha señalado con un asterisco), los sujetos que se identificaran con el estímulo relajado se identificarán también con el estímulo tenso, en mayor medida que con cualquier otro. Ello conduciría a la obtención de un gran número de patrones de error para las permutaciones tenso-incómodo-relajado (101) y tenso-tranquilo-relajado (101), que por otra parte, son ambas buenas representaciones de la escala J identificada.

Tabla 9.3. Matriz de adyacencia condicional esperada bajo el modelo de desplegamiento probabilístico.

| | tenso | incómodo | tranquilo | relajado |
|-----------|-------|----------|-----------|----------|
| tenso | - | | | * |
| incómodo | | - | * | |
| tranquilo | * | | - | |
| relajado | * | | | - |

Por otra parte, cada fila de esta matriz puede interpretarse como una representación aproximada de las CCI de los items. Si las filas muestran distribuciones no decrecientes, es posible que los estímulos a que pertenecen presenten CCI no decrecientes, en favor del modelo

acumulativo. En cualquier caso, bajo el modelo probabilístico de despliegamiento son posibles filas no decrecientes, no crecientes y unimodales. De modo que aún siendo éstas no decrecientes, el modelo de despliegamiento puede ser todavía el modelo adecuado.

La matriz de correlaciones entre los estímulos, bajo el modelo de despliegamiento probabilístico, es una matriz triangular en la que: 1) cada fila puede presentar 0, 1 o 2 cambios de signo, y 2) si presenta dos cambios, los signos centrales son positivos, y los extremos negativos. En este sentido, las celdillas centrales corresponden a items contiguos, o próximos y por lo tanto positivamente relacionados, y las celdillas de los extremos corresponden a items distantes, entre los que cabe esperar una relación inversa.

La matriz de coeficientes H de Loevinger es también una matriz triangular, que expresa la proporción entre el coeficiente de correlación observado y su valor máximo dadas las frecuencias del estímulo fila y columna. Los signos de esta matriz deben ser iguales a los de la matriz de correlaciones. Cuando los signos de la matriz de correlaciones y los de la matriz de Loevinger son todos positivos, puede que los datos se ajusten al modelo de escalamiento acumulativo de Guttman. Pero también es posible que los datos se ajusten a este modelo acumulativo y estas matrices presenten elementos negativos. Esto podría ocurrir en el caso de que la escala desplegada pudiera considerarse formada por dos escalas acumulativas. Más adelante volveremos sobre este punto, pero en definitiva, se puede concluir que al ser el modelo acumulativo un caso limítrofe del modelo de despliegamiento, resulta difícil discernir cuál de los dos modelos es el adecuado.

Una vez se conoce la configuración de la escala desplegada, se procede a asignar valores a los estímulos que la forman y a los sujetos. MUDFOLD no ofrece información métrica, por lo que la asignación de valores a los estímulos se realiza bajo el supuesto de que los estímulos adyacentes son equidistantes. Ello permite asignar a cada estímulo un valor igual al del orden que ocupa. Cuando los patrones de respuesta se han obtenido bajo la instrucción "toma cualquiera/n", el procedimiento por el que se asignan puntuaciones a los sujetos en la escala obtenida

consiste en sumar y promediar el valor de los estímulos que ha elegido como preferidos. Los sujetos que no han elegido ningún estímulo no pueden ser representados en la escala. Los sujetos que han elegido sólo un estímulo, o todos ellos, aunque son eliminados del análisis, sí pueden ser representados en la escala por el procedimiento de asignación de promedios señalado. En el caso de "toma k/n" las puntuaciones se obtienen de forma análoga, pero no son representables en la escala aquellos sujetos con un número de estímulos preferidos menor o mayor que k.

CAPITULO 10

RESULTADOS Y DISCUSION.

1. ESCALAMIENTO BAJO EL MODELO DE DESPLEGAMIENTO PROBABILISTICO.

Este primer conjunto de análisis se ha realizado bajo el supuesto de que las escalas bajo estudio se ajustan al modelo probabilístico de desplegamiento. En consecuencia, se ha empleado la modalidad de procedimiento "toma cualquiera/n", definiendo como valores en la escala original indicativos de un estímulo preferido los valores máximos de la escala de respuesta: 5 y 6. De este modo, el programa ha dicotomizado los datos transformando en 1 las respuestas a los items contestados con un 5 o un 6, y en 0 a las restantes. En esta modalidad cada estímulo es juzgado de forma independiente del resto de estímulos.

En esta modalidad de escalamiento desplegado los sujetos con patrones de respuesta no falsables son aquellos que han preferido todos los estímulos (111111), los que no han preferido ninguno (000000) y los que han preferido sólo 1, en cualquier posición (por ejemplo 100000). El programa permite al investigador determinar si desea realizar los análisis sobre los datos de todos los sujetos, o sólo sobre los datos de aquellos sujetos con patrones falsables. En este sentido, Van Schuur y Post (1991) advierten que cuando existe un número elevado de sujetos con patrones no falsables, los coeficientes de escalabilidad de las ternas en cada permutación ($H_{(i,k,j)}$) sobreestimarán la despleabilidad del conjunto de estímulos. Esto es así porque la frecuencia observada del patrón de error ($O_{(i,j,k)}$) se calcula a partir de los sujetos con patrones falsables, mientras que la frecuencia esperada del patrón de error bajo el modelo nulo ($EO_{(i,j,k)}$) se calcula sobre todos los sujetos, incluidos los que presentan patrones no falsables. Ello produce un incremento del error esperado respecto al error observado que no depende de la "calidad" de la permutación, sino de la discrepancia entre el número de sujetos considerados para uno y otro cálculo, que puede llevar a la sobreestimación de los coeficientes de escalabilidad, tal como se puede derivar de la expresión que se emplea para su cálculo:

$$H_{(i,j,k)} = 1 - \frac{O_{(i,j,k)}}{EO_{(i,j,k)}}$$

Por ello, cuando el número de sujetos con patrones no falsables es elevado, los autores recomiendan realizar un primer análisis considerando todos los sujetos, y contrastar los resultados obtenidos con un segundo análisis realizado sólo sobre los sujetos con patrones falsables.

A continuación es necesario indicar el criterio a seguir para encontrar y extender la escala desplegada. Esto se puede hacer de dos modos. En primer lugar, se puede identificar la escala J inicial en base al valor del coeficiente de escalabilidad de cada terna en cada permutación. En este caso, la búsqueda de la terna inicial finalizará cuando se identifique una terna para la que una de sus tres permutaciones muestre un coeficiente de escalabilidad positivo y mayor que un valor dado (.30), mientras que las dos restantes muestren coeficientes de escalabilidad negativos. La escala J inicial quedará definida por los estímulos que formen esta terna, en la ordenación que presentan en la permutación que ofrece un coeficiente de escalabilidad positivo. Una vez identificada la escala J inicial, comienza el proceso de extensión de la escala inicial de modo que la escala final incluya el mayor número de estímulos posibles. En este sentido, la adición de un nuevo estímulo a la escala se realiza también en función de los nuevos coeficientes de escalabilidad que produce cada uno de los estímulos candidatos, reteniéndose aquél que produce el coeficiente mayor.

El segundo procedimiento por el que puede guiarse la búsqueda y extensión de la escala J se realiza en base al valor del estadístico T, es decir, considerando la significación estadística y no simplemente la magnitud del coeficiente de escalabilidad. En este sentido los autores recomiendan emplear inicialmente el criterio del valor del coeficiente de escalabilidad, por ser el que mayores garantías reúne por el momento. Ello se debe a la reciente incorporación al programa del criterio de significación estadística. En consecuencia, hemos realizado los análisis siguiendo ese criterio.

El procedimiento de extensión de la escala se ha dado por terminado cuando no se ha encontrado ningún estímulo con un

coeficiente de escalabilidad igual o mayor que .30. En el caso de que más de un estímulo presente coeficientes mayores que .30, se puede dar prioridad a aquél que presente el menor número de posiciones posibles en la escala, o a aquél que incremente en mayor medida el coeficiente de escalabilidad de la escala. Siguiendo a Van Schuur y Post (1990), hemos empleado el primer criterio, que es el que garantiza en mayor medida la interpretabilidad del resultado.

Una vez establecidos los criterios seguidos en la realización de los análisis, a continuación se presentan los resultados obtenidos para cada una de las tres escalas.

2. ESCALA TENSO-CALMADO.

La descripción de la distribución de puntuaciones dicotomizadas se encuentra en los recuadros 1 y 2. El primero de ellos muestra en cada celdilla el número de sujetos que eligen el estímulo de la columna entre el número de estímulos que eligen en total, que viene dado por el número de la fila correspondiente. Recordaremos que se entiende que un sujeto elige un estímulo cuando emplea las alternativas de máximo acuerdo para responder, que en las escalas que nos ocupan corresponden a los valores 5 y 6. Se observa que más de un tercio de los sujetos no eligen ningún estímulo, y que de entre los que eligen alguno, la mayor parte elige 1 o 2 estímulos.

Tabla 10.1. Distribución número de estímulos elegidos

| N | A | B | C | D | E | F | TOT |
|---|---|---|---|----|----|----|-----|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 118 |
| 1 | 3 | 2 | 6 | 29 | 23 | 5 | 68 |
| 2 | 2 | 2 | 4 | 52 | 40 | 26 | 63 |
| 3 | 0 | 0 | 2 | 42 | 41 | 41 | 42 |
| 4 | 0 | 1 | 1 | 1 | 0 | 1 | 1 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Nota: N=número de estímulos elegidos; A=tenso, B=incómodo, C=preocupado, D=tranquilo, E=contento, F=relajado; TOT=número de sujetos en esa fila

En la tabla 10.2 se observa en primer lugar la frecuencia absoluta con que se ha elegido cada estímulo, y entre paréntesis la frecuencia relativa correspondiente. Destaca el hecho de que muy pocos sujetos han elegido los estímulos negativos de la escala.

Tabla 10.2. Frecuencias absolutas y relativas

| | | | |
|--------------------------|------------|-----|------------|
| NUMBER OF RESPONDENTS | | 292 | |
| NUMBER OF STIMULI | | 6 | |
| NUMBER OF PICKED STIMULI | | ANY | |
| A) TENSO | - MARGINAL | 5 | (1.71 %) |
| B) INCOMODO | - MARGINAL | 5 | (1.71 %) |
| C) PREOCUPADO | - MARGINAL | 13 | (4.45 %) |
| D) TRANQUILO | - MARGINAL | 124 | (42.47 %) |
| E) CONTENTO | - MARGINAL | 104 | (35.62 %) |
| F) RELAJADO | - MARGINAL | 73 | (25.00 %) |

La tabla 10.3 muestra la terna que cumple los requisitos para ser considerada "mejor terna", y consecuentemente ser el punto de partida de la escala desplegada. El primer par de permutaciones que aparece es EBF=FBE, el segundo par de permutaciones es BEF=FEB, y el tercer par es BFE=EFB. Los coeficientes de escalabilidad negativos muestran que las permutaciones que consideran como elemento central los estímulos B o E producen un número igual o mayor de errores que los esperados bajo el modelo nulo de independencia estadística, en tanto que cuando el estímulo F ocupa el lugar central, el número de errores observados es menor que el esperado, y por lo tanto el coeficiente de escalabilidad asociado resulta positivo. De este modo, la permutación BFE es tomada como punto inicial para la construcción de la escala desplegada. El orden adecuado BFE o EFB es determinado por el programa durante el proceso de extensión de la escala. En este momento del proceso de construcción de la escala ambas permutaciones son imágenes espejo una de la otra, por lo que no se puede determinar cuál es el mejor orden de los estímulos extremos. La adición de un cuarto estímulo permitirá al programa identificar la mejor ordenación.

Tabla 10.3. Terna y permutación inicial

| | EBF=FBE | | | BEF=FEB | | | BFE=EFB | | |
|-----|---------|-------|--------|---------|-------|--------|---------|------|--------|
| | H | T | S(jik) | H | T | S(ijk) | H | T | S(ikj) |
| BEF | -0.84 | -6.05 | 0.14 | -0.24 | -0.24 | 1.02 | 1.00 | 1.36 | 0.74 |

Nota: H=coeficiente de escalabilidad; T=valor del estadístico T; S=error típico de H.

La tabla 10.4 muestra el proceso seguido en la extensión de la permutación inicial. En primer lugar aparece la proporción de sujetos que han elegido el estímulo de la fila (P(I)). A continuación aparecen los coeficientes de escalabilidad de cada estímulo calculados en cada paso de la construcción de la escala (H(I)). La columna 3 muestra los coeficientes de escalabilidad de los estímulos que forman la permutación inicial, y constituye el paso inicial en la identificación de la escala J. En la columna 4 aparecen los coeficientes para estos tres estímulos así como el correspondiente al nuevo estímulo añadido a la escala, el estímulo tranquilo. Los coeficientes que aparecen en esta columna son los calculados para cada estímulo en la nueva escala, es decir, en la escala formada por los cuatro estímulos siguientes: incómodo, relajado, tranquilo y contento. A su vez, el orden que ocupan en las filas de la tabla es el que les corresponde en la dimensión latente. En la columna 5 aparecen los coeficientes de escalabilidad para la nueva escala formada ahora por cinco estímulos. Por último, en la última fila aparece el coeficiente de escalabilidad de cada una de las escalas obtenidas durante el proceso de extensión, es decir, para tres, cuatro y cinco estímulos. De acuerdo con Van Schuur y Post (1990), se admite que un estímulo es escalable cuando el coeficiente que presenta es igual o mayor que .30. Por consiguiente, tanto los estímulos incluidos en la escala como la propia escala muestran un ajuste aceptable al modelo de desplegamiento analizado. Únicamente un estímulo, preocupado, no puede ser incluido en la escala desplegada obtenida. No se ofrece su coeficiente de escalabilidad porque no ha logrado ser parte de la escala. No obstante, este estímulo resulta ser el estímulo negativo elegido con mayor frecuencia por los sujetos, de modo que su exclusión de la escala final cuestiona la continuidad entre los items negativos y positivos en la escala desplegada.

Tabla 10.4. Proceso de búsqueda de la escala desplegada final

| ITERATIVE SEARCH PROCEDURE 1 | | | | |
|--|------|-----------|-----------|-----------|
| P(I): PERCENTAGE OF RESPONDENTS WHO MENTIONED STIMULUS I | | | | |
| H(I): SCALABILITY VALUE FOR STIMULUS I | | | | |
| ESTIMULOS | P(I) | 3 H(I) | 4 H(I) | 5 H(I) |
| A) TENSO | 0.02 | --- | --- | 1.00 |
| B) INCOMODO | 0.02 | 1.00 | 1.00 | 1.00 |
| F) RELAJADO | 0.25 | 1.00 | 0.61 | 0.68 |
| D) TRANQUIL | 0.42 | --- | 0.60 | 0.69 |
| E) CONTENT | 0.36 | 1.00 | 0.60 | 0.67 |
| ----- | | | | |
| H-SCALE | | 1.00 | 0.63 | 0.75 |

Nota: P(I)=coeficiente de escalabilidad del estímulo de la fila; H(I)=valor del coeficiente de escalabilidad asociado al estímulo de la fila en cada paso seguido en el procedimiento de extensión de la escala.

La tabla 10.5 muestra la descripción de la escala desplegada final. Junto a la proporción de sujetos que eligen cada estímulo (P(I)) y el coeficiente de escalabilidad correspondiente (H(I)) aparecen el número de errores observados $O_{(i,j,k)}$, el número de errores esperados bajo el modelo nulo de independencia estadística $E O_{(i,j,k)}$, el error típico asociado a cada coeficiente de escalabilidad (SD(H)) y el valor T asociado a cada estímulo (T(H)). El número de errores observados respecto de los estímulos representativos del polo negativo que forman parte de la escala es cero. No obstante, también son muy pocos los patrones de respuesta analizados en los que aparecen. Sin embargo, hay que recordar que el número de errores esperado ha sido calculado sobre todos los sujetos, independientemente de si sus patrones de respuesta son empleados en el análisis o no, de modo que el coeficiente de escalabilidad correspondiente a estos estímulos puede resultar sobreestimado. Por lo demás, todos los coeficientes resultan significativamente distintos de cero ($p < 0.01$, dos colas) y elevados, así como el coeficiente de escalabilidad de la escala total (H=.75, T=4.39).

Tabla 10.5. Escala desplegada final

| FINAL SCALE 1 | | | | | | |
|---|------|------|----------------------|-----------------------|-------|------|
| STIMULI ABOVE DASHED PAGE FORM AN UNFOLDING SCALE | | | | | | |
| | P(I) | H(I) | O _(i,j,k) | EO _(i,j,k) | SD(H) | T(H) |
| A) TENSO | 0.02 | 1.00 | 0 | 9.0 | 0.33 | 2.97 |
| B) INCOMODO | 0.02 | 1.00 | 0 | 9.0 | 0.28 | 3.48 |
| F) RELAJADO | 0.25 | 0.68 | 7 | 22.0 | 0.17 | 4.02 |
| D) TRANQUILO | 0.42 | 0.69 | 7 | 22.3 | 0.17 | 3.92 |
| E) CONTENTO | 0.36 | 0.67 | 7 | 21.4 | 0.20 | 3.28 |
| ----- | | | | | | |
| C) PREOCUPADO | 0.04 | | | | | |

Nota: P(I)=proporción de sujetos que han elegido el estímulo de la fila; H(I)=coeficiente de escalabilidad del estímulo de la fila; O_(i,j,k)=error observado asociado al estímulo fila; EO_(i,j,k)=error esperado bajo el modelo nulo asociado al estímulo de la fila; S(H)=error típico del coeficiente de escalabilidad del estímulo de la fila; T(H)=valor del estadístico T asociado al coeficiente de escalabilidad estímulo de la fila.

Tabla 10.6. Matriz de adyacencia condicional

| CONDITIONAL ADJACENCY MATRIX 1 | | | | | |
|--|--------|---------|---------|--------|--------|
| IN CELL (I,J): proportion of respondents who mentioned row stimulus i among all respondents who mentioned column stimulus j. | | | | | |
| In brackets: the standard error belonging to the proportion | | | | | |
| | A | B | F | D | E |
| ----- | | | | | |
| A) TENSO | * | 0.20 | 0.00 | 0.00 | 0.00 |
| | | (.179) | (.000) | (.000) | (.000) |
| B) INCOMODO | 0.20 | * | 0.01 | 0.01 | 0.00 |
| | (.179) | | (.014) | (.008) | (.000) |
| F) RELAJ | 0.00 | 0.20 | * | 0.49 | 0.45 |
| | (.000) | (0.179) | | (.045) | (.049) |
| D) TRANQUIL | 0.00 | 0.20 | 0.84 | * | 0.70 |
| | (.000) | (.179) | (0.043) | | (.045) |
| E) CONTENT | 0.00 | 0.00 | 0.64 | 0.59 | - |
| | (.000) | (.000) | (.056) | (.044) | |
| ----- | | | | | |

Como se ha indicado anteriormente, la obtención de coeficientes de escalabilidad elevados no es indicativo por sí solo de que la escala se ajusta adecuadamente al modelo de desplegamiento. Pueden esperarse estos coeficientes tanto si los datos se ajustan al modelo de desplegamiento probabilístico, como si se ajustan al modelo acumulativo

de Guttman. Para determinar en la medida de lo posible cuál es el modelo de escalamiento adecuado, se indicaron ciertas pautas relacionadas con la matriz de adyacencia condicional que pasamos a considerar a continuación.

La tabla 10.6 muestra esta matriz. Bajo el modelo de despliegamiento probabilístico se espera que cada fila -que puede interpretarse como una aproximación de la curva característica del estímulo correspondiente- presente un máximo ubicado de tal modo que se encuentre a la izquierda, a la derecha o sobre el máximo de la fila siguiente. Esto es así porque se asume que los estímulos forman una escala desplegada, y están ubicados en la matriz en el orden con que aparecen en la dimensión latente. En consecuencia, se espera que los sujetos que eligen un estímulo columna dado, prefieran también sobre todos los demás un estímulo fila próximo al de la columna, en lugar de preferir otro que resulte más distante. Por ejemplo, se espera que los sujetos que eligen el estímulo columna "relajado" prefieran también el estímulo fila "tranquilo" (0.84), en lugar de preferir otro más distante, como "incómodo" (0.01), tal como sucede en este caso. Los datos que aparecen en esta tabla indican que se cumple el requisito enunciado. No obstante, no se puede descartar que los estímulos presenten CCI no decrecientes, ya que en los casos en que los valores a la derecha del máximo son inferiores a él, y considerando los errores típicos asociados a los mismos (entre paréntesis, bajo la celdilla correspondiente) se observa que pudieran pertenecer a una CCI no decreciente. Si sumamos y restamos dos errores típicos al valor máximo y a los valores contiguos, obtenemos un intervalo de confianza con una probabilidad del 95% de contener el valor poblacional de la proporción. Por este procedimiento se observa que en todos los casos podemos obtener CCI no decrecientes. Por ejemplo, si tomamos la fila correspondiente al estímulo "tranquilo", observamos que su máximo se ubica en la columna correspondiente al estímulo "relajado", con un valor de 0.84. Si multiplicamos por dos el error típico asociado a este estadístico para obtener el intervalo de confianza ($\pi < 0.05$) en el que se encontrará el valor del parámetro correspondiente, obtenemos el intervalo (0.754, 0.926). La celdilla siguiente muestra la probabilidad de que los sujetos que eligen el estímulo contento, prefieran sobre los demás estímulos el estímulo

"tranquilo". Esta probabilidad es de 0.70, y el intervalo de confianza correspondiente oscila entre 0.61 y 0.79. Ambos intervalos de confianza muestran cierto grado de solapamiento, de modo que cabe la posibilidad de que el valor poblacional de ambas probabilidades sea igual, es decir, que la CCI del estímulo tranquilo sea no decreciente. En la práctica, un ítem con una CCI no decreciente indica que los sujetos que eligen ese ítem también eligen todos los ítems que se encuentran ubicados a su derecha en la dimensión. En este caso, una CCI no decreciente para el ítem "tranquilo" indicaría que todos los sujetos que se identifican con este ítem también se identifican con el ítem contento, que ocupa un lugar más a la derecha en la dimensión. La misma conclusión se podría extraer de los estímulos "relajado" y "contento". No sucede así para los ítems "tenso" e "incómodo", que más bien presentarían CCI constantes, con una probabilidad condicional poblacional nula. En cualquier caso, Van Schuur y Post (1990) advierten de la falta de justificación teórica que fundamenta este tipo de inferencias acerca de las CCI sobre la base de la información que proporciona su programa. El razonamiento anterior queda pues circunscrito al ámbito de la especulación, a falta de comprobaciones más precisas. Su valor es meramente orientativo, por cuanto que permite formular hipótesis, pero no contrastarlas. Más adelante ilustraremos este punto.

En conclusión, los resultados indican que los datos se ajustan al modelo de despliegamiento probabilístico, formando la escala desplegada los ítems tenso, incómodo, relajado, tranquilo y contento. No obstante, ya indicamos al inicio del análisis que más de un tercio de los sujetos presentan patrones de respuesta que han quedado excluidos del recuento de errores observados, pero no de la estimación de los errores esperados bajo el modelo nulo. Cuando esto sucede, los autores del programa recomiendan un segundo análisis realizado únicamente sobre los sujetos con patrones falsables, para comprobar si los coeficientes de escalabilidad en el análisis inicial están sobreestimados.

En consecuencia, hemos realizado ese segundo análisis, del que sólo daremos la información necesaria para aceptar o rechazar la hipótesis de sobreestimación.

En la tabla 10.7 se observa que el número de sujetos que intervienen en este análisis es de 106, frente a los 292 del análisis anterior. En concreto, se han eliminado todos los sujetos con ninguno o un estímulo preferido. No obstante, las nuevas frecuencias no suponen cambios sustantivos respecto del análisis anterior para los estímulos negativos, aunque se aprecia un aumento en la frecuencia relativa de los estímulos positivos.

Tabla 10.7. Frecuencias absolutas y relativas

| | | | |
|--------------------------|------------|-----|------------|
| NUMBER OF RESPONDENTS | | 106 | |
| NUMBER OF STIMULI | | 6 | |
| NUMBER OF PICKED STIMULI | | ANY | |
| A) TENSO | - MARGINAL | 2 | (1.89 %) |
| B) INCOMODO | - MARGINAL | 3 | (2.83 %) |
| C) PREOCUPADO | - MARGINAL | 7 | (6.60 %) |
| D) TRANQUILO | - MARGINAL | 95 | (89.62 %) |
| E) CONTENTO | - MARGINAL | 81 | (76.42 %) |
| F) RELAJADO | - MARGINAL | 68 | (64.15 %) |

La escala desplegada final contiene un elemento menos que la anterior, el adjetivo relajado. En cuanto al número de errores observados para los estímulos negativos, son los mismos que para el análisis anterior. Si los coeficientes hubieran sido sobreestimados, aparecería una diferencia notable entre el número de errores esperados en este análisis y en el anterior. Sin embargo, no es ése el caso. Más bien al contrario, los estímulos en este análisis alcanzan coeficientes de escalabilidad perfecta ($H=1$), así como la escala final.

En consecuencia, los resultados obtenidos en el análisis inicial no han sido un artefacto debido a la diferencia entre el número de sujetos con patrones falsables y el número de sujetos total.

Tabla 10.8. Escala desplegada final

| FINAL SCALE 1 | | | | | | |
|---|------|------|---------------|----------------|--------|--------|
| STIMULI ABOVE DASHED PAGE FORM AN UNFOLDING SCALE | | | | | | |
| | P(I) | H(I) | $O_{(i,j,k)}$ | $EO_{(i,j,k)}$ | SD(H) | T(H) |
| A) TENSO | 0.02 | 1.00 | 0 | 3.4 | 0.2325 | 4.3015 |
| B) INCOMODO | 0.03 | 1.00 | 0 | 3.5 | 0.2735 | 3.6567 |
| D) TRANQUIL | 0.90 | 1.00 | 0 | 2.1 | 0.2644 | 3.7823 |
| E) CONTENT | 0.76 | 1.00 | 0 | 1.9 | 0.4786 | 2.0893 |
| ----- | | | | | | |
| C) PREOCUPAD | 0.07 | | | | | |
| F) RELAJ | 0.64 | | | | | |

Nota: P(I)=proporción de sujetos que han elegido el estímulo de la fila; H(I)=coeficiente de escalabilidad del estímulo de la fila; $O_{(i,j,k)}$ =error observado asociado al estímulo fila; $EO_{(i,j,k)}$ =error esperado bajo el modelo nulo asociado al estímulo de la fila; S(H)=error típico del coeficiente de escalabilidad del estímulo de la fila; T(H)=valor del estadístico T asociado al coeficiente de escalabilidad del estímulo de la fila.

3. ESCALA DEPRIMIDO-ENTUSIASMADO.

Como en la escala anterior, la descripción de la distribución de puntuaciones dicotomizadas se encuentra en las tablas 10.9 y 10.10. Se observa que más de la mitad de los sujetos no eligen ningún estímulo (puntuaciones de 5 o 6 en la escala de respuesta original), y que de entre los que eligen alguno, la mayor parte elige 1 o 2 estímulos.

Tabla 10.9. Distribución número de estímulos elegidos

| N | A | B | C | D | E | F | TOT |
|---|---|---|---|----|----|----|-----|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 177 |
| 1 | 0 | 3 | 0 | 30 | 3 | 17 | 53 |
| 2 | 3 | 3 | 4 | 29 | 7 | 30 | 37 |
| 3 | 0 | 0 | 0 | 27 | 27 | 27 | 27 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Nota: N=número de estímulos elegidos; A=deprimido, B=pesimista, C=desgraciado, D=alegre, E=entusiasmado, F=optimista TOT=número de sujetos en esa fila

En la tabla 10.10 se observa en primer lugar la frecuencia absoluta con que se ha elegido cada estímulo, y entre paréntesis la frecuencia relativa correspondiente. Nuevamente, muy pocos sujetos han elegido los estímulos negativos de la escala.

Tabla 10.10. Frecuencias absolutas y relativas

| | | | |
|--------------------------|------------|-----|------------|
| NUMBER OF RESPONDENTS | | 292 | |
| NUMBER OF STIMULI | | 6 | |
| NUMBER OF PICKED STIMULI | | ANY | |
| A) DEPRIMIDO | - MARGINAL | 3 | (1.02 %) |
| B) PESIMISTA | - MARGINAL | 6 | (2.04 %) |
| C) DESGRACIADO | - MARGINAL | 2 | (0.68 %) |
| D) ALEGRE | - MARGINAL | 86 | (29.25 %) |
| E) ENTUSIASMADO | - MARGINAL | 37 | (12.59 %) |
| F) OPTIMISTA | - MARGINAL | 74 | (25.17 %) |

En la tabla 10.11 se muestra la terna que cumple los requisitos para ser consideradas "mejor terna", y consecuentemente ser el punto de partida de la escala desplegada. De este modo, la permutación CFD es tomada como punto inicial para la construcción de la escala desplegada. Como ya se indicó, el orden adecuado CFD o DFC vendrá dado por la adición de un nuevo estímulo a la escala, tal como aparece en la tabla 10.12. En ella, se observa que este orden es BACFDE.

Tabla 10.11. Terna y permutación inicial

| | DCF=FCD | | | CDF=FDC | | | CFD=DFC | | |
|-----|---------|-------|--------|---------|-------|--------|---------|------|--------|
| | H | T | S(jik) | H | T | S(ikj) | H | T | S(ikj) |
| CDF | -1.42 | -9.00 | 0.16 | -1.81 | -1.19 | 1.52 | 1.00 | 0.75 | 1.33 |

Nota: H=coeficiente de escalabilidad; T=valor del estadístico T; S=error típico de H.

La tabla 10.12 muestra el proceso seguido en la extensión de la permutación inicial. En primer lugar aparece la proporción de sujetos que han elegido cada estímulo fila (P(I)). A continuación aparecen los coeficientes de escalabilidad de cada estímulo calculados en cada paso

de la construcción de la escala (H(I)). La columna 3 muestra los coeficientes de escalabilidad de los estímulos que forman la permutación inicial, y constituye el paso inicial en la identificación de la escala J. En la columna 4 aparecen los coeficientes para estos tres estímulos así como el correspondiente al nuevo estímulo añadido a la escala, el estímulo deprimido. Los coeficientes que aparecen en esta columna son los calculados para cada estímulo en la nueva escala, es decir, en la escala formada por los cuatro estímulos deprimido, desgraciado, optimista y alegre. A su vez, el orden que ocupan en las filas de la tabla es el que les corresponde en la dimensión latente. En la columna 5 aparecen los coeficientes de escalabilidad para la nueva escala formada ahora por cinco estímulos, tras la adición del estímulo "pesimista". La columna 6 muestra los coeficientes de escalabilidad para la escala final formada por los seis estímulos estudiados, tras la adición del estímulo "entusiasmado". Por último, en la última fila aparece el coeficiente de escalabilidad de cada una de las escalas obtenidas durante el proceso de extensión, es decir, para tres, cuatro, cinco y seis estímulos. Tanto los estímulos incluidos en la escala, que en este caso son todos los analizados, como la propia escala muestran un ajuste aceptable al modelo de desplegamiento analizado.

Tabla 10.12. Proceso de búsqueda de la escala desplegada final

| ITERATIVE SEARCH PROCEDURE 1 | | | | | |
|--|------|-----------|-----------|-----------|-----------|
| P(I): PERCENTAGE OF RESPONDENTS WHO MENTIONED STIMULUS I | | | | | |
| H(I): SCALABILITY VALUE FOR STIMULUS I | | | | | |
| | P(I) | 3 H(I) | 4 H(I) | 5 H(I) | 6 H(I) |
| B) PESIMISTA | 0.02 | --- | --- | 0.62 | 0.71 |
| A) DEPRIMIDO | 0.01 | --- | 1.00 | 0.82 | 0.86 |
| C) DESGRACIADO | 0.01 | 1.00 | 1.00 | 0.81 | 0.85 |
| F) OPTIMISTA | 0.25 | 1.00 | 1.00 | 0.84 | 0.64 |
| D) ALEGRE | 0.29 | 1.00 | 1.00 | 0.56 | 0.51 |
| E) ENTUSIASMADO | 0.13 | --- | --- | --- | 0.62 |
| ----- | | | | | |
| H-SCALE | | 1.00 | 1.00 | 0.72 | 0.67 |

Nota: P(I)=coeficiente de escalabilidad del estímulo de la fila; H(I)=valor del coeficiente de escalabilidad asociado al estímulo de la fila en cada paso seguido en el procedimiento de extensión de la escala.

posterioridad al criterio de 0.30. En conclusión, parece que en este caso los resultados obtenidos sobre el conjunto total de sujetos que componen la muestra acerca de la despleabilidad de los estímulos analizados puede ser un artefacto producido por la diferencia entre el número de sujetos contabilizados y el número de sujetos con patrones de respuesta falsables.

4. ESCALA FATIGA-VIGOR.

Las tablas 10.15 y 10.16 muestran la descripción de la distribución de las puntuaciones dicotomizadas. Se observa que más de un tercio de los sujetos no eligen ningún estímulo, y que entre los que eligen alguno, la mayor parte elige 1 o 2 estímulos.

Tabla 10.15. Distribución número de estímulos elegidos

| N | A | B | C | D | E | F | TOT |
|---|---|---|---|----|----|----|-----|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 106 |
| 1 | 0 | 2 | 0 | 48 | 1 | 12 | 63 |
| 2 | 1 | 0 | 1 | 48 | 18 | 36 | 52 |
| 3 | 7 | 4 | 9 | 60 | 56 | 59 | 65 |
| 4 | 3 | 1 | 3 | 2 | 1 | 2 | 3 |
| 5 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Nota: N=número de estímulos elegidos; A=fatigado, B=sin energía, C=cansado, D=despierto, E=lleño de energía, F=animado; TOT=número de sujetos en esa fila.

En la tabla 10.16 se observa en primer lugar la frecuencia absoluta con que se ha elegido cada estímulo, y entre paréntesis la frecuencia relativa correspondiente. Nuevamente, muy pocos sujetos han elegido los estímulos negativos de la escala.

La tabla 10.17 muestra la terna que cumple los requisitos para ser consideradas "mejor terna", y consecuentemente ser el punto de partida de la escala desplegada. De este modo, la terna CDE es tomada como punto inicial para la construcción de la escala desplegada.

Tabla 10.16. Frecuencias absolutas y relativas

| | | |
|--------------------------|------------|----------------|
| NUMBER OF RESPONDENTS | | 290 |
| NUMBER OF STIMULI | | 6 |
| NUMBER OF PICKED STIMULI | | ANY |
| A) FATIGADO | - MARGINAL | 12 (4.14 %) |
| B) SIN ENERGIA | - MARGINAL | 7 (2.41 %) |
| C) CANSADO | - MARGINAL | 14 (4.83 %) |
| D) DESPIERTO | - MARGINAL | 159 (54.83 %) |
| E) LLENO ENERGIA | - MARGINAL | 77 (26.55 %) |
| F) ANIMADO | - MARGINAL | 110 (37.93 %) |

Tabla 10.17. Terna y permutaciones candidatas a ser la terna inicial

| | DCE=ECD | | | CDE=EDC | | | DEC=CED | | |
|-----|---------|------|--------|---------|------|--------|---------|-------|--------|
| | H | T | S(jik) | H | T | S(ijk) | H | T | S(ikj) |
| CDE | -0.77 | 0.20 | 4.92 | 0.40 | 0.57 | 0.72 | -0.06 | -0.20 | 0.32 |

Nota: H=coeficiente de escalabilidad; T=valor del estadístico T; S=error típico de H.

La tabla 10.17 muestra el proceso seguido en la extensión de la permutación inicial. En primer lugar aparece la proporción de sujetos que han elegido cada estímulo fila (P(I)). A continuación aparecen los coeficientes de escalabilidad de cada estímulo calculados en cada paso de la construcción de la escala (H(I)). El proceso de despliegue puede ser seguido en esta tabla. Bajo la columna 3 se encuentran los estímulos que componen la permutación inicial. Bajo la columna 4 aparece además el estímulo incorporado en este paso a la escala, que es "sin energía". La columna 5 muestra el siguiente paso, que consiste en la adición del estímulo "fatigado". Finalmente, la columna 6 muestra la escala final, completada tras la adición del estímulo "animado", y en el orden adecuado. Tanto los estímulos incluidos en la escala, que en este caso son todos los analizados, como la propia escala muestran un ajuste aceptable al modelo de desplegamiento analizado.

Tabla 10.17. Proceso de búsqueda de la escala desplegada final

| ITERATIVE SEARCH PROCEDURE I | | | | | |
|--|------|-----------|-----------|-----------|-----------|
| P(I): PERCENTAGE OF RESPONDENTS WHO MENTIONED STIMULUS I | | | | | |
| H(I): SCALABILITY VALUE FOR STIMULUS I | | | | | |
| | P(I) | 3 H(I) | 4 H(I) | 5 H(I) | 6 H(I) |
| B) SIN ENERGIA | 0.02 | --- | 1.00 | 1.00 | 0.95 |
| A) FATIGADO | 0.04 | --- | --- | 0.94 | 0.86 |
| C) CANSADO | 0.05 | 0.40 | 0.86 | 0.94 | 0.87 |
| F) ANIMADO | 0.38 | --- | --- | --- | 0.72 |
| D) DESPIERTO | 0.55 | 0.40 | 0.84 | 0.89 | 0.69 |
| E) LLENO ENERGIA | 0.27 | 0.40 | 0.77 | 0.81 | 0.79 |
| ----- | | | | | |
| H-SCALE | | 0.40 | 0.87 | 0.92 | 0.80 |

Nota: P(I)=coeficiente de escalabilidad del estímulo de la fila; H(I)=valor del coeficiente de escalabilidad asociado al estímulo de la fila en cada paso seguido en el procedimiento de extensión de la escala.

Tabla 10.18. Escala desplegada final

| FINAL SCALE I | | | | | | |
|---|------|------|---------------|----------------|-------|------|
| STIMULI ABOVE DASHED PAGE FORM AN UNFOLDING SCALE | | | | | | |
| | P(I) | H(I) | $O_{(i,j,k)}$ | $EO_{(i,j,k)}$ | SD(H) | T(H) |
| B) SIN ENERGIA | 0.02 | 0.95 | 1 | 20.7 | 0.23 | 3.97 |
| A) FATIGADO | 0.04 | 0.86 | 4 | 29.5 | 0.15 | 5.43 |
| C) CANSADO | 0.05 | 0.87 | 4 | 30.6 | 0.14 | 5.93 |
| F) ANIMADO | 0.38 | 0.72 | 11 | 39.3 | 0.12 | 5.91 |
| D) DESPIERTO | 0.55 | 0.69 | 13 | 42.0 | 0.13 | 5.03 |
| E) LLENO ENERGIA | 0.27 | 0.79 | 6 | 29.2 | 0.20 | 3.79 |

Nota: P(I)=proporción de sujetos que han elegido el estímulo de la fila; H(I)=coeficiente de escalabilidad del estímulo de la fila; $O_{(i,j,k)}$ =error observado asociado al estímulo fila; $EO_{(i,j,k)}$ =error esperado bajo el modelo nulo asociado al estímulo de la fila; S(H)=error típico del coeficiente de escalabilidad del estímulo de la fila; T(H)=valor del estadístico T asociado al coeficiente de escalabilidad del estímulo de la fila.

La tabla 10.18 muestra la descripción de la escala desplegada final. Junto a la proporción de sujetos que eligen cada estímulo (P(I)) y el coeficiente de escalabilidad correspondiente (H(I)) aparecen el número

de errores observados $O_{(i,j,k)}$, el número de errores esperados $EO_{(i,j,k)}$, el error típico asociado al coeficiente de escalabilidad de cada estímulo (SD(H)) y el valor T correspondiente. Todos los coeficientes de escalabilidad resultan significativamente distintos de cero y elevados, así como el coeficiente de escalabilidad de la escala total (H=.80, T=6.60).

La tabla 10.19 muestra la matriz de adyacencia condicional. Nuevamente se han señalado en negrita los máximos de cada fila. Excepto la última fila, el resto cumple los requisitos esperados bajo el modelo probabilístico de desplegamiento.

Tabla 10.19. Matriz de adyacencia condicional

| CONDITIONAL ADJACENCY MATRIX 1 | | | | | | |
|--|--------|-------------|-------------|-------------|--------|-------------|
| IN CELL (I,J): proportion of respondents who mentioned row stimulus i among all respondents who mentioned column stimulus j. | | | | | | |
| In brackets: the standard error belonging to the proportion | | | | | | |
| | B | A | C | F | D | E |
| B) SIN ENERGIA | * | 0.42 | 0.36 | 0.00 | 0.01 | 0.00 |
| | | (.142) | (.128) | (.000) | (.006) | (.000) |
| A) FATIGADO | 0.71 | * | 0.86 | 0.04 | 0.03 | 0.03 |
| | (.171) | | (.094) | (.018) | (.014) | (.018) |
| C) CANSADO | 0.71 | 1.00 | * | 0.05 | 0.04 | 0.03 |
| | (.171) | (0.000) | | (.022) | (.016) | (.018) |
| F) ANIMADO | 0.00 | 0.33 | 0.43 | * | 0.58 | 0.79 |
| | (.000) | (.136) | (0.132) | | (.039) | (.046) |
| D) DESPIERTO | 0.14 | 0.42 | 0.50 | 0.85 | - | 0.94 |
| | (.132) | (.142) | (.134) | (.034) | | (.028) |
| E) LLENO ENERGIA | 0.00 | 0.17 | 0.14 | 0.55 | 0.45 | - |
| | (.000) | (.108) | (.094) | (.047) | (.039) | |

En conclusión, los resultados indican que los datos se ajustan en general al modelo de desplegamiento probabilístico, formando la escala desplegada los items sin energía, fatigado, cansado, animado, despierto y lleno de energía. Nuevamente, más de un tercio de los sujetos presentan patrones de respuesta que han quedado excluidos del recuento de errores observados, pero no de la estimación de los errores esperados bajo el modelo nulo. Por ello, hemos realizado un segundo análisis

únicamente sobre los sujetos con patrones falsables para comprobar si los coeficientes de escalabilidad en el análisis inicial están sobreestimados. En este caso, la escala resultante tampoco mostró un buen ajuste al modelo de desplegamiento. El coeficiente de escalabilidad de la escala presentó un valor de 0.42, con un valor $T=1.43$, es decir, estadísticamente no significativo. Por su parte, aunque los coeficientes de escalabilidad de los items resultaron mayores de 0.30, ninguno de ellos fue significativamente diferentes de cero. En conclusión, parece que los resultados obtenidos sobre el conjunto total de sujetos que componen la muestra, acerca de la despleabilidad de los estímulos analizados, puede ser un artefacto producido por la diferencia entre el número de sujetos contabilizados y el número de sujetos con patrones de respuesta falsables.

Los resultados obtenidos en el análisis de despliegue de estas dos últimas escalas, deprimido-entusiasmado y fatiga-vigor, cuestiona la explicación de Van Schuur y Kiers (en prensa) acerca de que la obtención de factores monopulares cuando se espera un factor bipolar se debe a la aplicación de un modelo de medida y un método de análisis inadecuados. En concreto a la aplicación del modelo de escalamiento tipo Likert, y del método de análisis factorial cuando el modelo de escalamiento adecuado es el de desplegamiento probabilístico, y el método de análisis el de desplegamiento.

5. ANALISIS CORRELACIONAL.

La segunda de las hipótesis que se han formulado acerca del modelo de desplegamiento es que las puntuaciones obtenidas bajo este modelo de medida presentarán coeficientes de correlación iguales o mayores con los items de diferencial semántico correspondientes, que los obtenidos bajo los modelos de medida anteriormente explorados.

En la tabla 10.20 aparecen las correlaciones correspondientes a la escala tenso-calmado. Como ya se mencionó anteriormente, no todos los sujetos que componen la muestra reciben puntuaciones bajo el modelo de desplegamiento. Aquellos sujetos con patrones de respuesta compuestos exclusivamente por ceros no pueden ser puntuados, de

modo que la tabla siguiente se ha calculado sólo con la submuestra de sujetos que recibe puntuaciones bajo los tres modelos estudiados hasta ahora.

La inspección de la tabla 10.20 revela que las puntuaciones que asigna el modelo de desplegamiento presentan una correlación con el criterio que estamos empleando -el ítem de diferencial semántico- que no alcanza el nivel de significación estadística ($r=0.188$, $p>0.05$, dos colas). Por el contrario, las correlaciones con el criterio correspondientes al modelo tradicional y dos mitades sí son estadísticamente significativas. Además, aunque la correlación entre las puntuaciones bajo el modelo tradicional y las obtenidas en el criterio (.534) es mayor que la obtenida entre las puntuaciones del modelo dos mitades con las obtenidas en el criterio (.478), esta diferencia no es estadísticamente significativa ($t=1.86$, $p>0.05$, dos colas). En cambio, la diferencia entre las correlaciones obtenidas bajo el modelo dos mitades con el criterio (.478) y bajo el modelo de desplegamiento con el criterio (.188) sí resulta estadísticamente significativa ($t=3.84$, $p<0.01$, dos colas). El contraste de diferencias entre la correlación calculada bajo el modelo tradicional con el criterio (.534), y la correlación calculada bajo el modelo de desplegamiento con el criterio (.188), también indicó que ambas correlaciones presentaban diferencias estadísticamente significativas ($t=3.96$, $p<0.01$, dos colas).

Tabla 10.20 Análisis correlacional. Escala tenso-calmado (N=165)

| | ds | mt | dm | md |
|----|-------|------|------|----|
| ds | 1 | | | |
| mt | .534 | 1 | | |
| dm | .478 | .884 | 1 | |
| md | .188* | .348 | .423 | 1 |

Nota: ds=puntuación en la escala criterio; mt=puntuación bajo el modelo tradicional; dm=puntuación bajo el modelo de "dos mitades"; md=puntuación bajo el modelo de desplegamiento probabilístico; $r>0.194$, $p<0.05$; $r>0.253$, $p<0.01$; * $p>0.05$, dos colas.

Este análisis revela que incluso en el subconjunto de sujetos para los que el modelo de desplegamiento ha resultado adecuado, son las puntuaciones calculadas bajo los modelos anteriores -tradicional y "dos mitades"- las que más se aproximan a la ubicación de los sujetos en un espacio explícitamente bipolar. También se obtiene evidencia de que ambos modelos producen puntuaciones que se relacionan por igual con el criterio.

En el caso de la escala deprimido-entusiasmado las puntuaciones que asigna el modelo de desplegamiento presentan una correlación con el criterio que estamos empleando -el ítem de diferencial semántico-estadísticamente significativa ($r=.366$, $p<0.01$, dos colas), al igual que las correlaciones con el criterio correspondientes al modelo tradicional ($r=.668$, $p<0.01$, dos colas), y dos mitades ($r=.496$, $p<0.01$, dos colas). En este caso la correlación entre las puntuaciones bajo el modelo tradicional y las obtenidas en el criterio (.668) es significativamente distinta ($t=4.26$, $p<0.01$, dos colas) de la obtenida entre las puntuaciones del modelo dos mitades con las obtenidas en el criterio (.496).

Tabla 10.21 Análisis correlacional. Escala deprimido-entusiasmado (N=115)

| | ds | mt | dm | md |
|----|------|------|-----|----|
| ds | 1 | | | |
| mt | .668 | 1 | | |
| dm | .496 | .834 | 1 | |
| md | .366 | .652 | .75 | 1 |

Nota: ds=puntuación en la escala criterio; mt=puntuación bajo el modelo tradicional; dm=puntuación bajo el modelo de "dos mitades"; md=puntuación bajo el modelo de desplegamiento probabilístico; $r>0.194$, $p<0.05$; $r>0.253$, $p<0.01$.

En cuanto al modelo restante, la diferencia entre las correlaciones obtenidas bajo el modelo de desplegamiento (.336) y bajo el modelo de dos mitades (.496) también resulta estadísticamente significativa ($t=2.48$, $p<0.05$, dos colas). Estos resultados coinciden con los obtenidos en la escala anterior en tanto que el modelo de desplegamiento muestra

la correlación más débil con el criterio. Sin embargo, en este caso el modelo tradicional presenta una relación con el criterio mayor que la del modelo de dos mitades.

Finalmente, en el caso de la escala fatiga-vigor, la tabla 10.22 presenta un patrón de correlaciones similar al que mostró la escala tenso-calmado. Las puntuaciones que asigna el modelo de despliegamiento presentan una correlación con el criterio estadísticamente significativa ($r=.533$, $p<0.01$, dos colas), al igual que las correlaciones con el criterio correspondientes al modelo tradicional ($r=.701$, $p<0.01$, dos colas) y dos mitades ($r=.673$, $p<0.01$, dos colas). La correlación entre las puntuaciones bajo el modelo tradicional y las obtenidas en el criterio (.701) no difiere significativamente ($t=1.120$, $p>0.05$, dos colas) de la obtenida entre las puntuaciones del modelo dos mitades con las obtenidas en el criterio (.673). En cuanto al modelo restante, la diferencia entre las correlaciones obtenidas bajo el modelo de despliegamiento con el criterio (.533) y bajo el modelo de dos mitades con el criterio (.673) sí resulta estadísticamente significativa ($t=2.81$, $p<0.05$, dos colas). En conclusión, aunque la correlación entre las puntuaciones bajo el modelo desplegado y el criterio son mayores que en los casos anteriores, continúan siendo superadas por las correlaciones obtenidas bajo los modelos alternativos, que en este caso no difieren significativamente entre sí.

Tabla 10.22 Análisis correlacional. Escala fatiga-vigor (N=184)

| | ds | mt | dm | md |
|----|------|------|------|----|
| ds | 1 | | | |
| mt | .701 | 1 | | |
| dm | .673 | .888 | 1 | |
| md | .533 | .596 | .616 | 1 |

Nota: ds=puntuación en la escala criterio; mt=puntuación bajo el modelo tradicional; dm=puntuación bajo el modelo de "dos mitades"; md=puntuación bajo el modelo de despliegamiento probabilístico; $r>0.194$, $p<0.05$; $r>0.253$, $p<0.01$.

CAPITULO 11

**ESCALAMIENTO ACUMULATIVO DE GUTTMANN BAJO EL
MODELO DE DESPLEGAMIENTO.**

1. INTRODUCCION.

Recientemente, Van Schuur y Kruijtbosch (1994) han desarrollado un enfoque alternativo al del modelo de despliegamiento probabilístico mencionado anteriormente. Partiendo del estudio de la escala de Equilibrio Emotivo de Bradburn (1969), llegaron a mostrar que ésta, como otras escalas desarrolladas para medir dimensiones bipolares, podía ajustarse al modelo de escalamiento acumulativo de Guttman (Guttman, 1944; 1950), que a su vez puede considerarse un caso limítrofe del modelo de despliegamiento.

El modelo de escalamiento acumulativo es, de acuerdo con la clasificación de Torgerson (1958), un modelo centrado en la respuesta, y de carácter determinista. Como se recordará, ello indica que se trata de un modelo que aborda el escalamiento conjunto de sujetos y estímulos, y que considera que toda la variabilidad en las puntuaciones se debe a diferencias en la ubicación de los sujetos y de los estímulos en la dimensión latente, y no a otras fuentes como el error de medida.

Imaginemos una escala constituida únicamente por items representativos de estados de ánimo negativos. En una escala desarrollada bajo este modelo, los items se ubican en el continuo de forma ordenada, en función del grado en que les caracteriza el atributo bajo estudio. La disposición de los items en la dimensión latente es análoga a la del escalamiento desplegado probabilístico. Así, en la matriz 11.1 se observa que el ítem extremadamente negativo sería el número 1, y el ítem neutro el número 3. Por otra parte, las respuestas de los sujetos a los items típicos de este tipo de escalas sólo presentan dos alternativas de respuesta: respuesta positiva o de aceptación - representada con un 1-, y respuesta negativa o de rechazo - representada con un 0-. Más adelante ilustraremos el significado de cada una de ellas. Una escala de tres items que se ajustara al modelo determinista acumulativo podría desplegarse en una matriz de modo que los items se ordenaran en las columnas de extremadamente negativo a neutro, y los sujetos en las filas, ordenados de forma que los que presentan estados de ánimo neutros se colocan en la parte superior

de la matriz, y los que presentan estados de ánimo negativos se ubican en la parte inferior. La representación de esta matriz sería la siguiente:

Matriz 11.1 Escalamiento de Guttman. Ejemplo 1

| | | estado de ánimo | | | |
|----------|---|---------------------|---|---|---|
| | | negativo.....neutro | | | |
| | | Items | | | |
| | | 1 | 2 | 3 | |
| Sujetos | | - | | + | |
| neutro | A | + | 0 | 0 | 0 |
| | B | | 0 | 0 | 1 |
| | C | | 0 | 1 | 1 |
| negativo | D | - | 1 | 1 | 1 |

Si calculamos la puntuación total de cada sujeto en la escala como el número de estímulos a los que el sujeto ha dado una respuesta positiva, podemos asignar al sujeto A la puntuación 0, al sujeto B la puntuación 1, al sujeto C la puntuación 2 y al sujeto D la puntuación 3. Bajo este modelo se asume que una respuesta positiva a un ítem indica que el punto ideal del sujeto en el continuo está a la izquierda del que corresponde al ítem, y que una respuesta negativa indica que el punto de ideal del sujeto está a la derecha del que corresponde al ítem (Guttman, 1950). Tomemos como ejemplo el patrón de respuestas del sujeto C. Este sujeto ha respondido afirmativamente a los ítems 2 y 3, y negativamente al ítem 1. Ello indica que su posición en el continuo se encuentra a la derecha del ítem con la respuesta 0 o negativa, y a la izquierda de los ítems con la respuesta 1 o positiva. Ello nos daría una ubicación para este sujeto comprendida entre las posiciones que ocupan los estímulos 1 y 2. Por el mismo procedimiento podemos determinar la ubicación del sujeto B en el espacio comprendido entre los ítems 2 y 3.

Esta matriz ilustra la característica fundamental del modelo de escalamiento acumulativo de Guttman: conocida la puntuación de un sujeto en la escala, es posible deducir a qué ítems el sujeto respondió

afirmativamente, y a qué ítems respondió negativamente. En otros términos, es posible reproducir su patrón de respuesta. Puesto que los ítems están ordenados de estado de ánimo extremadamente negativo a estado de ánimo neutro, y son acumulativos, la respuesta positiva de un sujeto a un ítem implica que también ha respondido positivamente a todos los ítems que se encuentran a su derecha. Veámoslo con otro ejemplo. Si elaboramos una escala acumulativa para medir el retardo en el crecimiento de un conjunto de niños, podríamos emplear los siguientes ítems: 1. altura menor que 1.10; 2. altura menor que 1.20; y 3. altura menor que 1.30. Esta escala ordenaría los estímulos en la dimensión retardo en el crecimiento de tal forma que los sujetos con ligero retardo responderían afirmativamente al ítem 3, los sujetos con retardo moderado responderían afirmativamente al ítem 3 y al 2, y sólo los sujetos con retardo severo responderían afirmativamente también al ítem 1. Por su parte, los sujetos sin retardo responderían negativamente a todos los ítems. Las respuestas positivas son acumulativas, ya que una respuesta positiva a un ítem necesariamente implica que también se ha respondido positivamente a los ítems que se encuentran a la derecha de éste. No puede suceder que un niño muestre un retardo moderado, y por lo tanto responda positivamente al ítem 2. altura menor que 1.20, y que al tiempo responda negativamente al ítem 3. altura menor que 1.30. Si su altura es menor que 1.20, también es menor que 1.30. De modo que si conocemos que un sujeto determinado ha obtenido una puntuación igual a 2 en esta escala, podemos reproducir directamente su patrón de respuesta. Este sólo puede haber sido 011. Es decir, si sólo ha contestado afirmativamente a dos estímulos, y los estímulos se encuentran ordenados en orden acumulativo decreciente -de mayor a menor retardo-, esos dos estímulos quedan directamente determinados por las propiedades de la escala, y sólo pueden ser los estímulos 2 y 3.

Volviendo al ejemplo anterior, la aplicación de esta lógica indica que los sujetos con estados de ánimo neutro responderán afirmativamente sólo al ítem 3, los sujetos con estados de ánimo moderadamente negativos responderán positivamente al ítem 3 y también al ítem 2, y los sujetos con estados de ánimo extremadamente negativos, responderán positivamente a los ítems anteriores y al ítem 1.

Los sujetos que no responden positivamente a ningún ítem presentarán estados de ánimo positivos, que no quedan recogidos en la escala.

En cualquier caso, la lógica que se sigue es identificar un conjunto de ítems tales que uno de ellos represente el nivel base de la cualidad a medir. En el ejemplo, el estímulo 3 correspondería a este estado base, común a todos los sujetos caracterizados por la dimensión objeto de estudio. A continuación se trata de identificar estímulos de forma que cada uno de ellos represente un nivel del atributo a medir mayor que el anterior. En nuestro ejemplo, el estímulo 2 representa un estado de ánimo más negativo que el 3, y el 1 es más negativo que todos los anteriores. De este modo, los sujetos que se caractericen por poseer el atributo medido en mayor grado que la línea base de la escala (el estímulo 3), responderán positivamente a los estímulos que también se alejen hacia el extremo negativo de esa línea base, y a todos los estímulos que se encuentren entre la línea base y el que mejor describa su posición ideal. Así, el sujeto D que es el que presenta la mayor desviación desde el origen de la escala o línea base ha respondido positivamente al estímulo más extremo, y también a todos los que ocupan posiciones intermedias. Finalmente, podríamos deducir el patrón de respuesta de cada sujeto a partir de la puntuación en la escala, por el procedimiento aplicado en el ejemplo anterior.

Imaginemos ahora otra escala constituida únicamente por ítems representativos de estados de ánimo positivos, también de tres ítems, que se ajustara al modelo determinista acumulativo. Esta escala podría desplegarse de modo similar a la anterior, de forma que los ítems se ordenaran en las columnas de neutro a extremadamente positivo, y los sujetos en las filas, ordenados de forma que los que presentan estados de ánimo más positivos se colocaran en la parte superior de la matriz, y los que presentan estados de ánimo neutros se ubicaran en la parte inferior. La representación de esta matriz sería la siguiente:

Matriz 11.2 Escalamiento de Guttman. Ejemplo 2

| | | estado de ánimo | | |
|----------|-----|---------------------|---|---|
| | | neutro.....positivo | | |
| Sujetos | | Items | | |
| | | 1 | 2 | 3 |
| | | - | + | |
| positivo | A + | 1 | 1 | 1 |
| | B | 1 | 1 | 0 |
| | C | 1 | 0 | 0 |
| neutro | D - | 0 | 0 | 0 |

La primera escala que se ha presentado como ejemplo constituiría una escala que se ajustaría al modelo de escalamiento acumulativo, con CCI no decrecientes, y la segunda constituiría otra escala del mismo tipo, pero con CCI no crecientes. La relación entre la escala y las CCI de los ítems que la componen puede ser ilustrada en los siguientes términos. Consideraremos nuevamente la matriz 11.1, pero en este caso centrándonos en sus columnas. Cada columna muestra la distribución de respuestas positivas y negativas al ítem, en función de la posición de los sujetos en la dimensión. Cada columna puede interpretarse como una aproximación a la CCI del ítem correspondiente a esa columna. Bajo esta perspectiva se observa que la probabilidad de obtener una respuesta positiva a cualquiera de los ítems no disminuye a lo largo de la columna que le corresponde. La probabilidad de dar una respuesta positiva a un ítem es nula para los sujetos con un nivel en el atributo medido inferior al del ítem, y distinta de cero y no decreciente para los sujetos con un grado del atributo medido igual o mayor al del ítem. En cualquier caso, lo que no puede ocurrir en este tipo de escala es que las columnas presenten, de arriba a bajo, respuestas positivas y luego respuestas negativas. Es decir, sus CCIs han de ser no decrecientes.

Matriz 11.1 Escalamiento de Guttman. Ejemplo 1

| | | | estado de ánimo | | |
|----------|---|---|---------------------|---|---|
| | | | negativo.....neutro | | |
| | | | Items | | |
| | | | 1 | 2 | 3 |
| Sujetos | | | - | | + |
| neutro | A | + | 0 | 0 | 0 |
| | B | | 0 | 0 | 1 |
| | C | | 0 | 1 | 1 |
| negativo | D | - | 1 | 1 | 1 |

Si observamos las columnas de la matriz 11.2, se aprecia que el patrón de respuestas positivas es inverso, de modo que lo que no puede ocurrir en este tipo de escala es que las columnas presenten, de arriba a abajo, respuestas negativas y luego respuestas positivas. Es decir, sus CCI's han de ser no crecientes.

Si cada una de estas escalas midiera una de las mitades de un continuo bipolar, y se presentaran juntas, obtendríamos la siguiente matriz:

Matriz 11.3 Escalamiento de Guttman. Ejemplo 3

| | | | estado de ánimo | | | | | |
|----------|---|---|-----------------------|---|---|---|---|---|
| | | | negativo.....positivo | | | | | |
| | | | Items | | | | | |
| | | | 1 | 2 | 3 | 4 | 5 | 6 |
| Sujetos | | | - | | | | | + |
| positivo | A | + | 0 | 0 | 0 | 1 | 1 | 1 |
| | B | | 0 | 0 | 1 | 1 | 1 | 0 |
| | C | | 0 | 1 | 1 | 1 | 0 | 0 |
| negativo | D | - | 1 | 1 | 1 | 0 | 0 | 0 |

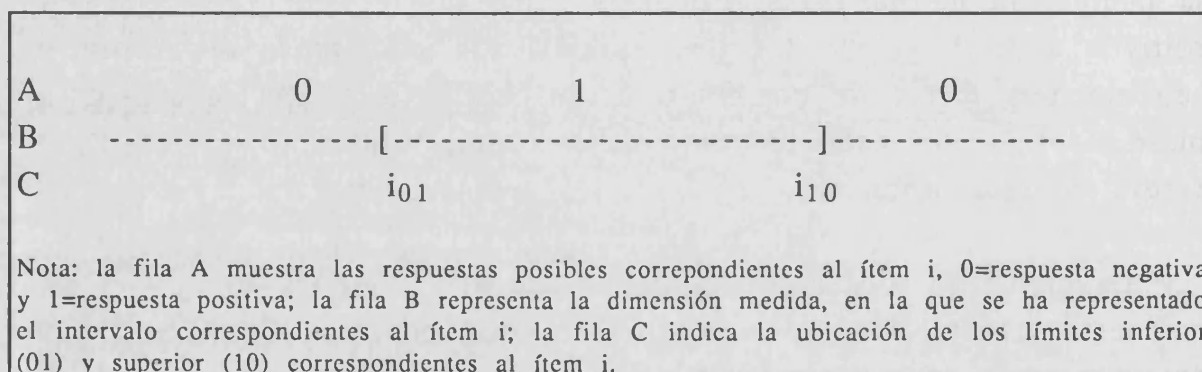
Ambas escalas tomadas de forma conjunta formarían un paralelogramo perfecto, que se ajusta al modelo determinista de despliegue de datos dicotómicos, también conocido como "análisis de paralelogramos" (Coombs, 1964; Hoijtink, 1990).

Uno de los problemas con que se enfrentaba esta conceptualización del modelo de medida subyacente a escalas bipolares era el de la obtención de puntuaciones en la escala conjunta. Como ya indicamos, las puntuaciones de los sujetos en cada una de las escalas acumulativas se calculan mediante el número de respuestas positivas a los estímulos correspondientes. El problema se plantea a la hora de asignar puntuaciones en la escala de despliegue paralelográfica. En este caso esa puntuación no puede ser calculada como la suma de los elementos positivos a lo largo de las dos escalas. Sin embargo, era necesario encontrar una forma de combinar las respuestas a ambas escalas para ubicar a los sujetos adecuadamente en el continuo latente. Veamos cómo se resolvió este problema.

La hipotética escala de despliegue que estamos considerando es un compuesto latente unidimensional, en el que cada uno de los ítems ocupa una posición diferente. En esta escala, la respuesta negativa 0 dada por un sujeto puede tener dos interpretaciones, dependiendo del resto de respuestas dadas por el mismo sujeto. Así, una respuesta 0 a la izquierda de una respuesta positiva o 1 indicaría que el sujeto ha superado el grado de la cualidad medida que caracteriza al ítem contestado con un 0. Por el contrario, si la respuesta 0 se ubica tras una respuesta positiva, lo que indicaría es que el sujeto no alcanza el grado de la cualidad medida que caracteriza al ítem con la respuesta negativa, y que el sujeto se identifica con el grado de la cualidad que caracteriza al ítem con la respuesta positiva. Si nos fijamos en la matriz 11.3, podemos observar que el sujeto B presenta el siguiente patrón de respuestas "001110". La respuesta negativa al estímulo 1 y al estímulo 2 indica que su afecto es más positivo que el que representan estos ítems, y su respuesta negativa al estímulo 6 indica que su afecto no es tan positivo como el que representa este ítem. Tomando en consideración todo ello, Van Schuur (1993a) propuso el siguiente método de asignación de puntuaciones. Si se considera que cada ítem delimita un intervalo

cerrado en la dimensión latente, entonces se puede esperar que el sujeto dé una respuesta positiva cuando su posición se encuentre dentro de ese intervalo, y negativa cuando su posición se ubique fuera. Concretando, dado un ítem "i", se puede designar como i_{01} el límite inferior del intervalo que le representa, y como i_{10} el límite superior del mismo. Los sujetos cuya posición se ubica dentro del intervalo darán una respuesta positiva a este ítem, en tanto que los sujetos que se encuentran a la izquierda o derecha de ese intervalo darán una respuesta negativa, tal como se ilustra en la figura siguiente.

Figura 11.4 Representación del intervalo de un ítem, y las respuestas posibles

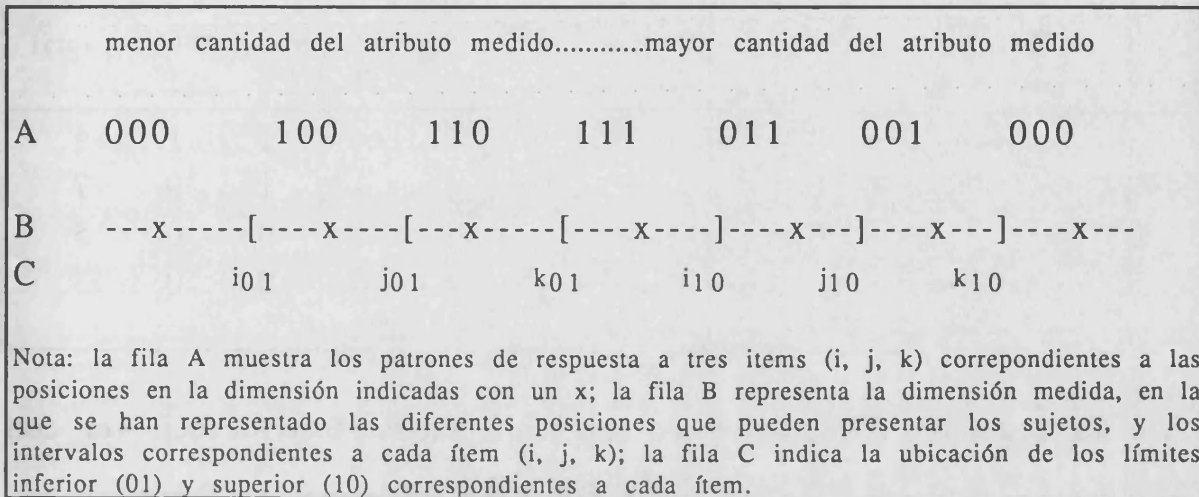


El modelo propuesto por Van Schuur (1993a) asume que todos los ítems presentan la misma ordenación para los límites inferiores y superiores de los intervalos que ocupan. Esto implica una ordenación inequívoca de los ítems, en la que tres ítems darían lugar a seis límites y siete áreas isotónicas, tal como aparece representado en la figura 11.5.

De este modo, las dos interpretaciones de la respuesta 0 de un sujeto quedan representadas en el modelo: si la respuesta 0 se da tras una respuesta positiva, como ocurre en el patrón de respuesta 100, ello indica que el sujeto supera o se ubica a la derecha del límite inferior del ítem i que presenta una respuesta positiva (i_{01}), pero a la izquierda del límite inferior del ítem j, que muestra una respuesta negativa (j_{01}). Sin embargo, si tras la respuesta 0 se da una respuesta positiva, por ejemplo 001, ello indica que el sujeto se ubica a la derecha del límite

superior del ítem j que presenta una respuesta negativa (j_{10}), pero a la izquierda del límite superior del ítem k, que muestra una respuesta positiva (k_{10}).

Figura 11.5 Representación de una escala de tres ítems (i, j, k), y sus respuestas posibles.



Finalmente, la puntuación correspondiente a un sujeto vendría dada por el número de límites de intervalos que supera. Por ejemplo, un sujeto con el patrón de respuesta 011 habría superado los dos límites del intervalo correspondiente al primer ítem (i) (i_{01} y i_{10}), y sólo el límite inferior de los dos ítems siguientes (j y k) (j_{01} y k_{01}), de modo que a este sujeto le correspondería la puntuación de 4. De forma similar, un sujeto con un patrón de respuesta 110 habría superado los dos límites inferiores de los ítems i y j, pero ninguno del ítem k, por lo que recibiría una puntuación de 2.

Si aplicamos este sistema de puntuación a la escala de despliegue mostrada con anterioridad, obtenemos las puntuaciones que aparecen a continuación:

Matriz 11.6 Asignación de puntuaciones en la escala paralelogramática.

| | | estado de ánimo | | | | | | | |
|----------|---|-----------------|---|---|----------|---|---|-------|---|
| | | negativo | | | positivo | | | | |
| | | Items | | | | | | | |
| | | 1 | 2 | 3 | 4 | 5 | 6 | total | |
| Sujetos | | - | | | | | + | | |
| positivo | A | + | 0 | 0 | 0 | 1 | 1 | 1 | 9 |
| | B | | 0 | 0 | 1 | 1 | 1 | 0 | 7 |
| | C | | 0 | 1 | 1 | 1 | 0 | 0 | 5 |
| negativo | D | - | 1 | 1 | 1 | 0 | 0 | 0 | 3 |

El proceso de obtención de estas puntuaciones se pueden ver con mayor claridad en la matriz siguiente:

Matriz 11.7 Asignación de puntuaciones en la escala paralelogramática por el procedimiento de recuento de intervalos.

| | | estado de ánimo | | | | | | | |
|----------|---|-----------------|---|---|----------|---|---|-------|---|
| | | negativo | | | positivo | | | | |
| | | Items | | | | | | | |
| | | 1 | 2 | 3 | 4 | 5 | 6 | total | |
| Sujetos | | - | | | | | + | | |
| positivo | A | + | 2 | 2 | 2 | 1 | 1 | 1 | 9 |
| | B | | 2 | 2 | 1 | 1 | 1 | 0 | 7 |
| | C | | 2 | 1 | 1 | 1 | 0 | 0 | 5 |
| negativo | D | - | 1 | 1 | 1 | 0 | 0 | 0 | 3 |

Nota: total= puntuacion total en la escala paralelogramática formada por las dos escalas acumulativas combinadas como muestra la matriz.

Comparando las matrices 11.6 y 11.7 se observa que las respuestas negativas a la izquierda de las respuestas positivas no tienen el mismo valor que las que aparecen a la derecha de las respuestas

positivas. Como hemos visto, las primeras implican haber superado los dos límites que las delimitan, y las segundas ninguno de ellos. En consecuencia, los ceros de la izquierda se transforman en 2, y los de la derecha mantienen su valor. En ambos casos, la puntuación refleja el número de límites superados. Por otra parte, esta escala permite definir el término de error de forma análoga al que empleamos bajo el modelo de desplegamiento probabilístico: se trata del patrón de respuesta 101, ya que la escala es acumulativa, y un sujeto no puede dar una respuesta positiva a dos estímulos dados, separados en la dimensión por un tercero, sin dar una respuesta positiva también a ese estímulo. De este modo es posible aplicar el método de desplegamiento anterior para análisis, con las modificaciones que detallaremos más adelante.

Cuando las puntuaciones objeto de estudio proceden de escalas dicotómicas, o bien han sido dicotomizadas -como sucede aquí- la comparación de las puntuaciones asignadas bajo este modelo con las asignadas bajo el modelo tradicional presenta la siguiente correspondencia:

Tabla 11.8. Comparación de puntuaciones bajo diferentes modelos.

| sub. negativa | | | N | sub. positiva | | | P | MT1 | MT2 | EA |
|---------------|---|---|---|---------------|---|---|---|-----|-----|----|
| 2 | 2 | 2 | 0 | 1 | 1 | 1 | 3 | 3 | 6 | 9 |
| 2 | 2 | 1 | 1 | 1 | 1 | 0 | 2 | 1 | 4 | 7 |
| 2 | 1 | 1 | 2 | 1 | 0 | 0 | 1 | -1 | 2 | 5 |
| 1 | 1 | 1 | 3 | 0 | 0 | 0 | 0 | -3 | 0 | 3 |

Nota: Subescala negativa= puntuaciones en los items de la subescala negativa, tal como se calcularon en la matriz 11.5; N=puntuaciones totales en la subescala negativa; Subescala positiva=puntuaciones en los items de la subescala positiva, tal como se calcularon en la matriz 11.5; P=puntuaciones totales en la subescala positiva; MT1=puntuaciones totales en la escala por el procedimiento de suma y diferencia, modelo tradicional; MT2= puntuaciones totales en la escala por el procedimiento de inversión y promedio, modelo tradicional; EA=puntuaciones totales en la escala bajo el modelo de escalas acumulativas.

Bajo el modelo bipolar tradicional, se han empleado dos formas de calcular la puntuación total en la escala: 1) invertir los items negativos y sumar a continuación las puntuaciones de todos los items (MT2) (Warr, 1992); y 2) sumar los items positivos y los negativos por separado, y calcular la diferencia entre la suma de los items positivos y la suma de los items negativos (MT1) (Bradburn, 1969). Como se puede deducir de la comparación de las puntuaciones de la tabla 11.1, las puntuaciones obtenidas por el segundo procedimiento son una combinación lineal de las obtenidas por el primer procedimiento. Podemos calcular las puntuaciones MT2 simplemente sumando la constante 3 a las puntuaciones obtenidas en MT1. En la muestra empleada en este estudio, las correlaciones entre estas dos puntuaciones en las escalas tenso-calmado, deprimido-entusiasmado y fatiga-vigor correlacionaron entre .998 y 1. En la tabla 11.1, estas puntuaciones son las designadas con MT1 (suma y diferencia), y MT2 (inversión y suma).

Bajo el modelo de escalamiento desplegado que considera la escala total como la combinación de dos escalas acumulativas, la puntuación final se calcula según el modo anteriormente expuesto. Esta puntuación se designa en la tabla 11.8 como EA.

Finalmente, las puntuaciones obtenidas en cada escala acumulativa por separado se ofrecen en la tabla 11.8 bajo las columnas N y P.

En la tabla 11.8 se observa la correpondencia que existe entre las puntuaciones bajo el modelo bipolar tradicional (MT 1 y2) y la calculada bajo el modelo acumulativo (EA). Como hemos indicado, entre las dos puntuaciones tradicionales la diferencia consiste en la suma de la constante 3. Pero aún resulta más sorprendente comprobar que entre éstas y la puntuación calculada bajo el modelo acumulativo la diferencia consiste en sumar la constante 6 o 3, según sea para MT1 o MT2 respectivamente. Es decir, mientras que era lógico esperar que las puntuaciones calculadas bajo el modelo tradicional fueran una combinación lineal la una de la otra, no era lógico esperar lo mismo entre puntuaciones calculadas bajo modelos de medida diferentes. Y sin embargo eso es precisamente lo que indican los datos de la tabla 11.8.

De acuerdo con Van Schuur y Kruijtboch (1994), llegados a este punto, cabe plantearse la siguiente hipótesis. Si las escalas bipolares empleadas habitualmente en la medida de dimensiones afectivas constituyeran en realidad dos escalas acumulativas del tipo descrito, entonces se explicaría:

1. por qué el uso del análisis factorial sobre estas escalas identifica factores monopolares en lugar de bipolares, y por lo tanto aporta evidencia empírica en favor de un modelo del afecto monopolar; y

2. por qué las puntuaciones obtenidas en estos factores monopolares presentan peores correlaciones con otros indicadores del afecto que la que ofrece la puntuación calculada bajo el modelo bipolar tradicional.

La discusión se centra en torno a la validez de constructo de estas escalas. Mientras la evaluación de la dimensionalidad de las mismas indica que miden factores monopolares, e invalida el modelo del afecto bipolar, la evaluación de las correlaciones que muestran estas escalas con otros indicadores de afecto ofrece la interpretación opuesta. Este es el problema que trata de resolver esta reconceptualización, que, de resultar confirmada por los datos, también explicaría la escasa relación con las escalas criterio que mostraron las puntuaciones calculadas bajo el modelo de desplegamiento probabilístico.

En consecuencia, el objetivo de los siguientes capítulos será contrastar empíricamente la hipótesis de que cada una de las escalas bipolares objeto de estudio constituye un par de escalas acumulativas del tipo descrito.

CAPITULO 12

METODO.

1. PROCEDIMIENTO DE RECOGIDA DE DATOS Y MUESTRA.

Esta parte del estudio ha sido realizada sobre la misma muestra que se empleó en la primera parte, en la que figura su composición y descripción.

2. MEDIDAS.

Las escalas empleadas en esta parte del estudio son las escalas desarrolladas por Warr (1990) para la medida de las dimensiones del bienestar afectivo tenso vs. calmado, deprimido vs. entusiasmado y fatigado vs. animado, que ya quedaron descritas en la primera parte del estudio. Recordaremos que estas escalas constan de seis ítems cada una, y su escala de respuesta, simétrica, oscila entre "1. nunca" y "6. siempre". Cada una de las escalas consta de tres de adjetivos típicos de uno de los polos del eje (positivo), y otros tres típicos del polo opuesto (negativo). También se incluyó un ítem en formato de diferencial semántico por cada una de las escalas a evaluar, contabilizando un total de tres ítems. Estos ítems presentaban en cada extremo el par de adjetivos opuestos que caracterizan cada una de las dimensiones medidas, entre los que se dispuso una línea continua dividida en 7 espacios, de entre los que el sujeto marcaba el que mejor definiera su estado de ánimo en el mismo referente temporal que en las escalas tipo likert.

3. HIPOTESIS.

Partiendo de la tesis de que las escalas estudiadas constituyen la combinación de dos escalas acumulativas que ofrecen datos interpretables bajo el modelo de medida de desplegamiento, se pueden formular una serie de hipótesis empíricamente contrastables, que confirmarán o no la tesis enunciada. Esto implica que:

1) cada escala mide un continuo bipolar, a partir del desplegamiento de dos escalas acumulativas, cada una de las cuales cubre una mitad de ese continuo, y

2) las puntuaciones de los sujetos en cada escala acumulativa pueden ser combinadas para obtener un indicador de la ubicación de los sujetos en ese continuo bipolar.

Las hipótesis a comprobar son las siguientes:

1) Para cada escala será posible encontrar una ordenación acumulativa de los ítems correspondientes en un continuo bipolar, de modo que las puntuaciones de los sujetos indiquen su posición relativa respecto de los estímulos que la componen. La contrastación de esta hipótesis se efectuará mediante análisis de desplegamiento de cada escala, bajo el supuesto de que cada una de ellas está formada por dos escalas acumulativas. Este análisis se llevará a cabo mediante el programa MUDFOLD (Multiple UniDimensional unFOLDding, Van Schuur y Post, 1990).

2) Las puntuaciones obtenidas bajo este modelo de medida correlacionarán con las variables criterio en la misma medida en que lo hacen las puntuaciones obtenidas bajo los modelos de medida tradicional o "dos mitades". Como ya hemos visto, las correlaciones que presentan las puntuaciones obtenidas bajo estos dos últimos modelos son tan altas como las ofrecidas por otra escala explícitamente bipolar. Ello indicará que el modelo de escalamiento acumulativo iguala la validez mostrada por los modelos anteriores.

4. ANALISIS.

El proceso de búsqueda y evaluación de las escalas bajo el modelo desplegado se ha llevado a cabo mediante el programa MUDFOLD. Este programa busca la representación del número máximo de estímulos entre el total de ellos (n), y del número máximo de sujetos (N) en una dimensión, usando para ello sólo los juicios de preferencia -o identificación- de cada sujeto. El método que emplea este programa requiere el uso de puntuaciones dicotomizadas: las respuestas ante los estímulos con los que se identifican los sujetos son codificadas como 1, y el resto se codifican como 0.

Como se indicó con anterioridad, se pueden distinguir dos aproximaciones básicas en la codificación de los datos:

1) emplear los n estímulos "preferidos" de cada sujeto, para lo que se indica al programa el valor o valores que definen un estímulo como preferido -en una escala tipo Likert de 6 puntos se emplearían las puntuaciones que expresan el mayor acuerdo, es decir, 5 y 6-; y

2) emplear los k estímulos preferidos de cada sujeto, siendo el investigador quien determina el número k de estímulos pertinentes para el análisis. Así, esta segunda aproximación no se basa en las preferencias absolutas expresadas por los sujetos. En lugar de ello, este procedimiento se basa en la ordenación que cada sujeto realiza de los estímulos, y codifica como 1 los k estímulos a los que el sujeto asigna la puntuación más alta, y como cero al resto. Los autores diferencian ambos procedimientos denominando "toma cualquiera/ n " al primer caso, y "toma k/n " al segundo. En este último caso, los sujetos presentan patrones de k unos y $n-k$ ceros, en cualquier orden.

Además de esta diferencia en la codificación de los datos, hay algunas particularidades más en los análisis que realizaremos a continuación que pasamos a detallar.

En primer lugar, es necesario determinar el número de estímulos preferidos, k , pertinentes para este análisis. Puesto que se trata de contrastar si las escalas conforman dos escalas acumulativas combinadas, es necesario fijar k al número de estímulos máximo que puede preferir un sujeto en una de esas escalas. Esto nos lleva a fijar $k=3$. De esta forma se contrasta la hipótesis de que ambas escalas forman una escala de despliegue paralelográfica, del tipo que presenta la matriz 12.1.

Matriz 12.1 Escala paralelográmica

| | | | estado de ánimo | | | | | |
|----------|---|---|-----------------|---|---|----------|---|---|
| | | | negativo | | | positivo | | |
| Sujetos | | | Items | | | | | |
| | | | 1 | 2 | 3 | 4 | 5 | 6 |
| | | | - | | | | | + |
| positivo | A | + | 0 | 0 | 0 | 1 | 1 | 1 |
| | B | | 0 | 0 | 1 | 1 | 1 | 0 |
| | C | | 0 | 1 | 1 | 1 | 0 | 0 |
| negativo | D | - | 1 | 1 | 1 | 0 | 0 | 0 |

Una vez obtenidas las puntuaciones dicotomizadas bajo la condición $k=3$, el programa analiza todas las combinaciones de tres estímulos que se pueden formar, en la forma indicada en el capítulo anterior, y ofrece la escala final junto con algunos indicadores de ajuste tanto para la escala como para cada uno de los estímulos incluidos en ella, así como las puntuaciones correspondientes de los sujetos y los estímulos.

No obstante, la asignación de puntuaciones en la escala que realiza el programa tanto para los estímulos como para los sujetos no se corresponde con el modelo de desplegamiento paralelográmico que se hipotetiza. Como se recordará, el programa asigna a cada estímulo un valor igual al del orden que ocupa, en tanto que la puntuación de los sujetos se obtiene de sumar y promediar el valor de los estímulos que han elegido como preferidos.

En el caso que nos ocupa, las puntuaciones de los sujetos serán calculadas de forma independiente a las que ofrece el programa, en función del número de estímulos incluidos en la escala, su ordenación, y el patrón de ceros y unos que presente cada sujeto. Todo ello bajo el modelo de límites de intervalos expuesto con anterioridad, ilustrado en la matriz 12.2.

Matriz 12.2 Asignación de puntuaciones en la escala paralelográmica por el procedimiento de recuento de intervalos superados.

| | | estado de ánimo | | | | | | | |
|----------|---|-----------------|---|---|----------|---|---|-------|---|
| | | negativo | | | positivo | | | | |
| | | Items | | | | | | | |
| | | 1 | 2 | 3 | 4 | 5 | 6 | total | |
| Sujetos | | - | | | | | + | | |
| positivo | A | + | 2 | 2 | 2 | 1 | 1 | 1 | 9 |
| | B | | 2 | 2 | 1 | 1 | 1 | 0 | 7 |
| | C | | 2 | 1 | 1 | 1 | 0 | 0 | 5 |
| negativo | D | - | 1 | 1 | 1 | 0 | 0 | 0 | 3 |

Nota: total= puntuacion total en la escala paralelográmica formada por las dos escalas acumulativas combinadas como muestra la matriz

CAPITULO 13

RESULTADOS Y DISCUSION.

1. ESCALAMIENTO BAJO EL MODELO DE DESPLEGAMIENTO DE DOS ESCALAS ACUMULATIVAS.

Este primer conjunto de análisis se ha realizado bajo el supuesto de que las escalas bajo estudio se ajustan al modelo de desplegamiento paralelográfico. En consecuencia, se ha empleado la modalidad "toma k/n ", siendo $k=3$. Bajo esta modalidad, el programa dicotomiza las respuestas de cada sujeto codificando como 1 las respuestas a los tres ítems frente a los que el sujeto ha respondido con las puntuaciones más elevadas. En esta modalidad cada estímulo es juzgado en relación con el resto de estímulos.

En esta modalidad de escalamiento desplegado los sujetos con patrones de respuesta no falsables son aquéllos que han preferido menos de tres estímulos, o bien más de tres estímulos. En el caso en que un sujeto designe como preferidos más de tres estímulos, no se puede determinar cuál puede ser eliminado, por lo que se elimina el sujeto.

El resto de pasos necesarios para la realización de los análisis, identificación de la mejor escala J inicial, extensión de la escala J inicial y evaluación del ajuste de la escala J final, son análogos a los enunciados en el capítulo referido al desplegamiento probabilístico, por lo que no los repetiremos aquí.

A continuación se presentan los resultados obtenidos para cada una de las tres escalas bajo este nuevo modelo.

2. ESCALA TENSO-CALMADO.

La descripción de la distribución de puntuaciones dicotomizadas se encuentra en las tablas 13.1 y 13.2. La primera de ellas muestra en cada celdilla el número de sujetos que eligen el estímulo de la columna entre los tres estímulos que eligen en total. Así, la fila 3 contiene los sujetos que han entrado en el análisis, que como se recordará, son los sujetos para los que es posible identificar los tres estímulos más preferidos, y cada celdilla de esta fila indica el número de sujetos que han preferido

ese estímulo columna. Se observa que más de la mitad de los sujetos eligen los estímulos positivos.

Tabla 13.1. Distribución del número de estímulos elegidos

| N | A | B | C | D | E | F | TOT |
|---|----|----|----|-----|-----|-----|-----|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 3 | 23 | 18 | 29 | 157 | 162 | 148 | 179 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | |

Nota: N=número de estímulos preferidos por los sujetos contabilizados en esa fila; Estímulos: A=tenso, B=incómodo, C=preocupado, D=tranquilo, E=contento, F=relajado; TOT= número de individuos total en esa fila.

En la tabla 13.2 se observa en primer lugar la frecuencia absoluta con que se ha elegido cada estímulo, y entre paréntesis la frecuencia relativa correspondiente.

Tabla 13.2. Frecuencias absolutas y relativas

| | |
|--------------------------|----------------|
| NUMBER OF RESPONDENTS | 179 |
| NUMBER OF STIMULI | 6 |
| NUMBER OF PICKED STIMULI | 3 |
| A) TENSO - MARGINAL | 23 (12.85 %) |
| B) INCOMODO - MARGINAL | 18 (10.06 %) |
| C) PREOCUPADO - MARGINAL | 29 (16.20 %) |
| D) TRANQUILO - MARGINAL | 157 (87.71 %) |
| E) CONTENTO - MARGINAL | 162 (90.50 %) |
| F) RELAJADO - MARGINAL | 148 (82.68 %) |

En la tabla 13.3 se muestra la terna que cumple los requisitos para ser considerada "mejor terna", y consecuentemente ser el punto de partida de la escala desplegada. La permutación AED es tomada como punto inicial para la construcción de la escala desplegada. El orden

adecuado ADE o EDA se determinará, como vimos, en el proceso de extensión de la escala inicial.

Tabla 13.3. Terna y permutación inicial

| | DAE=EAD | | | ADE=EDA | | | AED=DEA | | |
|-----|---------|-------|--------|---------|-------|--------|---------|------|--------|
| | H | T | S(jik) | H | T | S(ijk) | H | T | S(ikj) |
| ADE | -0.12 | -2.77 | 0.04 | -0.20 | -0.57 | 0.36 | 1.00 | 2.43 | 0.41 |

Nota: H=coeficiente de escalabilidad; T=valor del estadístico T; S=error típico de H.

La tabla 13.4 muestra la descripción de la escala desplegada final. Junto a la proporción de sujetos que eligen cada estímulo (P(I)) y el coeficiente de escalabilidad correspondiente (H(I)) aparecen el número de errores observados $O_{(i,j,k)}$, el número de errores esperados $EO_{(i,j,k)}$, el error típico asociado al coeficiente de escalabilidad de cada estímulo (SD(H)) y el valor T asociado a cada coeficiente de escalabilidad (T(H)).

Tabla 13.4 Escala desplegada final

| FINAL SCALE 1 | | | | | | |
|---|------|------|---------------|----------------|-------|------|
| STIMULI ABOVE DASHED PAGE FORM AN UNFOLDING SCALE | | | | | | |
| | P(I) | H(I) | $O_{(i,j,k)}$ | $EO_{(i,j,k)}$ | SD(H) | T(H) |
| A) TENSO | 0.13 | 0.75 | 15 | 60.1 | 12.47 | 3.61 |
| B) PREOCUPAD | 0.16 | 0.75 | 16 | 64.8 | 9.74 | 5.00 |
| F) CONTENT | 0.91 | 0.77 | 14 | 60.8 | 8.63 | 5.41 |
| D) TRANQUILO | 0.88 | 0.81 | 12 | 62.6 | 8.90 | 5.68 |
| E) RELAJADO | 0.83 | 0.80 | 12 | 59.2 | 10.30 | 4.57 |
| ----- | | | | | | |
| C) INCOMODO | 0.10 | | | | | |

Nota: P(I)=proporción de sujetos que han elegido el estímulo de la fila; H(I)=coeficiente de escalabilidad del estímulo de la fila; $O_{(i,j,k)}$ =error observado asociado al estímulo fila; $EO_{(i,j,k)}$ =error esperado bajo el modelo nulo asociado al estímulo fila; S(H)=error típico del coeficiente de escalabilidad del estímulo de la fila; T(H)=valor del estadístico T asociado al coeficiente de escalabilidad (H) del estímulo de la fila.

A diferencia de lo que sucedía bajo el modelo probabilístico, en este tipo de análisis el número de sujetos sobre los que se calcula el

error esperado y observado coincide, de modo que se elude el problema de la sobreestimación que afectaba a aquél. Por otra parte, se observa que en este caso el número de errores observado y esperado es bastante homogéneo para los items, independientemente de si son negativos o positivos. Señalar que el ítem "incómodo" ha sido descartado de la escala, lo que indica que no ha sido posible encontrar una relación de orden entre las respuestas de los sujetos a este estímulo y al resto. Finalmente, todos los coeficientes de escalabilidad de los estímulos incluidos en la escala final resultan significativamente distintos de cero ($p < 0.01$, dos colas) y elevados, así como el coeficiente de escalabilidad de la escala total ($H = .78$, $T = 5.21$).

En conclusión, los resultados obtenidos muestran que los items tenso, preocupado, contento, relajado y tranquilo pueden formar una escala desplegada. No obstante, desde el punto de vista de las dos escalas acumulativas propuestas, el hecho de eliminar un adjetivo de la escala desplegada conlleva serias repercusiones.

Matriz 13.5 Resultado de eliminar un ítem de la subescala negativa.

| | | | estado de ánimo | | | | | |
|----------|---|---|-----------------|---|---|----------|---|---|
| | | | negativo | | | positivo | | |
| Sujetos | | | Items | | | | | |
| | | | - | 2 | 3 | 4 | 5 | 6 |
| | | | - | | | | + | |
| positivo | A | + | - | 0 | 0 | 1 | 1 | 1 |
| | B | | - | 0 | 1 | 1 | 1 | 0 |
| | C | | - | 1 | 1 | 1 | 0 | 0 |
| negativo | D | - | - | 1 | 1 | 0 | 0 | 0 |

De hecho, la desaparición de un ítem conlleva que las escalas ya no se ajustan al modelo de desplegamiento paralelográfico, considerando el razonamiento seguido hasta aquí. En este sentido, el mismo patrón de respuesta para la subescala negativa caracteriza a sujetos con posiciones diferentes en la dimensión latente, como se aprecia en la matriz 13.1. En

esta matriz se observa que los sujetos con patrones de respuesta C y D tienen el mismo patrón de respuesta en la subescala negativa, pero no en la subescala positiva, y su ubicación en la dimensión latente tampoco es la misma.

3. ESCALA DEPRIMIDO-ENTUSIASMADO.

La descripción de la distribución de puntuaciones dicotomizadas se encuentra en las tablas 13.5 y 13.6. La primera de ellas muestra en cada celdilla el número de sujetos que eligen el estímulo de la columna entre los tres estímulos que eligen en total. Así, la fila 3 contiene los sujetos que se han utilizado en el análisis, que como se recordará, son los sujetos para los que es posible identificar los tres estímulos más preferidos, y cada celdilla de esta fila indica el número de sujetos que han preferido ese estímulo columna. Se observa que la mayoría de los sujetos eligen los estímulos positivos, mientras que sólo un 15% aproximadamente elige los estímulos negativos.

Tabla 13.6. Distribución número de estímulos elegidos

| N | A | B | C | D | E | F | TOT |
|---|----|----|----|-----|-----|-----|-----|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 3 | 32 | 37 | 13 | 172 | 137 | 161 | 184 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | |

Nota: N=número de estímulos preferidos por los sujetos contabilizados en esa fila; Estímulos: A=deprimido, B=pesimista, C=desgraciado, D=alegre, E=entusiasmado, F=optimista; N=número de sujetos contabilizados en esa fila.

En la tabla 13.6 se observa en primer lugar la frecuencia absoluta con que se ha elegido cada estímulo, y entre paréntesis la frecuencia relativa correspondiente.

Tabla 13.7. Frecuencias absolutas y relativas

| | | |
|--------------------------|------------|----------------|
| NUMBER OF RESPONDENTS | | 184 |
| NUMBER OF STIMULI | | 6 |
| NUMBER OF PICKED STIMULI | | 3 |
| A) DEPRIMIDO | - MARGINAL | 32 (17.39 %) |
| B) PESIMISTA | - MARGINAL | 37 (20.11 %) |
| C) DESGRACIADO | - MARGINAL | 13 (7.07 %) |
| D) ALEGRE | - MARGINAL | 172 (93.48 %) |
| E) ENTUSIASMADO | - MARGINAL | 137 (74.46 %) |
| F) OPTIMISTA | - MARGINAL | 161 (87.50 %) |

La tabla 13.8 se muestra la terna que cumple los requisitos para ser considerada "mejor terna", y consecuentemente ser el punto de partida de la escala desplegada. La permutación ABC=CBA es tomada como punto inicial para la construcción de la escala desplegada.

Tabla 13.8 Terna y permutación inicial

| | BAC=CAB | | | ABC=CBA | | | ACB=BCA | | |
|-----|---------|-------|--------|---------|------|--------|---------|-------|--------|
| | H | T | S(jik) | H | T | S(ijk) | H | T | S(ikj) |
| ABC | -0.02 | -0.02 | 1.01 | 1.00 | 0.92 | 1.09 | -3.77 | -6.02 | 0.63 |

Nota: H=coeficiente de escalabilidad; T=valor del estadístico T; S=error típico de H.

La tabla 13.9 muestra la descripción de la escala desplegada final. En este caso, la escala final incluye todos los estímulos analizados. Además, todos los coeficientes resultan significativamente distintos de cero ($p < 0.01$, dos colas) y elevados, así como el coeficiente de escalabilidad de la escala total ($H=0.76$, $T=6.29$).

En conclusión, los resultados obtenidos indican que los estímulos desgraciado, pesimista, deprimido, alegre, optimista y entusiasmado forman dos escalas acumulativas que se ajustan al modelo de desplegamiento paralelográfico.

Tabla 13.9 Escala desplegada final

| FINAL SCALE 1 | | | | | | |
|---|------|------|---------------|----------------|-------|------|
| STIMULI ABOVE DASHED PAGE FORM AN UNFOLDING SCALE | | | | | | |
| | P(I) | H(I) | $O_{(i,j,k)}$ | $EO_{(i,j,k)}$ | SD(H) | T(H) |
| B) DESGRACIADO | 0.07 | 0.80 | 11 | 54.4 | 15.14 | 2.86 |
| A) PESIMISTA | 0.20 | 0.70 | 34 | 112.3 | 14.81 | 5.28 |
| C) DEPRIMIDO | 0.17 | 0.67 | 36 | 109.8 | 12.05 | 6.11 |
| F) ALEGRE | 0.93 | 0.75 | 23 | 93.6 | 10.98 | 6.42 |
| D) OPTIMISTA | 0.87 | 0.76 | 23 | 95.1 | 11.17 | 6.44 |
| E) ENTUSIASMADO | 0.74 | 0.93 | 5 | 75.5 | 13.41 | 5.25 |

Nota: P(I)=proporción de sujetos que han elegido el estímulo de la fila; H(I)=coeficiente de escalabilidad del estímulo de la fila; $O_{(i,j,k)}$ =error observado asociado al estímulo fila; $EO_{(i,j,k)}$ =error esperado bajo el modelo nulo asociado al estímulo fila; S(H)=error típico del coeficiente de escalabilidad del estímulo de la fila; T(H)=valor del estadístico T asociado al coeficiente de escalabilidad (H) del estímulo de la fila.

4. ESCALA FATIGA-VIGOR.

Las tablas 13.10 y 13.11 muestran la distribución de las respuestas dicotomizadas. Al igual que en las escalas anteriores, los sujetos tienden a elegir con mayor frecuencia los estímulos positivos que los negativos.

Tabla 13.10 Distribución número de estímulos elegidos

| N | A | B | C | D | E | F | TOT |
|---|----|----|----|-----|-----|-----|-----|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 3 | 37 | 18 | 32 | 159 | 139 | 149 | 178 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | |

Nota: N=número de estímulos preferidos por los sujetos contabilizados en esa fila; Estímulos: A=fatigado, B=sin energía, C=cansado, D=despierto, E=lleño de energía, F=animado N=número de sujetos contabilizados en esa fila;

En la tabla 13. 11 se observa el número de sujetos retenidos para el análisis, la frecuencia absoluta con que se ha elegido cada estímulo, y entre paréntesis la frecuencia relativa correspondiente.

Tabla 13.11 Frecuencias absolutas y relativas

| | | | |
|--------------------------|------------|-----|------------|
| NUMBER OF RESPONDENTS | | 178 | |
| NUMBER OF STIMULI | | 6 | |
| NUMBER OF PICKED STIMULI | | 3 | |
| A) FATIGADO | - MARGINAL | 37 | (20.79 %) |
| B) SIN ENERGIA | - MARGINAL | 18 | (10.11 %) |
| C) CANSADO | - MARGINAL | 32 | (17.98 %) |
| D) DESPIERTO | - MARGINAL | 159 | (89.33 %) |
| E) LLENO ENERGIA | - MARGINAL | 139 | (78.09 %) |
| F) ANIMADO | - MARGINAL | 149 | (83.71 %) |

A diferencia de lo que ocurría en las escalas anteriores, en este caso no ha sido posible encontrar una terna la en que se diera la condición necesaria para iniciar el proceso de desplegamiento de los estímulos. Ninguna de las ternas posibles ha dado lugar a una permutación con un coeficiente de escalabilidad positivo, y las dos restantes con coeficientes de escalabilidad negativos. En otros términos, esto indica que no existe ninguna combinación de tres estímulos que pueda ser ordenada de una única forma, que es el indicador inicial de la escalabilidad de un conjunto dado de estímulos. En conclusión, los estímulos que forman esta escala no se ajustan al modelo de desplegamiento paralelográfico, tal como se hipotetizaba inicialmente.

5. ANALISIS CORRELACIONAL.

La segunda de las hipótesis que se han formulado acerca del modelo de desplegamiento acumulativo es que las puntuaciones obtenidas bajo este modelo de medida presentarán correlaciones iguales o mayores con los items de diferencial semántico correspondientes que las obtenidas bajo los modelos de medida anteriormente explorados.

A este respecto sólo es posible obtener y consecuentemente comparar las puntuaciones que este modelo ofrece para la escala

deprimido-entusiasmado. Esta es la única escala que se ajusta al modelo de desplegamiento bajo el supuesto de que es una combinación de dos escalas acumulativas.

Como ya se indicó anteriormente, las puntuaciones obtenidas directamente por el programa no se ajustan al modelo de medida hipotetizado, por lo que fue necesario acudir a los patrones de respuesta codificados como ceros y unos, identificar cada uno de ellos, asignarles el valor correspondiente al número de límites superados -en los términos expuestos en el capítulo 11- y dar a los sujetos las puntuaciones adecuadas en función de ese patrón. En concreto, los patrones que se detectaron en esta escala, así como su frecuencia aparecen representados en la tabla 13.12.

Tabla 13.12 Tipos de patrones de respuesta, y frecuencia observada.

| PATRONES AJUSTADOS | | | | | | | F | P |
|-----------------------|---|---|---|---|---|---|-----|---|
| 1 | 0 | 0 | 0 | 1 | 1 | 1 | 135 | 9 |
| 2 | 0 | 0 | 1 | 1 | 1 | 0 | 9 | 7 |
| 3 | 0 | 1 | 1 | 1 | 0 | 0 | 11 | 5 |
| 4 | 1 | 1 | 1 | 0 | 0 | 0 | 10 | 3 |
| TOTAL | | | | | | | 165 | |
| PATRONES NO AJUSTADOS | | | | | | | F | P |
| 5 | 0 | 1 | 0 | 1 | 1 | 0 | 14 | 7 |
| 6 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 7 |
| 7 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 9 |
| 8 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 5 |
| 9 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 9 |
| 10 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 7 |
| TOTAL | | | | | | | 19 | |

Nota: Patrón=patrón de respuesta; F=frecuencia observada; P=puntuación asignada al patrón de la fila.

En la tabla anterior se observa la distribución de los patrones de respuesta observados en esta escala. En total se contabilizaron 165 sujetos con patrones perfectamente ajustados al modelo de desplegamiento acumulativo, y 19 sujetos que se separaban en mayor o

menor medida de lo esperado bajo este modelo. Las frecuencias de cada uno de los patrones se muestran bajo la columna F -en negrita los que se ajustan al modelo, y en cursiva los que no-, y las puntuaciones bajo el modelo de escalas acumulativas se muestran bajo la columna P.

En la tabla 13.13 aparecen las correlaciones correspondientes a la puntuación total en esta escala calculadas bajo cada uno de los modelos analizados. Como ya se mencionó anteriormente, no todos los sujetos que componen la muestra reciben puntuaciones bajo el modelo de despliegamiento acumulado. Aquellos sujetos con patrones de respuesta con un número de estímulos preferidos distinto de 3 no pueden ser puntuados, de modo que la tabla siguiente se ha calculado sólo sobre la submuestra de sujetos que recibe puntuaciones bajo este modelo.

La inspección de la tabla revela que las correlaciones de las puntuaciones calculadas bajo todos los modelos estudiados con el criterio son estadísticamente significativas ($p < 0.01$, dos colas). La correlación entre las puntuaciones calculadas bajo el modelo de dos mitades con el criterio y la correlación entre las puntuaciones calculadas bajo el modelo de despliegamiento paralelográfico no son significativamente distintas ($t = 0.367$, $p > 0.05$, dos colas). Sin embargo, la correlación de las puntuaciones calculadas bajo el modelo tradicional con el criterio sí es significativamente distinta de las obtenidas para los modelos de dos mitades ($t = 2.52$, $p < 0.05$, dos colas) y paralelográfico ($t = 2.50$, $p < 0.05$, dos colas). Es decir, las puntuaciones finales obtenidas bajo el modelo de despliegamiento acumulativo o paralelográfico pronostican la ubicación de los sujetos en el continuo bipolar bajo estudio en la misma medida que las obtenidas bajo el de "dos mitades", pero ambos modelos presentan correlaciones significativamente menores que las presentadas por el modelo tradicional. Si junto a ello consideramos que estas correlaciones se han obtenido únicamente sobre el conjunto de sujetos para los que sí funcionaba el modelo de escalamiento acumulativo, la conclusión acerca de la utilidad de este modelo resulta aún más desfavorable.

Por su parte, la correlación entre las puntuaciones calculadas bajo el modelo de escalas acumulativas y las puntuaciones calculadas bajo el

modelo de dos mitades es alta (.878), y significativamente distinta ($t=2.13$, $p<0.05$, dos colas) de la que presenta con las puntuaciones calculadas bajo el modelo tradicional (.836).

Tabla 13.13 Análisis correlacional. Escala deprimido-entusiasmado (N=170)

| | ds | mt | dm | dp | da1 |
|-----|------|------|------|------|-----|
| ds | 1 | | | | |
| mt | .727 | 1 | | | |
| dm | .653 | .852 | 1 | | |
| dp | .412 | .542 | .372 | 1 | |
| da1 | .657 | .836 | .878 | .318 | 1 |

Nota: ds=puntuación en la escala criterio; mt=puntuación bajo el modelo tradicional; dm=puntuación bajo el modelo de "dos mitades"; dp=puntuación bajo el modelo de despliegamiento probabilístico; da1=puntuación bajo el modelo acumulativo. $r>0.194$, $p<0.05$; $r>0.253$, $p<0.01$.

Finalmente, cabe señalar la escasa relación que existe entre las puntuaciones de los sujetos bajo el modelo desplegado acumulativo y el probabilístico (0.35). Esto puede ser debido a que los sujetos que han tomado parte en un análisis no lo han hecho en el otro. Sólo el subconjunto de sujetos que respondieron con puntuaciones de 5 o 6 al menos a un estímulo recibieron puntuaciones bajo el modelo de despliegamiento probabilístico. Esto redujo la muestra a 115 sujetos. Los 180 sujetos restantes recibieron la puntuación constante -99, que es la puntuación que este programa asigna a los sujetos con "dato faltante o missing". Este programa presenta la particularidad de no admitir sujetos con datos faltantes en el conjunto de datos que constituyen el input para los análisis, y cuando los genera, ha de asignarles algún valor diferente de los tradicionales puntos o espacios en blanco. Por otra parte, sólo los sujetos con tres estímulos preferidos sobre los demás han formado parte del análisis bajo el modelo de escalas acumulativas, y por lo tanto tienen puntuaciones bajo este modelo. Esto reduce la muestra a 170 sujetos. Pero entre los 170 sujetos con puntuaciones bajo este último modelo, un número sin determinar puede presentar puntuaciones de -99 bajo el

modelo de despliegamiento probabilístico. En consecuencia, es posible que esta escasa relación se deba a que los sujetos considerados en ambos análisis 1) presenten puntuaciones dispares o 2) no hayan sido puntuados bajo el modelo probabilístico, por lo que tiene la puntuación constante de -99.

5. CONCLUSIONES.

El objetivo de este capítulo era determinar si las escalas bipolares estudiadas podían, a diferencia de lo generalmente establecido, ser interpretadas como una combinación de dos escalas acumulativas tal que pudiera desplegarse de acuerdo con las características del análisis de paralelogramos (Van Schuur y Kuitjbosh, 1994). De ser así, se podría atribuir la identificación de factores monopoles en escalas de estas características a la inadecuación del modelo de escalamiento empleado, y del método de análisis utilizado para su validación, en lugar de atribuir estos resultados a la falta de validez del modelo bipolar del afecto.

En consecuencia, la primera tarea que se abordó fue determinar si los items de estas escalas podían representarse en un continuo bipolar, de modo que las puntuaciones de los sujetos indicaran su posición relativa respecto de ellos. Los resultados obtenidos a este respecto indican que no siempre es posible interpretar las respuestas dadas por un subconjunto de sujetos de la muestra en base a una escala desplegada acumulativa. En concreto, sólo la escala correspondiente a la dimensión deprimido-entusiasmado pudo ser adecuadamente representada de esta forma.

La segunda tarea que se abordó fue determinar si las puntuaciones obtenidas bajo este modelo pronosticaban adecuadamente la ubicación de los sujetos en un continuo explícitamente bipolar. Puesto que sólo una escala resultó adecuadamente representada por este modelo, la correspondiente a la dimensión deprimido-entusiasmado, la comparación de las puntuaciones obtenidas bajo este modelo con las obtenidas bajo los modelos restantes sólo pudo realizarse para una de

las tres escalas estudiadas. El resultado del análisis correlacional reveló que la ubicación que obtienen los sujetos en la dimensión latente bajo este modelo no predice mejor que los otros modelos la ubicación de los sujetos en un continuo explícitamente bipolar. Este resultado es especialmente adverso para la utilidad del modelo de escalas acumulativas, ya que se ha obtenido sobre el subconjunto de sujetos máximamente representables bajo este modelo. Como se recordará, los análisis a que nos referimos sólo se efectuaron a partir de los sujetos con tres estímulos preferidos, que eran los únicos representables en una escala paralelográfica de las características requeridas aquí. En consecuencia, la explicación del desdoblamiento de los factores subyacentes a las escalas tradicionales en términos de este modelo de medida, aunque interesante y plausible, no resulta respaldada empíricamente por nuestros resultados.

Volviendo hacia atrás, el origen del replanteamiento de estas escalas como escalas acumulativas desplegadas fue la constatación de que las puntuaciones que este modelo asigna a los sujetos es una combinación lineal de las que les asigna el modelo tradicional. Tanto un modelo como el otro ubicarían a los sujetos en la dimensión latente en el mismo orden.

Sin embargo, esta relación que se pensó podría explicar la incongruencia entre la dimensionalidad y la validez de las escalas no es privativa de este tipo de escalamiento. Si consideramos las puntuaciones finales calculadas bajo cada uno de los modelos de medida estudiados, resulta que esta característica es compartida por todas ellas.

En la tabla 13.14 aparecen las puntuaciones finales calculadas bajo estos supuestos y dicotomizadas: MT1 corresponde a la puntuación calculada bajo el modelo tradicional, modalidad diferencia positivo-negativo; MT2 corresponde al modelo tradicional en su modalidad de inversión y suma; DM corresponde al modelo de "dos mitades"; EA corresponde al modelo de escalamiento acumulativo, y EP al modelo de escalamiento desplegado probabilístico. Si tomamos como punto de referencia la puntuación MT1, la puntuación MT2 se obtiene tras la adición de la constante +3, la puntuación DM coincide prácticamente sin

necesidad de transformación, la puntuación EA requiere la adición de la constante +6, y finalmente la puntuación EP requiere la transformación $x+n+1$, donde x es la puntuación en la escala MT1, y n es el número de orden del patrón que se está transformando. Todas estas transformaciones son válidas únicamente si empleamos puntuaciones dicotómicas o dicotomizadas. Queda por determinar la medida en que este razonamiento también es aplicable a puntuaciones directas o no dicotomizadas.

Tabla 13.14. Comparación de puntuaciones.

| sub. negativa | | | N | sub. positiva | | | P | MT1 | MT2 | DM | EA | EP |
|---------------|---|---|---|---------------|---|---|---|-----|-----|----|----|----|
| 2 | 2 | 2 | 0 | 1 | 1 | 1 | 3 | 3 | 6 | 3 | 9 | 5 |
| 2 | 2 | 1 | 1 | 1 | 1 | 0 | 2 | 1 | 4 | 2 | 7 | 4 |
| 2 | 1 | 1 | 2 | 1 | 0 | 0 | 1 | -1 | 2 | -2 | 5 | 3 |
| 1 | 1 | 1 | 3 | 0 | 0 | 0 | 0 | -3 | 0 | -3 | 3 | 2 |

Nota: N= puntuaciones en la subescala negativa por el procedimiento de recuento de intervalos; P=puntuaciones en la subescala positiva por el procedimiento de recuento de intervalos.

En otro orden de cosas, queda realizar la comparación de los resultados aquí obtenidos respecto de los que obtienen los autores que proponen este modelo (Van Schuur y Kuitjboch, 1994). Estos autores emplearon este modelo para analizar los datos ofrecidos por la escala de Equilibrio Emotivo (Bradburn, 1969) sobre un conjunto de 10 muestras de distintos países de Europa. De los 12.463 sujetos que integraron inicialmente las 10 muestras, los análisis de desplegamiento se relizaron únicamente sobre 1.221 sujetos. Los resultados de estos análisis ofrecieron como promedio coeficientes de escalabilidad para la escala final en torno a .45 ($p<0.01$, dos colas). Las correlaciones que obtuvieron entre la puntuación final calculada por el procedimiento de la diferencia entre el subconjunto de items positivos y negativos, de un lado, y la obtenida bajo su modelo mostró una media a través de las 10 muestras de 0.65 ($p<0.01$, dos colas). La correlación que obtuvieron entre la puntuación tradicional y dos criterios -un medida de felicidad y otra de satisfacción- fue como promedio de 0.41 y 0.43 respectivamente, mientras que las correlaciones que mostró la puntuación calculada bajo su modelo con los mismos criterios presentó un promedio de 0.37 y 0.40

respectivamente. Estos resultados condujeron a los autores a la conclusión de que la escala de Equilibrio estudiada podía considerarse *"para todos los efectos prácticos, como una escala de despliegue en la que las dos mitades forman, cada una, una escala acumulativa. Esta explicación justifica el uso de las puntuaciones de la escala como medición válida de la percepción de calidad de vida, y explica por qué, según los resultados del análisis factorial, las puntuaciones de los sujetos para emociones positivas y negativas no están relacionadas"* (Van Schuur y Kruijtbosch, 1994, p. 567).

Realizando una comparación entre los resultados obtenidos por Van Schuur y Kruijtbosch (1994) y los aquí obtenidos pueden señalarse varias diferencias. En primer lugar, estos autores comparan únicamente el funcionamiento de dos modelos: el modelo tradicional y el acumulativo propuesto por ellos. Sus resultados indican que ambos modelos funcionan de forma similar. Sin embargo, su modelo es el que presenta mayor coherencia interna, frente a las inconsistencias que plantea el modelo tradicional. En el presente trabajo se realiza la comparación de cuatro modelos, aunque aquí retendremos tres modelos: el modelo tradicional, el modelo de dos mitades y el modelo de desplegamiento acumulativo. En nuestros datos el modelo acumulativo muestra coeficientes de escalabilidad mayores -en torno a 0.75-. La correlación entre la puntuación obtenida por el modelo tradicional y la obtenida bajo este modelo es de 0.836. La correlación entre la puntuación final tradicional y el criterio es de 0.727, en tanto que la que presenta la puntuación bajo este modelo acumulativo es de 0.657. En principio, si no contáramos con más información, estos resultados deberían llevarnos a la misma conclusión que presentan Van Schuur y Kruijtbosch (1994). Sin embargo, la comparación de los resultados ofrecidos por este modelo con los ofrecidos por el modelo de dos mitades indica que:

1) las correlaciones que muestran las puntuaciones calculadas bajo ambos modelos con las escalas criterio son similares, pero

2) mientras que el modelo acumulativo funciona únicamente en una de las escalas, el modelo de dos mitades funciona en todas ellas;

3) frente al modelo acumulativo, bajo el modelo de dos mitades no se pierden sujetos -entre un tercio y la mitad-, ni tampoco estímulos, y

4) aunque ambos resuelven el problema de incoherencia que presenta el modelo tradicional, y resultan teóricamente consistentes,

5) el modelo de dos mitades es más sencillo conceptual y operacionalmente.

Ante estas consideraciones, parece preferible el modelo de las dos mitades al modelo de escalamiento acumulativo. Más aún, ponen en duda la afirmación hecha por Van Schuur y Kruijtbosch (1994) que hemos reproducido en el párrafo anterior: ¿qué hubiera sucedido si hubieran contrastado también el modelo de dos mitades? Nuestros resultados indican que en algunos casos el modelo acumulativo puede explicar las relaciones observadas en estas escalas, pero también indican que el modelo de dos mitades explica estas relaciones mejor en *todas* las escalas estudiadas.

PARTE IV

CONCLUSIONES.

Comenzamos este trabajo con el propósito de aclarar en la medida de lo posible la controversia existente acerca de la estructura del afecto. Una controversia que gira en torno a la polaridad de las dimensiones implicadas en la estructura del afecto, y que ha dado lugar a la elaboración de dos modelos teóricos alternativos: el que defiende que la estructura del afecto es bipolar (Russell, 1980), y el que defiende que esta estructura es monopolar (Watson y Tellegen, 1985).

Ambos modelos están muy relacionados, por cuanto que la desconfirmación de uno conduce a la confirmación del otro. Pero esta relación no es, por así decirlo, simétrica. La evidencia empírica necesaria para apoyar al modelo bipolar es de carácter mucho más restrictivo que la que necesita el modelo monopolar. Mientras que el modelo bipolar necesita apoyarse en factores bipolares, el modelo monopolar recibe apoyo tanto de factores monopolares como bipolares, estos últimos bajo el argumento de que son "conceptualmente monopolares pero descriptivamente bipolares" (Watson y Tellegen, 1985). ¿Cuál es la diferencia? Más adelante volveremos sobre este punto.

En cualquier caso, la debilidad del modelo bipolar ha "alimentado" la validez del modelo monopolar: la evidencia empírica muestra que cuando se esperan factores bipolares, en cambio aparecen factores monopolares, de modo que cada polo de la dimensión bipolar se constituye en un factor monopolar independiente. Sin embargo, la combinación de las puntuaciones en los ítems que saturan en los factores monopolares empíricamente obtenidos ofrecen relaciones mayores con otros indicadores de afecto que si estas relaciones se calculan para cada factor monopolar por separado.

Al margen del modelo monopolar, la revisión de la literatura acerca de la estructura del afecto muestra numerosos intentos de explicar las razones por las cuales escalas diseñadas bajo el modelo bipolar no miden las dimensiones bipolares pretendidas. Ciertamente existe un buen número de investigaciones con esta finalidad. Pero como se indica en la revisión mencionada, ninguna de las explicaciones exploradas hasta el momento ha logrado resolver esta cuestión.

Analizando las diferentes explicaciones propuestas en los estudios revisados en el primer capítulo, se observa que la mayor parte de ellas aducen diferentes debilidades metodológicas, como el sesgo de aquiescencia o la asimetría de la escala de respuesta. Las diferentes investigaciones han tenido pues un primer denominador común: no han cuestionado las premisas que permiten esperar que las escalas tipo Likert contrabalanceadas midan dimensiones bipolares. En concreto, estas premisas son:

A. si la dimensión subyacente es bipolar,

B. entonces aparecerán relaciones lineales inversas entre los items representativos de polos opuestos, que darán lugar a:

1) correlaciones elevadas y negativas entre items opuestos, y

2) factores bipolares, en los que los items representativos de polos opuestos presentarán saturaciones con signos opuestos.

Asimismo, estas investigaciones centradas en debilidades metodológicas también presentan un segundo denominador común: la insuficiencia de las explicaciones metodológicas mencionadas conduce a la invalidación del modelo bipolar del afecto.

La clave de esta argumentación está pues en la aparición de relaciones lineales e inversas entre subescalas opuestas, que son las que permiten la obtención de los esperados factores bipolares. Y, ¿qué es un factor bipolar?, ¿cómo se distingue de un factor monopolar? Los diferentes estudios revisados han esperado obtener factores definidos por items representativos de los polos opuestos de la dimensión bipolar hipotetizada, los unos con saturaciones positivas, y los otros con saturaciones negativas. Los factores con esta composición son denominados bipolares y constituyen la evidencia necesaria para validar la bipolaridad de la dimensión subyacente. Sólo si aparece esta configuración se valida la bipolaridad de la dimensión ¿Y a la inversa? ¿La aparición de esta configuración es posible sólo si la dimensión subyacente es bipolar? ¿La aparición de relaciones lineales e inversas es sólo posible si la dimensión subyacente es bipolar? No, ésa es exactamente la misma configuración que se espera cuando se mide una dimensión monopolar mediante una escala que contiene items

invertidos. En el caso de escalas monopolares, los items invertidos denotan la negación de la cualidad medida. Y se espera que las relaciones entre los items invertidos y los que no lo están sean lineales, negativas y elevadas. A su vez, se espera que la factorización de la escala muestre un factor común en el que los items presenten saturaciones positivas y negativas. De modo que la misma evidencia empírica puede apoyar la existencia de dimensiones bipolares como monopolares. Entonces ¿cuál es la diferencia? En el modelo monopolar, los extremos de la dimensión que se está midiendo son "ausencia de A" y "A", y la dimensión oscilaría desde la ausencia de A a la presencia de A. Puntuaciones bajas en los items indicativos de ausencia de A implicarían puntuaciones altas en los items indicativos de presencia de A. En el modelo bipolar, los extremos de la dimensión serían "A" y "B", y la dimensión oscilaría desde "muy A" a "muy B". Puntuaciones bajas en items representativos de presencia de A implicarían puntuaciones altas en items indicativos de presencia de B sólo si la escala de respuesta es *bipolar, sólo si ante items representativos de A o de B el sujeto se está ubicando en el mismo continuo bipolar que oscila desde "muy A" a "muy B"*. En consecuencia, el conjunto de premisas arriba detallado está incompleto. Todo ello sólo es posible si se asume implícitamente que la escala de respuesta es bipolar, y por lo tanto complementaria. Es la bipolaridad de la escala de respuesta, y consecuentemente la complementariedad, lo que permite hipotetizar relaciones lineales e inversas entre items de polos opuestos.

En consecuencia, la no aparición de la evidencia empírica esperada en favor de la bipolaridad de la dimensión medida puede deberse a que estos dos supuestos se incumplan. Sin embargo, ninguno de los estudios anteriormente mencionados ha considerado la existencia de estos supuestos.

A partir de aquí el problema del desajuste entre la estructura teórica hipotetizada bajo el modelo bipolar y la empíricamente obtenida cobra una nueva dimensión. Había que comprobar si los resultados contrarios a lo esperado se explicaban no en términos de la inadecuación del modelo teórico, sino en términos del desajuste del modelo de medida y del método de análisis empleado para la validación de ese modelo

teórico. El primer objetivo del presente trabajo se ha centrado en dilucidar esta cuestión. Y los resultados obtenidos han indicado que los sujetos no reconstruyen el continuo bipolar asumido a la hora de señalar una alternativa de respuesta. En otros términos, los supuestos de bipolaridad y complementariedad no han obtenido respaldo empírico. La obtención de factores monopolares no desconfirma la bipolaridad de las dimensiones objeto de estudio, sino la adecuación del modelo de medida y del método de análisis empleados para evaluarla.

En este punto, la validez del modelo bipolar del afecto quedaba en suspenso. Pero podía ser nuevamente puesta a prueba con los mismos datos, aunque bajo diferentes premisas. Con este fin se ha elaborado un modelo de medida alternativo (el modelo de las "dos mitades") basado también en la bipolaridad de las dimensiones medidas, pero que no implica los supuestos anteriormente identificados. Este modelo no asume la dependencia completa entre las puntuaciones obtenidas en ítems de polos opuestos, es decir, no requiere la presencia de relaciones inversas y lineales entre opuestos, aunque sí incluye este tipo de relación entre las varias posibles. Este modelo se basa en la idea de que los sujetos no reconstruyen implícitamente un continuo bipolar a la hora de responder a un ítem que únicamente es representativo de uno de los polos de esa dimensión, es decir, al responder a un ítem mediante una escala explícitamente monopolar. Bajo este modelo, las relaciones entre las respuestas a subescalas opuestas no son ni dependientes, ni independientes, son interdependientes. Sólo cuando un sujeto se identifica con el afecto descrito por una subescala se puede determinar que no le identifica el afecto descrito por la subescala opuesta. Por el contrario, cuando un sujeto no se identifica con el afecto que describe una subescala, no se puede determinar el grado en que se identificará con el afecto opuesto.

La evidencia empírica obtenida ha mostrado apoyo a esta interpretación de los datos. Las puntuaciones de los sujetos en las subescalas opuestas se relacionan según lo esperado bajo el supuesto de que cada subescala mide un polo de la dimensión bipolar, y la linealización de las puntuaciones propuesta predice tan

aproximadamente como es posible la ubicación de los sujetos en las escalas criterio empleadas. No obstante, hay otros acercamientos alternativos a este problema. Y aunque el modelo de "dos mitades" explica de forma satisfactoria el problema de medida abordado, era necesario determinar si estos acercamientos alternativos son mejores.

Estos acercamientos alternativos se diferencian claramente del de "dos mitades" tanto en su origen como en la definición del problema que proponen. En cuanto al origen, el modelo de "dos mitades" parte de un análisis detallado de las inferencias que realiza el investigador bajo el modelo bipolar tradicional, y de los supuestos en que se basan estas inferencias. Primero se han identificado las debilidades conceptuales que presenta el modelo bipolar tradicional. A continuación se ha estudiado el modo en que pueden ser soslayadas. Y finalmente se ha elaborado un modelo bipolar alternativo que no presenta estos problemas.

El origen de los acercamientos alternativos que presentamos en la tercera parte de este trabajo se basa en la constatación de que la aplicación del análisis factorial a datos que se ajustan a un modelo de desplegamiento producirá un factor "extra" (Coombs y Kao, 1964) y en la correspondencia entre las puntuaciones calculadas bajo el modelo bipolar tradicional y las ofrecidas por uno de estos modelos alternativos. Podría decirse que mientras el modelo de "dos mitades" surge de la reflexión teórica, los acercamientos centrados en el modelo de desplegamiento surgen de una reflexión técnica.

En cuanto a la definición del problema, el modelo de "dos mitades" cuestiona las relaciones entre escalas opuestas que se esperan bajo el modelo tradicional. Frente a este modelo, que sólo cuestiona una parte del modelo de medida tradicional -en definitiva, la consideración de que tanto los items positivos como negativos son réplicas unos de otros-, las nuevas explicaciones exploradas cuestionan el modelo de medida completo. Encuadradas entre las aproximaciones centradas en la respuesta, estas aproximaciones asumen que ningún ítem es réplica de ningún otro. Muy al contrario, proponen que cada ítem tiene su propia ubicación en la dimensión medida, en función del grado en que posee el



atributo medido. En concreto, parten de otro modelo de medida, el modelo de escalamiento desplegado, para explicar el desdoblamiento de los factores. De hecho, había evidencia empírica que indicaba que el desdoblamiento de dimensiones bipolares en factores monopolares podía deberse, precisamente, a la aplicación de métodos de análisis factorial sobre datos que se ajustaban, en cambio, a un modelo de desplegamiento.

En consecuencia, se ha analizado la plausibilidad de este nuevo enfoque, comenzando por el modelo de desplegamiento probabilístico. Y, efectivamente, este modelo ha mostrado un ajuste aceptable a las escalas estudiadas, aunque en algunos casos dejando algunos items por el camino, y en todos los casos, a más de la mitad de los sujetos que componen la muestra. A pesar de todos estos recortes, la capacidad de las puntuaciones obtenidas bajo este modelo para predecir la ubicación de los sujetos en un continuo explícitamente bipolar ha resultado inferior a la observada en el modelo de las "dos mitades".

Finalmente, se ha explorado una tercera alternativa, muy en relación con la anterior. Esta se basa en que las subescalas positiva y negativa de las escalas estudiadas pueden constituir, cada una, una escala acumulativa, y en conjunto, una escala paralelográfica desplegable. Esta hipótesis explica también la validez de criterio de las puntuaciones obtenidas en escalas como la de Equilibrio Emotivo de Bradburn (1969), que miden dos factores independientes de afecto positivo y negativo. A pesar de reducir la muestra a la mitad, los resultados obtenidos han indicado que sólo una de las tres escalas puede ajustarse a este modelo. En esa escala, y para el subconjunto de sujetos que pudieron ser puntuados bajo este modelo, las puntuaciones que ofrece este modelo predicen la ubicación de los sujetos en la escala criterio en la misma medida que el modelo de "dos mitades".

Tomados en conjunto, estos resultados ilustran con claridad lo que Popper (1991) indica cuando advierte que encontrar un modelo que se ajuste bien a un conjunto de datos no es prueba de que sea el único modelo adecuado, sino que puede haber más, y que es necesario continuar investigando a pesar de ello. Además, los resultados obtenidos

también ponen de manifiesto lo que se ha venido a denominar la "hipótesis de la desconfirmación" (Miller y Pollock, 1994). Trabajando bajo la hipótesis de la confirmación, cualquiera de estos modelos pudiera haber sido calificado de válido. En consecuencia, podría haberse detenido el proceso de búsqueda y contraste de hipótesis alternativas. En cambio, el procedimiento bajo la hipótesis de la desconfirmación es el contrario. No se busca confirmar la hipótesis de la que parte el investigador, sino desconfirmar las hipótesis alternativas.

En este caso, los resultados obtenidos en cuanto al ajuste de los datos a los sucesivos modelos analizados no permiten desconfirmar claramente ninguno de ellos, aunque sí permiten establecer diferencias. Si nos basamos en criterios prácticos, el modelo de "dos mitades" es más sencillo y presenta mejores resultados en la evaluación de su ajuste y en la evaluación de su relación con las escalas explícitamente bipolares que los modelos alternativos basados en el modelo de escalamiento desplegado.

A diferencia de los modelos desplegados, este modelo:

- 1) explica la variabilidad de las puntuaciones de todos los sujetos que componen la muestra;
- 2) en todos los ítems que componen cada una de las tres escalas;
- 3) propone una explicación sencilla del motivo por el que se desdoblan las escalas;
- 4) muestra un ajuste satisfactorio a las puntuaciones obtenidas en las tres escalas; y
- 5) reproduce la ubicación de los sujetos en las escalas criterio con la misma precisión con que lo hace otra escala explícitamente bipolar.

No obstante, el modelo de "dos mitades" también presenta algunas debilidades:

1. El proceso de transformación empleado para obtener la puntuación final puede ser cuestionado. En este sentido, las transformaciones empleadas sólo conservan la ordenación de las

puntuaciones en el continuo bipolar hipotetizado. Los sujetos con puntuaciones bajas en la subescala positiva, y con puntuaciones medias en la subescala negativa obtendrán puntuaciones menores que los sujetos con puntuaciones bajas en ambas subescalas, que a su vez presentarán puntuaciones menores que los sujetos con puntuaciones medias en la subescala positiva y bajas en la subescala negativa. Pero no hay evidencia de que conserven las distancias relativas entre las ubicaciones de los diferentes sujetos en la dimensión bipolar subyacente. No se puede asumir que esta transformación ofrece puntuaciones en una escala de intervalo. De hecho, algunas transformaciones asignan puntuaciones completamente arbitrarias excepto en el orden en que ubican a los sujetos. Así, los sujetos con puntuaciones medio bajas en la subescala positiva y medio altas en la subescala negativa reciben la puntuación -3, y los sujetos con la combinación inversa la puntuación 3. Por otra parte, los sujetos con puntuaciones muy bajas en la subescala positiva y bajas en la subescala negativa reciben la puntuación -2, y los sujetos con la combinación inversa la puntuación 2. Estos valores únicamente indican que los sujetos con puntuaciones de 3 tienen una ubicación más cercana al extremo positivo de la dimensión que los sujetos con puntuaciones de 2, pero igualmente se podían haber elegido las puntuaciones 1 y 2, en lugar de las puntuaciones 2 y 3. No existe una unidad de medida constante para todos los valores de la escala que pueda reflejar las distancias relativas de los sujetos en la dimensión subyacente.

2. La bondad de ajuste se ha evaluado de forma indirecta. No se han podido aplicar métodos de evaluación de ajuste directos puesto que este modelo predice diferentes tipos de relaciones entre las puntuaciones de las subescalas, en función de las puntuaciones de los sujetos en cada una de ellas.

3. Este modelo deja sin resolver una cuestión fundamental ¿por qué el modelo bipolar tradicional, aplicando una serie de premisas incorrectas, ofrece sin embargo puntuaciones que ubican adecuadamente a los sujetos en las escalas criterio? Van Schuur y Kruijtbosch (1994) explicaban esta paradoja argumentando que bajo ciertas condiciones (puntuaciones dicotómicas o dicotomizadas) las

puntuaciones que ofrecía el modelo tradicional bipolar eran una transformación lineal de las que ofrecía el modelo propuesto por ellos - el de las dos escalas acumulativas desplegadas-. Sin embargo ya se indicó que esto ocurría también con las puntuaciones ofrecidas por otros modelos. Y, en cualquier caso, no explica por qué esa transformación es posible.

Respecto a este tercer punto, Van Schuur y Kruijtbosch recurren a un proverbio holandés que dice "*se puede dar un golpe (cor)recto con un palo torcido...*" (Van Schuur y Kruijtbosch, 1994, p. 569). Parece que éso es lo que ha sucedido con las puntuaciones calculadas bajo el modelo bipolar tradicional. Pero ¿por qué? Una cosa es la constatación de este hecho, y otra muy distinta es su explicación. No se ha ofrecido todavía una explicación a este fenómeno, por lo que es necesario que futuras investigaciones se ocupen de ello.

Este trabajo no pretende en forma alguna ser considerado el punto final de la polémica que presentamos al inicio del mismo. Muy al contrario, lo que pretende es dar un punto de referencia alternativo al que hasta ahora ha dominado en el debate acerca de la estructura del afecto. El desdoblamiento de los factores bipolares esperados en factores monopulares independientes no debe ser atribuido de forma directa y exclusiva a la falta de validez del modelo bipolar del afecto. Lo que indica este trabajo es que ello puede deberse al desajuste entre el modelo de medida, el método de evaluación de la dimensionalidad y el constructo objeto de medida. En este sentido, este trabajo no se interpretaría como un "punto y seguido" en la investigación actual en este ámbito, por cuanto que presenta un enfoque cualitativamente distinto del problema. Más bien sería un "punto y aparte" con el que comenzar una nueva línea de investigación. Por ejemplo, sería necesario incorporar en este ámbito de estudio nuevas aproximaciones metodológicas de gran valor para discernir el problema que aquí tratamos, como los estudios de simulación. El uso de estos estudios permitiría, dado un modelo teórico de partida conocido, observar el comportamiento de las puntuaciones que produce. Ello posibilitaría comparar los patrones de puntuaciones observados con los ofrecidos

bajo los diferentes modelos, a fin de determinar con mayor seguridad si el modelo de "dos mitades" es el más adecuado. Junto a ello, también sería de gran interés retomar los métodos de investigación más tradicionales, pero no por ello menos valiosos, como los diseños experimentales.

Para concluir, parece claro que tanto el modelo de medida como el método de análisis utilizados tradicionalmente en la validación de la estructura del afecto no han sido los más adecuados. Esto ha llevado a la invalidación del modelo teórico hipotetizado, cuando en realidad el problema residía en su propia inadecuación para apresar las relaciones que se pretendía estudiar. Y este problema, que no está ligado al contenido de las dimensiones medidas, sino a su condición de dimensiones bipolares, es generalizable a otros contenidos que como éste, se suponen bipolares, entre los que destacan el ámbito de las actitudes políticas, la satisfacción laboral (Herzberg, Mausner y Snyderman, 1959; Kerlinger, 1984), y de algunas dimensiones de personalidad como la masculinidad o feminidad (Bem, 1974, 1981; Van Schuur, 1993b).

Si tuviéramos que ofrecer una conclusión global de todo lo anterior esta sería volver la vista hacia lo obvio, porque a veces no lo es tanto.

Si se han empleado con tanta profusión las escalas tipo Likert, es porque resulta obvio que son las más sencillas. Si se han invertido items representativos de polos opuestos, es porque resulta "obvio" que hay que hacerlo así. Si se ha empleado el análisis factorial para evaluar la dimensionalidad de estas escalas, es porque resulta "obvio" que es el método más empleado para este fin. Y, finalmente, si se han desdoblado los pretendidos factores bipolares en factores monopulares independientes, resulta "obvio" que se debe a que el modelo teórico de partida era inadecuado.

Sólo una cosa no era obvia: cómo era posible que escalas que medían dimensiones monopulares independientes ofrecieran buenos

Los resultados obtenidos en este estudio -a lo largo de tantos enfoques como nos ha sido posible encontrar- parecen indicar que la explicación de la paradoja anterior se encuentra, precisamente, en que lo obvio no es siempre lo correcto.

A modo de epílogo, se podría añadir que en investigación, el método hipotético-deductivo-experimental generalmente empleado parte de la identificación del objeto de estudio, continúa con la formulación de una teoría y la deducción de hipótesis empíricamente contrastables, sigue con la selección cuidadosa del método apropiado para ello, y finaliza con la contrastación de las hipótesis que el método permite. Sin embargo, esto no es más que una simplificación de la realidad. Tendemos, en el estilo más puro de razonamiento lógico, a linealizar la realidad, en busca de la preciada parsimonia que domina la ciencia. Pero a veces la excesiva simplificación, el pasar por alto los "detalles" en favor de lo "importante", lleva a la desfiguración del propio objeto de estudio, o del propio proceso de investigación. El método, que no es más que el instrumento utilizado para apresar el objeto, puede llegar a configurar el objeto, a dar de él una representación parcial, sesgada, sobre la que se basarán futuras conclusiones. En ocasiones, el objeto puede llegar a ser un artefacto producido por el método.

Este estudio es un ejemplo de ello.

ANEXO

Escalas de bienestar psicológico (Warr, 1987).

A continuación le mostramos una serie de adjetivos ante los que tendrá que dar una respuesta. Su tarea consiste en indicar cuántas veces se ha sentido como dice cada uno de los adjetivos. Para ello piense en cómo se ha sentido durante las últimas semanas.

Para contestar a cada adjetivo tiene que elegir una de las siguientes alternativas de respuesta:

| | |
|-------------------|---------------------|
| 1. nunca. | 4. bastantes veces. |
| 2. rara vez. | 5. casi siempre. |
| 3. algunas veces. | 6. siempre. |

Después de elegir la alternativa de respuesta que más se ajusta a su caso particular, escriba el número de la alternativa escogida entre los paréntesis que siguen a cada adjetivo.

ADJETIVOS:

- 1. TENSO()
- 2. INCOMODO()
- 3. PREOCUPADO()
- 4. TRANQUILO.....()
- 5. CONTENTO.....()
- 6. RELAJADO.....()
- 7. DEPRIMIDO()
- 8. PESIMISTA()
- 9. DESGRACIADO()
- 10. ALEGRE()
- 11. ENTUSIASMADO()
- 12. OPTIMISTA.....()
- 13. FATIGADO()
- 14. SIN ENERGIA()
- 15. CANSADO()
- 16. DESPIERTO()
- 17. LLENO DE ENERGIA.....()
- 18. ANIMADO.....()

BIBLIOGRAFIA

- Abelson, R. P. y Sermat, V. (1962). Multidimensional scaling of facial expressions. *Journal of Experimental Psychology*, 63, 546-554.
- Arce, C. (1994). *Técnicas de construcción de escalas psicológicas*. Madrid: Síntesis.
- Askevis-Leherpeux, F. y Matalon, B. (1990). Une limite à l'usage des échelles bipolaires de jugement. *Revue Internationale de Psychologie Sociale*, 3, 233-248.
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42, 115-162.
- Bem, S. L. (1981). The BSRI and gender schema theory: a reply to Spence and Helmreich. *Psychological Review*, 88, 369-371.
- Bentler, P. M. (1969). Semantic space is (approximately) bipolar. *Journal of Psychology*, 71, 33-40.
- Borgatta, E. I. (1961) Mood, personality and interaction. *Journal of General Psychology*, 64, 105-137.
- Bowling, A (1991). *Measuring health*. Philadelphia: Open University Press, Milton Keynes.
- Bradburn, N. M. (1969). *The structure of psychological well-being*. Chicago: Aldine.
- Bush, L. E. (1973). Individual differences multidimensional scaling of adjectives denoting feelings. *Journal of personality and social psychology*, 25, 50-57.
- Campbell, N. O. y Grissom, S. (1979). Influence of item direction on student responses in attitude assessment. Paper presented at the 63 annual meeting of the American Educational Research Association, San Francisco, CA.
- Cliff, N. y Young, F. W. (1968). On the relation between unidimensional judgments and multidimensional scaling. *Organizational Behavior and Human Performance*, 3, 269-285.
- Clyde, D. J. (1963). *Manual for the Clyde Mood Scale*. Coral Gables, Florida: University of Miami.
- Cohen, J. M. y Cohen, H. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. London: Lawrence Erlbaum associates INC.
- Coombs C. H. (1964). *A theory of data*. New York: John Wiley.
- Coombs C. H. y Kao, R. C. (1960). On a connection between factor analysis and multidimensional unfolding. *Psychometrika*, 25, 219-31.
- Coombs C. H. (1950). Psychological scaling without a unit of measurement. *Psychological Review*, 57, 145-158.
- Coombs C. H. (1952). A theory of psychological scaling. *Engineering Research Institute Bulletin*, 34. Ann Arbor, MI: University of Michigan Press.
- Costa, P. T. y McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology*, 38, 668-678.
- Davison, M. L. (1977). On a metric, unidimensional unfolding model for attitudinal and developmental data. *Psychometrika*, 42, 523-548.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 3, 542-575.
- Girelli, S. T. y Stake, J. E. (1993). Bipolarity in Jungian type theory and the Myers-Briggs type indicator. *Journal of Personality Assessment*, 60, 290-301.
- Gotlib, I. H. y Meyer, J. P. (1986). Factor analysis of the Multiple Affect Check List: A separation of positive and negative affect. *Journal of Personality and Social Psychology*, 50, 1161-1165.
- Green, R. F. y Goldfried, M. R. (1965). On the bipolarity of semantic space. *Psychological Monographs*, 79, (nº 6 completo).
- Guildford, J.P. (1954). *Psychometric methods*. New York: McGraw Hill.
- Guttman, L. L. (1944). A basis for scaling qualitative data. *American Sociological Review*, 9, 139-159.

- Guttman, L. L. (1950). The basis for scalogram analysis. En S. A. Stouffer et al. (Eds.), *Mesurement and prediction*. Princeton, NJ: Princeton University Press.
- Hall, C. A. (1977). Differential relationships of pleasure and distress with depression and anxiety over a past, present and future time framework. Tesis Doctoral no publicada. Universidad de Minnesota.
- Hendrick, C. y Lilly, R. S. (1970). The structure of mood: A comparison between sleep deprivation and normal wakefulness conditions. *Journal of Personality*, 38, 453-465.
- Hendrickson, A. E. y White, P. O. (1966). A method for the rotation of higher-order factors. *British Journal of Mathematical and Statistical Psychology*, 19, 97-103.
- Herzberg, F.; Mausner, B. y Snyderman, B. B. (1959). *Motivation to work*. New York: Wiley.
- Hoijtink, H. (1990). A latent trait model for dichotomus choice items. *Psychometrika*, 55, 641-656.
- Izard, C. I. (1972). *Patterns of emotions*. New York: Academic Press.
- Jöreskog, K. G. 1969. A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34, 183-202.
- Jöreskog, K. G. 1971. Statistical analysis of sets of congeneric tests. *Psychometrika*, 36, 106-133.
- Kerlinger, F. N. (1984). *Liberalism and Conservatism*. Hillsdale: Lawrence Erlbaum Associates.
- Lorr, M; Shi, A. Q. y Youniss, R. P. (1989). A bipolar multifactor conception of mood states. *Personality and Individual Differences*, 10, 155-159.
- Lorr, M; McNair, D. M. y Fisher, S. (1982). Evidence for bipolar mood states. *Journal of Personality Assessment*, 43, 468-72.
- Lloret, S. y González-Roma, V. (1994). La medición del bienestar psicológico. En V. González-Romá (director), *La medición del bienestar psicológico y otros aspects de la salud mental*. Valencia: Nau Llibres.
- Lloret, S.; González-Roma, V.; Ferreres, A.; Espejo, B.; Tomás, I. y Hernández, A. (1995). La medición de constructos afectivos bipolares mediante escalas tipo Likert: un análisis crítico. Comunicación presentada al IV Simposio de Metodología de las Ciencias del Comportamiento, La Manga del Mar Menor, Murcia, 4-7 abril 1995.[
- McLachlan, J. F. C. (1976). A short adjective check list for the evaluation of anxiety and depression. *Journal of Clinicl Psychology*, 32, 195-197.
- Meddis, R. (1972). Bipolar in mood adjective check lists. *British Journal of Sociala and Clinical Psychology*, 11, 178-184.
- McIver, J. P. y Carmines, E. G. (1986). *Unidimensional Scaling*. Beverly Hills: Sage Publications.
- McDowell, I. y Newell, C. (1987). *Measuring health*. BNew York: Oxford University Press.
- McNair D. M. y Lorr, M. (1964). An analysis of mood in neuroticas. *Journal of Abnormal and Social Psychology*, 69, 620-627.
- McNair D. M.; Lorr, M. y Droppleman, L. F. (1971). *Manual: Profile of mood states*. San Diego: Educational and Industrial Testing Service.
- Mehrabian, A. y Russell, J. A. (1974). *An approach to environmental psychology*. Cambridge: MIT Press.
- Miller, T. R. y Clearly, T. A. (1993). Direction of wording effects in balanced scales. *Educational and Psychological Measurement*, 53, 51-61.
- Miller, N. y Pollock, V. I. (1994). Met-analysis and some science-compromising problems of social psychology. En Shadish, W. R. y Fuller, S. (Eds.) *The social psychology of science*. New York: Guldford Press.

- Mook, J.; Kleijn, W. C. y van der Ploeg, H. M. (1991). Simptom positively and negatively worded items in two popular self-report inventories of anxiety and depression. *Psychological Reports*, 69, 551-560.
- Mook, J.; van der Ploeg, H. M. y Kleijn, W. C. (1992). Simptom-positive and symptom-negative items in the State-Trait Anxiety Inventory: a comparison and replication. *Anxiety, Stress and Coping*, 5, 134-145.
- Mook, J.; Kleijn, W. C. y van der Ploeg, H. M. (1992). Positively and negatively worded items a self-report measure of dispositional optimism. *Psychological Reports*, 71, 275-278.
- Neufeld, R. W. J. (1975). Simultaneous processing of multiple stimulus dimensions among paranoid and non paranoid schizophrenics. *Multivariate behavioral research*, 11, 425-441.
- Nowlis, V. (1965). Research with the Mood Adjective Check List. In S. S. Tomkins & C. E. Izard (Eds.), *Affect, cognition, and personality*. New York: Springer, 1965.
- Nowlis, V. y Nowlis H. H. (1956). The description and analysis of mood. *Annals of the New York Academy of Sciences*, 65, 345-355.
- Osgood, C. E., Suci, G. J. y Tannenbaum, P. H. (1976). *La medida del significado*. Madrid: Gredos.
- Popper, K. (1991). Selections from the logic of scientific discovery. En Boyd, R.; Gasper, P. y Trout, J. D. (Eds.) *The philosophy of science*. Cambridge: MIT Press.
- Purcell, A. T. (1982). The structure of activation and emotion. *Multivariate behavioral research*, 17, 221-251.
- Ray, J. J. (1990). Acquiescence and problems with forced-choice scales. *The Journal of Social Psychology*, 130, 397-399.
- Ross, R. T. (1938). A statistic for circular scales. *Journal of educational psychology*, 29, 384-389.
- Ross, J. y Cliff, N. (1964). A generalization of the interpoint distance model. *Psychometrika*, 29, 167-176.
- Russel, J. A. (1978). Evidence of convergent validity on the dimensions of affect. *Journal of Personality and Social Psychology*, 36, 1152-1168.
- Russel, J. A. (1979). A circumplex model of affect. *Journal of Personality and Social Psychology*, 37, 345-356.
- Russel, J. A. (1980). Affective space is bipolar. *Journal of Personality and Social Psychology*, 39, 1161-1178.
- Russel, J. A. y Mehrabian, A. (1977). Evidence for a three factor theory of emotions. *Developmental Psychology*, 19, 795-804.
- Russel, J. A. y Ridgeway, D. (1983). Dimensions underlying children's emotion concepts. *Journal of Research on Personality*, 11, 273-294.
- Sevastos, P.; Smith, L. y Cordery, J. L. (1992). Evidence on the reliability and construct validity of Warr's (1990) well-being and mental health measures. *Journal of occupational and Organizational Psychology*, 65, 33-49.
- Scheier, M. F. y Carver, C. S. (1985). Optimism, Coping and Health: assessment and implications of generalized outcome expectancies. *Health Psychology*, 4, 219-247.
- Scholsberg, H. (1952). The description of facial expressions in terms of two dimensions. *Journal of Experimental Psychology*, 61, 81-88.
- Scholsberg, H. (1954). Three dimensions of emotion. *Psychological Review*, 44, 229-237.
- Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrika*, 27, 125-139; 219-246.
- Tellegen, A. (1984). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis in self-report. En Tuma A. H. y Maser, J. D. (Eds.) *Anxiety and the anxiety disorders*. Hillsdale, NJ: Erlbaum.
- Thayer, R. E. (1967). Measurement of activation through self-report. *Psychological Reports*, 20, 663-678.

- Torgerson, W. S. (1958). *Theory and methods of scaling*. New York: John Wiley.
- Van Schuur, W. H. y Kruijtbosch, M. (1994). El bienestar subjetivo: despliegue de la escala de Equilibrio Emotivo de Bradburn. En J. Díez & R. Inglehart (Eds.), *Tendencias mundiales de cambio en los valores sociales y políticos*. Madrid: FUNDESCO, 1994.
- Van Schuur, W. H. (1988). Stochastic unfolding. En Saris W. E. y Gallhofer, I. N. (Eds.) *Sociometric research, Volume 1, Data Collection and Scaling*. London: McMillan.
- Van Schuur, W. H. y Post, W. (1990). *User's manual MUDFOLD, a program for multiple unidimensional unfolding*. Groningen i.e.c.ProGamma.
- Van Schuur, W. H. (1993a). Nonparametrical unidimensional unfolding for multicategory data. *Political Analysis*, 4, 41-74.
- Van Schuur, W. H. (1993b). Masculinity and Femininity -or androgyny? A comparison of the three unfolding models. En J. Oud y R. van Blokland-Vogeleang(Eds.), *Advances in Longitudinal and multivariate analysis in the behavioral sciences*. Proceedings of the SMABS conference 1992 en Nimegen, Holanda.
- Van Schuur, W. H. y Kiers, H. A. L. (en prensa). Why factor analysis is often the wrong model for analyzing bipolar concepts and what model to use instead. *Applied Psychological Measurement*.
- Warr, P. (1990). The measurement of well-being and other aspects of mental health. *Journal of Occupational Psychology*, 63, 193-210.
- Warr, P. (1987). *Work, unemployment and mental health*. Oxford: Oxford University Press.
- Watson, D. y Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, 98, 219-235.
- Watson, D. y Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin*, 96, 465-490.
- Watson, D.; Clark, L. A. y Tellegen, A. (1984). Cross-cultural convergence in the structure of mood: A Japanese replication and a comparison with US findings. *Journal of Personality and Social Psychology*, 47, 127-144.
- Watson, D.; Clark, L. A. y Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS Scales. *Journal of Personality and Social Psychology*, 54, 1063-1070.
- Westbrook, M. T. (1976). Positive affect: a method of content analysis for verbal samples. *Journal of Consulting and Clinical Psychology*, 44, 715-719.
- Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology*, 37, 395-412.
- Zevon, M. A. y Tellegen, A. (1982). The structure of mood change: an idiographic/nomothetic analysis. *Journal of Personality and Social Psychology*, 43, 111-122.
- Zuckerman, M. y Lubin, B. (1965). *The multiple affect adjective check list*. San Diego CA: Educational and Industrial Testing Service.

