

**UNIVERSITAT DE VALENCIA**

**Facultat de Psicologia**

-----



**EL USO DE LOS MODELOS DE ECUACIONES ESTRUCTURALES  
Y DEL ANALISIS FACTORIAL CONFIRMATORIO EN EL  
ANALISIS PSICOMETRICO DE CUESTIONARIOS: UNA BATERIA  
DE SEGURIDAD LABORAL.**

**TESIS DOCTORAL**



**PRESENTADA POR:**

**JOSE MANUEL TOMAS MIGUEL**

**DIRIGIDA POR :**

**Dr. JOSE LUIS MELIA NAVARRO**

**València, Noviembre 1.993**

UMI Number: U607341

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 U607341

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



BID. T 1236

UNIVERSIDAD DE VALENCIA
FACULTAD DE PSICOLOGIA
BIBLIOTECA
Reg de Entrada n° 6629
Fecha: 7-2-94
Signatura P7379

D. 173.538

L. 173.566

*Data do not speak for themselves but through the medium of the analytic techniques applied to them. It is important to realize that the analytic techniques not only set limits to the scope and the answers one may obtain from data, but also affects the type of questions a researcher asks and the manner in wich the questions are formulated. "It comes as no particular surprise to discover that a scientist formulates problems in a way wich requires for their solution just those techniques in which he himself is especially skilled"*  
(Kaplan, 1964, p. 28).

*Pedhazur, 1982, p. 4.*





*Per a les tres dones que més vullc:  
la meua mare, Maria  
la meua iaia, Maria  
i la meua dona, Amparo.*



# Indice

## Capítulo 1. Introducción

1. El problema de los accidentes de trabajo, definición 1
2. La evolución de la investigación sobre accidentes laborales 2
3. Clasificación de los estudios psicológicos sobre accidentes laborales 5
  - 3.1. Estudios descriptivos 6
  - 3.2. Estudios correlacionales 9
  - 3.3. Estudios teóricos 22
    - 3.3.1. Modelos de secuencia de dominó 22
    - 3.3.2. Modelos de liberación de energía 23
    - 3.3.3. Modelos cognitivos 24
    - 3.3.4. Modelos de estado del sistema 24
    - 3.3.5. Modelos de factores situacionales y de factor humano 25
4. Medición de factores implicados en la seguridad 25
  - 4.1. Clima de seguridad 26
  - 4.2. Respuesta de los superiores 31
  - 4.3. Respuesta de los compañeros 32
  - 4.4. Medida de los riesgos laborales 32

## Capítulo 2. Metodología

1. Descripción de la muestra 38
2. Métodos y técnicas 44
  - 2.1. Coeficientes de fiabilidad y homogeneidad 44
  - 2.2. Coeficientes de correlación y regresión 46
  - 2.3. Modelos causales 46
  - 2.4. Comprobación del supuesto de normalidad multivariable. 64
  - 2.5. Validación de constructo y modelos causales. 66
    - 2.5.1. Algunas aplicaciones de los modelos causales en la validación de constructo. 69
3. Variables y su medida 82
4. Objetivos 86
  - 4.1. Objetivos generales 86
  - 4.2. Objetivos específicos 87
5. Hipótesis 87

## **Capítulo 3. Clima organizacional hacia la seguridad**

1. Introducción 89
2. Formación-Información 91
  - 2.1. Medida 91
  - 2.2. Estructura factorial confirmatoria 91
  - 2.3. Fiabilidad y homogeneidad de la escala y los items 99
  - 2.4. Validez de los items 102
  - 2.5. Conclusiones para la escala de formación-información. 103
3. Política de seguridad de la empresa 105
  - 3.1. Medida 105
  - 3.2. Estructura factorial confirmatoria 105
  - 3.3. Fiabilidad y homogeneidad de la escala y los items 116
  - 3.4. Validez criterial de los items 119
  - 3.5. Conclusiones para la escala de política. 121
  - 3.6. Presentación de una escala de política depurada 123
4. Estructura de seguridad 125
  - 4.1. Medida 125
  - 4.2. Estructura factorial confirmatoria 125
  - 4.3. Fiabilidad y homogeneidad de la escala y los items 130
  - 4.4. Validez de los items 133
  - 4.5. Conclusiones para la escala de estructura de seguridad. 134

## **Capítulo 4. Medición de los riesgos laborales**

1. Introducción 136
2. Riesgo Basal 138
  - 2.1. Medida 138
  - 2.2. Estructura factorial confirmatoria 139
  - 2.3. Fiabilidad y homogeneidad de la escala de riesgo basal y sus items 157
  - 2.4. Validez criterial de los items 158
  - 2.5. Conclusiones sobre la escala de riesgo basal. 160
3. Riesgo Real 162
  - 3.1. Medida 162
  - 3.2. Estructura factorial confirmatoria 162
  - 3.3. Fiabilidad y homogeneidad de la escala y los items 174
  - 3.4. Validez criterial de los items 177
  - 3.5. Conclusiones sobre la escala de riesgo real. 179
4. Validez convergente-discriminante de las escalas de riesgo basal y real 180

## **Capítulo 5. Medidas de contexto social**

1. Respuesta de los Superiores en seguridad 183
  - 1.1. Introducción y medida 183
  - 1.2. Estructura factorial confirmatoria 185
  - 1.3. Estructura factorial de segundo orden 196
  - 1.4. Fiabilidad de la escala y de los items 198
  - 1.5. Validez criterial 200
  - 1.6. Validez convergente-discriminante 202
  - 1.7. Conclusiones 204
2. Respuesta de los Compañeros 205
  - 2.1. Introducción y medida 205
  - 2.2. Estructura factorial confirmatoria 206
  - 2.3. Fiabilidad de la escala y de los items 213
  - 2.4. Validez criterial 214

## **Capítulo 6. Respuesta y efectos sobre la seguridad del trabajador**

1. Introducción 217
2. Tensión psicológica 219
  - 2.1. Medida 219
  - 2.2. Estructura factorial confirmatoria 219
  - 2.3. Fiabilidad de la escala y de los items 228
  - 2.4. Validez criterial 230
3. Conducta hacia la seguridad de los trabajadores 233
  - 3.1. Medida 233
  - 3.2. Estructura factorial confirmatoria 235
  - 3.3. Fiabilidad de la escala y de los items 242
  - 3.4. Validez criterial 246
  - 3.5. Conclusiones 248
4. Indices para la medición de la siniestralidad 250

## **Capítulo 7. Modelos explicativos de los accidentes y conclusiones**

1. Relaciones bivariadas 259
2. Modelo hipotético 262
3. Modelos causales 265
4. Modelo considerando los errores de medida de las variables 276

5. Conclusiones 284

**REFERENCIAS METODOLOGIA i**

**REFERENCIAS SEGURIDAD LABORAL vii**

# 1. Introducción.

## **1. El problema de los accidentes de trabajo, definición**

Los accidentes laborales u ocupacionales han venido siendo estudiados por numerosos investigadores, debido a que es un problema de enorme magnitud y cuantiosas pérdidas.

Si nos remitimos a los datos de la Comunidad Valenciana para el año 1992, observamos un total de 71.896 accidentes laborales con baja laboral. De estos siniestros, 70.452 fueron leves, pero 1361 fueron graves y 99 mortales. Los costes personales de los accidentes son pues elevados, pero también son elevados los económicos. Así, Harris (1992) cita como costes de un accidente laboral los siguientes:

- Costes directos: como los costes de los empleados -costes de compensación-; costes legales; costes médicos; daños al material o a los productos.

- Costes indirectos: como pérdida de producción; tiempo de investigaciones sobre los accidentes; interrupciones y pérdidas de tiempo en otros trabajadores que rodean al accidentado; aumento en los costes de las aseguradoras; reentrenamiento del nuevo trabajador que reemplaza al accidentado; tiempo extra del supervisor;



- Y, por último, considera algunos costes intangibles, difíciles de cuantificar en términos económicos, pero importantes como: daños a la imagen de la compañía; descenso de la moral; efectos psicológicos en los trabajadores; pérdida de negocios; daños a las relaciones industriales, etc.

Harris gráficamente compara los gastos asegurados frente a los no asegurados usando la alegoría del iceberg, donde la parte de gastos asegurada no es más que la punta. Cuantificando, y siempre según este autor, de cada libra esterlina gastada en costes asegurados, un accidente supone de ocho a treinta y seis libras de costes no asegurados.

A la vista de la variedad de costes que se desprenden de un accidente, podemos hacernos una idea de la pérdidas económicas, al margen de las humanas, que se han producido sólo en la Comunidad Valenciana en este último año. Además, hay que tener en cuenta que las estadísticas oficiales sólo evalúan como tales a los accidentes que se han producido con baja. Esto es, sólo se reflejan como accidentes de trabajo los que producen un daño físico con consecuencia de pérdida, al menos de un día de trabajo.

La definición de accidente de trabajo que manejan la mayoría de investigadores actuales es mucho más amplia. Así se puede definir accidente como:

"Un evento no planeado e incontrolado, en el cual la acción o reacción de un objeto, sustancia, persona o radiación, resulta en una lesión o en la probabilidad de que ello ocurra" (Heinrich et al., 1980).

Lo que pone de manifiesto este autor es que no es necesaria la existencia de lesión para que un accidente se considere como tal. Efectivamente, la lesión en sí puede considerarse tan sólo como una consecuencia del propio accidente. Usualmente a los accidentes sin lesión también se les ha llamado accidentes blancos o incidentes. Los costes de los accidentes aumentarían considerablemente si se incluyeran todos los accidentes y no solo los que causan baja laboral.

## **2. La evolución de la investigación sobre accidentes laborales**

Los estudios sobre accidentes laborales han pasado a lo largo de su corta historia por diferentes fases en las que puede apreciarse un mayor interés uno u otro enfoque, o por ciertos factores característicos. Así, Harris (1992) marca tres fases en esta evolución de los estudios sobre accidentabilidad en las empresas químicas inglesas: en los setenta la mayor atención la ocupaban los factores técnicos, de ingeniería, especialmente en los aspectos de equipos de protección o en general de equipos de producción; en los ochenta el interés se fué desplazando hacia los aspectos de procedimiento, de ergonomía, de cómo se hacían los procesos productivos; por último, en los años noventa los esfuerzos se concentran en la "seguridad a través de las personas", haciendo notar la importancia del factor humano en la seguridad.

Más detallada es la división en cuatro fases que presenta el profesor Peña Baztan (1990), basada en un trabajo del profesor Borrás París al IV Congreso Nacional de Psicología de Barcelona de 1973.

### PRIMERA FASE: Factor técnico.

Arranca en la creación en 1876 de la Asociación de Higiene y Prevención de la región de Alsacia por parte del industrial alemán Engel Dollfus.

Toda esta fase se caracteriza por una prevalencia del factor técnico. Se considera que los accidentes tienen una causa concreta, y es la deficiencia de los equipos e instalaciones industriales y que si se eliminan estas deficiencias se eliminan los accidentes 'salvo aquéllos puramente fortuitos'. Así, como factor desencadenante tenemos los aspectos técnicos. La proporcionalidad que se atribuía a las causas de los accidentes era, en prácticamente un 100%, a la inadecuación del equipo e instalaciones industriales.

Por esa época, y en nuestro país se sigue también esta línea y aparece la primera Ley de Accidentes de trabajo (30 de Enero de 1900).

La experiencia demuestra pronto que las medidas de tipo técnico realizadas en su mayor parte por los ingenieros consiguen una rebaja en el número de los accidentes. Esta rebaja no es, sin embargo, lo espectacular que cabía esperar teniendo en cuenta que se consideraba la única causa de los accidentes y se estabiliza pronto (Denton, 1986).

## *Capítulo 1. Introducción.*

---

### *SEGUNDA FASE: Susceptibilidad individual a accidentarse.*

El profesor Marbe de la Universidad de Wurtzburg plantea a partir de los años 20 una teoría que a partir de estudios con escolares generalizados a la industria sitúa el factor humano como causa de accidentes añadida al factor técnico.

Lo que plantea esta teoría es que existe en los trabajadores una mayor o menor predisposición individual al accidente. Así, lo que se pretende no es que los individuos tengan un rasgo de, digamos, capacidad para accidentarse, sino que un sujeto está predispuesto al accidente cuando se le coloca en un puesto para el cual no tiene condiciones. Por ejemplo, no se podría poner a trabajar a una persona a cierta altura en un andamio si tiene poco sentido del equilibrio o vértigo. Lo que se plantea, por tanto, es una relación entre aptitud y las condiciones de trabajo.

Desde este enfoque la atribución porcentual de las causas de los accidentes se planteaba aproximadamente en estos términos: un 75% debido factor humano -trabajadores situados en puestos para los que no tienen las aptitudes adecuadas-; y un 25% debido a factor técnico.

Desde este enfoque del problema, la prevención de los accidentes es evidente. Por un lado, y como primer paso, cabe hacer todo tipo de modificaciones técnicas. Posteriormente se realizaría una adecuada selección del personal para adecuar a los candidatos a aquellos puestos para los que estuvieran adecuadamente predispuestos.

### *TERCERA FASE: La actitud del trabajador.*

A partir de los años sesenta empieza a parecer claro que la imputación de los accidentes a causas técnicas y de aptitud no daba cuenta de la mayor parte de los siniestros. Hacía falta buscar nuevas causas. Estas causas se buscaron en una mejor delimitación del concepto de error humano. Mientras en la anterior fase éste se entendía como la asimilación de las aptitudes del trabajador al puesto requerido, poniéndose el énfasis en la aptitud o propio valer del trabajador para el puesto. En esta fase se introduce un nuevo concepto en la idea de factor humano, la actitud del trabajador, y "con ello el querer hacer las actividades encomendadas correctamente". Esta actitud se supone que afecta a la realización concreta en el puesto de trabajo, de forma que puede estar determinando la realización de actos de inseguridad, de imprudencias.

En esta etapa a los conocidos factores técnicos y de factor humano, asimilado a aptitud, se añaden la actitudes hacia el trabajo y la seguridad dentro del factor humano. En resumen, se añade un

nuevo macrofactor a los determinantes de los accidentes y el individuo, el trabajador, aparece como el máximo responsable de la seguridad.

En esta línea, la UNESCO en 1961 evalúa, para los países industrializados, el posible impacto de estos factores. Las cifras ofrecidas consideran que un 12.5% de los accidentes pueden ser debidos a factores técnicos, un 7.5% son totalmente fortuitos, y un 80% son debidos a factor humano. No ofrecen cifras, siquiera aproximativas, sobre el porcentaje que de éste 80% se corresponde más con problemas aptitudinales o cuál con actitudinales, pero estiman que el factor actitudinal tiene una mayor importancia en la explicación de los accidentes.

La importancia de la actitud de los trabajadores se traduce en el campo de la intervención en un mayor énfasis en la formación y entrenamiento de los trabajadores.

#### CUARTA FASE: Seguridad integral.

Según el profesor Borrás, la cuarta fase puede considerarse que arranca del informe de 1967 de la Comunidad Europea del Carbón y del Acero titulado "Los factores Humanos y la Seguridad".

Desde esta nueva concepción se sigue haciendo hincapié en el factor humano, pero extendiendo su significado. El factor humano no sólo incluye las aptitudes y actitudes del trabajador accidentado, sino que se complementa con la incorporación del grupo y de las variables organizativas, de entorno y de estructura organizativa. Desde este enfoque, se considera que las aptitudes y sobre todo las actitudes del trabajador se ven modificadas o determinadas en parte por el grupo y también por importantes variables de la organización.

Lógicamente esta concepción más plural del problema hace mucho más difícil la evaluación de proporción de causas de accidente y por otro lado, también señala múltiples vías de prevención. Así como en la primera fase se hablaba de soluciones de ingeniería y ergonomía, en la segunda de selección y en la tercer fase de formación, ahora las posibilidades de prevención se multiplican y pueden combinarse. Todas las anteriores, junto con el control organizacional, el cambio de actitudes, la formación de supervisores y del grupo de compañeros, y las técnicas de modificación de conducta pueden considerarse.

Aunque la distinción en cuatro fases, además de los momentos de inicio de cada una de ellas, sean discutibles y tan sólo un intento clasificador, sí son interesantes para describir la progresiva complejización que de hecho se ha producido en la mayoría de los estudios sobre accidentes laborales. Efectivamente, de la

## *Capítulo 1. Introducción.*

---

consideración de un único factor en la explicación de los accidentes, se ha ido progresivamente pasando hacia concepciones más multivariadas, donde se consideran procesos complejos de causación. Las investigaciones actuales sobre accidentes son coincidentes con la opinión de Chockar (1990) de que los accidentes son la consecuencia final de una secuencia de eventos.

### **3. Clasificación de los estudios psicológicos sobre accidentes laborales**

Basándonos en el trabajo de Islas y Meliá (1991) podemos clasificar los estudios psicosociales sobre accidentes ocupacionales en cuatro categorías: 1) descriptivos; 2) correlacionales; 3) de intervención; y 4) teóricos. De ellos, para este trabajo son de especial interés los descriptivos y correlacionales y los de tipo teórico. Dado que nuestro trabajo es de tipo correlacional y llevado a cabo en un ambiente natural, los datos de los estudios descriptivos y, sobre todo, los correlacionales son interesantes por que ayudan a entender qué factores o variables pueden estar relacionados con la ocurrencia de los accidentes. Y los de tipo teórico por que ayudan a definir las posibles vías de causación y el orden de las variables, a la hora de la implementación de modelos causales complejos. Los estudios de intervención han sido generados en su mayoría desde una óptica de modificación de la conducta insegura del trabajador y no serán revisados aquí. Los porcentajes de éxitos conseguidos en algunos de estos estudios sirven no obstante para poner de manifiesto que la conducta del trabajador es un antecedente relevante de los accidentes. Desgraciadamente, estos trabajos no explican el porqué de estos actos seguros o inseguros.

#### **3.1. Estudios descriptivos**

Los estudios de tipo descriptivo son los primeros en aparecer y hasta ahora los de mayor número. Entran dentro de un campo exploratorio, donde se trata de describir las variables descriptivas de los trabajadores, de la organización y de los equipos de seguridad, que pueden parecer importantes en la ocurrencia de accidentes.

Se trata de conocer las condiciones de trabajo, especialmente de aquellas empresas con alta siniestralidad o trabajadores con accidentes. En general los instrumentos de medida utilizados en este tipo de estudios son los cuestionarios y entrevistas a los trabajadores.

El trabajo de Fernández, Nogareda y Oncins (1988) es una encuesta sobre condiciones de trabajo, con especial interés en la salud y seguridad laboral. En este trabajo, de considerable magnitud y pretensiones, se utiliza una muestra de 4000 trabajadores para analizar las siguientes variables:

- De individuo: edad, sexo, estado civil, número de hijos, y nivel de estudios.

## *Capítulo 1. Introducción.*

---

- De tarea: tipo de trabajo, carga de trabajo, nivel de comunicación, status y nivel de autonomía.

- De organización: actividad de la empresa, tamaño, tipo de contrato, salario, categoría profesional, estabilidad, promoción, horario, participación, prevención y formación.

- De entorno: temperatura, iluminación, ruido y vibraciones, contaminantes químicos, y condiciones de seguridad.

- Consecuencias: accidentes, enfermedades comunes y profesionales, y molestias.

Aunque los resultados son descriptivos, de frecuencia, y no se calculan estadísticos de relación, pueden destacarse algunas conclusiones:

Por un lado, y en cuanto a las estructuras de seguridad, a pesar de que los comités de seguridad e higiene son obligatorios en empresas de más de 100 trabajadores, hay un porcentaje importante que oscila entre el 15.9% y el 26.4% que no tienen. Entre las empresas que no es obligatorio un 16.7% sí lo tienen. Un elevado porcentaje de empresas no tienen a nadie encargado de la seguridad laboral.

En cuanto a condiciones físicas de trabajo: Un 19.1% declara sentirse cansado al terminar la jornada; el 23.5% reconoce que el ruido al que está sometido es perjudicial para su salud; y un 18% sufre problemas con la iluminación; un 17% consideran que respiran sustancias contaminantes en proporción peligrosa.

Los riesgos más frecuentes en la muestra son los golpes, caídas, cortes, quemaduras y contactos eléctricos. Por su parte la atribución de causalidad que hacen los trabajadores hace referencia básicamente a exceso de confianza por la costumbre (76.5%), menos importante es la atribución a la peligrosidad intrínseca del trabajo (30.7%). Pocos consideran que los factores técnicos o de diseño tengan una influencia importante en la ocurrencia de accidentes.

En cuanto a los resultados para algunas variables psicosociales, destaca que la comunicación de problemas o sugerencias de seguridad o de mejora de las condiciones en general se realiza generalmente por vía informal, a través de los compañeros, no existiendo un adecuado nivel de comunicación formal. En cuanto a la formación, sólo un cuarto de los trabajadores reciben cursos, y de éstos sólo un 20% ha recibido algún curso de seguridad e higiene en el trabajo.

El trabajo de Walter y Haines de (1988) utiliza una muestra de 452 trabajadores de la fundición del metal, textil, y hospitales. El trabajo, mediante cuestionarios estructurados trata de estudiar las percepciones, conocimientos y acciones posibles de los trabajadores ante los riesgos profesionales a que están sometidos. Los resultados apuntan a un elevado nivel de conocimiento de cómo el trabajo puede dañar la salud y de la probabilidad de tener accidentes. Sin embargo, carecían de la suficiente información sobre los resultados de sus exámenes médicos, de sus derechos legales ante problemas de salud y seguridad y no tenían conocimiento adecuado de la forma de reducir los riesgos.

La investigación realizada por Pekkarinen y Antonen (1989) estudia las condiciones de seguridad y los tipos de accidentes y riesgos en dos complejos industriales, uno situado en Finlandia y otro situado en Rusia, pero también con trabajadores finlandeses. La muestra total la componían 3500 trabajadores. Se realizaron comparaciones descriptivas entre los trabajadores finlandeses que trabajaban en su país y los que trabajaban en Rusia.

Las principales conclusiones apuntaban a que había más accidentes en Rusia, pero que eran menos severos. Los autores lanzan como hipótesis que estos resultados de mayor número de accidentes al trabajar en el extranjero pueden deberse en parte a importantes aspectos del clima organizacional, como los cambios que se producen en la instrucción y la planificación, a falta de servicios de seguridad e higiene en el trabajo, la falta de coordinación, y la falta de instrucciones claras de seguridad.

Otro resultado importante fue el análisis de las causas de accidentes, mayoritariamente debidos en Rusia a la falta de orden y limpieza en el lugar de trabajo que, según los autores, podía también estar relacionado en cierto modo con factores de clima de seguridad como la falta de inspecciones, la poca claridad de las órdenes y la división del trabajo, etc. En resumen, numerosos fallos en el clima laboral, como los comentados y falta de entrenamiento y adecuada selección fueron detectados en los lugares de trabajos con mayor número de accidentes. Otros factores más específicos, como el mayor consumo de alcohol de los trabajadores finlandeses que trabajaban fuera de casa y dejaron a sus familias, también se consideró que afectaban al 6% de los accidentes.

Estos tres trabajos revisados incluían muestras de tamaño elevado y el rango y tipo de empresas era amplio, por lo que sus resultados pueden verse como más generales. Otros trabajos han evaluado muestras más pequeñas de ambientes laborales más concretos.



Así, el trabajo de Islas, Vera y Ocaño (1986) se lleva a cabo con una muestra de 25 trabajadores, mecánicos de una fábrica de construcción de material. El objetivo del trabajo era analizar posibles factores históricos y situacionales disposicionales hacia los accidentes de trabajo. Para evaluar la ocurrencia de estas posibles variables disposicionales se emplearon la observación sistemática del lugar de trabajo y cuestionarios contestados por los trabajadores. Los resultados más interesantes del trabajo son:

- En cuanto a condiciones ambientales, muchos trabajadores trabajan en posturas inadecuadas, existe falta de limpieza, exceso de polvo, maquinaria ruidosa y calor excesivo. Todo ello son, según los trabajadores y los observadores de conducta, factores basales que condicionan la realización de actos inseguros que pueden llevar a accidentes.

- En cuanto a la política organizativa, un 60% de los trabajadores no había recibido formación suficiente para el puesto, y la formación específica de seguridad era casi inexistente. Además, los trabajadores presentaban una actitud negativa hacia los cursos.

- Por último, la evaluación que hacen los trabajadores de los accidentes, la mayoría consideran que la negligencia y descuido por su parte es la principal causa. No obstante creen que estos descuidos se encuentran en muchos casos relacionados con la presión para trabajar que realizan los supervisores, y con las condiciones inseguras de la planta. Otra causa de accidentes era el no utilizar o no tener disponibles los equipos de seguridad, que eran considerados además en muchos casos como incómodos o innecesarios.

Otros factores evaluados, como la actitud familiar sobre los accidentes, no mostraron indicios de relación alguna con la ocurrencia de los accidentes.

El trabajo de Vojtecky y Schmitz (1986) es también muy específico. En una muestra de 124 encargados de evaluar los programas de salud y seguridad, se trataba de valorar los propios programas de seguridad. Se pretendía extraer la opinión de los expertos en evaluación de programas sobre cómo se estaban llevando a cabo los entrenamientos, la formación y sobre qué tipo de riesgos estaban actuando y los niveles de satisfacción que producían. Los resultados, a decir de los autores no permiten evaluar estos programas a través de la muestra tratada.

Walter y Haines (1988) evalúan las relaciones que pueden existir entre las percepciones sobre seguridad el trabajador, sus conocimientos sobre salud y seguridad, y su implicación en las

medidas de seguridad con la ocurrencia de accidentes. La hipótesis de los autores consiste en que la participación de los trabajadores en la toma de decisiones y su conocimiento de las fuentes de riesgo y de sus derechos legales estarán negativamente relacionados con los accidentes. Se buscaba apoyo a la hipótesis partiendo de la base de que los trabajadores en empresas con sindicalistas encargados del tema de seguridad estarían más implicados y conocerían más. Sus resultados, aunque no concluyentes indican que, efectivamente, los niveles de conocimiento y la participación son mayores en esta situación.

### **3.2. Estudios correlacionales**

La característica fundamental de este bloque de estudios correlacionales es que tratan de ofrecer medidas de asociación univariadas o multivariadas entre alguna medida de accidentabilidad y factores que se asumen relacionados, dentro de un contexto de campo. Otros trabajos no se relacionan directamente con una medida de ocurrencia de accidentes, sino que se trata de estudios sobre percepciones de los trabajadores sobre la seguridad o de clima de seguridad, presentando medidas y estudiando sus características psicométricas.

El estudio de Porter y Corlett (1989) pretendía evaluar la existencia de un rasgo de propensión a accidentarse, y en qué podía consistir éste. Para la investigación, se administró un cuestionario a una muestra de estudiantes, y se recogieron 171 cumplimentados. Este cuestionario sirvió para dividir a los sujetos en dos grupos muy extremos: aquéllos con una tendencia a accidentarse; y aquéllos sin tendencia o propensión a accidentarse, siempre según el cuestionario. Tras esta clasificación, que pretendía poder controlar un factor como la exposición diferencial a riesgos de accidente y considerado por los autores como un posible factor confundente, se procedió a la fase experimental. Consistía ésta en una tarea de concentración y respuesta motora rápida. Se encontraron diferencias estadísticamente significativas entre ambos grupos, con los sujetos propensos a accidentarse teniendo una peor realización.

Melamed, Luz, Najenson, Jucha y Green (1989) realizan una investigación buscando la relación entre características personales y nivel ergonómico de estrés, con el nivel de ocurrencia de accidentes y/o de bajas por enfermedad.

El factor que estos autores denominan como nivel ergonómico de estrés consiste en un indicador compuesto de diversas

condiciones ambientales, en concreto: postura de trabajo y movilidad del cuerpo; nivel de esfuerzo físico; riesgos físicos, como maquinaria o herramientas peligrosas, etc; y estresores ambientales, entre los que se incluye temperatura (calor, frío o cambios bruscos de temperatura), ruido, polvo, humo, vibraciones, exposición a agentes químicos y pobre iluminación o iluminación inadecuada. En definitiva puede considerarse un indicador de riesgos físicos de accidente y/o enfermedad. Este índice compuesto mostró en diversas regresiones logísticas afectan positivamente a la ocurrencia de accidentes.

En cuanto a las características personales se probaron ciertas medidas somáticas (electrocardiogramas), nivel de satisfacción con el trabajo y la edad. Sólo el nivel de satisfacción laboral presentaba una relación estadísticamente significativa en una regresión logística para predecir ocurrencia de accidentes.

Estos dos predictores, nivel ergonómico de estrés y satisfacción laboral, también eran predictores al cambiar la variable dependiente, de ocurrencia de accidentes a nivel de bajas laborales.

Lógicamente, a pesar de que teóricamente parece claro que el nivel ergonómico de estrés es un antecedente a la ocurrencia de accidentes y a las ausencias por enfermedad, no es tan claro este supuesto orden causal en el caso de la satisfacción laboral, que puede más bien verse como un consecuente.

En cualquier caso, hacia lo que apuntan los resultados de Melamed et al. (1989) es hacia la integración en un momento dado de ciertas características del trabajo, niveles de riesgo y estresores ambientales con características personales para aumentar la probabilidad de tener accidentes y/o bajas laborales.

Otros factores específicos han sido relacionados con la probabilidad de accidentarse. Muchos autores han presentado la edad como un predictor importante de los accidentes. Este resultado de relación puede deberse, no obstante, a múltiples factores confundentes, como por ejemplo que las personas con más antigüedad en la empresa, y por tanto mayores, son colocados en los puestos de mayor responsabilidad y riesgos. Un reciente trabajo de Salminen (1993) trataba de ahondar en posibles explicaciones de la relación edad-accidentes. Este autor recoge los estudios sobre edad y accidentes realizados hasta la fecha para poner de manifiesto los resultados contradictorios. Así, algunos trabajos señalan que los trabajadores más jóvenes tienen menos accidentes, mientras otros no encuentran relación entre ambas variables y otros encuentran la relación contraria. Los resultados que sí son consistentes son los que se refieren a la gravedad de los accidentes, parece que los trabajadores mayores tienen accidentes más severos. En este contexto el autor pretendía encontrar posibles variables que explicaran el porqué de esta mayor severidad. En una muestra de

102 accidentados de todas las ramas industriales encontró que los accidentes severos de las personas mayores de 50 años estaban en una proporción muy elevada con problemas del ritmo de trabajo y de vehículos en movimiento. Parecía que los trabajadores mayores tenían una peor acomodación a ritmos rápidos de trabajo e infraestimaban la velocidad de los vehículos rápidos.

Un trabajo de Goodman y Garber (1988) explora las relaciones entre absentismo y ocurrencia de accidentes en una cadena de cinco minas de carbón con un total de 2238 mineros, por tanto en un contexto de alto riesgo de accidente. Los autores encuentran una relación entre los niveles de absentismo y el número de accidentes. Plantean que es posible que los accidentes sean en parte una consecuencia del nivel de absentismo. No obstante, plantean un modelo más complejo con una variable mediadora que explicaría la relación. En concreto proponen el concepto de falta de familiaridad como variable mediadora. La hipótesis es que un elevado absentismo lleva a una falta de familiaridad con la tarea a desempeñar y consecuentemente a un aumento de la probabilidad de accidentarse. El concepto de falta de familiaridad de hecho tiene mucho que ver con el de experiencia y formación que otros muchos investigadores han apuntado como involucrado en los accidentes laborales. Así, operacionalizan la falta de familiaridad como una falta de conocimiento sobre las máquinas y herramientas, sobre el ambiente físico de trabajo y sobre los hábitos de los mineros. En definitiva una formación inadecuada, aunque informal en este caso. Los resultados apuntan a que la hipótesis propuesta es razonable, pero únicamente para niveles altos de absentismo y de falta de familiaridad. En cualquier caso, los propios autores indican que a pesar de haber encontrado cierto apoyo a la hipótesis no consideran que el absentismo sea uno de los factores primariamente relacionados con los accidentes, sino un factor más. La falta de formación puede estar más relacionada con otros aspectos organizacionales, como la planificación, los tipos de contratación comunes, el clima de la empresa que con el nivel de absentismo, y por ello sólo a elevados niveles de absentismo se encuentran relaciones importantes.

El trabajo de Leather (1988) atiende al estudio de las actitudes de los trabajadores en torno a la seguridad, y en la comparación de estas actitudes en dos contextos laborales diferentes, uno formado por empresas de construcción privadas y otras públicas. La muestra del estudio consistía en un total de 132 trabajadores, 65 de sector privado y 67 de sector público. Para la medición de las posibles diferencias entre las dos poblaciones se empleó un cuestionario que medía el *locus de control* de los accidentes. Los resultados indican hacia diferencias significativas entre la percepción de la causalidad

de los accidentes, con el sector público con un mayor locus de control externo, y el privado con locus de control interno. Además se realizaron análisis ítem a ítem, de forma que se podían evaluar algunas diferencias finas entre ambas. Los factores más discriminativos entre ambas poblaciones eran el referido al efecto de las primas en la seguridad y la importancia de la experiencia para evitar los accidentes. Los trabajadores de sector público consideran que el abono de primas por trabajar aprisa no tiene porqué incrementar la probabilidad de accidentarse, mientras los trabajadores del sector privado tienden a conceder más importancia a este factor en la ocurrencia de siniestros. Con respecto a la importancia de la experiencia para evitar los accidentes, los trabajadores del sector privado le conceden mayor importancia que los del sector público. Ambos sectores se muestran de acuerdo en la importancia que tienen el grupo de trabajadores y los supervisores para aprender las formas de conducta segura, y para evitar riesgos. Además, consideran que los factores organizacionales, como los antes mencionados y la política de la empresa, junto con la percepción de los riesgos laborales contribuyen fuertemente a la generación de las actitudes y conductas de seguridad del trabajador.

Algunas investigaciones correlacionales han tratado de definir y medir el clima de seguridad, que se supone es uno de los factores organizacionales más importantes afectando a la ocurrencia de los accidentes. Estos trabajos, que arrancan del artículo de Zohar de 1980, suponen que hay que definir un constructo de clima específico de seguridad, ya que estas percepciones del trabajador de los diversos factores que incluye el concepto se supone guía su conducta de seguridad en la organización, y por tanto "cuando se operacionalice y valide puede servir como una herramienta útil para entender la conducta organizacional" (Zohar, 1980).

Lo que hace Zohar (1980) es aplicar el concepto de clima recogido en definiciones como la de Schneider (1973) al campo específico de la seguridad. Así, el clima de seguridad se referiría a las percepciones del trabajador sobre los diversos aspectos de seguridad que se dan en su ámbito de trabajo. Lo que propone Zohar es definir como dimensiones de seguridad los diversos aspectos de seguridad y medirlos mediante cuestionario. Definido de esta forma, el concepto de clima es holístico, en el sentido de que incluye todos los aspectos relevantes para la seguridad en una empresa y se convierte de hecho en un macro factor con aspectos diversos en su interior. El cuestionario total de 40 ítems fue sometido a validación, pero desafortunadamente no se buscaron relaciones con otras medidas de seguridad, ni con la frecuencia de ocurrencia de accidentes, ni con otros, no porque al autor no le pareciera relevante sino porque no pudo disponer de medidas de estas variables.

Posteriormente, el concepto de clima de seguridad empleado por Zohar y su cuestionario ha sido utilizado de diversas formas para generar nuevos cuestionarios de este mismo aspecto. Así, Brown y Holmes (1986) presentan un estudio sobre el cuestionario de Zohar. Este trabajo tiene un triple objetivo. Primero, realizar una replicación de la estructura factorial encontrada por Zohar. Segundo, establecer una estructura factorial válida, caso de que la encontrada por Zohar no se sostuviera. Y, tercero, dado que Zohar no había relacionado su medida de clima con ninguna variable externa, intenta relacionarlo con algún criterio externo.

Por lo que hace al primer propósito, y aunque los autores dicen que la estructura hallada por Zohar para los 40 ítems es teóricamente adecuada, no encuentran apoyo empírico al someterla a un análisis factorial confirmatorio.

El segundo propósito, encontrar una solución factorial que ajuste a los datos, se consigue sólo tras la realización de un algoritmo exploratorio que reduce a tres factores, pero con una reducción paralela en el número de ítems, hasta llegar a retener tres factores que explican diez ítems. De esta forma, el concepto de clima permanecía igual, pero era operacionalizado mediante una escala de longitud mucho menor.

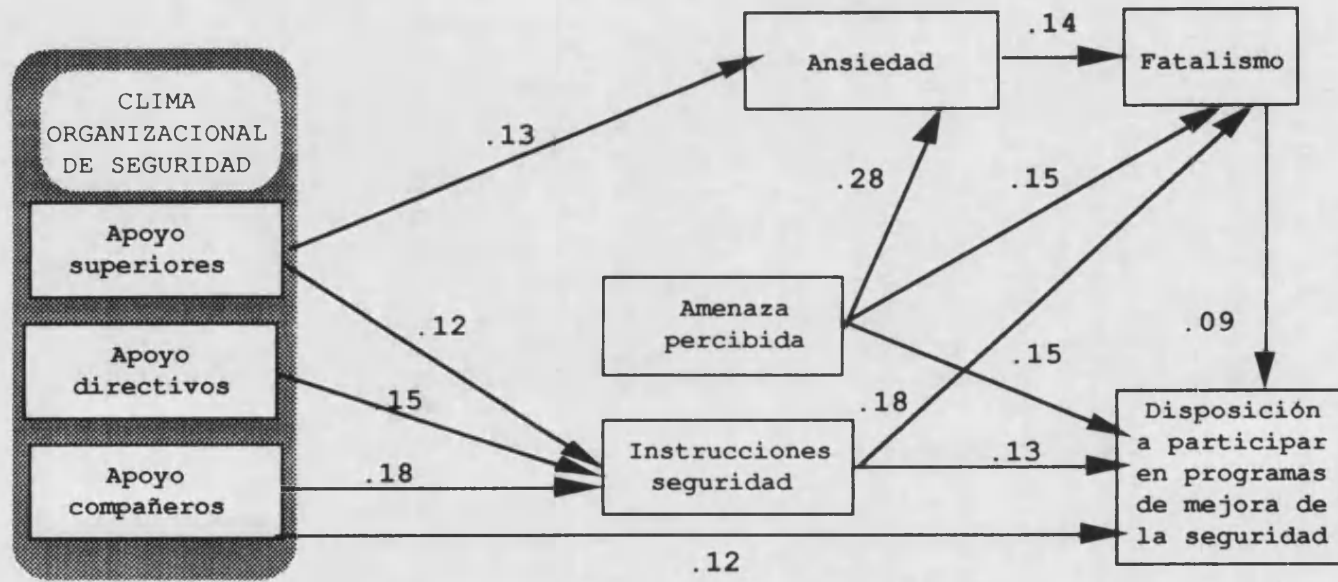
En cuanto al tercer objetivo, buscar relaciones con variables externas, los autores proponen el siguiente análisis: La muestra de sujetos se divide en dos grupos, uno de sujetos accidentados y otro sin accidentar. La hipótesis que planteaban era que los sujetos accidentados habían sufrido un proceso traumático que afectaría a su percepción del clima haciéndolo más negativo que la percepción de los no accidentados. Los resultados no daban apoyo a esta hipótesis. Por nuestro lado, opinamos que la hipótesis puede no ser cierta. No necesariamente la percepción del clima de los sujetos debe verse afectada en forma significativa por haber tenido un accidente. Es posible que al evaluar los aspectos de seguridad del entorno el sujeto sea lo suficientemente afectado como para no variarlos por haberse accidentado.

Basándose en los ítems considerados por Brown y Holmes (1986), Dedobbeleer y Béland (1991) presentan una replicación de la estructura factorial de cuestionario en una muestra de trabajadores de la construcción. Estos autores modificaron algunos contenidos de los ítems para adaptarlos a la industria de la construcción. Además, se elimina un ítem estadísticamente problemático. El resultado es que se replica mediante análisis factorial confirmatorio la estructura encontrada por Brown y Holmes, pero sólo mediante un tipo de estimación. Este trabajo sobre el clima de seguridad tampoco relaciona el constructo con ningún criterio externo.

El estudio de Leigh (1986) es un intento por descubrir y evaluar diversos predictores de los accidentes de trabajo. Las medidas que relaciona Leigh con la ocurrencia de accidentes son ciertas características personales y características del trabajo. La medida de accidentes era un indicador objetivo de pagos por lesión y enfermedades profesionales. Esta medida de accidentabilidad presenta dos problemas. Por un lado, no todos los accidentados cobran por daños. En segundo lugar, se incluye también dentro de la medida las enfermedades profesionales, un aspecto diferente a los accidentes y con causas propias. Mediante regresión logística se evaluó el impacto de 33 características personales y del puesto de trabajo sobre la ocurrencia de accidentes. De éstas, los factores personales que aparecen como relacionados son: sexo, las mujeres tienen menor número de accidentes; y estar recién divorciado, en menos de un año, que aumenta también la ocurrencia de accidentes. Obviamente, la relación de la variable sexo con ocurrencia de accidentes se deba a terceras variables, como el tipo de trabajo que desempeñan. En cuanto a las características de trabajo que se encontraron relacionadas tenemos: la sobrecarga de trabajo, positivamente relacionada con los accidentes; los trabajos de cuello azul frente a los de cuello blanco, con mayor número de accidentes los primeros; los trabajos de granja; y el nivel de riesgos del trabajo, positivamente relacionados con los accidentes.

Goldberg, Dar-El y Rubin (1991) presentan un interesante trabajo de modelos causales, a pesar de no tener como variable dependiente la ocurrencia de accidentes. El objetivo del trabajo es evaluar el impacto de diversas variables organizacionales y personales, así como cognitivas, en la intención de los trabajadores de participar en un programa de seguridad laboral. Las hipótesis a la base del trabajo es que la detección de los riesgos es una importante variable para controlar la ocurrencia de accidentes, y que si se da una percepción incorrecta de éstos los resultados son los accidentes. Así, los autores proponen que diversas variables organizacionales, que ellos denominan en conjunto clima de seguridad, pueden explicar la percepción de los riesgos, y que esta percepción, junto con el efecto de otras variables, ser un importante predictor de la participación en programas de mejora de la seguridad. Todas estas hipótesis globales y específicas se probaron mediante regresiones múltiples, para finalmente poner a prueba un modelo de 'path analysis', que finalmente quedó definido como se muestra en la gráfica 1.1.

Gráfica 1.1. Modelo causal planteado por Golberg, Dar-El y Rubin (1991).





Puede verse que las relaciones del modelo pueden considerarse en general como moderadas o bajas. Es importante darse cuenta de cómo queda definido el clima de seguridad, como apoyo a la seguridad de la dirección, de los supervisores y de los compañeros del trabajador. Curiosamente, estas tres variables que los autores consideran como aspectos o factores del clima no se hipotetiza correlacionadas. Todas ellas tienen un impacto directo sobre el nivel de cumplimiento de las normas de seguridad, resultado éste muy relevante, a pesar de que los coeficientes estructurales son bajos. Por otro lado, ninguna de las variables de clima de seguridad son capaces de predecir la percepción de los riesgos del trabajador, hipótesis manejada por los autores y para la que no hallan apoyo empírico. La percepción de los riesgos se convierte de esta forma en una variable exógena, que no es explicada por el modelo, pero con efecto sobre otras. Así, tienen una relación importante y positiva con el nivel de ansiedad y el fatalismo de los trabajadores, pero también con la participación en programas de seguridad. Aunque en la gráfica pueden consultarse los valores, destacar que la participación en programas de seguridad queda básicamente predicha directamente por los riesgos percibidos, las instrucciones de seguridad y el apoyo de los compañeros. Otros aspectos de clima afectan, pero más bien indirectamente.

A nuestro entender, y aunque los resultados de Goldberg et al. (1991) son interesantes y van en la línea de que riesgos percibidos y factores organizacionales son importantes predictores de la conducta de seguridad del trabajador, importantes deficiencias metodológicas están presentes en el trabajo. Las deficiencias metodológicas giran en torno al uso de los modelos causales y en concreto del análisis path. Aunque se han estimado los coeficientes estructurales del modelo presentado, no se ofrece ninguna medida de ajuste. Por un lado, no se ofrecen medidas de ajuste global del modelo a los datos, de forma que no sabemos si las relaciones planteadas son una buena representación de las covarianzas que existen entre las variables. Por otro lado, no se ofrecen tampoco índices de ajuste analítico, de forma que no sabemos nada sobre la significatividad estadística de los coeficientes estructurales. En este sentido, el modelo pierde una parte de su importancia y se hace difícil decidir sobre la importancia de la predicción de las variables dependientes.

Otros dos trabajos que buscan relaciones de diversas variables con los accidentes son de especial interés. Ambos suponen un paso más en la búsqueda de relaciones, no simplemente intentando una relación bivariada, sino intentando relaciones complejas teóricamente fundadas y en un marco multivariado. El primero de ellos lo presentan dos sociólogos, Dwyer y Raftery (1991), y el segundo es, a decir del autor (Hansen, 1989), el único modelo causal probado para la predicción de la ocurrencia de accidentes, y presenta importantes detalles a tener en cuenta.

En primer lugar analizaremos el trabajo de Dwyer y Raftery (1991). Estos autores presentan un trabajo que igual puede considerarse correlacional como puede considerarse teórico, dado que en ambos aspectos ofrece resultados relevantes.

Los autores parten de una conceptualización teórica que puede resumirse en una hipótesis simple: los accidentes están producidos por relaciones sociales en el trabajo. Esta hipótesis supone muchos aspectos específicos. En concreto, los autores afirman que las relaciones sociales tienen niveles, e identifican tres: nivel de recompensas, de control directo y organizacional. Además de factores sociales, relativamente externos a los trabajadores también incluyen como variable importante de la predicción de los accidentes el factor individual, en el que ellos consideran los actos inseguros de falta de cuidado, de despreocupación, etc.

Para operacionalizar estas hipótesis se midieron 31 variables. Entre ellas tenemos:

- Turno de trabajo, diario o nocturno.
- Tipo de turno: rotativo o fijo.
- Tipo de proceso: semiartesanal, semimecanizado y de proceso continuo.
- Estilo de dirección.
- Productividad.
- Coste potencial de los accidentes.
- Materiales de riesgo.
- Peligro físico.
- Dispositivos de seguridad.
- Una medida de peligro del proceso automatizado.
- Planificación de seguridad.
- Existencia de inspecciones de seguridad.
- Existencia de comités de seguridad.
- Existencia de encargados de seguridad.
- Formación.
- Nivel de recompensas.
- Nivel organizacional.
- Nivel de control.
- Nivel individual.
- Nivel de control de los trabajadores.
- Incentivos.
- Cualificación profesional.
- Primeros auxilios.

Los resultados de la predicción de la ocurrencia de los accidentes tras diversas regresiones logísticas ofrecieron el mejor ajuste con un modelo de tres variables. Este modelo con una capacidad predictiva de 85%, según los autores, incluía: peligro de los materiales, una variable compuesta que intentaba ser un indicador de autocontrol de los trabajadores, y un indicador compuesto que

intenta medir nivel de control organizacional por medio de recompensas y control directo. Veamos exactamente qué miden estas variables:

Las tres variables que consiguen una adecuada explicación de los accidentes son indicadores compuestos extraídos en base a las cuatro variables que en análisis previos se muestran más relacionadas con la accidentabilidad. Estas variables son:

Variable 1: peligro de los materiales, con siete niveles de peligro, desde 1 a 7. Como siete eran las plantas industriales que componían la muestra se hizo un ranking de peligro entre las siete.

Variable 2: Nivel de autocontrol del trabajador, con dos niveles 1= alto y 2= bajo.

Variable 3: Peso del nivel organizacional de recompensas, con otros dos niveles: 1= bajo y 2= alto.

Variable 4: Nivel de control organizacional, con dos niveles: 1= bajo y 2= alto.

Se combinaron estas variables para dar lugar a las tres variables que predecían el 85% de la ocurrencia de los accidentes. La primera variable de la regresión logística coincide con la variable 1 citada justo arriba, no es un compuesto. La segunda variable de la regresión logística es un indicador compuesto, con dos niveles y que se forma así: 1 si la variable 2 es 1, la variable 3 es 1 y la variable 4 es 1; 0 en cualquier otro caso. De esta forma, queda definida como amplio autocontrol del individuo, ya que es alta si el autocontrol es elevado y el control de la organización en recompensas o control directo es bajo. La tercera variable es justo el contrario, se forma de manera inversa. Es uno si el autocontrol es bajo (variable 2= 2), y el control de la organización es alto (variables 3 y 4= 2).

Así definidas las variables predictoras, la primera impresión es realmente buena. Se consigue una capacidad predictiva nunca vista en estudios sobre predicción de accidentes e incluso, a nuestro entender, difícilísima de obtener en trabajos realizados desde las ciencias sociales. Aunque el procedimiento metodológico parece correcto, tenemos algunas críticas que podrían explicar en cierta forma este aparatoso resultado.

La operacionalización de las variables no es adecuada desde nuestro punto de vista. Por un lado, parece que la medida de riesgos es un poco débil. Más importante es, no obstante, la operacionalización de las otras tres variables, de la 2 a la 4. Se restringe a dos valores, cuando a nuestro entender estos conceptos son complejos y presentan una variabilidad mucho mayor. Se han hecho extremos los valores, forzando a dos valores además, procedimiento no muy recomendable metodológicamente. Esta mala

operacionalización se ha extremado considerablemente al definir las variables que son indicadores compuestos. Su propia definición define dos grupos muy extremos en alto y bajo control de la organización y en alto y bajo autocontrol. Se ha agudizado la tendencia a la dicotomía, expresando variables complejas aún más restringidamente.

Por último, las dos variables últimas de la regresión logística final, esto es, los indicadores compuestos es muy posible que estén altamente correlacionados, por lo que puede producirse un problema de multicolinealidad. Efectivamente, los trabajadores con alto autocontrol difícilmente tendrán un control excesivo por parte de la organización y al contrario, por lo que el uso de las dos variables puede considerarse hasta redundante.

El último trabajo que revisaremos en el apartado de metodología correlacional es un modelo causal de los accidentes laborales presentado por Hansen (1989). El objetivo del trabajo de Hansen es evaluar la relación entre ciertos factores personales, cognitivos y de personalidad con la accidentabilidad. Aunque esto no es novedoso, lo que sí es nuevo es el uso por primera vez en la literatura psicológica de los modelos de path analysis para construir un modelo sobre la causalidad de los accidentes.

Las variables consideradas en este trabajo y que entraron en el modelo final de causación de accidentes fueron:

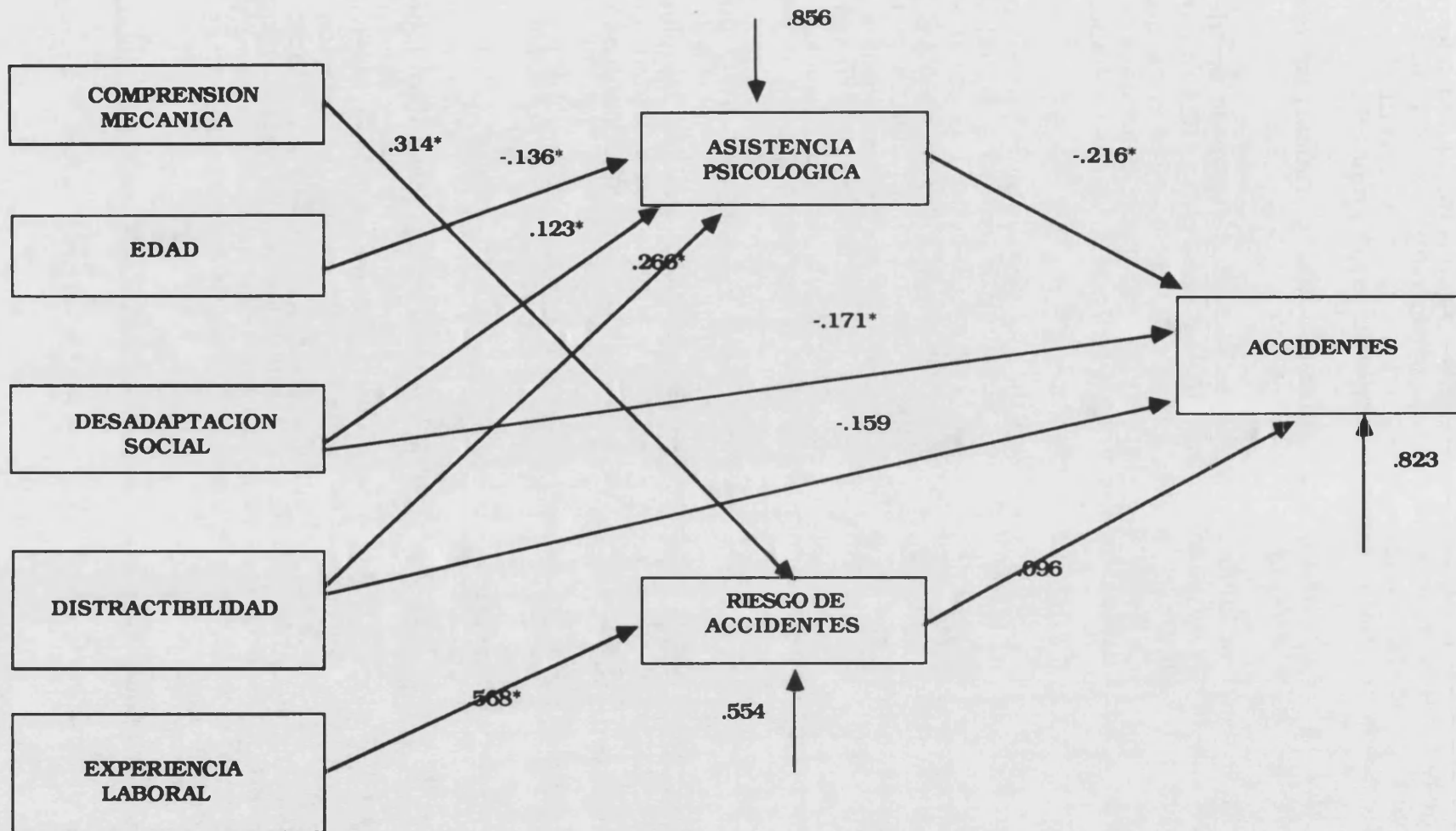
- El test de Bennett sobre *comprensión mecánica*, como medida de esta capacidad cognitiva de los trabajadores.
- *Edad*. La hipótesis de los autores es que los más jóvenes tendrán mayor número de accidentes.
- Una escala de *mal ajuste social general*. Esta escala la componen ciertos items extraídos del MMPI.
- Una escala de *distractibilidad*, que se extrae de items del MMPI. En opinión de los autores la incapacidad de atención a la tarea, síntoma de neurosis puede estar implicada en la ocurrencia de accidentes.
- *Experiencia laboral*. Medida como el número de meses que el trabajador lleva trabajando para la empresa.
- *Asistencia psicológica*. Es una medida sobre si se está en tratamiento psicológico de algún tipo, se codificó como 0 ó 1, con sin asistencia psicológica.

- *Riesgo de accidente.* Para la medición utiliza una medida de riesgo llevada a cabo por las empresas, que en realidad evalúa no los riesgos del puesto sino la responsabilidad individual en el accidente.

- Como variable dependiente final, se plantea la *consistencia de los accidentes*. Es un indicador de *accidentabilidad* compuesto de número de accidentes más el número de años en que se ha tenido al menos un accidente.

El autor realizó una serie de hipótesis de relación entre variables que fueron puestas a prueba, algunas variables quedaron eliminadas y de las ocho que mantuvo el mapa de relaciones hallado puede verse en la gráfica 1.2.

Gráfica 2.1. Modelo causal planteado por Hansen (1989). (\*= p<0.05)



En cuanto al ajuste global, puede decirse que el modelo representa adecuadamente a los datos, ya que la ji-cuadrado no es estadísticamente significativa. Las medidas de ajuste analítico que indican si los coeficientes estructurales son estadísticamente significativos uno a uno también se ofrecen en la gráfica 1.2.

Los resultados son muy interesantes y muestran ciertas tendencias sorprendentes.

El primer resultado sorprendente es que el riesgo de accidentes no sea un predictor significativo de los accidentes. Resulta difícil aceptar que los riesgos del medio no estén relacionados en alguna medida con la siniestralidad. Parece que este resultado puede deberse a una medida de los riesgos deficiente, como ya habíamos apuntado.

Es también sorprendente que los dos únicos predictores directos de la ocurrencia de los accidentes sean dos medidas de desajuste emocional o social. Es también llamativo el consistente signo negativo de la relación de la escala de desajuste social y la escala de distractibilidad, ambas extraídas del MMPI, con la variable de accidentabilidad. Tal y como han sido medidas, esto indica que cuanto más distractibilidad se tiene, menor es el riesgo de accidentarse; y por lo que hace a la escala de desajuste social, que cuanto mayor es éste, menor el número de accidentes.

Cuando se controla la experiencia y otros factores, puede apreciarse cómo la edad no tiene relación con los accidentes. Sin embargo, la variable experiencia que vendría a explicar la relación bivariada entre edad y accidentes tampoco se muestra como predictor de la siniestralidad, contrariamente a lo encontrado en otras investigaciones.

### **3.3. Estudios teóricos**

La mayoría de los estudios teóricos sobre seguridad laboral realizados desde las ciencias sociales pueden inscribirse dentro de ciertos modelos. Entre otros se pueden considerar (Meliá, Ricarte y Arnedo, 1993): modelos de secuencia de dominó; modelos de liberación de energía; modelos cognitivos; modelos de estado del sistema; modelos de factores situacionales y de error humano; modelos de perspectiva sociológica y modelos de factores cognitivos y de personalidad. Repasaremos brevemente las características de cada uno de ellos.

### **3.3.1. Modelos de secuencia de dominó**

La hipótesis básica de estos modelos es que un accidente se produce por una serie de factores concatenados, siendo el resultado final de éstos. En analogía con las fichas de dominó, cada factor es potencialmente un causante de accidentes, pero también puede ser un punto en que la cadena se corta y se evita el accidente.

El primer modelo en esta línea lo presenta Heinrich (1931). Considera seis factores en el modelo: la herencia y entorno social; el fallo de la persona; el acto inseguro o riesgo físico-mecánico; el accidente; y, como consecuencia del accidente la lesión.

Este modelo ha sufrido diversas revisiones y ampliaciones, como la de Bird (1969) o Weaver (1971), pero quizá la de Adams (1976) es la de mayor interés para nosotros. Incluye nuevamente cinco pasos, pero la naturaleza de estos factores se centra más en variables organizacionales o de entorno laboral. Así, el primer factor sería organizacional, incluyendo metas y políticas de la organización, medidas de seguridad que plantea la organización, etc. El segundo factor es también organizacional, pero basado más en errores operacionales, no tan a largo plazo sino errores de supervisión, más concretos. El tercer paso de la cadena son los errores tácticos causa directa de los accidentes, concretamente las conductas inseguras del trabajador y los riesgos. El resultado es un accidente, que puede tener como consecuencia una lesión.

En otros términos, lo que el modelo de Adams plantea es un primer nivel de clima de seguridad, un segundo nivel de respuesta de los superiores, un tercer nivel de actos inseguros y riesgos laborales y, como consecuencia de todo un posible efecto, el accidente.

### **3.3.2. Modelos de liberación de energía**

Los modelos de liberación de energía son básicamente como los de secuencia de dominó, pero definiendo el accidente como consecuencia de una liberación de energía. El modelo de Zabetakis incluye también una delimitación más clara de los diversos factores que se dan en secuencia temporal. Este autor incluye tres factores causantes de los accidentes por orden temporal:

1) Causas básicas. Que tiene a su vez tres niveles: política de seguridad (producción y metas, selección de personal, entrenamiento, mantenimiento y gestión interna, etc); factores personales (motivación, capacidades, estado físico y psíquico, experiencia, formación); y factores del entorno (temperatura, presión, iluminación, ruido).



2) Causas indirectas. Que incluye actos inseguros y condiciones inseguras, factores muy similares a los propuestos por Adams como factores inmediatos al accidente.

3) A diferencia de otros modelos incluye un factor de causa directa que se define como la liberación de una energía eléctrica, química, termal o mecánica, que sería lo que propiamente conduce al accidente.

Como vemos este modelo puede considerarse como de secuencia de dominó y no es muy diferente al planteado por Adams, al que más bien complementa.

### **3.3.3. Modelos cognitivos**

Los modelos cognitivos intentan aplicar el paradigma de procesamiento de información al estudio de los accidentes laborales. Entienden estos modelos que en el desempeño laboral el sujeto es un procesador activo de la información, y que es un fallo en algún punto de este procesamiento el que provoca un accidente, o al menos aumenta la probabilidad de accidentarse. Las fases que se consideran son: percepción, cognición, valoración, decisión y ejecución (Anderson et al., 1978; Hale y Hale, 1970; Lawrence, 1974; Snyder y Knoblauch, 1971). Estos modelos confieren una, a nuestro entender, desmesurada importancia a la porción cognitiva de las acciones del sujeto. Los otros factores sólo se consideran como afectando a las cogniciones del sujeto, como factores que cambian las actitudes y percepciones de los trabajadores y sólo a través de éstos aumentan la probabilidad de accidentarse.

Más flexible es el modelo de Leather (1987). Aunque en el centro del modelo se coloca nuevamente al sujeto potencialmente accidentado como procesador activo de información. En cualquier caso, el autor considera que las actitudes, percepciones y conductas del trabajador interactúan con un número de factores organizacionales, sociales y los riesgos.

Así, la conducta del trabajador se ve regulada por sus actitudes y conductas, pero también por los agentes de su lugar de trabajo, los supervisores, los compañeros, la organización y los aspectos del ambiente laboral. En resumen, el modelo de Leather propone una multiplicidad de factores afectando a las cogniciones y conductas del trabajador que interactúan con el medio para generar un accidente.

### **3.3.4. Modelos de estado del sistema**

Los modelos parten de una concepción de los accidentes como suceso consecuencia de una desestabilización en el sistema, en la producción laboral, que normalmente se sitúa en homeostasis (McDonald, 1972). Las variables de sujeto, organizacionales y de ingeniería están en interacción y con equilibrio. Si uno de estos factores se desequilibra y el sistema no es capaz de efectuar algún movimiento compensatorio que lo vuelva a equilibrar se producirá el accidente (Kjellén, 1984).

El modelo de Smillie y Ayoub (1976) considera dos tipos de factores, primarios y secundarios. Los primarios son la información presentada y la percibida, las expectativas, la acción y la retroalimentación a la situación. Los factores secundarios comprenden aspectos como la edad, la experiencia y formación, fallos del equipo, etc. Cualquiera de estos factores puede desequilibrar el sistema.

### **3.3.5. Modelos de factores situacionales y de factor humano**

Estos modelos consideran el error humano como factor último, determinante inmediato de los accidentes laborales. No obstante esta observación, consideran que en la ocurrencia del error humano influyen un número de factores.

Dejoy (1990) considera tres grandes tipos de causas como determinantes del error humano:

Por un lado, la comunicación entre persona y máquina.

En segundo lugar el entorno, que incluiría aspectos tales como el ambiente físico, la adaptación a la tarea, etc. En tercer lugar la toma de decisiones. En este último gran factor da cabida a un número heterogéneo de variables como las predisposiciones del sujeto (actitudes, creencias, expectativas), la estimación del peligro, la capacitación profesional del trabajador, los refuerzos o castigos que se producen por el desempeño laboral, etc.

El autor presenta, junto a las causas del error humano, unas estrategias de control, que son verdaderas técnicas de prevención de accidentes. Las tres estrategias consideradas son las de ingeniería, las de autoprotección -como el reforzamiento de las conductas seguras-, y las técnicas organizativo-directivas -como la implicación de los supervisores en la seguridad o la mejora del clima-.

Los modelos teóricos ponen de manifiesto, al igual que la investigación empírica, la multiplicidad de factores que actúan antes y en el momento inmediatamente anterior de la generación de un accidente. Llamémosle accidente, liberación de energía, error humano o desequilibrio del sistema, la mayoría de los estudios revelan la posibilidad de que estos factores se conjuguen en cierto orden temporal. La investigación debería optar por modelos complejos donde se consideren efectos directos e indirectos sobre la probabilidad de accidentarse.

#### **4. Medición de factores implicados en la seguridad**

Hemos revisado hasta aquí algunos de los hallazgos empíricos y teorías propuestas para entender las causas de los accidentes y los factores relacionados. La lista de variables propuestas es enorme y difícilmente pueden introducirse todas en un diseño de investigación. No obstante aquí vamos a revisar diversas operacionalizaciones que se han dado en la literatura de variables que después vamos a utilizar en nuestro diseño de investigación.

##### **4.1. Clima de seguridad**

Es éste sin duda uno de los factores que siempre aparece definido de una u otra forma en casi todos los trabajos empíricos. A veces considerado en su totalidad o a veces en partes de su contenido, pero es frecuentemente utilizado. Esto es en parte porque la misma definición de clima de seguridad, que arranca del trabajo de Zohar (1980) es altamente inclusiva. Efectivamente este autor asimila el concepto de clima a las percepciones del trabajador de todos los aspectos de seguridad de su entorno laboral.

La operacionalización de una definición tan amplia pasa por la determinación de las distintas dimensiones de seguridad de la empresa. Tras una revisión de la literatura hasta ese punto propuso una escala de 40 items que intentaban recoger esas dimensiones que se habían mostrado relacionadas en alguna forma con los accidentes. Los cuarenta items respondían a siete dimensiones teóricas, y tenían una escala de respuesta de cinco puntos. Se sometieron los items a un análisis de componentes principales con rotación varimax. Los factores encontrados por orden de importancia en el porcentaje de varianza explicada:

1) Importancia percibida de los *programas de entrenamiento en seguridad*. Este factor incluye seis items y explica un 40.9% de la varianza total del cuestionario.

## Capítulo 1. Introducción.

---

2) *Actitudes percibidas de la dirección en seguridad.* Incluye nueve items y explica un 19.3% de la varianza total.

3) *Efecto percibido de la conducta segura en la promoción.* Lo definen siete items que explican un 10.6% de la varianza.

4) *Nivel de riesgos percibidos.* Cinco items que explican un 9.7% de la varianza.

5) *Efectos percibidos del ritmo de trabajo sobre la seguridad.* Medido mediante tres items, que explican en su conjunto un porcentaje de la varianza del 6.9%.

6) *Estatus percibido del encargado de seguridad.* Son cinco items que explican un 4.8% de la varianza total.

7) *Efectos percibidos de la conducta de seguridad sobre el estatus social.* Incluye dos items que explican un 4.4% de la varianza.

8) *Estatus percibido del comité de seguridad.* Incluye tres items con un porcentaje de varianza explicada del 3.4%.

Al margen del análisis de componentes principales el autor llevó a cabo otros análisis para validar el cuestionario. Para testar que el clima de seguridad es relativamente homogéneo dentro de cada compañía, con respecto al resto de compañías de la muestra se realizó un ANOVA que ofrecía apoyo a ésta hipótesis. Se comparaba la varianza entre compañías frente a la varianza intra, llegando a la conclusión de que la variabilidad entre compañías era significativamente mayor.

Se intentó también testar la siguiente hipótesis "el clima de seguridad percibido por los sujetos estará alta y negativamente relacionado con la frecuencia de accidentes y su severidad", pero el autor no dispuso de estos datos.

Por último, Zohar llevó a cabo un análisis discriminante para probar qué dimensiones eran máximamente discriminativas entre los niveles de clima de seguridad de las distintas factorías. Los resultados apuntan a que las siguientes dimensiones son las más discriminantes: importancia de los programas de entrenamiento; efecto del ritmo de trabajo en la seguridad; estatus del encargado; y estatus del comité de seguridad.

Las operacionalizaciones del clima empiezan en Zohar, pero no acaban en él. Basándose en la escala de Zohar se han generado dos cuestionarios más, depuraciones de éste.

El primero de ellos es el propuesto por Brown y Holmes. En un primer momento, estos autores pretendieron replicar en una muestra estadounidense la estructura factorial encontrada por Zohar en una muestra israelí. Cambiaron de método de análisis factorial, empleando factorial confirmatorio estimado por máxima verosimilitud, dado que la estructura era conocida aún cuando no había sido convenientemente replicada. El proceso de replicación llevó a un ajuste claramente inadecuado de esta estructura y supuso que los autores pusieran en marcha un algoritmo exploratorio llevado a cabo con LISREL para encontrar un ajuste adecuado.

El algoritmo encontró un buen ajuste de tres factores, pero sólo tras una reducción en el número total de ítems a un cuarto. Estos diez ítems que quedaron en la escala venían explicados por tres factores:

1) Percepción del trabajador de que la dirección está preocupada por la seguridad. Explica a seis ítems, uno de ellos también satura en el factor tres.

2) Percepción del trabajador de que la dirección toma parte activa en la seguridad. Explica dos ítems.

3) Percepción de los riesgos laborales. Incluye tres ítems, uno compartido con el factor 1.

La estructura permite la correlación entre los factores, pero sólo alcanzan la significatividad estadística la covariación entre los dos primeros, que indica que actitudes y conductas de la dirección se perciben por el trabajador como relacionadas, y la covarianza entre el factor uno y el tres.

Desde nuestro punto de vista, el trabajo de Brown y Holmes tiene una importante deficiencia. Se ha utilizado un algoritmo puramente estadístico para encontrar un buen ajuste. Este algoritmo es absolutamente exploratorio, pero se lleva a cabo mediante una técnica eminentemente confirmatoria. No es que nos parezca mal el uso exploratorio de los modelos causales, pero creemos necesaria una orientación teórica. La justificación de la eliminación de los ítems no tiene ninguna base teórica, y no se justifica desde ningún análisis de contenido. Es esto lo que no nos parece justificable, el operar únicamente en base a criterios estadísticos.

El segundo trabajo que recoge la aportación de Zohar es el de Dedobbeleer y Béland (1991). Realmente recoge los ítems de Brown y Holmes e intenta replicar la estructura factorial hallada por éstos. La escala original sufrió, no obstante, unos ligeros cambios de contenido para adaptarla a las empresas de la construcción a que iban destinados y además se eliminó un ítem. En estos términos

difícilmente puede hablarse de replicación de la estructura de Brown y Holmes. Los autores ponían también en entredicho el tipo de estimación utilizado por los últimos, dado que tratándose de escalas de respuesta tipo Likert donde difícilmente se cumple la normalidad multivariada parecía más conveniente usar mínimos cuadrados ponderados, no basados en este supuesto, que máxima verosimilitud.

Los resultados de Dedobbeleer y Béland apuntan a que la estructura trifactorial que se pretendía replicar ajustaba a los datos correctamente, tanto mediante máxima verosimilitud como mediante mínimos cuadrados ponderados ( $\chi^2=16.45$ ;  $gl=24$ ;  $p=0.871$ ). No obstante, otra estructura factorial hipotetizada por los autores del trabajo, bifactorial, ajustaba estadísticamente igual cuando se aplicaban mínimos cuadrados ponderados ( $\chi^2=21.19$ ;  $gl=26$ ;  $p=0.732$ ), y como resultaba más parsimoniosa se retuvo. Estos dos factores eran:

1) *Compromiso de la dirección en la seguridad.* Que incluye cinco items.

2) *Implicación de los trabajadores en seguridad.* Explica cuatro items.

Vamos a repasar los contenidos de los nueve items de la escala. De esta forma apreciaremos qué tipo de operacionalización de la definición del clima se está utilizando y cuál es el nivel de concreción de los items.

Item 1. ¿Cuán importantes son para la dirección de la empresa las prácticas seguras de los trabajadores? Es una medida de la actitud de la dirección sobre seguridad.

Item 2. ¿Cuánto se preocupan los supervisores y otros directivos por su seguridad? Otro indicador de la actitud de la dirección y los supervisores inmediatos hacia la seguridad.

Item 3. ¿Cuánto énfasis hace el supervisor en la prácticas seguras en el trabajo? Es un indicador de conducta de seguridad del supervisor. En un sentido amplio mide la respuesta de los supervisores en materia de seguridad.

Item 4. ¿Al entrar en el puesto de trabajo se le ofrecieron instrucciones sobre política de seguridad de la empresa? Es un indicador del nivel de información que se ofrece sobre seguridad.

Item 5. ¿Hay reuniones regulares sobre seguridad? Es una medida de la formación y entrenamiento sobre seguridad que se ofrece y su periodicidad.

Item 6. ¿Está disponible el equipo adecuado para realizar la tarea? Es una medida global de si existen deficiencias en el equipamiento, de forma que pueda suceder un accidente.

Item 7. ¿Cuánto control piensa usted que tiene sobre su seguridad en el trabajo?. Es una medida de control percibido sobre los accidentes.

Item 8. ¿Es tomar riesgos una parte de su trabajo? Es una medida de los riesgos que se ve obligado a afrontar el trabajador.

Item 9. ¿Cuál es la probabilidad de accidentarse que tiene usted en los próximos 12 meses? Es una medida de riesgo.

Como puede verse la operacionalización de clima es extensiva, en el sentido de que se incluyen aspectos muy diversos de la seguridad, como la política de la empresa, la existencia de equipos de seguridad, la respuesta de los supervisores, el control percibido, la conducta del trabajador y los niveles de riesgo. Además, es global en otro sentido. Las preguntas son muy generales, los indicadores son muy globales, de forma que algunas preguntas muestrean un concepto amplio, como es el equipo de trabajo en base sólo a un sí-no.

Un acercamiento nuevo a la medición del clima, es el que presenta Goldberg et al. (1991). Definen grandes factores del clima de seguridad, que miden aparte y los relacionan con el resto de variables de un modelo causal en base a hipótesis diferentes. Estos tres factores de clima se corresponden con tres grupos de referencia del trabajador en el entorno laboral. Así, se define un primer factor como acciones de seguridad de la dirección, un segundo de acciones de los supervisores inmediatos, y un tercer factor de acciones de seguridad de los compañeros del trabajador. Como los contenidos de los items de los factores dos y tres serán revisados en apartados separados (respuesta de los superiores y los compañeros, respectivamente), listaremos aquí los items de la escala que presentan Golberg et al. (1991) para medir la primera variable de clima, el apoyo de la dirección a la seguridad. Son tres items, con un coeficiente alfa de 0.68: (1) la dirección hace cosas sobre seguridad; (2) la dirección está dispuesta a emplear esfuerzos y dinero para aumentar la seguridad; (3) la dirección considera como muy serias las violaciones de las normas de seguridad aún cuando no conlleven una lesión.

Las medidas que hemos revisado hasta aquí pueden considerarse globales, pero de alguna forma intentan abarcar todos los aspectos que según los autores entran en la definición de clima. Este no es el

único tipo de medidas del clima que se han venido utilizando. En muchos trabajos se usan aspectos específicos de clima, muchas veces sin darle este nombre, como la formación de seguridad, etc.

Por ejemplo, Vojtecky y Schmitz (1986) incluyen en su trabajo una medida de formación-entrenamiento que se corresponde en su contenido con un factor de clima.

Algunos ejemplos de cuestiones de la escala son:

- ¿Recibe por parte de la compañía entrenamiento formal en salud y seguridad laboral?
- ¿Contiene el programa los aspectos legales sobre seguridad y salud laboral?

Y sigue el cuestionario con medidas aún más específicas como:

- ¿Es adecuado el material que se le ofrece en estos cursos?
- ¿Se utiliza evaluación formal en el curso?
- ¿Se utiliza la evaluación para cambiar los contenidos de los cursos?

Y así hasta un total de diecisiete preguntas. Como puede verse ésta es una aproximación más extensiva a uno de los factores del clima que en anteriores escalas revisadas queda medido en una pocas cuestiones generales.

Del mismo modo, diversas variables que se manejan en el trabajo de Dwyer y Raftery (1991) pueden considerarse como relativas a aspectos del clima, aunque en los análisis no se utilizan como tal medida de clima. Por ejemplo:

- Frecuencia de inspecciones de seguridad llevadas a cabo por el gobierno, con escala de respuesta de tres puntos desde por debajo de la media a por encima de la media.
- Existencia de encargados de seguridad: Como todas las empresas tenían, se asignó 1= especialista y 2= no especialista.
- Existencia de comité de seguridad. Con la siguiente escala de respuesta: 1= existe y es considerado efectivo; 2= existe y es considerado inefectivo; y 3= no existe.
- Existencia de incentivos. Con las siguientes alternativas: 0= no existen incentivos; 1= existen incentivos pero los trabajadores no aceptan trabajar peligrosamente para conseguirlos; 2= se ofrecen incentivos y los trabajadores aceptan trabajar peligrosamente para conseguirlos, pero sus acciones se supervisan por los supervisores;



3= los incentivos se ofrecen y los trabajadores actúan de forma insegura para conseguirlos.

Además presenta dos variables de formación y cualificación profesional.

- Nivel de cualificación profesional. Con dos niveles: 1= se requiere alta cualificación para el trabajo; 2= no se requiere alta cualificación.

- Entrenamiento, formación: 1= existen esquemas formales de entrenamiento; 2= no existen.

Como puede verse, de una u otra forma el concepto de clima aparece, bien en su totalidad, bien como variables específicas en numerosos trabajos empíricos.

#### **4.2. Respuesta de los superiores**

Sólo el trabajo de Goldberg ofrece una medida separada de la actividad y el apoyo que los supervisores directos ofrecen en materia de seguridad, aunque considerándolo un factor del clima.

La escala de apoyo del supervisor incluye 11 items, con un coeficiente alfa de 0.89. Los contenidos son: (1) conocimientos del supervisor sobre seguridad; (2) hace algo sobre la seguridad, comparado con otros; (3) sigue las reglas de seguridad; (4) intenta activamente de aumentar la seguridad; (5) actúa para mantener en condiciones el equipo; (6) entrena a los trabajadores en seguridad; (7) para el trabajo hasta que pasa la situación de riesgo; (8) es estricto sobre la limpieza y orden del entorno; (9) dice a los trabajadores qué actos inseguros están haciendo para evitarlos; (10) hace algo sobre las situaciones inseguras; (11) está preparado a hablar con los trabajadores que hacen actos inseguros.

El resto de trabajos suelen considerar la respuesta de los supervisores como una parte de clima que está identificada con la dirección y por tanto no se mide aparte.

En la escala de Zohar (1980) un factor coincide parcialmente con el contenido que puede tener la variable respuesta de los superiores. Este es el factor 2, que mediante nueve items mide las actitudes de la dirección en seguridad. Los contenidos engloban a los supervisores inmediatos y a la dirección sin distinguirlos.

Lo mismo ocurre en los factores 1 y 2 de la escala de Brown y Holmes (1986), que no diferencian a los supervisores inmediatos como vehículo causal diferente a la dirección.

En la escala de Dedobbeleer y Béland (1991) el factor 1, compromiso de la dirección en la seguridad, incluye dos ítems que hacen hincapié en el supervisor más que en la dirección. Los ítems son el dos y el tres, el ítem dos pregunta acerca de la actitud del supervisor y la dirección (haciendo, como se ve, la distinción entre ambos ámbitos) en seguridad, mientras el tres hace lo mismo pero cuestionando acerca de las acciones de seguridad.

Poco más específico se ha hecho para separar operacionalmente al importante grupo de supervisores de otros grupos de referencia como los compañeros o la dirección. Y ello a pesar de que muchos autores ponen de manifiesto su importante papel de refuerzo, castigo, control y modelado sobre los trabajadores (De Joy, 1985; Leather, 1987; Walter y Haines, 1988).

#### **4.3. Respuesta de los compañeros**

Al igual que ocurría en el caso de la respuesta de los superiores, ningún trabajo la ha medido como separada de otros aspectos de clima. Únicamente el citado trabajo de Golberg et al. (1991) incluye una escala para su medición. Esta escala incluye tres ítems y presenta un alfa de 0.54. Los contenidos de los ítems son:

(1) Los mejores compañeros tienen cuidado con la seguridad.

(2) Hay una aceptación general entre los compañeros hacia las normas de seguridad.

(3) Los trabajadores están dispuestos a parar a un compañero que realiza prácticas peligrosas.

Junto a esta escala, en diversos trabajos aparecen índices sueltos de medida de la conducta de los compañeros, pero aún así son poco numerosos. Por ejemplo, en el trabajo de Dwyer y Raftery (1991) una variable puede considerarse próxima a la respuesta de los compañeros, aunque en este caso se refiere a los trabajadores que forman parte de un sindicato. El contenido pregunta acerca de la frecuencia de actividades de los sindicalistas en la seguridad de la empresa.

#### **4.4. Medida de los riesgos laborales**

El análisis de los riesgos laborales es un campo extensísimo y multidisciplinar. No podemos abarcar ni una parte de los estudios que se han realizado para evaluar los riesgos, así que centraremos la atención sobre las medidas de riesgos percibidos más importantes que se han generado desde un punto de vista psicológico.

El tipo de instrumento de medición más común para la evaluación de riesgos en el entorno laboral ha sido el inventario de riesgos. La mayoría de estos inventarios no son siquiera una medida exclusiva de riesgos, sino que consideran las más de las veces la conducta hacia la seguridad e inclusive otras medidas de formación, etc. Especialmente la confusión de contenido entre riesgo y conducta es constante, lo que es relativamente normal dado que un acto inseguro genera a su vez un riesgo.

En una revisión de estos instrumentos en Alemania, Ruppert y Hoyos (1991) no encuentran ningún cuestionario o inventario exclusivo de riesgo. Sí encuentra factores de riesgos en escalas de amplio uso en el análisis de puestos.

El inventario TAI (Frieling, Kannheiser, Facaoaru, Wöcherl y Dürholt, 1984) contiene unos 2200. De ellos 22, un 1% cuestionan acerca de aspectos de seguridad. De entre ellos los contenidos más tratados son entrenamiento de conductas seguras, responsabilidad del trabajador en la seguridad, grado de cuidado del trabajador (que puede considerarse medida de conducta de seguridad) y fuentes de peligro, de riesgo.

El cuestionario elaborado por Frieling y Hoyos (1978) contiene 221 items sobre análisis del puesto, de los cuales el 6% se refieren a seguridad: existencia de medidas de seguridad y algunos items de exposición al riesgo.

En el cuestionario de Landau, Luczak, y Rohmert (1975) se presentan 15 items, de un total de 390, que se refieren a seguridad. Entre ellos algunos de exposición a riesgos, probabilidad de accidentarse o sufrir enfermedades y conductas inseguras.

El cuestionario de Bürgi (1976) "Analysis of Jobs and Job Demand" mide una única dimensión de seguridad, la exposición a los riesgos laborales.

Nohl y Thiemecke (1988) presentan un inventario que mide exclusivamente riesgos, pero los combina con operaciones sencillas, para dada la combinación poder evaluar la probabilidad de una lesión. Como puede verse también se puede apreciar aquí cómo la conducta segura está presente en la medición de los riesgos.

En el contexto anglosajón también se dispone de inventarios que incluyen análisis de riesgos laborales. Tuttle, Wood, Grether y Reed (1974) presentan un cuestionario, el "Diagnostic Safety Form", para medir diversos aspectos de la seguridad, incluyendo los riesgos.

Hale y Glendon (1987) presentan un inventario de riesgos que intenta evaluar cada fuente de riesgo potencial en base a tres dimensiones psicológicas: detección o percepción de riesgo, evaluación del riesgo, y acción contra el riesgo.

Uno de los intentos más actuales de presentar una medida de riesgo es el de Ruppert y Hoyos (1991). Presentan el cuestionario "Safety Diagnosis Questionnaire", que mide diversas dimensiones de seguridad, tres de los cuales son medidas de diversos aspectos de los riesgos laborales: la organización; cooperación y comunicación en el trabajo; actuación ante problemas de seguridad; planificación y prevención; riesgos y peligros; percepción y atención a los riesgos; juicio y previsión de riesgos.

Un último trabajo que revisaremos es la medición de los riesgos que hacen Melamed et al. (1989). Se incluyen en un cuestionario medidas de los siguientes riesgos y condiciones ambientales posibles causantes de accidentes:

- Postura corporal
- Esfuerzos físicos, como carga de materiales pesados, esfuerzos dinámicos, etc.
- Riesgos activos, como herramientas o máquinas peligrosas, tránsitos en el trabajo desprotegidos o mal preparados, escaleras, objetos no asegurados, derrumbamientos, etc.
- Estresores ambientales, como ruido, iluminación, calor, frío, polvo, etc.

Es muy importante valorar los intentos que desde la psicología se hace por la medición de la percepción de los riesgos que tienen los trabajadores, dado que se puede mostrar cómo estas percepciones no tienen excesiva relación con la evaluación que los expertos en riesgos hacen de éstos (Kraus y Slovic, 1988).



## 2. Metodología.

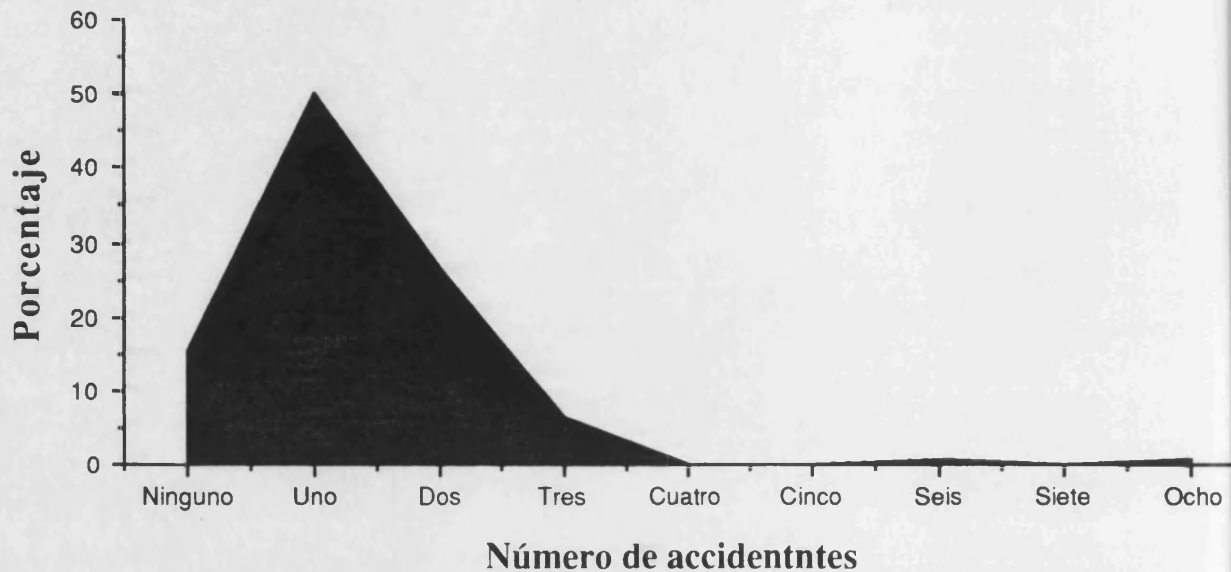
### **1. Descripción de la muestra**

La muestra está compuesta por un total de 124 trabajadores que acudían a los servicios médicos de Mutua General. Allí fueron entrevistados por personal de la línea de investigación de Seguridad Laboral, entre Julio y Septiembre de 1991. Los entrevistadores ofrecían la Bateria para ser contestada, pudiendo ser contestada sin ayuda, con ayuda o administrada directamente por los entrevistadores.

De los 124 protocolos válidos (algunos contestaron muy pocos items o dejaron la Bateria muy incompleta y fueron desechados) el 83.9% eran sujetos con accidentes y sólo 20 (el 16.1% restante) no habían sufrido ningún accidente.

La *gráfica 2.1* indica la distribución (en porcentajes) del número de accidentes sufridos por los entrevistados. Como puede apreciarse en la gráfica, la distribución es asimétrica positiva, oscila desde cero accidentes hasta ocho como puntuación máxima. La mayoría de la muestra ha tenido uno o dos accidentes.

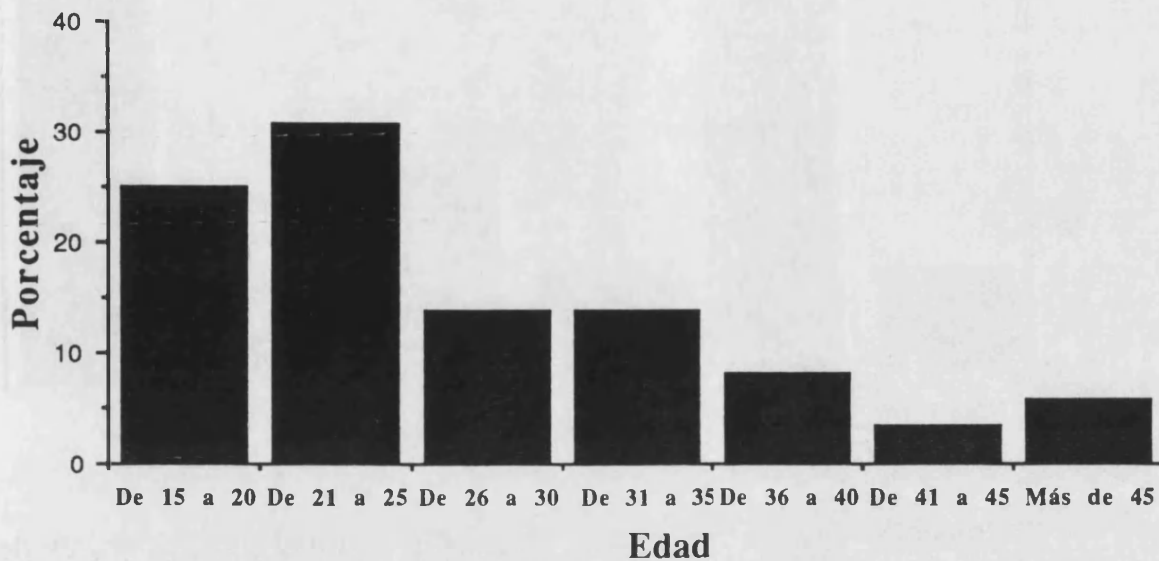
Gráfica 2.1. Porcentajes en la muestra para la variable número de accidentes.



En cuanto al sexo de los entrevistados, un 21.8% son mujeres, mientras que el 78.2% restante son hombres.

Las edades de las personas que componen la muestra se observan en la gráfica 2.2. Aquí las edades se presentan agrupadas en intervalos de cinco años desde el valor mínimo hallado de 15 y con el último intervalo abierto.

Gráfica 2.2 Edades de los trabajadores de la muestra.

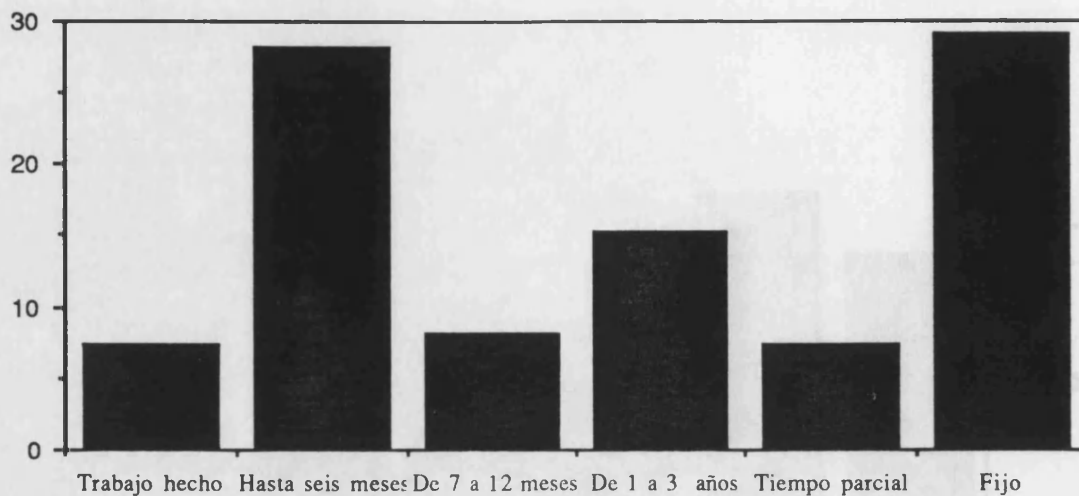


Su nivel de estudios fue como sigue: un 3.2% de los trabajadores no tiene ninguno o sólo lee o escribe; un 42.4% tiene estudios primarios; un 16.9% ha cursado formación profesional de primer grado; un 6.5% formación profesional de segundo grado; y un 20.2% tiene bachiller, BUP o estudios universitarios. Un sujeto no contestó la pregunta (0.8%).

Las personas que componían la muestra trabajaban bajo diversos tipos de contratos laborales. Los tipos de contrato pueden consultarse en la gráfica 2.3, que se presenta a continuación.



Gráfica 2.3. Tipo de contrato para las 124 personas de la muestra.



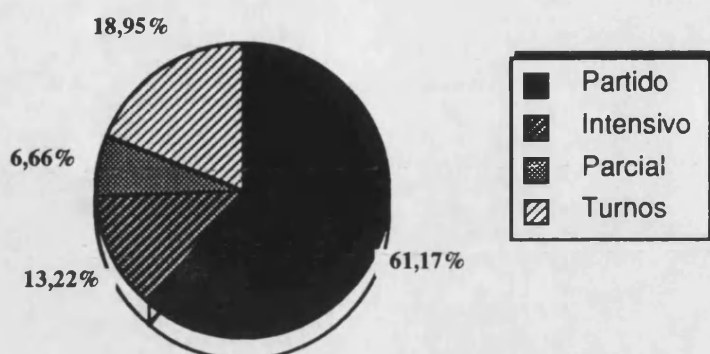
La posición que ocupan en la empresa, su nivel jerárquico, es de empleado en el 87.9% de los casos, siendo superiores o supervisores el 9.7% restante.

Los trabajos que realizaban antes tenían características muy diferentes a las del actual para un 44.4% de los entrevistados; este es el primer trabajo para un 19.4% de la muestra; es un trabajo igual que el anterior para el 13.7%; tenían un trabajo parecido anteriormente el 12.9%; un 4% ha estado siempre en la misma empresa pero en puestos diferentes; y siempre ha trabajado en el mismo trabajo otro 4%.

En los puestos de trabajo que actualmente ocupan los trabajadores era necesaria experiencia previa para un 33.1% de ellos.

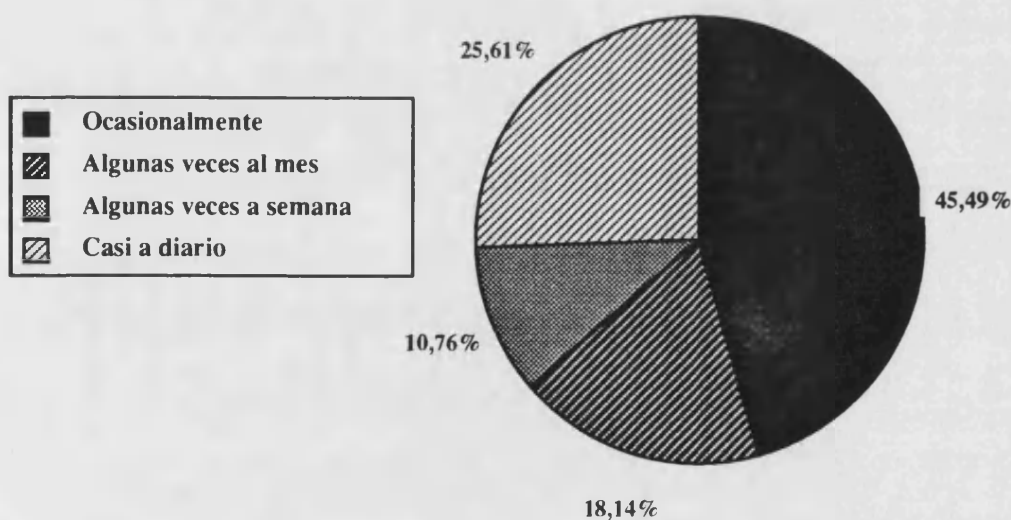
Los horarios de trabajo partidos (mañana/tarde) eran los mayoritarios, aunque la distribución para los diferentes tipos de horarios se presenta en la *gráfica 2.4*.

Gráfica 2.4. Tipos de horario para la muestra de trabajadores



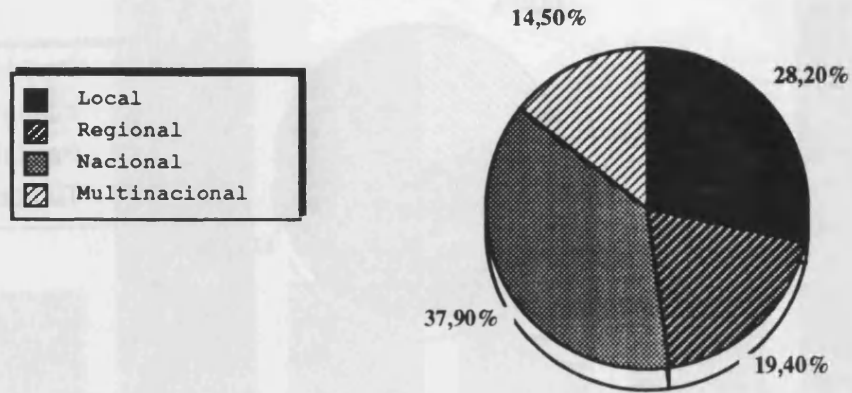
Un 44.4% no realizaban horas extraordinarias, mientras que un 25% las realiza casi a diario. La frecuencia de horas extraordinarias puede consultarse en la gráfica 2.5.

Gráfica 2.5. Frecuencia de horas extraordinarias en la muestra.



En cuanto a las características de las empresas en que trabajan los entrevistados tenemos el tipo de empresa y el ámbito de localización geográfica. El 24.2% de las empresas son públicas y un 73.4% privadas. El ámbito de localización puede consultarse en la gráfica 2.6.

Gráfica 2.6. Ambito de localización geográfica de las empresas en que trabajan los sujetos de la muestra.



## 2. Métodos y técnicas

En este punto se expone el conjunto de métodos y técnicas estadísticas y psicométricas que se han utilizado para el estudio de las variables y las relaciones entre ellas. No se intenta explicar las técnicas, que por otro lado son estándar y bien conocidas. Tan sólo se comentará el uso que de ellas se ha hecho. Se dedicará más espacio a las técnicas de modelado causal, y esto por varias razones. Porque son métodos más novedosas, sobre todo en el campo concreto de Seguridad Laboral. Porque estas técnicas son poderosas herramientas de contraste de hipótesis, con amplias ventajas frente a los métodos exploratorios para el análisis de cuestionarios, básicamente en lo que a fiabilidad, estructura factorial y validación de las medidas se refiere. Es en este sentido en el que se pretende utilizarlas aquí. Además, el modelado causal presenta algunos inconvenientes que es necesario matizar para evitar interpretaciones erróneas muy corrientes en la literatura.

### 2.1. Coeficientes de fiabilidad y homogeneidad

Para las variables utilizadas en el modelo que se miden mediante cuestionarios, se ha estudiado la fiabilidad desde el modelo clásico de tests, y desde los resultados hallados en los análisis factoriales confirmatorios. Los cálculos se han realizado para las escalas y para los items.

#### Fiabilidad de las escalas

La fiabilidad presenta dos acepciones, estabilidad temporal y consistencia interna de los items de la escala. En este trabajo, por no tener los datos necesarios para establecer la estabilidad temporal de las medidas, se hará un estudio únicamente de la consistencia interna. Los índices utilizados han sido:

a) *Alpha*. El coeficiente alpha es una medida de cuánto covarían los items de una escala entre sí. A resaltar que alpha suele aumentar al crecer el número de items de la escala, por lo que hay que observar el número de items al valorar un determinado índice alpha en una escala. Además de calcular el alpha para la escala total se calculará para los items pares y los items impares.

b) *Guttman-Rulon*. Los coeficientes de Guttman (1945), también conocido como Guttman-Flanagan, y de Rulon (1939) son algebraicamente equivalentes (Wilkinson, 1986) y de aquí el nombre de Guttman-Rulon. En el paquete estadístico SYSTAT se calcula como

la correlación entre los ítems pares e impares de la escala. Se asume para el cálculo que los ítems que forman las dos mitades son paralelos, excepto por tener distintas varianzas. Se considera un caso especial del alpha de Cronbach (Meliá, 1990; Muñiz, 1993; Santisteban, 1990).

c) *Spearman-Brown*. Las formulas de Spearman (1910) y Brown (1910) llegaban a una fórmula para conocer la fiabilidad de tests paralelos y predecir el nivel de fiabilidad que se podría alcanzar modificando la longitud del test, el número de ítems. Un caso particular de esta fórmula se empleará en este trabajo, calculando la fiabilidad de las dos mitades del test. Para el cálculo de este índice se asume que las dos mitades del tests (pares-impares) son paralelas.

d) *Correlación de las dos mitades*. Además, como coeficiente de la consistencia de las medidas, se calcula la correlación entre los ítems pares y los ítems impares del test (Wilkinson, 1986).

### **Fiabilidad y homogeneidad de los ítem**

Para los cálculos de la consistencia interna de los ítems se calculan los siguientes índices: a) la correlación de la puntuación del ítem con la puntuación total de la escala, un índice de discriminación del ítem; b) la correlación ítem total excluyendo del total el ítem estudiado, con lo que se evita la inflación que en el índice anterior aparece por dejar en el total de la escala el ítem bajo estudio; c) índice de fiabilidad del ítem; y d) el alpha excluyendo el ítem, que muestra el coeficiente alpha que tendría el test si ese ítem fuera excluido de la escala.

Además, para calcular la fiabilidad de los ítems se han utilizado dos procedimientos a través de los resultados de los análisis factoriales confirmatorios para las escalas. Long (1983) propone que la saturación entre el ítem y el factor latente al que pertenece elevada al cuadrado es una medida de la fiabilidad del ítem. Más depurado es el índice de fiabilidad propuesto por Bagozzi (1981), apoyándose en el trabajo de Werts, Linn y Jöreskog (1974). Este índice se calcula mediante la siguiente fórmula:

$$\rho_i = (\lambda_i^2 \text{ var } A) / (\lambda_i^2 \text{ var } A + \psi_i)$$

donde

$\rho_i$  = a la fiabilidad del ítem i;

$\lambda_i$  = la saturación (carga factorial) del ítem i con el factor A;

var A = a la varianza del factor A;

$\psi_i$  = a la varianza de error o factor único para el ítem i.

Esta fórmula puede también generalizarse tomando sumatorios de las saturaciones factoriales desde  $i=1$  hasta  $n$  para calcular la fiabilidad de los factores confirmatorios de la escala (Bagozzi, 1981).

## 2.2. Coeficientes de correlación y regresión

Para el estudio de la validez criterial, tanto de la escala como de los ítems, se acudirá al cálculo de correlaciones y regresión simple o múltiple, además de las técnicas de modelado causal.

## 2.3. Modelos causales

Siguiendo a Bentler (1980), podemos decir que la metodología que genéricamente se denomina modelos causales tiene un propósito, explicar las propiedades estadísticas (covarianzas o correlaciones) entre una serie de variables observables por medio, o a través de, una estructura causal hipotética. Esta estructura causal hipotetizada debe dar cuenta del máximo posible de las varianzas-covarianzas de las variables observables. Además, las técnicas de modelado causal se han venido denominando análisis confirmatorios en contraposición a los exploratorios dada su naturaleza de contraste de hipótesis, en que las relaciones hipotetizadas (el modelo) son o no confirmadas, se pueden o no sostener para los datos (Bentler, 1980; Jöreskog, 1969; Tukey, 1977).

No obstante, lo que se ha venido denominando modelos causales incluye diversos subtipos de métodos, que pueden incluir sólo variables observables o bien hipotetizar factores latentes (variables inobservables) subyaciendo a los datos. La división más utilizada distingue entre: a) modelos de ecuaciones estructurales, también conocidos como 'path analysis'; b) análisis factorial confirmatorio; y c) modelos de estructuras de covarianzas. Todos ellos conservan, no obstante, unas características comunes que iremos señalando. Entre estas características se incluye una nomenclatura homogénea que denomina a las variables predictoras o independientes como *exógenas*. Estas son aquellas a las que no se dirige ninguna flecha, aquellas que no son explicadas por ninguna otra variable del modelo, su explicación es externa. Las variables *endógenas*, por su parte, son aquellas a las que se dirige al menos una flecha desde otra variable del modelo. Esto es, son las variables dependientes, que pueden a su vez explicar otras.

### Modelos de ecuaciones estructurales

Desarrollados por el genetista Wright (1934) (En Duncan, 1966) y aplicados en los estudios de ciencias sociales por Duncan (1966), también se conocen como 'path analysis'. Esta técnica explica una o más variables observables mediante otras variables también observables. No incluye, por tanto factores latentes. Los supuestos del path analysis (Visauta, 1986) son:

a) Ser modelos o sistemas de variables cerrados o completos. Esto es, cada variable a endógena se asume completamente determinada por los predictores (causas) del modelo. Dado que este supuesto es muy restrictivo, una formulación más realista añade un término de error o parte no explicada a cada variable endógena.

b) *Linealidad*. Las relaciones que unen a las variables que lo forman deben ser lineales.

c) *Aditividad*.

d) *Medidas de intervalo*. Esta condición demasiado restrictiva no se mantiene. Existen técnicas diseñadas para variables en escala ordinal o nominal y, además, si se parte de la medición en base a un número razonable de items puede aplicarse cualquier tipo de estimación (Boyle, 1970).

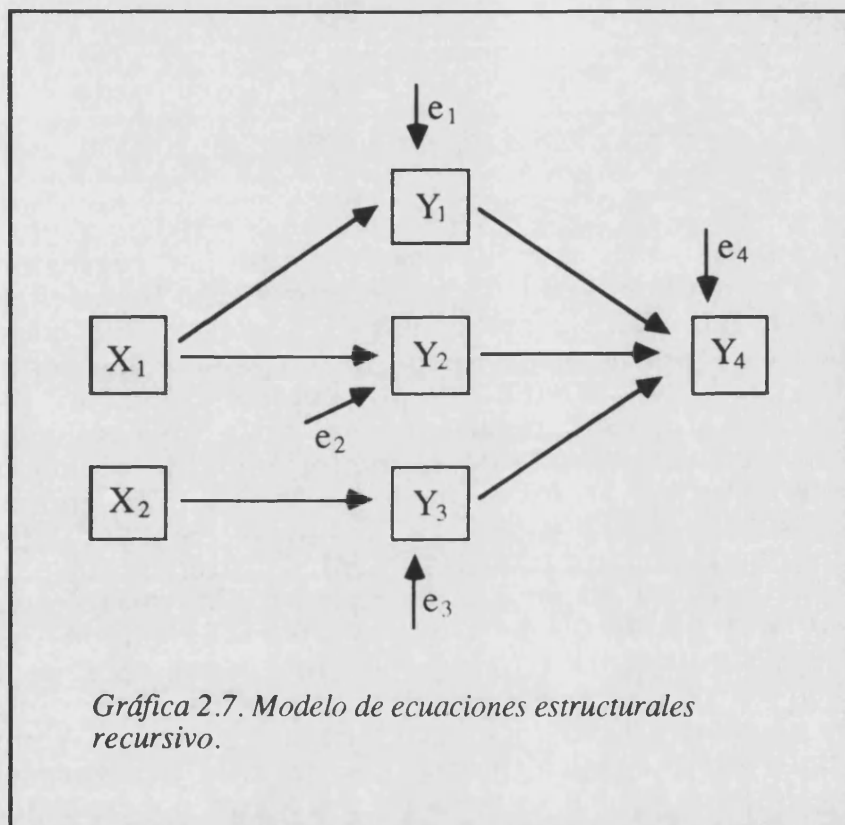
e) *Los errores deben estar incorrelacionados*. Su incumplimiento afecta a la estimación, pudiendo producir sobreestimaciones o subestimaciones.

f) *Las variables están medidas sin error* (Long, 1983). Este supuesto es incumplido prácticamente siempre al utilizar variables psicosociales medidas mediante escalas, precisamente el tipo de medidas utilizadas en este trabajo.

A este número de supuestos generales se añaden otros en función de la técnica de estimación que se utilice para calcular los coeficientes del modelo. Así, con la estimación máximo verosímil o de mínimos cuadrados generalizados se asume la normalidad multivariable. Otras técnicas de estimación han sido generadas para distribuciones elípticas o para estimar bajo cualquier tipo de distribución (estimadores arbitrarios) (Bentler, 1988). Sin embargo, también se desarrollan investigaciones sobre robustez frente a violaciones de los supuestos (Hausman, 1978; Kadane y Anderson, 1977).

La diferencia del path analysis con la regresión lineal múltiple es que en esta última tenemos una única variable observable predicha (criterio). Por su parte, el path analysis 'concatena' diversas regresiones múltiples y/o simples en una cadena ordenada de

causación. En palabras de Visauta (1986) se añade a la regresión una o más variables criterio y una dirección causal. A esto se suma, al igual que en el resto de modelos causales, su carácter confirmatorio de contrastación de hipótesis. Así que puede considerarse que las ecuaciones de regresión lineal son la pieza básica del path analysis (Bentler, 1980). Para ilustrar este punto, observemos la *gráfica 2.7*, donde se plantea un diagrama path que representa un posible modelo de ecuaciones estructurales con seis variables observables.



Gráfica 2.7. Modelo de ecuaciones estructurales recursivo.

En el diagrama path, las seis variables aparecen en el interior de un cuadrado para indicar que son variables observables. Las dos variables indicadas con una X son variables exógenas al modelo. No son predichas por ninguna otra variable del modelo, o de otro modo, sólo actúan en el path analysis como variables independientes. Las cuatro variables indicadas mediante Y son variables que sí son predichas por al menos una variable del modelo, o en términos gráficos al menos una flecha se dirige hacia ellas. Son, por tanto, variables endógenas. Además,  $Y_1$ ,  $Y_2$  e  $Y_3$  son a la vez predictoras (independientes), mientras que  $Y_4$  sólo actúa como variable dependiente. Cada variable endógena presenta un error asociado, la parte que no se puede explicar de cada variable en función de la o las variables predictoras. Como se aprecia en el diagrama, la cadena de causación fluye en un único sentido, sin ninguna conexión causal bidireccional, por lo que el modelo es recursivo.



El anterior modelo de path analysis puede descomponerse en cuatro ecuaciones de regresión, aquí llamadas ecuaciones estructurales. Tendremos una ecuación para explicar cada variable endógena del modelo:

$$Y_1 = b_1 X_1 + e_1$$

$$Y_2 = b_2 X_1 + e_2$$

$$Y_3 = b_3 X_2 + e_3$$

$$Y_4 = b_4 Y_1 + b_5 Y_2 + b_6 Y_3 + e_4$$

El punto de corte se elimina, ya que las regresiones se calculan puntuaciones diferenciales. Los parámetros  $b$  son los coeficientes de regresión que optimizan la predicción de las VD desde las variables independientes o predictoras. En path analysis se les conoce como coeficientes path o coeficientes estructurales. Por su parte, los términos  $e$  son errores o residuales, la parte no explicada de las variables endógenas. Además de estos dos tipos de coeficientes ( $b$  y  $e$ ), las varianzas y covarianzas de las variables independientes son también parámetros posibles a estimar (Bentler, 1988). Los coeficientes path estandarizados pueden interpretarse como la cantidad de variación en desviación típica de la variable endógena, cuando la independiente varía en una desviación estándar, permaneciendo invariable el resto (Visauta, 1986). Son coeficientes de regresión entre dos variables eliminando los efectos que sobre ellas puedan tener las otras variables del modelo (Wonnacott y Wonnacott, 1981). Por su parte, los residuales representan la proporción de varianza de la variable endógena no explicada por el modelo.

Ahora, para resolver el análisis path sólo queda estimar todos los parámetros y comprobar estadísticamente si el modelo hipotetizado da cuenta adecuadamente de las varianzas y covarianzas asociadas a las variables observadas. Como en los demás modelos causales esto se consigue mediante las técnicas de estimación, lo primero, y los índices de ajuste, lo segundo. Todo ello se tratará en sus puntos correspondientes.

### *Análisis factorial confirmatorio*

El segundo de los modelos causales es el modelo de medida, más conocido como análisis factorial confirmatorio (AFC). El AFC

surge de los trabajos de Jöreskog (1967, 69), y como mejor se entiende es por contraste frente al modelo de análisis factorial exploratorio. Los dos acercamientos comparten el objetivo general. Sirven para, mediante técnicas estadísticas, intentar describir un número de variables observables en función de un número (generalmente menor) de variables no observables o factores latentes.

En la siguiente gráfica puede verse un ejemplo de análisis factorial exploratorio. En este diagrama hay dos tipos de variables: unas denominadas con V y encerradas con un cuadrado, que son las variables observables; otras etiquetadas con F y rodeadas por un círculo, que son las variables no observables o factores. Como puede verse en la figura, a cada variable observable le llegan flechas. Esto es porque cada variable observada actúa como variable dependiente y se trata de explicarlas mediante un número más reducido de factores. Siempre queda una parte de varianza de las variables no explicada que se denomina factor único o error. Además de las relaciones entre F y V y V y E, tenemos en la figura las flechas de dos sentidos entre los factores. Estas flechas corresponden con la correlación o covariación entre los factores latentes. Los supuestos que el investigador debe asumir cuando usa análisis factorial exploratorio son (Long, 1983):

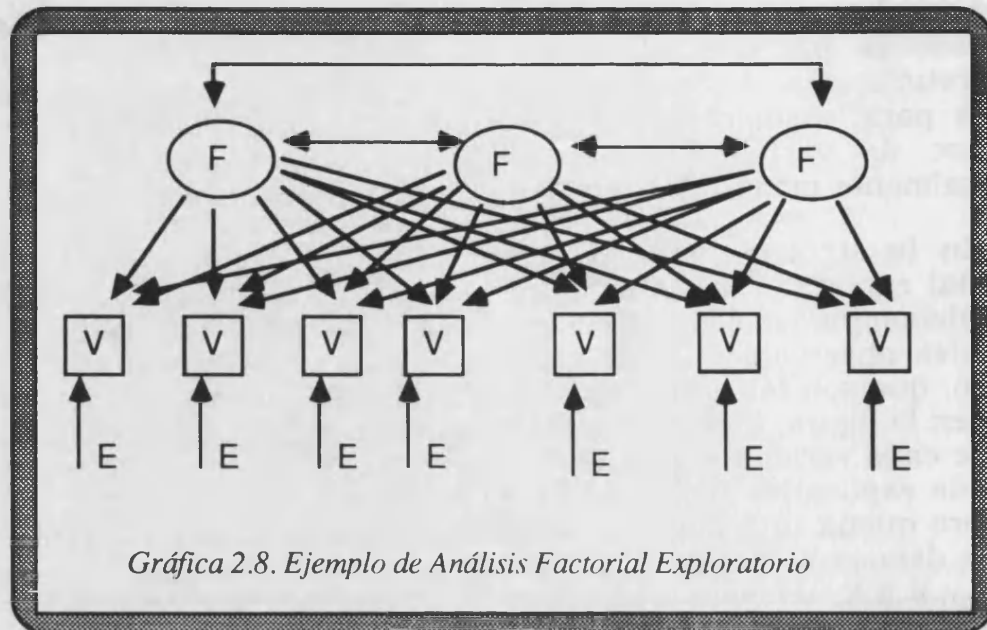
a) que todos los factores latentes, tras la rotación, están correlacionados entre sí, o en ciertos modelos de análisis factorial exploratorio que todos están no correlacionados entre sí (por ejemplo, al aplicar rotación varimax). No es posible que unos estén correlacionados y otros no.

b) que todas las variables observables se ven afectadas por todos los factores latentes.

c) que los factores únicos o errores no están correlacionados entre sí.

d) que todas las variables observables están afectadas por un error.

e) y que todos los errores están incorrelacionados con los factores latentes.



Gráfica 2.8. Ejemplo de Análisis Factorial Exploratorio

Aplicar la técnica implica calcular el modelo de esa forma. No importa qué conocimiento teórico tenga el investigador, no pueden hacerse *constricciones substantivamente motivadas*. Por ejemplo, no puede decirse que el primer factor explicará por sí solo la variable observable primera, o que el factor tres no está relacionado con el factor uno y sí con el dos.

La situación es radicalmente diferente en AFC, donde el investigador impone constricciones substantivamente motivadas (Long, 1983), teórica o prácticamente motivadas y las somete a un contrato estadístico. Las constricciones pueden determinar:

a) qué factor o factores afectan a cada variable observable. No necesariamente todas las variables observables se ven afectadas por todos los factores, como era el caso en análisis exploratorio.

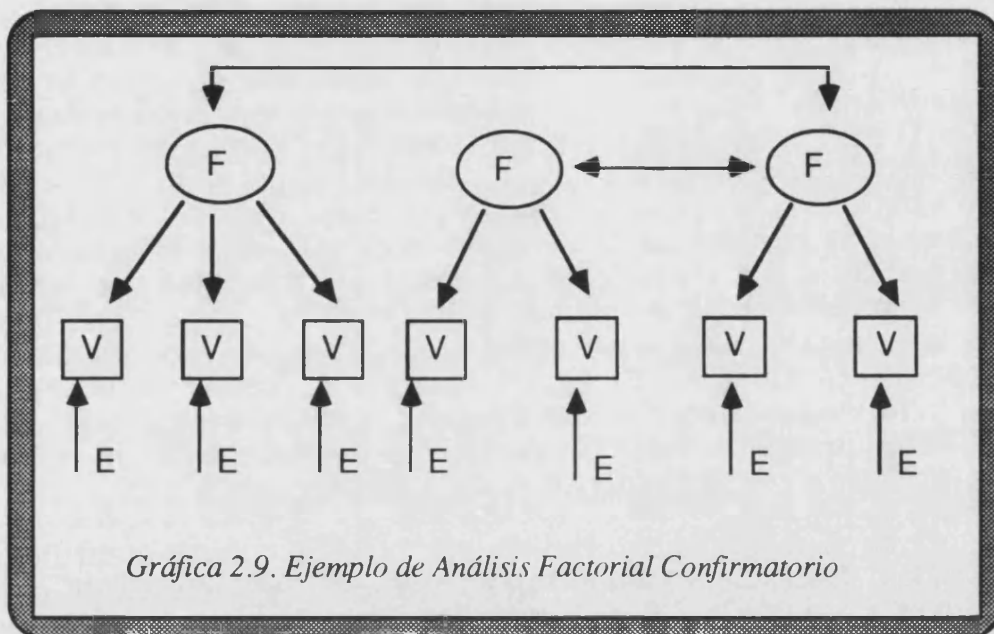
b) qué variables observables están afectadas por error y qué variables no lo están.

c) qué pares de factores covarían. Puede que algunos se planteen como correlacionados y otros no.

d) qué pares de errores están correlacionados entre sí (por ejemplo, debido a ser errores de la misma variable medidos en dos momentos temporales).

Un posible modelo de análisis factorial confirmatorio se presenta en la gráfica 2.9.

El análisis factorial confirmatorio, frente al exploratorio, usa del contraste estadístico para determinar si los datos observados en una muestra son o no consistentes con las constricciones impuestas o modelo hipotetizado.



Gráfica 2.9. Ejemplo de Análisis Factorial Confirmatorio

Antes de estimar cualquier modelo de análisis factorial confirmatorio hay que eliminar el problema de la *indeterminación de la escala* y establecer una métrica. Si la escala de los factores no se ha establecido, existe una indeterminación entre la varianza de un factor y las saturaciones de las variables observables en ese factor. Esto hace imposible distinguir entre el caso de que el factor tenga una gran varianza y las saturaciones sean pequeñas, o el caso en que las varianzas sean pequeñas y las saturaciones elevadas (Long, 1983). En otras palabras, tenemos un conjunto de pesos o saturaciones factoriales a estimar, que significan 'cuanto' cambio en la variable observable se produce por una unidad de cambio en un factor. Si el valor de esa saturación depende de las varianzas de los factores y éstas no están fijadas de antemano, es fácil entender que los dos tipos de valores están por estimar. La solución a este problema es doble (Bentler, 1988): a) puede fijarse la varianza de los factores (usualmente a la unidad); o b) fijar un peso o saturación de cada factor a un valor distinto de cero (generalmente la unidad).

Las saturaciones factoriales tienen, además, otra utilidad. En el contexto del análisis factorial confirmatorio los pesos de las variables pueden usarse como medida de la fiabilidad de los ítems (Long, 1983). La fiabilidad de un ítem se define como el cuadrado de la correlación con el factor en que satura. Si las variables observables se han tipificado y la varianza de los factores se ha fijado a uno, esta correlación coincide con la saturación factorial. Si elevamos al

cuadrado la saturación obtendremos, pues, una medida de su fiabilidad.

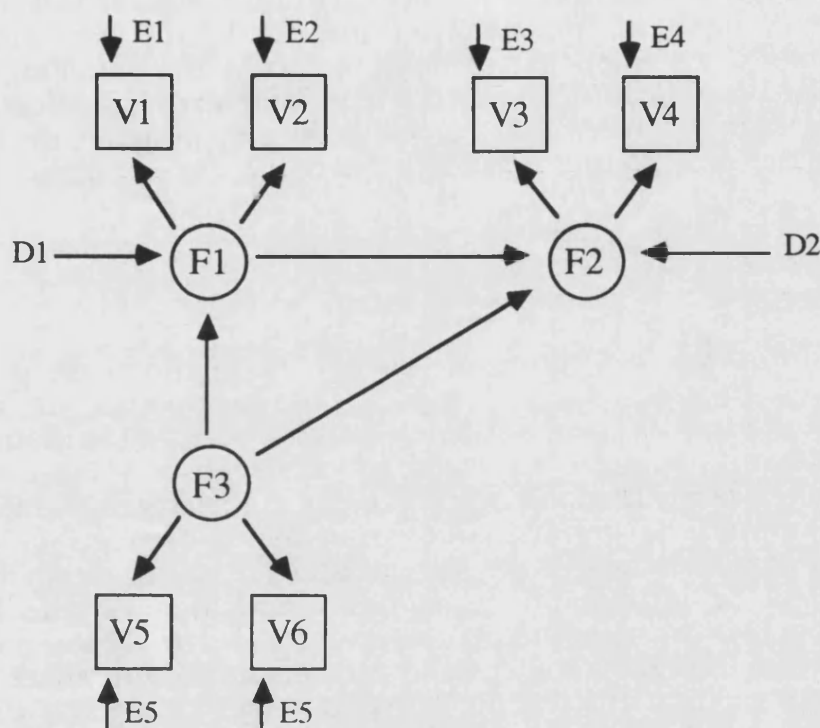
*Análisis Factorial Confirmatorio de segundo orden.* Igual que en análisis factorial exploratorio, en las técnicas confirmatorias de modelado causal es posible definir modelos factoriales más complejos, en los que subyaciendo a dos o más factores latentes se propone un factor de segundo orden que explica la covariación de los de primer orden. Esta técnica estadística arranca de los trabajos de Bentler (1976) y Bentler y Weeks (1980). De hecho, el número de orden de los factores (segundo, tercero, etc) a plantear sólo está en función de las suposiciones teóricas y/o empíricas del investigador, por ejemplo el número de factores latentes y las correlaciones entre ellos (Bentler, 1988). Si recordamos los supuestos de análisis factorial confirmatorio de primer orden, vemos que no era posible plantear una relación causal entre factores, sólo se podía plantear covariación entre ellos. El hecho de que el factorial de segundo orden plantee justamente este tipo de relación indica que está a mitad de camino entre el AFC que ya hemos visto y los modelos de estructuras de covarianza que veremos a continuación.

Para eliminar la indeterminación de la escala e identificar el modelo, las saturaciones factoriales entre los dos factores de primer orden y el de segundo deben hacerse iguales y la varianza del factor de segundo orden fijarse (usualmente a uno), como ya se hacía en AFC de primer orden con los factores. La restricción de igualdad de pesos es necesaria solo en los casos en que el factor de segundo orden tenga dos o tres indicadores (factores de primer orden) como en la figura (Bentler, 1988; Rindskopf y Rose, 1988).

### *Modelos de estructuras de covarianza*

Hemos visto cómo el análisis factorial confirmatorio utilizaba constructos latentes o factores en su formulación. No obstante, hemos visto también que no permitía una estructura de regresiones o ecuaciones estructurales entre estos factores (salvo el caso de AFC de segundo orden, a mitad de camino entre el AFC y los modelos de estructuras de covarianzas). La idea de permitir el cálculo de regresiones entre factores la introdujeron Jöreskog (1973) y Wiley (1973). No era más que conectar los dos tipos de técnicas conocidos hasta entonces en un modelo único. Se reunió en un solo modelo el path analysis y el análisis factorial confirmatorio, dejando que los factores se relacionaran no solo por flechas bidireccionales sino también unidireccionales. Lógicamente en el contexto de análisis confirmatorio la unidireccionalidad indicaba relación de causa-efecto. En otras palabras, factores que presentan sus respectivas variables observables como indicadores (AFC) pueden explicarse entre sí (path analysis). En la *gráfica 2.10* puede verse un ejemplo que aparece en

Bentler (1988) y que está extraído del trabajo de Wheaton, Muthén Alwin y Summers (1977).



Gráfica 2.10. Modelo de estructuras de covarianzas (en Bentler 1988)

Como se ve en la figura, el factor 3 actúa siempre como predictor, mientras que los factores 2 y 3 son predictores de sus indicadores empíricos como en AFC (las variables observables), pero al mismo tiempo actúan como criterio o dependientes. El factor 1 viene explicado por el factor 3 más un residual o parte no explicada, mientras que el factor 2 se explica por dos predictores, los factores 1 y 3 más un residual o parte de error.

Para identificar el modelo y eliminar la indeterminación de la escala es necesario fijar una saturación factorial de cada factor, habitualmente a uno.

El modelo de estructuras de covarianza presentado es recursivo. No presenta ninguna dirección causal bidireccional, en dos sentidos. Aún cuando no se usarán aquí, los modelos de estructuras de covarianza no recursivos son también posibles.

Tan solo se han presentado los tres modelos causales básicos. Existen depuraciones estadísticas y modelos más complejos, como los 'FASEM models' y los modelos de estructuras de medias y covarianzas que no trataremos aquí (Bentler, 1988).

Todos los modelos causales presentados tienen dos problemas estadísticos básicos (Bentler y Bonnett, 1980). Por un lado la estimación de los coeficientes o parámetros del modelo, como los efectos de una variable sobre otras, los errores asociados, las relaciones entre factores, etc. Y por otro la evaluación de la bondad del ajuste del modelo hipotetizado a los datos observados.

### **Estimación**

Los datos de una muestra de sujetos medidos en  $p$  variables observables pueden resumirse para los propósitos de modelado causal en una matriz ( $p \times p$ ) de varianzas-covarianzas ( $S$ ). Esta matriz  $S$  incluye las varianzas de las variables observables y las covarianzas entre estas.  $S$  es la entrada para cualquier modelo causal, bien en esta forma de matriz de varianzas-covarianzas, o bien en la forma de matriz de correlaciones entre las variables. Un supuesto básico en que se basa la estimación de modelos causales es que la matriz poblacional ( $\Sigma$ ) correspondiente se genera a partir de los parámetros poblacionales desconocidos que se pueden emplazar en el vector  $\theta$ . La matriz  $\Sigma$  es una función de los parámetros de  $\theta$  bajo un modelo causal hipotetizado (Bentler, 1988). Así, cada valor particular de  $\Sigma$ , digamos la covarianza entre las variables  $i$  y  $j$  es función de  $\theta$ :

$$\text{cov}_{ij} = f_{ij}(\theta)$$

La función  $f_{ij}$  puede verse como la función que describe la estructura particular bajo investigación, que relaciona los parámetros del modelo con los elementos de  $\Sigma$  (Bentler y Bonnett, 1980).

Los modelos, como hemos visto, tienen algunos parámetros libres (a estimar) y otros que se tratan como fijos (se les adjudica un determinado valor de antemano). Estos últimos, por convención, no se incluyen en el vector de parámetros  $\theta$ . De modo que una mejor representación de la función sería  $f(\theta, \omega)$ , donde  $\omega$  representa el vector de parámetros a estimar. Los parámetros a estimar son típicamente coeficientes de regresión, covarianzas de variables exógenas, varianzas de error, etc, y están emplazados en el vector  $\theta$ . El modelo concreto de que se trate -AFC, ecuaciones estructurales, o estructuras de covarianzas- determinará la función ( $f$ ) que permitirá calcular a partir de estos parámetros  $\Sigma$ .

Si la estructura teórica que planteaba el modelo es correcta y la muestra es grande se supone que  $S$  y  $\Sigma$  serán similares, convergiendo conforme aumenta el tamaño de la muestra.

Desgraciadamente, en situaciones reales,  $\theta$  y  $\Sigma$  son desconocidas, siendo necesario acudir a la estimación. Es preciso calcular un vector de parámetros estimados  $\theta'$  que lleve a una matriz poblacional estimada  $\Sigma'$ . Si  $\Sigma'$  se acerca elemento a elemento a  $S$  el modelo hipotetizado se convierte en un candidato plausible para representar la estructura subyacente a los datos.

Para comprobar el grado en que  $\Sigma'$  y  $S$  están cercanos se utiliza una función, que se denomina función de ajuste y que representaremos por  $F$  (Bentler y Bonnett, 1980; Marsh, Balla y McDonald, 1988). Los valores de  $\Sigma'$  que hacen mínimas las diferencias con  $S$  son los valores estimados de los parámetros del modelo. De este modo esta función sirve para estimar los parámetros del modelo hipotetizado (Visauta, 1986). El tipo de estimación que realicemos determinará la forma de la función de ajuste, pero la lógica es siempre la misma. A modo de ejemplo, si el método de estimación fuera mínimos cuadrados no ponderados tendríamos:

$$F = U = \frac{1}{2} \text{tr} [(S - \Sigma')^2]$$

Los valores de  $\Sigma$  se estiman de forma que hagan mínima esta función, de forma que cuanto más se acerquen los valores estimados a los datos observados menor será la función de ajuste. En el límite cuando  $\Sigma' = S$ ,  $F = 0$  y el ajuste es perfecto.

Actualmente, los métodos de estimación más comúnmente utilizados (Bentler, 1988) son mínimos cuadrados generalizados y máxima verosimilitud. El primero en aparecer en el campo de los modelos causales fue el de máxima verosimilitud (ML, maximum likelihood) (Jöreskog, 1967, 1969; Jöreskog y Sörbom, 1979). El propio Jöreskog junto a Golberger (1972) y Anderson (1973) pusieron las bases para el cálculo de estimadores mediante mínimos cuadrados generalizados (GLS, generalized least squares). Estos dos métodos se basan en el supuesto de distribución normal multivariada de las variables observadas. Si se cumple este supuesto, GLS y ML presentan una serie de propiedades estadísticas asintóticas: consistencia, normalidad, eficiencia e insesgancia aproximada (Browne, 1974; Visauta, 1986). Además, Lee (1977) mostró que ML y GLS eran asintóticamente iguales. No obstante, parece que las estimaciones mediante GLS pueden ser más sesgadas que con ML (Browne, 1974). Evidentemente, estas propiedades se mantienen únicamente cuando se mantiene el supuesto de distribución normal multivariada de las variables. Hay investigaciones que demuestran la robustez de los estimadores a violaciones del supuesto (Satorra y Bentler, 1986).



Otras vías de aumentar la idoneidad en la estimación han sido la de generar métodos que no se basen en el supuesto de normalidad. Los métodos de estimación elíptica establecidos por Bentler y Dijkstra (1985) están basados en un supuesto menos fuerte. Son aplicables si la distribución es simétrica, aunque no sea normal. Las distribuciones pueden tener colas más grandes o pequeñas que la normal (Bentler, 1988). Se supone, no obstante la simetría. Aún menos exigente es la estimación arbitraria, que permite la estimación sin importar el tipo de distribución (Bentler, 1988; Browne, 1984; Satorra, 1989). Estos métodos arbitrarios y elípticos son todavía novedosos y son de mucho menor uso aplicado que los más clásicos ML y GLS. Su uso debe, por tanto, ser cuidadoso. Tanaka (1984) indica que los resultados de estimación de los parámetros que se producen mediante ML o GLS cuando se han producido incumplimientos del supuesto de normalidad multivariable son razonables. Los problemas del incumplimiento del supuesto se dan más bien en el índice de ajuste chi-cuadrado y en las estimaciones de los errores estándar. En estos casos de incumplimiento de los supuestos otros índices de ajuste funcionarán mejor que chi-cuadrado y los errores utilizados para las pruebas de ajuste analítico no serán fiables (Bentler, 1988).

### **Indices de ajuste**

Los modelos causales presentan dos tipos discernibles de comprobación del ajuste, una de carácter analítico y otra de carácter global (González-Romá, 1990; Long, 1983). El ajuste de carácter analítico responde a la pregunta de qué efectos (saturaciones, coeficientes path,...) o covarianzas (entre factores, entre variables,...) resultan una a una significativas. El ajuste global responde, para todo el modelo, si la estructura propuesta para los datos es plausible. A continuación se revisarán ambos tipos de ajuste.

#### *Indices de ajuste global*

La función de ajuste (F) que sirve para propósitos de estimación tiene una importante propiedad (Bentler y Bonnett, 1980). Si multiplicamos la función de ajuste por el tamaño de la muestra obtenemos:

$$F \times (N-1) = \chi^2$$

que presenta una distribución chi-cuadrado asintótica, con grados de libertad iguales a las varianzas-covarianzas de las variables observables (p) menos el número de parámetros a estimar (q):  $p(p+1)/2 - q$  (Bentler, 1988).

Lo que tenemos es un estadístico que depende de N, el tamaño muestral y, lo que le confiere su valor como estadístico de bondad de ajuste, de F. Esta función de ajuste toma valores, como hemos visto, en función de la cercanía entre los valores estimados desde el modelo hipotetizado ( $\Sigma'$ ) y los datos observados (S). Esto significa que el test chi-cuadrado es una medida de lo bien que ajustan los valores generados desde el modelo causal a los datos reales.

En rigor, puede apreciarse cómo la misma F es una medida de ajuste. A mayor F menor ajuste. ¿Por qué es entonces necesario el test chi-cuadrado?, Chi-cuadrado presenta una distribución conocida y puede, por tanto, contrastarse con un valor extraído de las tablas para unos determinados grados de libertad, que indica la probabilidad de que por azar los datos difieran en la cantidad que representa el valor de chi-cuadrado. En otras palabras, es susceptible de probarse mediante una prueba de hipótesis. Si el test ofrece resultados significativos esto indicará que es tan grande que no puede retenerse la hipótesis nula (generalmente a un nivel máximo del .05 de probabilidad). Esta hipótesis indica en el contexto de modelos causales la igualdad de  $\Sigma'$  y S y, consecuentemente, la plausibilidad del modelo. Si el modelo es adecuado o plausible, si es una adecuada representación de los datos observados, la prueba resultará no significativa.

Desgraciadamente, y como puede apreciarse fácilmente en la fórmula, chi-cuadrado no depende tan solo de F sino también de N, el tamaño de la muestra. Conforme aumenta el N de la muestra se hace más fácil encontrar diferencias significativas, de forma que con muestras grandes prácticamente ningún modelo puede sostenerse (Bentler y Bonnett, 1980; Bentler, 1988; Mulaik, James, Van Alstine, Bennett, Lind y Stilwell, 1989). Estos problemas generaban que multitud de modelos fueran rechazados en las ciencias sociales, dada la potencia del ajuste mediante ji-cuadrado.

Queda abierta la cuestión de cuántos sujetos son un adecuado tamaño muestral, aunque Boosma (1983) y Anderson y Gerbing (1984) encuentran que 100 o más sujetos resulta bastante adecuado.

Marsh y Hocevar (1985) hacen notar que la mayoría de las aplicaciones del análisis factorial confirmatorio requieren una evaluación subjetiva, donde se trate de determinar cuándo un test de chi-cuadrado significativo es suficientemente pequeño para indicar (pese a su significatividad) un adecuado ajuste. Mejor que una evaluación subjetiva, los estadísticos empezaron a generar nuevos índices de ajuste.

Junto al índice chi-cuadrado y, con el propósito de comparar dos modelos que compiten para representar unos datos se ha desarrollado una prueba de las *diferencias entre las pruebas de chi-cuadrado* de los dos modelos (Bentler y Bonnett, 1980).

Si se tienen dos modelos hipotéticos, cada uno de ellos con una chi-cuadrado que los resume en su ajuste global, puede conocerse qué modelo de los dos es 'mejor', o más concretamente, qué modelo explica significativamente más de las varianzas-covarianzas de los datos.

Para ello se calcula una prueba de las diferencias entre las dos ji-cuadrado:

$$\chi^2_1 - \chi^2_2 = \Delta\chi^2$$

que se demuestra (Bentler y Bonnett, 1980) que se distribuye como una chi-cuadrado con grados de libertad iguales a la diferencia entre los grados de libertad de los dos modelos:

$$gl_1 - gl_2 = \Delta gl$$

Si la prueba de las diferencias resulta significativa se podrá rechazar la hipótesis nula de que ambos modelos explican la misma proporción de varianza. Deberá concluirse que aquel modelo con menor chi-cuadrado explica significativamente más parte de la varianza de los datos observables, es mejor predictor de los datos.

A partir del índice chi-cuadrado se han ido presentando otros índices de ajuste que pretendían subsanar los problemas de éste, problemas básicamente de potencia estadística. La lista se ha ido acrecentando con el tiempo (Bentler, 1988b; Marsh, Balla y McDonald, 1988), pero sólo presentaremos los que nos parecen más novedosos, más justificados y /o que mejores resultados empíricos ofrecen. Todos ellos se utilizarán para evaluar los modelos que se presenten en esta tesis.

Primero se presentarán los índices que se basan únicamente en la chi-cuadrado del modelo analizado, y que no la comparan con cualquier otro. Son los índices que Marsh, Balla y McDonald (1988) llaman 'stand alone indexes' por este motivo de que no se comparan con otros. Veremos entre éstos: a) F; b) LHR; c)  $\chi^2/gl$ ; d) CAK; y e) CSK.

a) *F*: La función de ajuste que sirve para estimar los valores es, como ya hemos comentado, una medida del ajuste. Su problema es que vale cero cuando  $\Sigma'$  y S son iguales, pero no presenta un límite superior. Se define como:  $F = \chi^2/N-1$ .

b) *LHR*, la razón de verosimilitud escalada. Se define como (Cudeck y Browne, 1983):

$$LHR = e^{-1/2(F)}$$

presenta límites definidos (0-1), indicando 0 un ajuste totalmente inadecuado, mientras 1 indica ajuste perfecto.

c)  $\chi^2/gl$ : Este índice divide la prueba chi-cuadrado entre sus grados de libertad. Con este cálculo se pretende introducir una función penalizadora del uso de más parámetros libres o a estimar. Es, por tanto, un índice de parsimonia de los modelos. Cuantos más grados de libertad aparecen para unos determinados datos quiere decir que existen menos parámetros a estimar y, en conclusión, que el número de relaciones hipotetizadas es menor. El modelo de más grados de libertad es más parsimonioso y suponiendo constante chi-cuadrado, el cociente será menor conforme más grados de libertad haya. Se penalizan así los modelos con muchos parámetros.

Este índice, no obstante, presenta la misma dependencia de N que tiene chi-cuadrado y no hay acuerdo en qué valores del índice representan un adecuado ajuste. Wheaton, Muthén, Alwin y Summers (1977) consideran 5 como límite máximo de ajuste y Carmines y McIver (1981) rebajan el mínimo aceptable a 3 ó 2. Por su parte Brooke, Russell y Price (1988) consideran que 2 (o menos) representa un ajuste muy adecuado.

d) CAK. Akaike (1974) propone el *Akaike Information Criterion* que incorpora también una función penalizadora basada en el número de parámetros a estimar. Lo planteaban como un indicador útil para comparar modelos que intentan representar unos mismos datos y que varían en el número de parámetros a estimar, modelos competitivos. Para eliminar o, al menos, hacer mínimo el efecto del tamaño de la muestra sobre este índice Cudeck y Browne (1983) proponen una versión escalada de este índice expresada matemáticamente como:

$$CAK = F + 2q/N$$

donde q es igual al número de parámetros a estimar.

e) CSK. Para los mismos propósitos Schwarz (1978) añade un nuevo índice, el *Schwarz Information Criterion*. Nuevamente Cudeck y Browne (1983) formulan una versión escalada del índice:

$$CSK = F + (q \times \ln N) / N$$

donde q = número de parámetros a estimar.

Hasta aquí los índices que utilizan una sola chi-cuadrado o una sola función de ajuste, la del modelo que se pretende poner a prueba.

Cuanto menores son sus valores (tanto de CAK y CSK), mejor es el ajuste y son útiles en la comparación de modelos. Son también en algún sentido índices de parsimonia, dado que incluyen ambos una

función penalizada para los modelos con más parámetros a estimar, para los modelos menos parsimoniosos.

Bentler y Bonnett (1980) propusieron comparar el modelo propuesto con un modelo nulo, típicamente aquél en que se asumen todas las variables como independientes o no relacionadas. Este modelo nulo sirve como punto de partida a partir del cual se demuestra si el modelo propuesto aporta más estadísticamente a la explicación de la covarianza de las variables. Desde entonces un amplio número de índices se han generado siguiendo esta lógica. Nosotros trataremos los tres que Bentler (1988) considera más relevantes: f) NFI; g) NNFI y g) CFI.

f) *NFI, normed fit index*: Bentler y Bonnett, ante las dificultades del ajuste mediante chi-cuadrado, propusieron que más que evaluar si las discrepancias entre  $\Sigma'$  y  $S$  son estadísticamente significativas, habría que proceder a la comparación de lo que aporta el modelo frente a otro modelo que actuara como modelo basal o punto de partida. Este modelo basal era, generalmente, un modelo nulo donde las variables se asumían como no relacionadas. Ambos modelos se comparaban poniendo en relación mediante la siguiente expresión (que es el mismo índice):

$$NFI = \chi^2_{\text{nulo}} - \chi^2_{\text{hipo}} / \chi^2_{\text{nulo}}$$

siendo  $\chi^2_{\text{nulo}}$  = el valor de la chi-cuadrado para el modelo nulo;

y  $\chi^2_{\text{hipo}}$  = el valor de la chi-cuadrado para el modelo hipotetizado.

Este índice es igual a 0 cuando los dos modelos son iguales, cuando el modelo basal no aporta nada frente al modelo nulo. El máximo valor que puede tomar es 1 indicando el máximo de ganancia del modelo hipotetizado frente al nulo, de hecho por oscilar entre 0 y 1 Bentler y Bonnett le llamaron índice de ajuste normativo. A mayor valor, mejor será el ajuste. No está claro a partir de qué valor puede considerarse adecuado el ajuste, aunque Bentler (1980) y Bentler y Bonnett (1980) afirman que .90 representa ya un buen ajuste.

Presenta las desventajas de verse afectado por el tamaño de la muestra y que en muestras pequeñas puede no alcanzar el valor 1 aún cuando el modelo es correcto (Bearden, Sharma y Teel, 1982; Bentler, 1988). En muestras pequeñas puede ser anormalmente bajo (indicando ajuste inadecuado) cuando otros índices muestran un adecuado ajuste (Anderson y Gerbing, 1984).

g) *NNFI, non-normed fit index*: Para eliminar el problema del índice NFI en cuanto al rango 0-1, y basándose en el índice Tucker-Lewis (1973) generado en el campo del factorial exploratorio Bentler y Bonnett (1980) generaron el NNFI. En este índice se realiza un ajuste en las chi-cuadrado para mejorar el funcionamiento en los valores cercanos al 1, y no con intenciones de generar un índice de parsimonia (Bentler, 1988). Se define como (Bentler y Bonnett, 1980; Bentler, 1988):

$$NNFI = (\chi^2_{nulo} - gl_{nulo} / gl_{hipo} \chi^2_{hipo}) / \chi^2_{nulo} - gl_{nulo}$$

donde todos los términos indican los mismo que en NFI, y los que aparecen nuevos son:  $gl_{nulo}$  = grados de libertad para el modelo nulo, y  $gl_{hipo}$  = grados de libertad para el modelo hipotetizado.

O alternativamente puede definirse como (Marsh, Balla y McDonald, 1988):

$$NNFI = (\chi^2_{nulo} / gl_{nulo} - \chi^2_{hipo} / gl_{hipo}) / \chi^2_{nulo} / gl_{nulo}$$

fórmulas que ofrecen idénticos resultados.

Como problemas de este índice tenemos que puede exceder de los límites inferior o superior. Además, en estudios muestrales se comprueba que NNFI tiene una varianza sustancialmente mayor que NFI (Bentler, 1988).

Como principal ventaja, tenemos que el NNFI representa adecuadamente el ajuste en casi todos los niveles muestrales (Anderson y Gerbing, 1984; Marsh, Balla y McDonald, 1988).

h) *CFI, comparative fit index*. Para superar los problemas de los índices NFI y NNFI Bentler (1988b) generaron el índice CFI. Este índice siempre oscila entre 0 y 1. No excede los límites. Es adecuado para todo tipo de tamaños de muestra, según comprueba empíricamente Bentler (1988b) en un estudio muestral y su variabilidad es menor que la de NFI y NNFI.

Hocevar, Khattab y Michael (1987) rechazan fuertemente este tipo de índices de comparación con un modelo nulo en el ámbito del análisis factorial confirmatorio. Según estos autores, la dificultad estriba en que a un nivel práctico "virtualmente ninguna matriz de medidas en las ciencias de la conducta es completamente independiente, y consecuentemente, uno no puede esperar encontrar soporte para ese modelo sin importar qué otro modelo alternativo se hipotetiza". Esto es, los autores afirman que el modelo nulo es tan poco probable que hasta otras estructuras de mal ajuste son mejores que él. Para demostrar esta crítica probaron el ajuste de

modelos factoriales aleatorios mediante los índices de comparación con el modelo nulo encontrando un adecuado ajuste. Como solución a este problema, proponen la redefinición del modelo nulo. Así, piensan que el modelo propuesto no debe compararse en análisis factorial confirmatorio frente a un modelo de completa independencia, sino frente a un modelo unifactorial. Proponen que el modelo nulo de referencia se defina como un modelo de un solo factor. Debe resultar evidente que este procedimiento resulta viable sólo cuando el modelo teórico propuesto es de más de un factor. Por otro lado, en Bentler (1988) se recoge un procedimiento mucho más general de modelos anidados, en que se comprueban los índices primero con un modelo nulo, y después con otros modelos de creciente complejidad para evaluar el modelo propuesto sustantivamente. Por nuestro lado, creemos que el problema se reduce a la comparación de modelos competentes. Sin olvidar que cuando las muestras son razonables en su tamaño el ajuste de chi-cuadrado no debe de menospreciarse.

Aparte de estos índices de ajuste global que se presentan para cualquier tipo de estimación, el programa EQS ofrece otro tipo de índices cuando el tipo de estimación es arbitraria.

*i) Fit index y adjusted fit index.* Cuando se realiza la estimación mediante métodos arbitrarios el programa EQS ofrece estos dos índices de ajuste. El fit index es más conocido en la literatura con el índice GFI (goodness of fit index, índice de bondad de ajuste). Tiene amplio uso en la literatura dado que se implementa en el programa LISREL 7 (Jöreskog y Sörbom, 1988). No se basa en chi-cuadrado, ni es una comparación frente a un modelo nulo. Se basa en el tamaño de los residuales comparados con los datos observados. Puede interpretarse como la proporción de la matriz de varianzas-covarianzas iniciales que se consigue explicar con el modelo propuesto. Oscila, lógicamente, entre 0 y 1 indicando el uno ajuste perfecto, o lo que es lo mismo el 100% de varianza explicada. Por su parte el adjusted fit index es más conocido como índice AGFI (adjusted goodness of fit index, índice bondad de ajuste ajustado) y también viene implementado en el LISREL 7. Este índice se diferencia del anterior en que tiene en cuenta los grados de libertad del modelo dando, por tanto, una idea de la parsimonia del modelo. También oscila entre cero y uno, indicando el valor máximo ajuste perfecto.

#### *Indíces de ajuste analítico*

Para cada parámetro a estimar del modelo, se puede generar una prueba z de contraste de hipótesis para comprobar si el valor encontrado es significativamente distinto de cero. De este forma puede conocerse estadísticamente la significatividad de los

coeficientes path, las saturaciones factoriales, las varianzas de error, y las covarianzas entre factores y entre variables. El valor del test z se consigue dividiendo el valor del estimador entre el error típico de estimación, acudiendo después a la tabla de las z para comprobar su probabilidad asociada.

Además del test z, se han generado dos pruebas estadísticas para estudiar analíticamente (parámetro a parámetro) los modelos propuestos. Estos dos tests son el LM test (Lagrangian Multiplier test) y el W test (Wald test) (Bentler, 1988) que tienen su réplica en los índices de modificación de LISREL (Jöreskog y Sörbom, 1988). No entraremos en la compleja trama estadística que sustenta estos dos tests sino sólo en el funcionamiento.

El LM test es en realidad un conjunto de tests aplicados a cada uno de los parámetros fijos en el modelo. Nos ofrece información sobre el cambio que se producirá en el ajuste global del modelo (chi-cuadrado) si un parámetro fijo se dejara libre a estimar. Además de facilitar la ganancia numérica en chi-cuadrado nos ofrece una información estadística sobre si esa ganancia será o no significativa, producirá un mejor ajuste global. En sentido contrario opera el W test, ofreciendo información sobre qué ocurriría si algún parámetro libre se convirtiera en fijo. Informa de aquellos parámetros a estimar que se podrían tratar como fijos sin merma significativa en el ajuste global mediante chi-cuadrado. El LM test y el W test no se implementan en el programa EQS cuando se utiliza la estimación arbitraria.

#### **2.4. Comprobación del supuesto de normalidad multivariable.**

Los métodos de estimación más comúnmente aplicados en modelado causal, máxima verosimilitud y mínimos cuadrados generalizados, se basan en el supuesto de distribución multivariable. Usualmente este supuesto es demasiado fuerte y, aunque, relativamente robustos a violaciones del mismo en ocasiones es difícilmente sostenible. Para evitar este problema el programa EQS monta otros dos tipos de estimación que no se basan en el supuesto de distribución multivariable, la estimación elíptica y la arbitraria. Antes de decidirse a aplicar uno u otro método debe comprobarse si los datos se alejan significativamente de la normalidad multivariable.

Las características que cumple una distribución multivariable son (Bisquerra, 1989; Cuadras, 1981; Tatsuoka, 1971): 1) todas las variables del análisis siguen una ley normal, son normales una a una; 2) todas las variables se distribuyen normalmente para valores dados de las otras variables; y 3) la distribución es exactamente conocida y se puede describir en base a las medias y desviaciones típicas de cada variable y a la matriz de correlaciones entre ellas.



De las características de la distribución normal multivariable se sigue que es condición necesaria, pero no suficiente para la normalidad multivariable.

Se puede rechazar la hipótesis de distribución multivariable si alguna de las variables implicadas en el análisis se aleja significativamente de la normalidad univariada. No obstante, que todas las variables sean normales una a una no implica que cumplan una ley multivariable. Para comprobar la normalidad univariada de cada variable puede utilizarse (Cuadras, 1981; Siegel, 1983) una prueba no paramétrica de Kolmogorov-Smirnov de bondad de ajuste. Esta prueba permite rechazar la hipótesis de normalidad multivariable, pero no aceptarla. Además el programa EQS ofrece para cada variable un índice de asimetría y otro de apuntamiento. La asimetría se calcula con el índice basado en el momento de tercer orden respecto a la media y la curtosis con un índice basado en el momento de cuarto orden (Amón, 1986; Botella, León, San Martín, 1992; Cuadras, 1991; Mardia, 1980). Ambos índices presentan errores estándar conocidos, y se distribuyen normalmente en muestras grandes, permitiendo el contraste de hipótesis. El error típico para el índice de asimetría viene dado por la expresión  $(6/N)^{1/2}$  que puede usarse con muestras mayores de 200. Para muestras de más de 1000, el error para la curtosis viene dado por la expresión  $(24/N)^{1/2}$ . Desgraciadamente dado el tamaño de nuestra muestra se hace desaconsejable el calcularlo.

Existen diversas pruebas para comprobar directamente el supuesto de normalidad multivariable, como la prueba de normalidad de las componentes principales, aplicable preferiblemente con muestras grandes (Cuadras, 1981; Krishnaiah, 1973). El paquete estadístico EQS ofrece el coeficiente de Mardia para la curtosis multivariable, que puede usarse para contrastar la normalidad multivariable (Mardia, 1974, 1980). Se rechaza la hipótesis nula de normalidad multivariable si el coeficiente de Mardia se aleja bastante, por encima o por debajo, de la distribución esperada bajo la distribución normal multivariable, que también ofrece el programa. Este uso del coeficiente presenta, no obstante, dos severos problemas. Para muestras grandes ofrece resultados de rechazo de la hipótesis nula para pequeñas desviaciones de la normalidad, y en muestras de tamaño medio la distribución del estadístico es asimétrica. Por todo ello, Bentler (1988) propone que se evalúe también la normalidad univariada para asegurar las conclusiones.

Una vez se encuentra que el supuesto de normalidad multivariable no se ajusta a los datos con que se trabaja la siguiente cuestión es elegir entre estimación elíptica y arbitraria. Bentler (1988) plantea que la estimación elíptica es adecuada sólo si la curtosis que presentan las variables es homogénea, de forma que observando la curtosis univariada puede, a grosso modo, elegirse

entre estimación elíptica o arbitraria. Ningún prueba se presenta en el prorama EQS, ni en los paquetes estadísticos más habituales, para comprobar estadísticamente la hipótesis de homogeneidad en las curtosis univariadas.

### 2.5. Validación de constructo y modelos causales.

La propiedad de validez consiste en que el test mida lo que se supone que debe medir. O dicho de otro modo, dado que los tests son utilizados para realizar inferencias, la validez puede definirse como el conjunto de datos recogidos para mostrar la pertinencia de tales inferencias (Muñiz, 1992). Esta propiedad, aparentemente fácil de entender, ha dado lugar a múltiples subtipos de validez y un largo número de diferentes operacionalizaciones.

La división tradicional de la validez ha sido recogida por la American Psychological Association en las publicaciones normativas sobre tests educativos y psicológicos, incluyendo:

1) *Validez criterial o predictiva*. Este tipo de validez es el más clásico en la teoría de tests. Se refiere a la capacidad del test para poder predecir una variable criterio externa al test con la que se espera una relación de cierto signo e intensidad. Si existen diversos criterios, la operacionalización de la validez criterial se obtiene mediante uno o varios coeficientes de correlación entre el test y los criterios. Estas correlaciones entre test y criterios se conocen como *coeficientes de validez*.

Dentro de la validez criterial puede también tratarse la *validez nomológica* (Campbell, 1960; Cronbach y Meehl, 1955). Se refiere al grado en que se confirman las predicciones en una red teórica formal. La diferencia entre validez predictiva y validez nomológica no consiste en que sean casos diferentes de validez, sino que es una distinción de grado (Bagozzi, 1981). *La validez criterial se referiría más bien a la relación del test con un criterio, antecedente o consecuente. La validez nomológica tiene en consideración múltiples antecedentes y/o consecuentes en sistemas teóricos complejos.*

2) *Validez de contenido*. Alude al grado en que el test presenta una muestra representativa de aquellos contenidos que representan lo que mide el test. Generalmente ha sido evaluada mediante análisis racionales realizados por expertos en el área que cubre el test.

3) *Validez de constructo*. Puede definirse como la recogida de evidencia empírica que garantice la existencia de un constructo psicológico por debajo de las medidas del test (Muñiz, 1991). Consistiría en un verdadero proceso de validación científico de la teoría subyacente al constructo que mide el test, a las relaciones de

los items entre sí y a las relaciones con otras variables. Consiste en una búsqueda de significado de las puntuaciones del test. Algunos autores incluyen dentro de la validez de constructo básicamente dos subtipos de validez, la validez factorial y la validez convergente-discriminante.

3.1) *La validez factorial.* Introducida por Guilford consiste en la comprobación de qué factores o constructos están subyaciendo a los items de cierto test. Se habla de validez factorial de un determinado factor o constructo cuando los diversos items definidos para medirlo muestran coherencia con ese factor latente. Suele realizarse esta comprobación mediante las técnicas de análisis factorial y análisis de clusters o, cuando se tiene algún conocimiento previo o hipótesis de la estructura de factores de la escala, mediante análisis factorial confirmatorio.

3.2) *Validez convergente-discriminante.* La validez convergente hace referencia al grado en que diversas medidas de un mismo constructo presentan altas relaciones entre sí. La validez discriminante se refiere a la ausencia o disminución de la relación entre medidas de distintos constructos. En definitiva, se espera que distintas medidas de un mismo constructo, sean items o métodos diferentes, deben mostrar relaciones más fuertes entre sí que con items o métodos que miden otros constructos. Un diseño habitual para la comprobación de la validez convergente-discriminante ha sido la matriz multirasgo-multimétodo sistematizada por Campbell y Fiske (1959).

Esta división de la validez en tres subcomponentes ha sido criticada y puesta en duda. Es fácil apreciar cómo la definición o acercamiento a la validez de constructo que aquí se ha manejado es muy amplia, más allá de la validez factorial y convergente discriminante. La validez de constructo implicaría un análisis, por todos los métodos disponibles, del significado de las puntuaciones de un test y de a qué concepto o constructo psicológico hace referencia (Cronbach y Meehl, 1955). Por lo tanto podría considerarse la validez de constructo como un concepto general que abarcaría los otros tipos de validez y, según algunos autores, los análisis de items y de fiabilidad. Loevinger (1965) plantea los siguientes procedimientos para comprobar la validez de constructo.

1) *Análisis de contenido.* Se correspondería con el concepto de validez de contenido.

2) *Análisis de la estructura interna del test.* Incluiría:

a) análisis de items, para conociendo características de los items como dificultad, homogeneidad o poder

discriminativo, entender mejor qué mide el test compuesto por esos ítems;

b) análisis de la fiabilidad del test, estabilidad temporal y consistencia interna.

c) análisis de la estructura factorial del test, de los factores latentes que subyacen a la realización en el test.

3) *Análisis de la relación con variables externas.* Incluiría:

a) Estudios de *validez criterial o predictiva*;

b) Estudios de *validez convergente-discriminante*;

c) *Estudios experimentales* que puedan apoyar las hipótesis que se deriven del constructo medido con el test;

d) *Análisis diferenciales.* Donde se trata de comprobar si las diferencias reales de grupos se traducen en diferencias en las puntuaciones en el test. Por ejemplo, dado que la inteligencia aumenta con la edad, en la infancia sería previsible que los tests aumentaran de puntuación significativamente en niños de diversas edades.

Lo que propone Loevinger es absolutamente concordante con la opinión de Cronbach y Meehl (1955) de que la validación de constructo consiste en una comprobación del significado de las puntuaciones en el test, en todo igual a la validación de cualquier teoría científica. El estudio intensivo por todos los medios posibles del test, incluyendo la fiabilidad, puede entenderse como validación de constructo.

Aunque Loevinger no hable explícitamente de ello, una de las propiedades especialmente relevantes para cualquier teoría científica o constructo científico, y por tanto para su medida, es la capacidad de generalización que tienen sus resultados. Interesa elaborar medidas para constructos que sean válidas y fiables, pero es muy importante conocer hasta qué punto los resultados que estamos comprobando hoy en una muestra determinada, permanecen invariantes cuando cambiamos de muestra, de población o de situación. ¿La teoría que explica la toma de decisiones empresariales predice igualmente bien en tiempos de crisis que en situaciones normales?; ¿el rendimiento académico en la Universidad puede predecirse en base a la inteligencia, la motivación, el rendimiento académico anterior, o las horas de estudio, de la misma forma en una muestra de españoles que en una muestra de estadounidenses?; ¿es igualmente fiable y válido un instrumento para medir inteligencia social en niños de seis años sordos que en niños de seis años no-sordos?. Todo este tipo de preguntas que permiten conocer la

consistencia de los resultados a pesar de que cambien aspectos de las muestras o de las situaciones de evaluación, se estudian desde la *validez cruzada*. Se dice que hay validez cruzada de un instrumento de medida cuando los resultados de validez no son estadísticamente diferentes entre dos o más muestras o situaciones.

Los métodos estadísticos aplicables a la validación de constructo son múltiples, dada la amplitud del concepto. Correlación y regresión, análisis discriminante, análisis de varianza, pruebas t, pruebas no paramétricas, análisis factorial, análisis de clusters (análisis de conglomerados), técnicas de estadística descriptiva, ... tienen cabida en la evaluación completa de un test. Entre la metodología empleada para validar un test, nuevos métodos muy sofisticados están empezando a cobrar preeminencia y a complementar a los ya existentes. Entre ellos se encuentran los modelos causales que presentan algunas ventajas frente a algunos tipos más habituales de análisis, y son aplicables a una gran cantidad de situaciones de estudio de la validez.

### **2.5.1. Algunas aplicaciones de los modelos causales en la validación de constructo.**

En este apartado se verán distintos usos de los modelos causales para comprobar diversos aspectos de la validez de constructo. Se tratarán los aspectos de: a) fiabilidad; b) validez factorial; c) validez criterial y nomológica; d) validez convergente-discriminante; e) validación cruzada; y f) análisis diferenciales o validez diferencial.

#### ***Fiabilidad***

La técnica del análisis factorial exploratorio ha sido utilizada clásicamente para evaluar la fiabilidad de las escalas y los ítems. Así, los coeficientes theta y omega son dos de los más conocidos como medida de la consistencia interna de una escala. Por otro lado, la saturación factorial elevada al cuadrado es utilizada a menudo como medida de la fiabilidad de los ítems, la consistencia interna con el factor del ítem.

Para calcular la fiabilidad de los ítems se han utilizado dos procedimientos a través de los resultados de los análisis factoriales confirmatorios para las escalas. Long (1983) propone que la saturación del ítem en el factor latente al que pertenece elevada al cuadrado, es una medida de la fiabilidad del ítem.

Más depurado es el índice de fiabilidad propuesto por Bagozzi (1981), apoyándose en el trabajo de Werts, Linn y Jöreskog (1974). Este índice se calcula mediante la siguiente fórmula:

$$\rho_i = (\lambda_i^2 \text{ var } A) / (\lambda_i^2 \text{ var } A + \psi_i)$$

donde

$\rho_i$  = fiabilidad del ítem i;

$\lambda_i$  = saturación (carga factorial) del ítem i en el factor A;

var A = varianza del factor A;

$\psi_i$  = varianza de error o factor único para el ítem i.

Esta fórmula puede también generalizarse tomando sumatorios de las saturaciones factoriales desde  $i=1$  hasta  $n$  y dividiendo entre el número de ítems para calcular la fiabilidad de los factores confirmatorios de la escala (Bagozzi, 1981).

### **Validez factorial**

Aunque para comprobar la validez factorial de la escala se pueden utilizar diversas técnicas estadísticas, como el 'cluster analysis' o los métodos de escalamiento multidimensional, las más utilizadas desde siempre han sido el análisis factorial exploratorio o el análisis de componentes principales. Cuando se tienen hipótesis teóricas claras sobre qué factores están midiendo los ítems de un test o se han obtenido ya resultados empíricos sobre la estructura factorial de la escala por alguno de los anteriores procedimientos, pueden someterse a prueba estas hipótesis mediante análisis factorial confirmatorio.

Con respecto a la validez factorial poco hay que añadir a lo ya expuesto al presentar el análisis factorial confirmatorio. Las principales ventajas del análisis factorial confirmatorio frente al exploratorio puede considerarse que son dos:

a) En primer lugar, pueden hacerse constricciones sustantivamente motivadas, teórica o empíricamente justificadas, de los parámetros. Esto indica, en concreto, que frente al factorial exploratorio:

- los factores pueden estar unos correlacionados y otros no.

- algunas variables observables pueden estar adscritas a sólo un factor, o a un subconjunto de factores, no necesariamente a todos los factores.

- los errores pueden o no estar correlacionados entre sí, y unos pueden estar correlacionados y otros no.

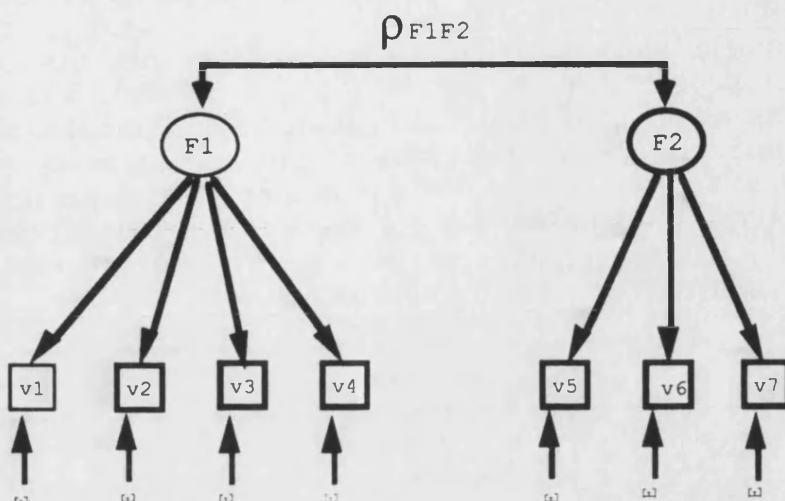
- no todas las variables observables tienen que estar sometidas a error.

b) En segundo lugar, pueden someterse a contraste estadístico los resultados. Los modelos pueden ser evaluados en su bondad de ajuste, tanto de forma global, todo el modelo, como de forma analítica, parámetro a parámetro.

Si la utilidad en la validación factorial del análisis factorial confirmatorio frente al exploratorio tuviera que resumirse en una frase sería que resulta más flexible en cuanto a posibilidades de especificación del modelo.

### ***Validez criterial y nomológica***

La validez criterial de un instrumento de medida psicológico se resume a menudo en uno o varios coeficientes de validez. Para calcular estos coeficientes usualmente se opera extrayendo, para una muestra de sujetos, una puntuación en el test y otra en el criterio externo y relacionándolas mediante una correlación de Pearson u otro tipo de coeficiente de correlación al uso. Al operar así, se está asumiendo de algún modo que todos los items miden de igual modo el constructo o factor que mide el test (Carmines y Zeller, 1987). Una forma alternativa de operar desde los modelos causales puede verse en la *gráfica 2.11*.



donde :

F1= factor que mide el test

F2= factor que mide el criterio

v1 a v4= indicadores (items) del constructo que mide el test

v5 a v6= indicadores del constructo que mide el criterio

$\rho_{F1F2}$ = coeficiente de validez del test

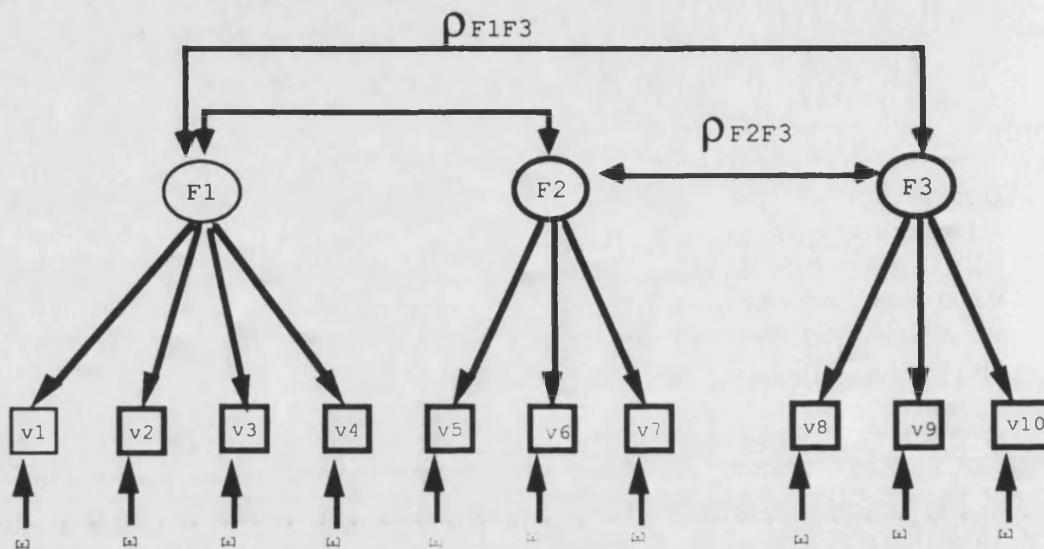
Gráfica 2.11. Coeficiente de validez criterial de un test.

En el procedimiento habitual, disponíamos de una correlación entre la puntuación global en el test y la puntuación global en el criterio. La puntuación global en el test y el criterio suelen elaborarse mediante la suma de las puntuaciones en los items, o mediante algún procedimiento de ponderación externo al análisis factorial. Esto quiere decir que todos los items aportan lo mismo a la hora de 'definir' el factor que mide el test. Pero todos sabemos que en realidad no ocurre así; muy al contrario, los diversos items no cargan igual en su factor no pudiendo considerarse igualmente relevantes al definir la puntuación en el factor. Una situación más común es que los diversos items tengan una importancia desigual al definir el factor, que no todos los items carguen igual en el factor que mide el test. El análisis factorial confirmatorio sirve para poner de manifiesto esta última desigualdad, de forma que las saturaciones factoriales indican qué items 'pesan' más en el factor.

En la gráfica 2.11 puede verse otra forma de actuar, a través del análisis factorial confirmatorio. Se plantea un análisis factorial confirmatorio del test y el criterio, cada uno medido a través de sus indicadores, items en este caso. Así, cada factor (lo que mide el test y lo que mide el criterio) ya no es la suma de las puntuaciones en los items sin ponderar. En esta ocasión, cada factor queda definido teniendo en cuenta la saturación de cada ítem en él, de forma que no



todos pesan lo mismo. La correlación entre los dos factores se corresponderá con la correlación entre el test y el criterio constituyendo el coeficiente de validez del test (igual que en el modo habitual de proceder) pero atendiendo a la importancia diferencial de los items en la medida de cada factor. Si el modelo de análisis factorial confirmatorio presenta un ajuste adecuado, la correlación entre los dos factores es, a nuestro entender, una mejor medida de la validez criterial que la simple correlación entre las puntuaciones globales en el test y en el criterio, siempre que proceda definir el criterio como un factor obtenido de indicadores.



donde :

F1= primer factor que mide el test

F2= segundo factor que mide el test

F3= factor que miden los items del criterio

v1 a v4= indicadores (items) del primer factor que mide el test

v5 a v6= indicadores del segundo factor que mide el test

v8 a v10= indicadores del criterio

$r_{F1F3}$ = coeficiente de validez del primer factor del test

$r_{F2F3}$ = coeficiente de validez del segundo factor del test

Gráfica 2.12. Coeficientes de validez criterial de un test.

El ejemplo de la gráfica 2.11 se correspondía con la situación en que los items del test miden únicamente un factor. Este no es siempre el caso. Muchas veces el test mide varios factores y conviene saber cómo se relaciona cada uno de ellos con el criterio. Lo comentado para la validez criterial, puede fácilmente generalizarse para este supuesto. En lugar de tener un único factor del test tendríamos varios, pero la forma de actuar sería la misma. En la gráfica 2.12 puede verse un ejemplo de los coeficientes de validez

riterial que resultan de un test que mide dos factores y un criterio que mide un factor.

Hemos visto el uso el análisis factorial confirmatorio para calcular coeficientes de validez. *Cuando se tiene evidencia empírica o existen razones teóricas para pensar que el o los factores que miden el test y la variable utilizada como criterio tienen una relación de causa-efecto, puede utilizarse en lugar de análisis factorial confirmatorio un modelo de estructuras de covarianza.* La lógica sería la misma, excepto que el coeficiente de validez en lugar de ser un parámetro de correlación entre dos factores sería un coeficiente path entre dos factores, uno que actuaría como causa y otro como efecto.

Hay que recordar nuevamente que los coeficientes de validez obtenidos mediante el análisis factorial confirmatorio son válidos en tanto el ajuste del modelo sea adecuado. Un análisis confirmatorio con ajuste inadecuado indica que el modelo propuesto no representa adecuadamente a los datos, de forma que cualquier conclusión extraída a partir de los parámetros estimados puede resultar inadecuada.

En la validez criterial se trata de la relación entre la medida que tratamos de validar y uno o varios criterios, pero considerados uno a uno. La *validez nomológica* también se refiere a la relación entre la variable que tratamos de validar y criterios, pero dentro de un marco teórico complejo de variables mutuamente implicadas (Campbell, 1960; Cronbach y Meehl, 1955). En este contexto, los modelos de ecuaciones estructurales y los de estructuras de covarianza tienen un uso claro. Tanto una como otra técnica permiten contrastar redes complejas entre variables o factores.

*Si el modelo, ya sea de ecuaciones estructurales o de estructuras de covarianza, presenta un adecuado ajuste y las hipótesis planteadas para las variables se cumplen y muestran también un buen ajuste analítico, la validez nomológica se habrá comprobado.*

Puede apreciarse cómo estos modelos causales proporcionan un marco más flexible para la comprobación de estas hipótesis teóricas complejas que otras técnicas. Frente a la prueba por separado de todas esas regresiones, el modelo de ecuaciones estructurales permite contrastarlas simultáneamente, indicando además, el grado de ajuste del modelo.

### **Validez convergente-discriminante**

La validez convergente puede definirse como la cantidad de acuerdo entre múltiples medidas de un mismo constructo. La lógica

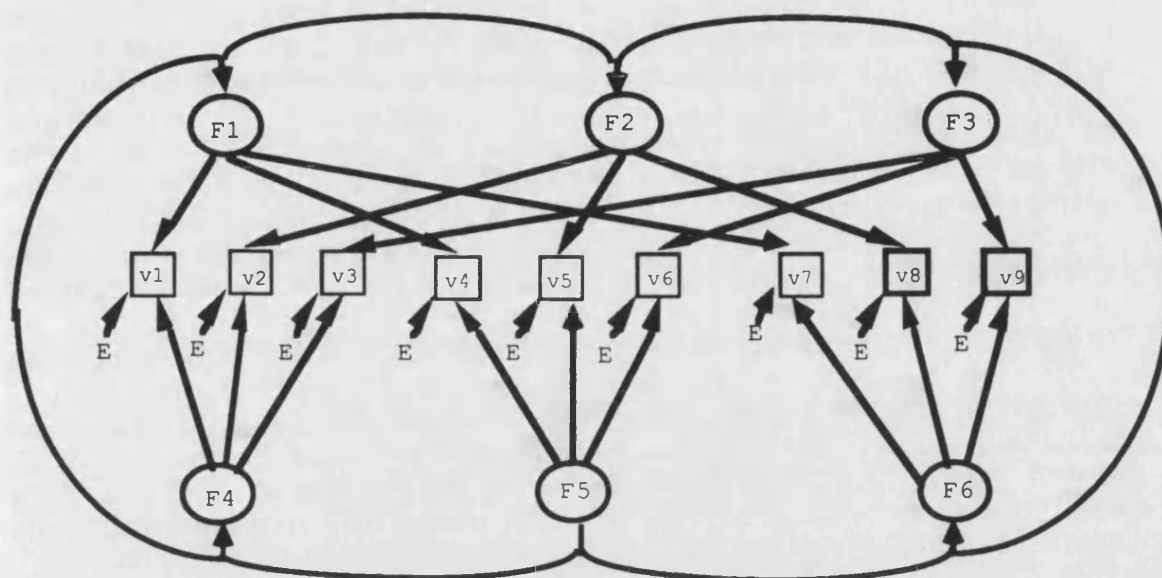
que la sustenta es que si se mide la misma variable o constructo (incluso con diferentes métodos o ítems) las puntuaciones para los sujetos deben estar fuertemente correlacionadas. Por su parte, la validez discriminante puede definirse como la unicidad entre medidas de diferentes variables, como la falta de relación o baja relación entre medidas de variables diferentes. Esas múltiples medidas pueden ser ítems o métodos diferentes. Es la ausencia de confusión, la falta de relación, entre medidas de distintas variables.

El estudio de las correlaciones entre las distintas medidas de las diferentes variables ha sido el método usado para comprobar la validez convergente y discriminante. Así por ejemplo, en la matriz multirasgo-multimétodo tenemos dos o más constructos cada uno de ellos medido por dos o más métodos. Se calculan las correlaciones entre esas medidas y se presentan en una matriz.

Se dice que hay validez convergente si las medidas de la misma variable correlacionan altamente entre sí, y existe validez discriminante si las medidas de las diferentes variables correlacionan más bajo aún cuando sean medidas por el mismo método.

Long (1983) presenta un uso del análisis factorial confirmatorio para comprobar la validez convergente-discriminante. Se tienen tres constructos psicológicos (por ejemplo: neuroticismo, psicoticismo y extraversión) medidos cada uno de ellos por tres métodos. Puede plantearse un estudio clásico de validez convergente-discriminante a través de las correlaciones entre estas medidas y con la matriz multirasgo-multimétodo, o bien puede plantearse mediante un análisis factorial confirmatorio como el que se muestra en la gráfica 2.13.

Gráfica 2.13. Análisis factorial confirmatorio que sustituye a los cálculos de una matriz multirasgo-multimétodo en la comprobación de la validez convergente-discriminante.



donde:

v1= medida del constructo neuroticismo con el método 1  
v2= medida del constructo neuroticismo con el método 2  
v3= medida del constructo neuroticismo con el método 3  
v4= medida del constructo psicoticismo con el método 1  
v5= medida del constructo psicoticismo con el método 2  
v6= medida del constructo psicoticismo con el método 3  
v7= medida del constructo extraversión con el método 1  
v8= medida del constructo extraversión con el método 2  
v9= medida del constructo extraversión con el método 3

F1= factor que representa el método 1  
F2= factor que representa el método 2  
F3= factor que representa el método 3

F4= factor de neuroticismo  
F5= factor de psicoticismo  
F6= factor de extraversión

Si observamos el diagrama causal podemos apreciar que se propone lo que se comprueba en una matriz convergente-discriminante. Al igual que la validez convergente, indica que las medidas de un mismo constructo tienden a coincidir aún cuando

procedan de distintos métodos de medida. Aquí se comprueba esa convergencia haciendo que todas las variables observables que miden un mismo constructo formen un factor latente que constituye el mismo constructo. Si estos resultados se sostienen al contrastar el modelo para los tres constructos (factores 4, 5 y 6 del diagrama) esto indicará que existe validez convergente. Por el contrario, la validez discriminante indica que otros constructos que no están muy relacionados no deben entrar a formar parte del constructo, o sea en nuestro ejemplo, que una medida de neuroticismo no tiene porqué saturar fuertemente en el factor psicoticismo ni en el de extraversión. En este sentido puede apreciarse cómo en el digrama causal se hipotetiza que ninguna de las medidas de otro constructo satura ni en el factor 4, ni en el 5, ni en el 6. Por ejemplo, todas las variables que miden neuroticismo (variables 1, 4, y 7) no saturan ni en el factor 5 (psicoticismo), ni en el 6 (extraversión). Si se comprueba que efectivamente las medidas de los distintos constructos no están relacionadas con constructos que no sean el que miden se dice que existe validez discriminante. Todas estas hipótesis se comprueban mediante el ajuste global y el analítico. El global nos indica si la representación de los datos mediante esta estructura es correcta, y el analítico si una a una todas las hipótesis son correctas.

En el diagrama causal puede apreciarse cómo los factores están unidos por una flecha continua. Generalmente se pueden dejar a estimar las correlaciones entre los seis factores. Se ha evitado poner todas esas relaciones bidireccionales por razones de simplicidad. Los errores generalmente se asumen como no correlacionados.

Ventajas del estudio de la validez convergente-discriminante mediante análisis factorial confirmatorio son:

1) que se tienen medidas de los factores no afectadas por los métodos de medición (Long, 1983).

2) que se pueden comprobar las correlaciones entre los constructos medidos sin que esta correlación esté afectada por el método de medición.

3) que se puede estudiar las relaciones que guardan los métodos de medida sin que se vean afectados por los constructos. Por ejemplo puede verse si las entrevistas a padres y profesores presentan altas o bajas correlaciones.

4) que se posee un potente test estadístico para comprobar el grado de ajuste del modelo a los datos, la prueba de chi-cuadrado (Bagozzi, 1981). En la matriz multirasgo-multimétodo esto se realiza mediante la simple observación de los datos, no mediante ninguna prueba estadística con niveles de probabilidad asociados.

5) que las pruebas de ajuste analítico nos permiten comprobar si existe algún problema localizado en algún constructo o método y no en el resto.

La matriz multirasgo-multimétodo no es la única forma de contrastar la validez convergente y discriminante. No siempre se dispone de diversos métodos de medición de distintos constructos, condición para poder construir una matriz de este tipo.

Un caso típico de comprobación convergente-discriminante es cuando se quiere conocer el grado en que convergen en uno o varios factores los items de un test, tal y como estudiábamos en la validez factorial. Junto a esto, puede estudiarse si estos items tienen validez discriminante con respecto a otros factores medidos por otro test.

### **Validez cruzada**

*Se entiende por validez cruzada la consistencia de los resultados a través de distintas muestras. Se dice que existe validez cruzada cuando los resultados obtenidos son estadísticamente iguales en muestras diferentes. Un resultado del tipo la inteligencia predice un 60% del rendimiento académico puede mantenerse en una muestra de alumnos de preescolar estadounidenses y otra muestra de preescolares españoles. En ese sentido se diría que se ha comprobado la validez cruzada. En resumen la validez cruzada no es más que la capacidad de generalización de los resultados obtenidos a nuevas muestras de una misma o diferentes poblaciones. Si estudiamos la validez de una escala en una muestra determinada mediante el uso de la regresión múltiple, se dice que existe validez cruzada si, calculando una regresión con las mismas variables, los coeficientes son similares y la capacidad predictiva no difiere estadísticamente (mediante el contraste del valor del coeficiente de determinación).*

*Todas las técnicas de modelado causal son susceptibles de ser sometidas a validación cruzada, igual que si se tratara de una regresión. Por un lado, se puede ver si los coeficientes hallados para las distintas muestras son similares en cuanto al sentido de la relación, a la intensidad y a su ajuste analítico; y por otro si el ajuste global del modelo es similar.*

Además de esta 'inspección visual' de los resultados para comprobar la validez cruzada, el programa EQS proporciona una herramienta estadística, el *análisis multimuestra*, para contrastar la igualdad entre muestras diferentes, la validez cruzada de un determinado modelo causal. Las técnicas de análisis multimuestra fueron desarrolladas principalmente en el campo del análisis factorial confirmatorio por Jöreskog (1971), quien elaboró el marco

estadístico para la estimación simultánea en diversos grupos. Aún cuando Jöreskog planteaba este tipo de análisis en el marco del factorial confirmatorio su uso es extensible a los modelos de estructuras de covarianzas.

Dentro de un análisis multimuestra las hipótesis de desigualdad-igualdad entre las muestras presentan diversos niveles y son, por tanto, más o menos restrictivas. Jöreskog (1971) proponía una estrategia que iba desde la hipótesis menos restrictiva de relativa igualdad de matrices de covariación sin especificar un modelo, hasta la más restrictiva de igualdad de todos los parámetros a estimar en las diversas muestras.

Se plantearían diversos grados de constricciones, hipótesis de igualdad, en los parámetros y se podrían contrastar todas estas igualdades o sólo las que interesaran (Anderson y Gerbing, 1988; Bentler, 1989; Kaplan, 1989):

a) Primero, sería la igualdad de varianzas-covarianzas entre las muestras sin explicitar un modelo concreto, ya propuesto por Box (1949) en el contexto de análisis de varianza multivariado.

b) Segundo, se probaría la igualdad de las saturaciones factoriales y de las relaciones entre las variables latentes. Diversos autores (Bentler, 1989; Jöreskog, 1971) plantean que es un paso necesario para poder plantear hipótesis más restrictivas, como la igualdad de varianzas y covarianzas entre factores. Se arguye que encontrar apoyo a una restricción de los valores de varianzas-covarianzas entre los factores es difícil de encontrar en la práctica si las cargas factoriales no son razonablemente iguales, y que aún cuando ajustara sería difícil concluir que se trata de los mismos factores latentes. Muthen y Christofferson (1981) sin embargo, proponen la conveniencia y plausibilidad de probar hipótesis de igualdad de factores incluso cuando las saturaciones factoriales no son iguales para todas las variables observables que conforman los factores. Esto es, plantean la posibilidad de lo que ellos denominaron una invarianza de medida parcial (partial measurement invariance).

c) El tercer paso sería constreñir a los mismos valores las varianzas y covarianzas de los factores latentes.

d) El último paso sería probar si los errores presentan las mismas varianzas y covarianzas a través de las muestras. Este último tipo de análisis es el menos realizado (Bentler, 1989).

Desde una perspectiva amplia, lo que se trata con un análisis factorial confirmatorio multimuestra es plantear hasta qué punto pueden mantenerse ciertas hipótesis precisas de igualdad intermuestras sobre ciertos parámetros claves (o todos) de un modelo factorial, testando mediante las técnicas estadísticas oportunas la

plausibilidad de cada una de esas constricciones. Bentler (1989) presenta una modificación del Lagrange Multiplier Test (LM) para evaluar en cada estimación qué igualdades inter-muestras son adecuadas, esto es, si se dejaran libres a estimar no producirían mejora en el  $c^2$  resultante final. Se presenta un LM test por cada restricción que se realiza entre grupos, que contrasta la hipótesis nula de que la restricción impuesta es cierta en la población. Resultados negativos contradicen esta hipótesis indicando que la igualdad inter-muestras para ese parámetro no puede sostenerse desde los datos. Si alguna de las igualdades planteadas no es plausible desde el LM test debe reestimarse el modelo sin esa restricción. Los resultados con la nueva estimación pueden mostrarnos que alguna otra restricción que estuviera cercana a la significatividad puede resultar ahora significativa, dando pie a una nueva estimación.

Antes de realizar análisis multimuestra habría que realizar análisis separados muestra a muestra (Kaplan, 1989). Los estudios separados pueden poner de manifiesto relaciones muy diferentes, con lo que el análisis multimuestra puede quedar descartado. La afirmación anterior es cierta desde dos puntos de vista alternativos. Por un lado, porque es necesario un adecuado ajuste al modelo hipotetizado en las diversas muestras para poder encontrar un ajuste global adecuado. Dado que el estadístico  $c^2$  global a calcular es, simplificando, una suma de las diversas funciones minimizadoras de cada submuestra ponderada por el tamaño de las muestras (Bentler, 1989), es fácil de entender que un adecuado ajuste al modelo hipotetizado (una función a minimizar lo más pequeña posible) es bastante necesario para encontrar un buen ajuste global. Alternativamente, un muy buen ajuste en algunas muestras podría hacer sostenible la igualdad de las muestras aún cuando el ajuste en una de ellas no fuera en absoluto justificable, lo que nos llevaría a un absurdo (Tomás, Meliá, Oliver y Sesé, 1993).

En resumen, un análisis multimuestra debe llevarse a cabo tras un conocimiento por separado de las muestras, y realizado en diversos grados, en un procedimiento iterativo que puede ir desde hipótesis menos restrictivas a más restrictivas.

Las ventajas de realizar los estudios en varias muestras a través del análisis multimuestra frente a la simple inspección visual de los datos son básicamente: a) se dispone de procedimientos para la comprobación del ajuste, tanto global como analítico, del modelo; y b) se dispone de una herramienta estadística, en este caso una modificación del LM test, que evalúa con rigurosidad estadística lo razonable de las diversas constricciones inter-muestras. La simple inspección visual de los datos puede llevar a menudo a errores gruesos en las conclusiones, como han analizado Meliá et al. (1993).



**Validez diferencial**

Cualquier hipótesis de diferencias entre grupos puede tratarse mediante los modelos que hemos visto hasta aquí. Además, se dispone de los análisis multimuestra para llevar a cabo de forma estadísticamente ajustada estas comparaciones.

### 3. Variables y su medida

Para los estudios empíricos de este trabajo se ha utilizado la Bateria de Seguridad Laboral (versión 3) de Meliá e Islas (1991). A continuación se presentarán todas las variables que la componen.

Tras una entrada que sirve para motivar al entrevistado, garantizarle el anonimato, enseñarle la mecánica del cuestionario e informar del interés y fines de la investigación se presenta la primera variable. Una pregunta abierta recoge, al principio de la batería, el *oficio o profesión* de la persona que contesta el cuestionario. Tras esta pregunta la batería recoge todas estas variables:

A- Como *características de los informantes o de sus trabajos* se recogen las siguientes:

A.1) *Sexo*.

A.2) Una variable que recoge *si el informante ocupa ahora el mismo puesto de trabajo de cuando sufrió su último accidente*. Esta variable es importante, dado que si el entrevistado hubiera cambiado de puesto estaría contestando a variables que tendrían que ver con su trabajo actual y no aquél en que sufrió el accidente.

A.3) *Edad*. Se recogen los años de edad en números, es una pregunta abierta.

A.4) *Nivel de estudios*. Con cinco categorías: ninguno; primarios; formación profesional primer grado; formación profesional segundo grado; y BUP o universidad.

A.5) *Antigüedad*. Esta variable viene medida por tres indicadores, uno para medir la antigüedad en la empresa, otro para la antigüedad en el puesto de trabajo y el tercero para la antigüedad en el oficio. Todos los indicadores se contestan en años y meses (pregunta abierta).

A.6) *Experiencia*. Se mide con dos indicadores. El primero pregunta acerca del tipo de trabajo que se desempeñaba antes. Con una escala de respuesta que incluye las siguientes alternativas: primer trabajo; otro trabajo diferente al que ahora realiza; otro trabajo parecido al actual; otro trabajo igual al actual; siempre ha trabajado en el mismo sitio pero con diferentes puestos; y ha trabajado siempre en el mismo puesto. El segundo indicador pregunta si era necesario (si/no) tener experiencia previa similar para entrar en el puesto de trabajo actual.

A.7) *Horas extraordinarias*. Mide la frecuencia de horas extraordinarias con una escala de cuatro puntos: nunca u

ocasionalmente; alguna vez al mes; alguna vez a la semana; casi todos los días. Tradicionalmente esta variable se ha considerado como relacionada con los accidentes, sobre todo por la sobrecarga y cansancio que puede suponer.

A.8) *Tipo de horario*. Con cuatro categorías: jornada partida; intensiva; horario parcial (sólo algunas horas); y sistema de turnos.

A.9) *Tipo de contrato*, es una variable medida con seis niveles: contratos por terminación de tareas o prácticas; contrato hasta seis meses; entre 7 y 12 meses; entre 1 y 3 años; contrato a tiempo parcial; y fijos.

A.10) Una variable dicotómica cuestiona si el entrevistado es empleado o trabajador, o bien es de *nivel jerárquico superior* (supervisor o superior).

A.11) Necesidad de *equipo de protección personal* en el desempeño laboral por motivos de seguridad. Ciertos trabajos requieren de equipo de seguridad. La escala que pretende medir el uso y condiciones de los equipos está compuesta por un total de doce items. 8 de los doce se contestan con una escala dicotómica, 1 mediante una escala de cuatro puntos (desde pésimas condiciones a perfectas condiciones), y los 3 restantes mediante una escala de tres puntos (desde nunca u ocasionalmente hasta siempre). Previamente a los doce items del cuestionario se presenta una pregunta filtro que pregunta sobre si el trabajador necesita utilizar equipo, en caso negativo se salta el cuestionario.

B- Como *variables descriptivas de las empresas* a las que pertenecen los integrantes de la muestra se integran en la batería las siguientes:

B.1) *Propiedad de la empresa* (pública o privada).

B.2) *Ambito de localización* de la empresa, con cuatro anclajes: local, regional, nacional y multinacional.

B.3) *Ramo de actividad* de la empresa.

C- Variable *clima de seguridad*.

El clima de seguridad se define como todas aquellas acciones realizadas por la empresa en materia de seguridad, tanto para aumentarla como para disminuirla.

El constructo clima puede ser dividido teóricamente en tres variables relativamente independientes, con el objeto de estudiar sus relaciones con otras variables de seguridad por separado. Esas tres variables teóricamente discernibles son:

C.1) *Formación.* Medida mediante una escala de seis items. El primer indicador mide la formación para el puesto que se dió al entrar en la empresa. El segundo y tercer indicadores miden la formación específica de seguridad que ofrece la empresa y la frecuencia de ésta. Los items cuatro, cinco y seis de la escala miden otros aspectos de formación o información ofrecidos por la empresa.

C.2) *Estructura de seguridad.* Por ley, y dependiendo del número de trabajadores de la empresa, es obligatoria la existencia de un comité y/o de un encargado de seguridad. Una escala de cuatro items de respuesta dicotómica (sí/no) mide la existencia de departamentos, comités y/o encargados de seguridad en las empresas de la muestra, así como el conocimiento que los trabajadores tienen de sus funciones y las personas que las desempeñan.

C.3) *Política organizacional en seguridad.* Recoge en ocho items el interés, la prioridad que supone la seguridad para la empresa, y la posibilidad de informar adecuadamente a la gerencia de los problemas y aspectos relacionados con la seguridad. Los items se contestan con escalas de respuesta que van desde la respuesta dicotómica a los cuatro anclajes.

D- *Variables de contexto social.* Dos grupos se han definido como relevantes para aumentar o disminuir la seguridad laboral del informante dentro de su contexto social. Por un lado los superiores o supervisores inmediatos del trabajador. Por otro, los propios compañeros tomados como un grupo. Para cada variable se dispone de un cuestionario en la batería de seguridad.

D.1) *Respuesta de los supervisores o mandos intermedios en seguridad.* La respuesta de los superiores o supervisores en seguridad puede ser un importante determinante del comportamiento seguro o inseguro de los subordinados. Se define esta respuesta como: a) el comportamiento seguro o inseguro que los supervisores realizan (que puede afectar al comportamiento de los empleados porque sirve como regla o porque actúa como modelo); b) las contingencias que ofrece al subordinado por su conducta segura o insegura; y c) la actitud en seguridad de los supervisores. Estos contenidos se recogen en un cuestionario de siete items con escala de respuesta con escala de respuesta tipo Likert, de los cuales cinco presentan 3 anclajes, y 4 anclajes los dos items restantes.

D.2) *Respuesta de los compañeros en seguridad.* Presenta los mismos tres contenidos básicos que la anterior variable. Esto es: contingencias, comportamientos y actitud, pero referidos esta vez al grupo de compañeros. Para medir estos contenidos se utilizó una escala de diez items, cuatro de ellos dicotómicos, otros cuatro de tres anclajes, uno de cuatro y otro más de cinco anclajes.

E- *Variables de respuesta de seguridad de los entrevistados.*

Los entrevistados son el foco central de análisis, dado que son los que informan sobre los accidentes ocurridos. Dos tipos de respuestas del sujeto se operacionalizan en la batería de seguridad. Por un lado una respuesta subjetiva, la actitud hacia la seguridad. Por otro, una respuesta comportamental, la conducta hacia la seguridad de los empleados.

E.1) *Actitud hacia la seguridad* de los entrevistados. Para la medición de esta actitud se utilizó una escala propuesta por Leather (1988). Este cuestionario presenta un total de doce ítems, que cuestionan al trabajador sobre las actitudes que tiene sobre diferentes aspectos de la seguridad en el trabajo. La escala de respuesta es de cinco anclajes, desde muy en desacuerdo a muy de acuerdo.

E.2) *Conducta hacia la seguridad* de los entrevistados. La escala de conducta de seguridad está compuesta por diez ítems de cinco anclajes. Cada ítem supone una conducta segura o insegura, y el sujeto valora si la realiza o no y con qué frecuencia.

F- *Variables de riesgo laboral.* La identificación y cuantificación de los riesgos laborales ha presentado en la literatura sobre seguridad un problema, con numerosas vías de operacionalización. No obstante, la distinción que aquí se propone es novedosa. Distingue y trata de medir diferencialmente dos aspectos del riesgo, el riesgo basal y el real.

F.1) *Riesgo basal.* El riesgo basal es el riesgo intrínseco a cada puesto de trabajo. Es más peligroso un puesto de montador de grúa que el de profesor universitario. Estos riesgos propios de cada puesto laboral se tratan de medir mediante una escala de diecisiete ítems de respuesta dicotómica. Preguntan acerca de la presencia y gravedad de algunos de los riesgos más comunes en ambientes laborales, riesgos de toxicidad, de accidente mecánico, etc.

F.2) *Riesgo Real.* Aún cuando cada profesión y puesto o incluso lugar de trabajo presenta unos riesgos propios, éstos pueden amortiguarse o incrementarse mediante unas oportunas medidas de seguridad. En otras palabras, pueden prevenirse o potenciarse. Los riesgos que aún permanecen son el riesgo real, que se intenta operacionalizar en la batería de seguridad mediante un cuestionario de siete ítems con cinco anclajes.

G- *Variables de efectos sobre el trabajador.* Todas las variables expuestas pueden tener, al menos, dos tipos de efectos sobre el trabajador que trata de recoger la batería. Por un lado pueden generar tensión psicológica. Por otro, y más importante por ser foco

de estudio específico de este trabajo pueden desencadenar accidentes laborales.

G.1) *Tensión psicológica.* Un cuestionario de ocho items se ha utilizado para medir el grado de tensión psicológica a que está sometido el trabajador. La escala que se utiliza es una versión reducida de la presentada por Meliá, Peiró, Zornoza, Sanz, Morte y Zurriaga (1987). Se han eliminado siete items intentando conservar el contenido teórico del constructo, pero haciendo la escala de más rápida aplicación. Se contesta con una escala de respuesta de verdadero falso.

G.2) *Accidentabilidad.* Para medir la accidentabilidad la escala presenta diversos indicadores, que pueden utilizarse por separado o bien de forma compuesta. El primer indicador es una pregunta para distinguir entre accidentados y no accidentados. El segundo indicador recoge el número de accidentes que ha sufrido cada entrevistado. Luego se recoge para los últimos tres accidentes el año y mes de ocurrencia y la gravedad.

## **4. Objetivos**

### **4.1. Objetivos generales**

a) *Presentar y validar cuestionarios para la medición de las variables que se asumen relacionadas con los accidentes directa o indirectamente.*

b) *Construir modelos causales de los accidentes laborales para la muestra bajo estudio.* Partiendo de la concepción de accidente como resultado (efecto) de una cadena de eventos, se plantea estudiar posibles cadenas de causas-efectos que puedan llevar a accidentes. El modelado causal (path analysis, estructuras de covarianza) es una técnica que precisamente pone a prueba hipótesis de causación entre variables, por lo que es especialmente recomendable su utilización para este propósito.

No obstante, han de considerarse algunas dificultades en este campo. La variable siniestralidad, efecto último de los modelos causales de accidentes es especialmente difícil estadísticamente por su poca variabilidad y extrema asimetría. La gran mayoría de los sujetos de nuestra muestra han sufrido un accidente, unos pocos dos y muy pocos tres o más. A pesar de que se generan índices compuestos para evitar este problema, lo cierto es que la precisión de la predicción de los accidentes podría verse menguada por este motivo.

c) *Evaluar la capacidad predictiva sobre las conductas seguras o inseguras de los trabajadores, antecedente inmediato de los accidentes.* La concepción más extendida actualmente en la literatura sobre seguridad laboral es la de que la mayoría de accidentes son resultado de algún tipo de acto inseguro. Esta hipótesis debe ser probada, intentando conectar conductas inseguras con accidentes. No obstante, la anterior concepción no responde a una pregunta básica: ¿Qué provoca la realización de un acto inseguro por parte de un trabajador?. Por otro lado, la conducta es el foco principal de análisis de los estudios sobre prevención de accidentes, en el sentido de utilizar técnicas para disminuir las conductas inseguras y aumentar las seguras. Por último, los problemas estadísticos que se presentan en la variable accidentabilidad pueden no presentarse o, al menos, no en tanta intensidad con la variable conducta. En resumen, por estas razones creemos que es importante predecir los comportamientos seguros e inseguros que se producen en el desempeño laboral, y dado que la hipótesis de causación múltiple es probablemente la más plausible, nuevamente el uso del modelado causal ofrece especial interés.

#### **4.2. Objetivos específicos**

a) El objetivo general de presentación de cuestionarios (objetivo general c) supone una serie de objetivos psicométricos específicos. Para cada variable medida a través de cuestionarios se deberán calcular los siguientes datos psicométricos, que se presentan como objetivos:

a.1) cálculos de la fiabilidad de la escala.

a.2) Cálculo de la fiabilidad y homogeneidad de los items.

a.3) Cálculo de la validez de la escala

a.4) Cálculo de la validez de los items.

a.5) Dimensionalidad de las escalas. Es objetivo metodológico prioritario de este trabajo el estudio de esta dimensionalidad mediante análisis factorial confirmatorio probando diversos métodos de estimación y calculando diversos índices de ajuste.

b) Ubicar mediante técnicas de modelado causal a la tensión laboral como antecedente o consecuente de los accidentes laborales.

## **5. Hipótesis**

a) Las variables psicológicas, interpersonales y sociales medidas tienen un efecto directo e indirecto sobre los accidentes laborales, de forma que pueden predecir en cierta medida éstos.

b) Los accidentes están básicamente producidos por conductas inseguras de los trabajadores, junto a riesgos presentes en toda actividad profesional. Esta hipótesis puede plantearse de otra forma más. A igualdad de riesgos, es tanto más probable un accidente cuanto mayor número de conductas inseguras o menor número de conductas seguras se den. En otras palabras todavía más específicas. Si se controla la variable riesgos debe haber una correlación negativa entre conducta segura y accidentes y una correlación positiva entre conductas inseguras y accidentes.

c) Aún cuando riesgos y conductas seguras/inseguras del propio trabajador son los directos predictores de accidentes, un accidente se produce por una serie de eventos (causas) que de forma concatenada o sumativa llevan a que crezca la probabilidad de accidentarse y como consecuencia a que aumenten los siniestros.

d) La tensión psicológica es más un consecuente que un antecedente de los accidentes.

e) Hay un número de hipótesis específicas de relación entre las variables que se tratan en este trabajo que se tratarán en cada capítulo específico en cierta profundidad y que serán puestas a prueba básicamente mediante modelos causales en el correspondiente capítulo.





3.

## Clima organizacional hacia la seguridad.

### **1. Introducción**

La empresa, sobre todo a nivel de dirección, es el máximo responsable de generar lo que ha venido denominándose 'clima' o 'cultura'. El clima de una empresa tiene diversos aspectos discernibles, entre ellos el clima de seguridad. Puede definirse este clima de seguridad como las percepciones que tienen los trabajadores sobre las acciones que lleva a cabo la empresa para aumentar o disminuir la seguridad laboral, como la percepción del interés o prioridad que la empresa confiere al tema de la seguridad entre los objetivos empresariales. Estas acciones, determinadas desde niveles elevados de la empresa, desde la dirección, pueden jugar un importantísimo papel sobre variables que están afectando directa o indirectamente a la probabilidad de que se produzcan accidentes. Así, parece claro que cómo entienda la empresa la seguridad y qué acciones lleve a cabo en esa dirección determina primariamente las acciones en seguridad de los supervisores y mandos intermedios. Estos superiores, que en muchas ocasiones trabajan controlando, dirigiendo, colaborando directamente con el trabajador han sido muchas veces considerados desde la literatura como el eslabón entre empresa-trabajador y como uno de los puntos clave en la prevención de accidentes. Estos supervisores, parece claro, actúan con referencia a una política diseñada por la empresa,

una política que, al tratar aspectos de seguridad, puede denominarse política de seguridad. También se ha considerado que no tienen las mismas armas de prevención de accidentes los trabajadores bien formados para el puesto y bien formados para tomar las medidas de seguridad oportunas, que aquellos que no han recibido indicaciones, formación, información sobre cómo prevenir accidentes y cómo tomar las adecuadas medidas de seguridad. Finalmente la existencia dentro de la estructura de la empresa de uno o varios encargados de seguridad, incluso de un comité o departamento de higiene y seguridad en el trabajo activo y que ejerza medidas de control de riesgos puede jugar un papel relevante en la prevención de accidentes. Todos los aspectos enumerados hasta aquí han sido considerados en este trabajo como clima organizacional hacia la seguridad.

Hemos distinguido dentro de este amplio constructo de clima tres aspectos o variables muy relacionadas, pero teóricamente discernibles. La idea es conseguir evaluar más analíticamente estos tres grandes aspectos, estudiando diferencialmente sus relaciones y posibles efectos sobre otras variables implicadas en la cadena causal que desemboca con cierta frecuencia en accidentes laborales. Los tres aspectos, que merecerán medición aparte son: a) formación de seguridad; b) política de seguridad; y c) estructura de seguridad.

Para cada variable se ha construido una escala de medida que será sometida a un análisis psicométrico. Se estudiará la fiabilidad de la escala, se analizarán los ítems -homogeneidad, fiabilidad-, se analizará la validez factorial de la escala y se realizarán estudios de validez criterial de cada ítem.

## 2. Formación-Información

### 2.1. Medida

La escala de formación-información está formada por seis ítems, los tres primeros con escala de respuesta de seis anclajes y los tres últimos dicotómicos. Los contenidos, junto con la escala de respuesta, se presentan a continuación:

#### ESCALA:

1. ¿Se le dió formación para el puesto de trabajo al entrar en la empresa?

0                      1                      2                      3                      4                      5  
Ninguna    Muy Poca    Alguna    Bastante    Mucha    Muy completa

2. ¿Ha recibido Ud. en su empresa formación sobre seguridad para la prevención de accidentes laborales?

0                      1                      2                      3                      4                      5  
Ninguna    Muy Poca    Alguna    Bastante    Mucha    Muy completa

3. ¿Con qué frecuencia ha recibido Ud. formación (cursillos, charlas, reuniones...) sobre seguridad en su empresa?

0                      1                      2                      3                      4                      5  
Nunca            Muy            Una vez            Dos o tres            Una vez            Casi  
                          ocasionalmente    al año            veces al año            al mes            constantemente.

4. En mi empresa hay carteles sobre seguridad en el trabajo ..... Si    No

5. Mi Empresa realiza cursillos o charlas sobre seguridad ..... Si    No

6. Hacemos reuniones de trabajo específicas sobre problemas de seguridad ..... Si    No

### 2.2. Estructura factorial confirmatoria

La estructura factorial ha sido probada mediante análisis factorial confirmatorio estimado mediante mínimos cuadrados generalizados y mediante máxima verosimilitud. Se han probado diversas estructuras teóricas, modelos competitivos para representar los datos. En concreto se han probado cuatro estructuras factoriales:

a) La primera es unifactorial.

b) La segunda es una estructura bifactorial. El primer factor incluye los ítems 1, 2, 3, y 5. Todos ellos teóricamente referidos a formación formal recibida en la empresa. El segundo factor incluye

los items 4 y 6 que suponen una formación menos formal o, más bien, acciones de información sobre seguridad.

c) La tercera se plantea tras calcular el ajuste de la segunda y por sugerencia estadística del Lagragian Multiplier Test, que informa como necesaria la relación del factor 5 con el factor 2, además de con el factor 1.

d) la cuarta estructura teórica elimina la relación del ítem 5 con el factor 1 y deja una estructura compensada de tres items por factor. El factor 1 incluye los items 1, 2, 3. El segundo factor los tres restantes (4, 5, y 6).

La tabla 3.1 muestra todos los índices de ajuste global calculados para los distintos modelos factoriales, que nos permitirán escoger la estructura más adecuada a los datos.

Tabla 3.1. Índices de ajuste global de las cuatro estructuras hipotetizadas para la escala de formación-información. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; CAK=Akaike information criterion; CSK=Schwarz information criterion; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: GLS=mínimos cuadrados generalizados; ML=máxima verosimilitud.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
UN	GLS	38.0	9	<001	.33	.846	4.22	.609	.949	.983	.979	.987
FACTOR	ML	50.9	9	<001	.44	.800	5.66	.722	1.06	.805	.716	.830
BIFACTO-	GLS	38.0	8	<001	.33	.846	4.75	.609	1.00	.983	.975	.987
RIAL (1)	ML	49.4	8	<001	.43	.804	6.18	.709	1.10	.811	.684	.832
BIFACTO-	GLS	12.3	7	.088	.10	.947	1.76	.430	.826	.995	.995	.998
RIAL (2)	ML	13.9	7	.052	.12	.940	1.99	.444	.840	.947	.940	.972
BIFACTO-	GLS	12.7	8	.121	.11	.945	1.59	.410	.778	.994	.996	.998
RIAL (3)	ML	13.9	8	.082	.12	.940	1.74	.421	.789	.947	.955	.976

A la vista de los índices, y como primera apreciación, cabe hablar de un mejor ajuste mediante la estimación mínimo cuadrática que mediante la máximo verosimil. Y esto a través de los cuatro modelos y a través de todos los índices evaluados.

Todos los índices de ajuste presentados para los cuatro modelos en competencia ofrecen resultados en la misma dirección. Los dos primeros modelos, el unifactorial y el bifactorial con cuatro items cargando en un factor y dos en el restante son inadecuadas representaciones de los datos, no aportando un grado significativo de explicación de las varianzas-covarianzas de los datos.

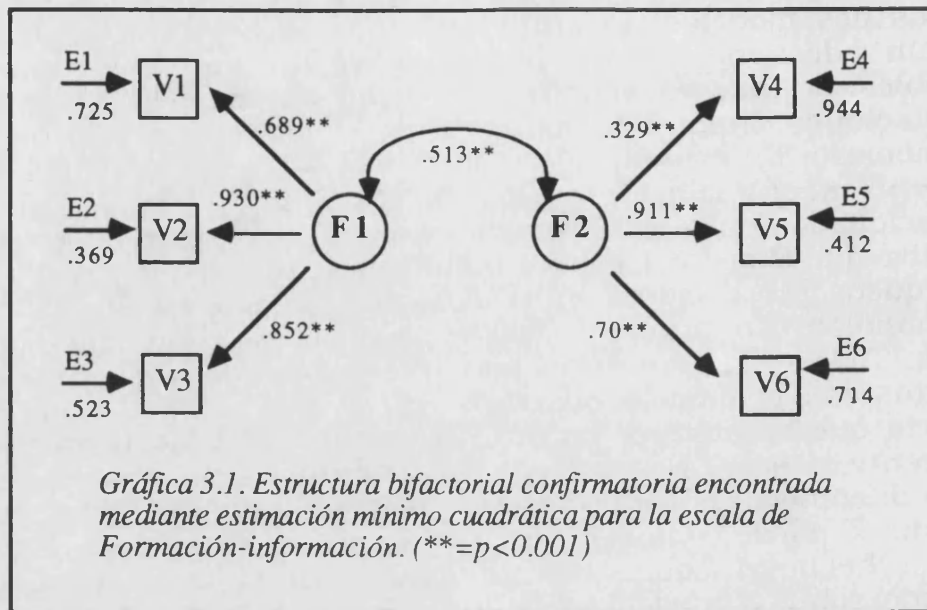
Los dos modelos restantes (bifactorial (2) y bifactorial (3)) sí resultan adecuadas representaciones de los datos, son modelos plausibles. La única diferencia entre estos dos modelos en términos de relaciones hipotetizadas es que el bifactorial (3) es más parsimonioso. En este modelo el factor 1 explica las variables 1, 2 y 3, mientras el factor 2 explica la 4, 5 y 6, con los factores correlacionados entre sí. El mismo esquema se sigue en el modelo tres salvo que el factor 1 explica también a la variable 5, el ítem 5. Así pues, queda por decidirse entre el modelo bifactorial (2) y (3), más parsimonioso. En primer lugar, cabe observar detenidamente los índices. Todos ellos muestran una leve mejora de la explicación de los datos con el modelo bifactorial (3), a excepción de una ligera ganancia que se produce en el CAK para el modelo bifactorial (2) - .947 frente a .945-. No obstante, cabe evaluar hasta que punto las ligeras diferencias encontradas son realmente significativas. A esta pregunta se puede contestar desde dos pruebas estadísticas. Por un lado, el hecho de que el Wald Test para eliminar parámetros calculado sobre el modelo bifactorial (2) ofreció como resultado que la pérdida de la flecha entre el factor 1 y el ítem 5 no suponía una pérdida estadísticamente significativa de explicación en el modelo. Recordemos que es precisamente esta relación la que diferencia entre el bifactorial (2) y el (3). Por otro lado, puede computarse un estadístico de diferencia entre las dos chi-cuadrado. Realizado este análisis los resultados son para mínimos cuadrados generalizados

$\Delta\chi^2=0.334$ , con  $\Delta gl=1$  que no es estadísticamente significativo,

y para máxima verosimilitud

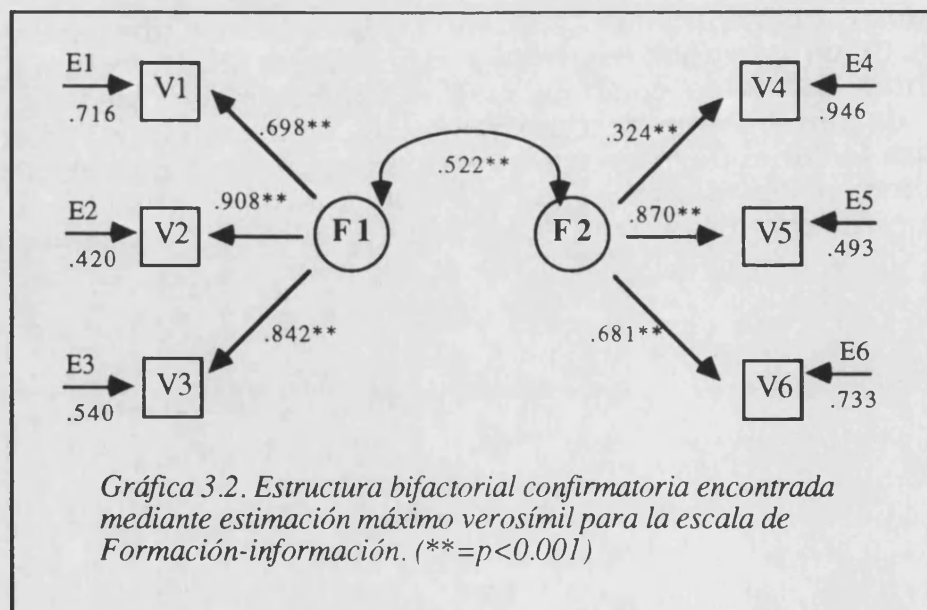
$\Delta\chi^2=0.029$ , con  $\Delta gl=1$  que tampoco resulta significativo.

Dados estos resultados apuntando a que la estructura bifactorial con tres ítems cargando en cada factor resulta tan adecuada para representar los datos como la más compleja bifactorial (2), y por razones de parsimonia, se opta por el modelo bifactorial (3) como estructura factorial de la escala en esta muestra. En las gráficas 3.1 y 3.2 se ofrecen los valores de las relaciones para el modelo escogido para las estimaciones mínimo cuadráticas y las máximo verosímiles.



Puede apreciarse en la gráfica 3.1 correspondiente a la estimación mínimo cuadrática cómo todas las relaciones hipotetizadas han resultado estadísticamente significativas. Los ajustes analíticos mediante el test z han resultado satisfactorios, indicando que las relaciones encontradas no son debidas al azar. La correlación entre los dos factores es alta y significativa.

Muy similares son los valores estimados para las relaciones en la estimación máximo verosímil cuyos datos se ofrecen en la gráfica 3.2. Los resultados de ajuste analítico son también concordantes con los hallados mediante GLS.



Se plantea ahora si los dos métodos de estimación aplicados, que se basan en el supuesto de distribución multivariada, eran procedentes. Puede aplicarse una prueba de bondad de ajuste de Kolmogorov-Smirnov para comprobar la normalidad univariada. Como se vió en el capítulo de metodología, para que se cumpla la distribución normal multivariada, supuesto a la base de la mayor parte de técnicas multivariadas, es necesario que cada variable por separado siga una distribución aproximadamente normal.

*Tabla 3.2. Prueba de Kolmogorov-Smirnov de una muestra para comprobar el ajuste de los items de la escala de formación-información a la distribución normal univariada.*

ITEM	N muestral	Diferencia máxima	Probabilidad
1	123	0.541	0.000
2	123	0.500	0.000
3	122	0.500	0.000
4	122	0.500	0.000
5	119	0.500	0.000
6	124	0.500	0.000

Los resultados de la prueba de bondad de ajuste apuntan a que las variables no cumplen una ley normal. Esto es altamente previsible de los tres últimos items, dado que son dicotómicos y difícilmente pueden distribuirse normalmente, pero sí es relevante para los tres primeros, que tienen seis anclajes.

También se presentan los datos descriptivos de asimetría y curtosis univariada en la tabla 3.3. Cuanto más alejados de cero se encuentren los valores más asimetría y apuntamiento muestran, respectivamente, para cada variable.

*Tabla 3.3. Datos de asimetría y curtosis para los items de la escala de formación-información.*

	Asimetría	Curtosis
Item 1	0.5505	-0.725
Item 2	0.9621	-0.445
Item 3	1.244	-0.162
Item 4	0.7632	-1.417
Item 5	0.942	-1.112
Item 6	1.500	0.250

Todas las variables presentan cierto grado de apuntamiento o de curtosis, o ambas tendencias a la vez.



COEFICIENTE DE MARDIA MULTIVARIADO=15.0886  
 Coeficiente esperado para la normal  
 multivariada=8.2572

En cuanto al apuntamiento propiamente multivariado, el programa EQS calcula el coeficiente de Mardia multivariado. Este coeficiente indica el grado de curtosis multivariado, que puede compararse con el esperado si las variables cumplieran la distribución normal multivariada. Como puede verse en el párrafo que presenta estos dos índices el alejamiento del coeficiente de Mardia frente al esperado bajo la multinormalidad es importante, indicando que el supuesto de normalidad multivariable no es sostenible.

A la vista de los datos de asimetría y curtosis de los items que componen la escala de formación-información se optó por estimar la estructura factorial que mejor ajuste presentaba mediante métodos arbitrarios. En concreto se utilizó AGLS.

En la tabla 3.4 se presentan los datos de ajuste global del modelo retenido mediante mínimos cuadrados y máxima verosimilitud como mejor representación de los datos, el modelo bifactorial (3). Este modelo incluye dos factores con tres items adscritos cada uno, el primer factor de formación y el segundo de información en seguridad. Se ofrecen los datos de ajuste global para los tres métodos de estimación. Los índices calculados en base a los métodos de máxima verosimilitud y mínimos cuadrados generalizados ya se han ofrecido anteriormente, pero se vuelven a presentar para facilitar la comparación entre métodos.

Tabla 3.4. Índices de ajuste global de la estructura bifactorial retenida para la escala de formación-información. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; CAK=Akaike information criterion; CSK=Schwarz information criterion; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: GLS=mínimos cuadrados generalizados; ML=máxima verosimilitud; AGLS= estimación arbitraria.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
Modelo de dos fac- tores (3)	GLS	12.7	8	.121	.11	.945	1.59	.410	.778	.994	.996	.998
	ML	13.9	8	.082	.12	.940	1.74	.421	.789	.947	.955	.976
	AGLS	8.77	8	.361	.07	.962	1.09	.303	.613	.926	.986	.992

Tanaka (1984) comprobaba que en las técnicas de modelos causales el uso de estimaciones basadas en la distribución

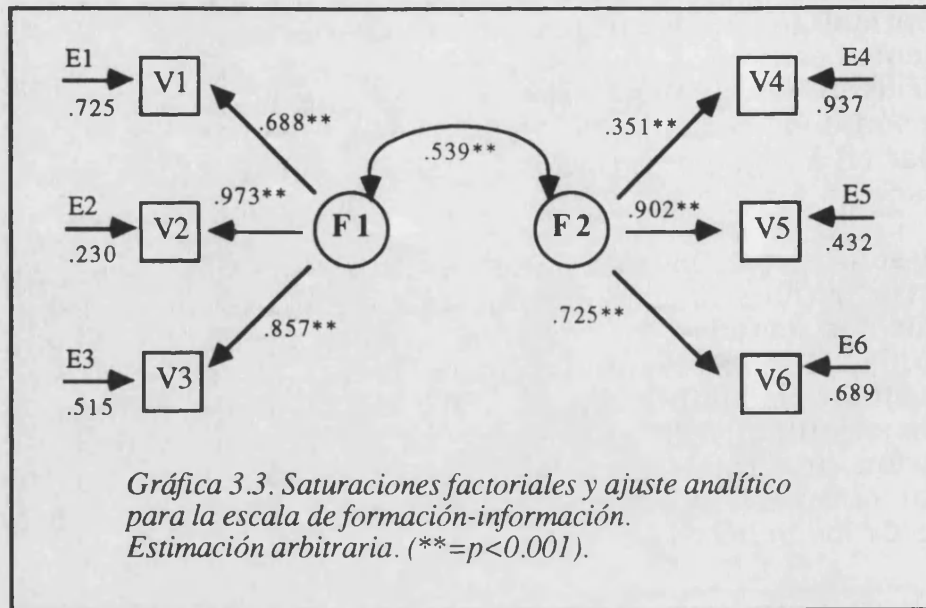
multinormal generaba errores. Cifraba los errores no al nivel de los coeficientes estimados, sino en la evaluación de las probabilidades de significación estadística. Esto es, más que los coeficientes era la evaluación de la probabilidad de la prueba chi-cuadrado y los errores estándar en los que se basan las pruebas z donde se encontraban los errores.

Si se observan los resultados de ajuste global se apreciarán las diferencias en la evaluación del ajuste entre la estimación arbitraria y los otros dos métodos. La probabilidad de la prueba chi-cuadrado ha pasado de .12 y .08, respectivamente, para ML y GLS a un valor de mejor ajuste mediante AGLS (.36). El resto de índices ofrecen también resultados de ajuste algo superiores, en general, para la estimación arbitraria, pero las diferencias son menores. Hay que recordar que para la estimación arbitraria el programa EQS ofrece, además de los índices habituales, otros dos, GFI y AGFI.

INDICE DE BONDAD DE AJUSTE, GFI= .975 Indice de bondad de ajuste ajustado, AGFI= .935
--

Estos dos coeficientes comparan, como hemos visto, la matriz de residuales frente a la matriz de datos. Pueden interpretarse como proporción de varianza de las variables explicada por la estructura factorial. Para la estructura bifactorial hipotetizada en este caso, los dos índices ofrecen una altísima proporción de varianza explicada, muestra de un adecuado ajuste. Los resultados de estos dos índices son, pues, concordantes con el resto de los calculados.

Además del ajuste global, otro centro de interés son las estimaciones de las saturaciones factoriales y varianzas de error que ofrece la estimación arbitraria. Y qué evaluación de la significatividad de estos coeficientes ofrece. En la gráfica 3.3 se ofrecen estos datos de ajuste analítico para la estructura bifactorial hipotetizada.



Como puede verse en la gráfica 3.3, los resultados de ajuste analítico son satisfactorios. Todos los ítems se muestran significativamente relacionados ( $p < 0.01$ ) con sus respectivos factores latentes. El ítem cuatro es el único que, aún estando relacionado con cierta fuerza, presenta una saturación algo más baja, mostrando menos homogeneidad con el factor.

En la siguiente tabla se presentan juntas las saturaciones encontradas para el modelo de dos factores retenido como mejor ajuste. Interesa comprobar si, como afirmaba Tanaka (1984), las estimaciones de los tres métodos son consistentes y ofrecen resultados similares.

Tabla 3.5. Saturaciones factoriales estimadas mediante mínimos cuadrados generalizados, máxima verosimilitud y estimación arbitraria.

	GLS	ML	AGLS
Item 1	.689	.698	0.688
Item 2	.930	.908	0.973
Item 3	.852	.842	0.857
Item 4	.329	.324	0.351
Item 5	.911	.870	0.902
Item 6	.70	.681	0.725
Relación factores	.513	.522	0.539

Efectivamente, los resultados a través de métodos de estimación son no sólo consistentes sino prácticamente iguales. Las diferencias se encuentran al nivel de las centésimas. Esta similitud entre estimaciones es tanto a nivel de las saturaciones factoriales como de

correlación entre los factores. Lógicamente los errores son, también, muy similares.

### 2.3. Fiabilidad y homogeneidad de la escala y los ítems

La escala de formación-información presenta una adecuada consistencia interna medida por el coeficiente alpha de Cronbach, que alcanza el valor de 0.757. Otros coeficientes de fiabilidad calculados pueden consultarse en la tabla 3.6. Todos se sitúan por encima del .75 y el más elevado de Spearman-Brown alcanza el valor de 0.863. Sólo son más bajos los coeficientes alpha para los ítems pares e impares, resultado lógico por haberse calculado para la mitad de ítems de la escala, tan sólo tres ítems, en cada caso.

Tabla 3.6. Coeficientes de fiabilidad, consistencia interna, de la escala de formación-información.

<i>Consistencia interna</i>	
Correlación dos mitades	0.759
Coeficiente Spearman-Brown	0.863
Coeficiente Guttman-Rulon	0.817
Coeficiente alpha total	0.757
Alpha ítems impares	0.648
Alpha ítems pares	0.358

Por su parte los datos de fiabilidad y homogeneidad de los ítems pueden consultarse en la tabla 3.7, junto con la media y desviación típica. Los coeficientes de homogeneidad y los coeficientes de homogeneidad corregidos (excluyendo el ítem) ofrecen, en general, resultados satisfactorios superiores al .48. No obstante, los ítems cuatro y seis presentan una baja homogeneidad y baja fiabilidad, indicando un grado menor de consistencia con el resto de ítems. El ítem cuatro pregunta sobre la existencia de carteles sobre seguridad en el trabajo, y el ítem seis sobre si se hacen reuniones sobre seguridad. Estos dos tipos de acciones pueden no ser vistos como propiamente formativos, sino más bien como solamente informativos. Son, no obstante, y en nuestra opinión contenidos relevantes teóricamente, ya que en muchas empresas ésta es prácticamente la única acción formativa o, al menos, informativa que se pone en marcha.

Tabla 3.7. Media, desviación típica, coeficiente de homogeneidad (correlación ítem-total), fiabilidad, homogeneidad excluyendo el ítem y coeficiente alpha que quedaría en la escala si se excluyera cada ítem, para cada ítem de la escala de formación-información.

ITEM	Media	D. T.	Homogeneidad	Fiabilidad	Homogen. sin el ítem	Alpha sin el ítem
1	1.783	1.619	0.814	1.317	0.641	0.680
2	1.296	1.684	0.907	1.527	0.802	0.614

3	1.235	1.716	0.875	1.502	0.737	0.644
4	0.322	0.467	0.368	0.172	0.282	0.770
5	0.287	0.452	0.550	0.249	0.481	0.753
6	0.200	0.400	0.411	0.164	0.340	0.767

Un dato importante que nos ofrece la tabla 3.7 es el alpha que quedaría en la escala si se eliminara el ítem foco de estudio. Como puede verse se produciría una pérdida sustancial si se eliminaran los ítems 1, 2 y 3. Sin embargo, no se produciría casi pérdida si se excluyeran los ítems 4, 5, y 6, tampoco se producirían ganancias significativas.

Además de como estudio de la dimensionalidad el análisis factorial confirmatorio se puede utilizar para medir la fiabilidad de los ítems. Las cargas factoriales al cuadrado son una medida de la fiabilidad de los ítems. En la tabla 3.8 se encuentran las cargas factoriales elevadas al cuadrado.

*Tabla 3.8. Saturaciones factoriales estimadas mediante mínimos cuadrados generalizados, máxima verosimilitud y estimación arbitraria elevados al cuadrado, como medida de la fiabilidad de los ítems. Los resultados indican la proporción de la varianza de cada ítem explicada por los factores latentes.*

	GLS	ML	AGLS
Item 1	0.474	0.487	0.473
Item 2	0.864	0.824	0.946
Item 3	0.725	0.708	0.734
Item 4	0.108	0.104	0.123
Item 5	0.829	0.756	0.813
Item 6	0.490	0.463	0.525

El ítem cuatro es el único que no se consigue explicar adecuadamente desde su factor latente. Si recordamos este ítem tampoco mostraba un buen comportamiento en los análisis de fiabilidad ni de homogeneidad. La consistencia interna de este ítem con el resto de la escala no es clara.

Puede calcularse el índice de fiabilidad de los ítems desde el análisis factorial confirmatorio desde la fórmula propuesta por Bagozzi (1981). Esta fórmula, más compleja tiene en cuenta no solo la saturación del ítem con el factor, sino que incluye también la varianza del factor y la varianza de error. La fórmula, ya presentada en el capítulo II, es la siguiente:

$$\rho_i = (\lambda_i^2 \text{ var } A) / (\lambda_i^2 \text{ var } A + \psi_i)$$

que, aplicada a las estimaciones mínimo cuadráticas para los seis ítems de la escala, ofrece los resultados de la tabla 3.9.

Tabla 3.9. Índices de fiabilidad de los ítems calculados desde los resultados del análisis factorial confirmatorio por la fórmula propuesta por Bagozzi (1981). Estimación mínimo cuadrática.

	Factor I	Factor II
Item 1	0.395	-
Item 2	0.700	-
Item 3	0.581	-
Item 4	-	0.102
Item 5	-	0.668
Item 6	-	0.406

Y aplicada desde las estimaciones máximo verosímiles ofrece los resultados de la tabla 3.10.

Tabla 3.10. Índices de fiabilidad de los ítems calculados desde los resultados del análisis factorial confirmatorio por la fórmula propuesta por Bagozzi (1981). Estimación por máxima verosimilitud.

	Factor I	Factor II
Item 1	0.404	-
Item 2	0.662	-
Item 3	0.567	-
Item 4	-	0.099
Item 5	-	0.605
Item 6	-	0.387

Por último se ha aplicado la fórmula propuesta por Bagozzi a las saturaciones factoriales estimadas desde la estimación arbitraria.

Tabla 3.11. Índices de fiabilidad de los ítems calculados desde los resultados del análisis factorial confirmatorio por la fórmula propuesta por Bagozzi (1981). Estimación arbitraria.

	Factor I	Factor II
Item 1	0.394	-
Item 2	0.804	-
Item 3	0.587	-
Item 4	-	0.116
Item 5	-	0.653
Item 6	-	0.432

Los índices hallados mediante mínimos cuadrados, estimación arbitraria y máxima verosimilitud son muy similares, como es lógico dado que las estimaciones del confirmatorio se diferenciaban poco.

Los valores de fiabilidad así hallados son similares a los que se obtenían elevando simplemente las saturaciones al cuadrado. Los resultados son, no obstante, algo menores, sobre todo para los valores

superiores. Se mantiene la consistencia con los resultados de métodos clásicos de medición de la fiabilidad (dos mitades, alpha, etc). Los ítems que muestran menor fiabilidad son, del primer factor el primer ítem, y en el segundo factor el ítem seis y especialmente el ítem cuatro con bajísima fiabilidad.

## 2.4. Validez de los ítems

Especialmente importantes como criterios son las otras dos variables extraídas del constructo de clima, estructura de seguridad (ES) y política organizacional de seguridad (PO). También deben ser altas y positivas las relaciones con la respuesta de los superiores, dado que la forma de dirigir a los trabajadores que exhiban los supervisores debe ser hasta cierto punto acorde con la cultura o clima de la empresa. A medida que nos vamos alejando en el modelo hacia comportamientos de los trabajadores y actitudes y conductas de los sujetos la relación debe ser más débil. Puede hipotetizarse que la correlación será positiva, pero más débil y mediatizada por la respuesta de los superiores que son los que más directamente canalizan la 'cultura' o clima empresarial a los trabajadores. Más complejas serán las relaciones con los riesgos, dado que puede haber empresas en que por tener un alto nivel de riesgos se tenga más en cuenta la seguridad y la formación sea mayor. Pero también es posible que por haber menos prácticas de seguridad, menor clima de seguridad y formación en seguridad el nivel de riesgo percibido sea mayor, lo que nos ofrecería resultados de correlación contrarios. En general se espera que la correlación con riesgo basal sea cercana a nula, puesto que el riesgo intrínseco al trabajo no se ve afectado por otras condiciones más que las del propio proceso de trabajo, aunque esta hipótesis es un poco débil y pueden encontrarse ejemplos contra ella. Por el contrario, aunque débiles y mediatizadas por otras variables, se espera que las correlaciones sean negativas con el riesgo real, puesto que éste se define como la unión en un momento dado del riesgo basal y de las medidas puestas (o no puestas) para disminuirlo.

Tabla 3.12. Correlaciones de los ítems de la escala de formación-información con los principales criterios. ES=estructura de seguridad; PO=política organizacional de seguridad; RS=respuesta de los superiores en seguridad; RC=respuesta de los compañeros en seguridad; L=actitud hacia la seguridad; CHS=conducta hacia la seguridad; RB=riesgo basal; RR=riesgo real; T=tensión psicológica; y AC=sinestralidad. Símbolos de significatividad estadística de las correlaciones: n.s.= no significativo; #= marginalmente significativo; \*= probabilidad <0.05; y \*\*= probabilidad <0.01, extraídos de tablas de significatividad de las correlaciones (Welkowitz, Ewen y Cohen, 1981).

	ES	PO	RS	RC	L	CHS	RB	RR	T	AC
Item 1	.014 ns	.347 **	.341 **	.200 *	-.159 #	.056 ns	-.125ns	-.122ns	-.058ns	-.161 #
Item 2	.163 #	.584 **	.484 **	.267 **	-.269**	.122 ns	-.052ns	-.076ns	-.132ns	-.137ns
Item 3	.025 ns	.557 **	.452 **	.261 **	-.243**	.129 ns	-.085ns	-.089ns	-.104ns	-.160 #



### Capítulo 3. Clima organizacional hacia la seguridad.

Item 4	.324 **	.288 **	.190 *	.021 ns	-.025ns	-.136ns	.156 #	.097 ns	.030ns	.005 ns
Item 5	.522 **	.546 **	.367 **	.242 **	-.041ns	.167 #	.015 ns	-.048ns	-.128ns	-.037ns
Item 6	.499 **	.527 **	.416 **	.177 #	-.073ns	.058 ns	-.027ns	-.016ns	-.073ns	-.008ns

Si comentamos los resultados de validez por criterios podemos ver que la primera hipótesis de correlación positiva con las otras variables de clima se cumple de modo general. Las correlaciones son, en su mayoría, altas y positivas. No obstante es curioso cómo la estructura de seguridad correlaciona relativamente alto con los tres ítems últimos y menos con los tres primeros. Esto puede deberse a que poner carteles, hacer charlas y reuniones (tres últimos ítems) sean las labores realizadas por los encargados o comités de seguridad, mientras que la formación general o específica de seguridad se ofrece por otro tipo de personal. Además de los tres ítems primeros el que pregunta si se dió formación específica de seguridad es el de correlación más elevada. Por su parte, con política de seguridad las correlaciones son positivas y elevadas en su mayoría, excepto el ítem cuatro que pregunta sobre la existencia de carteles de seguridad, ítem que ya daba resultados pobres de homogeneidad y fiabilidad y que tampoco ofrece excesivos buenos resultados en la validez criterial.

Las correlaciones de los ítems de formación con la escala de respuesta de los superiores son como se hipotetizaron, positivas y elevadas. Nuevamente, excepto el ítem 4. Más bajas, aunque también positivas, son las relaciones con la respuesta de los compañeros, actitud y conducta hacia la seguridad de los trabajadores.

No hay diferencias apreciables entre las correlaciones obtenidas para los dos tipos de riesgos. Todo y que son en su mayoría negativas no presentan una fuerza apreciable, por lo que puede hipotetizarse que el clima no tiene un efecto directo en el nivel de riesgos. Tampoco se presentan correlaciones importantes con el nivel de accidentes ni con el de tensión

#### 2.5. Conclusiones para la escala de formación-información.

Las conclusiones más relevantes de la escala de formación-información ya se han ido extrayendo al hilo de cada análisis psicométrico y estadístico. No vamos a repetir ahora todas las apreciaciones ya cometadas hasta aquí. Se trata de perder de vista por un momento los detalles concretos de cada análisis particular y resumir, si es posible, aquello que a través de métodos aparece como comportamiento psicométrico general de la escala.

De modo global la escala presenta un buen comportamiento psicométrico, con adecuada validez criterial de sus ítems, elevada consistencia interna ( $\alpha=.757$ ) y estructura factorial confirmatoria simple y exhaustiva. No obstante esta conclusión excesivamente



global, hay un número de conclusiones más concretas que pueden hacerse. La escala parece medir básicamente dos factores, que hemos etiquetado como 'formación' (ítems 1, 2, y 3) e 'información' (ítems 4, 5, y 6). El comportamiento de los tres primeros ítems es excelente. Saturan altamente en el factor, definiéndolo claramente, son homogéneos y fiables, no importa el método de análisis empleado, y, finalmente, muestran una elevada validez criterial comportándose conforme a hipótesis para todos los criterios sin excepción. La situación no es la misma para el segundo factor, menos claramente delimitado. Los tres ítems saturan significativamente en él, pero el ítem cuatro se halla relativamente desligado. Los resultados de homogeneidad y fiabilidad de los ítems repiten este efecto, con dos ítems psicométricamente fiables y consistentes con el total de la escala, y el ítem cuatro muy poco fiable, sobre todo en términos de su saturación factorial. La validez criterial es conforme a hipótesis en todos ellos, pero aquí puede apreciarse cómo los tres están muy relacionados con otros aspectos del clima, en el caso del ítem cuatro incluso más que con la propia escala de formación. La fuerte correlación, insistimos, sobre todo del ítem cuatro hace difícil su adscripción empírica a una u otra variable de clima (política o estructura de seguridad), por lo que se mantiene la decisión teórica de incluirlo e la escala de formación-información.

### 3. Política de seguridad de la empresa

#### 3.1. Medida

##### ESCALA:

1. ¿Considera usted que los aspectos de seguridad son una prioridad dentro de los objetivos de su empresa?..... Si No
2. ¿Conoce Ud. algún procedimiento en su empresa para informar a la Dirección de las condiciones inseguras que pudieran existir en su área de trabajo? ..... Si No
3. En el desempeño de su trabajo ¿está primero la rapidez que la seguridad? .... Si No
4. ¿Se le ha dado a conocer las políticas o normativas con respecto a la seguridad e higiene de su empresa?..... Si No
5. ¿Su empresa tiene algún sistema para recompensar, incentivar o premiar el trabajar en forma segura? ..... Si No
6. ¿Qué interés muestra su Empresa en la seguridad e higiene en el trabajo?
  - No muestra ningún interés especial ..... 1
  - Cumple las normas mínimas de seguridad ..... 2
  - Muestra interés porque el trabajo se haga con seguridad para los trabajadores ..... 3
  - Muestra mucho interés y realiza campañas promoviendo la seguridad ..... 4
7. Recibimos indicaciones e instrucciones orales o escritas sobre seguridad ..... Si No
8. En su empresa se realizan inspecciones para el control de las condiciones de seguridad?
  - Nunca u ocasionalmente ..... 1
  - Algunas veces ..... 2
  - Muy frecuentemente ..... 3

#### 3.2. Estructura factorial confirmatoria

Se han planteado dos estructuras factoriales hipotéticas para reproducir la matriz de varianzas-covarianzas observada entre las variables.

1) Por un lado una *estructura unifactorial*. Siempre se presentará una estructura unifactorial que se utiliza en ocasiones como modelo competitivo a partir del cual observar la ganancia que se produce en la explicación de los datos con estructuras de más factores, y en ocasiones por su mismo interés teórico. En este caso se plantea como posibilidad teórica la estructura de un factor, más que como elemento de comparación en la prueba del ajuste.

2) Se plantea, además, una *primera estructura bifactorial*. El primer factor incluiría los ítems 1, 4, 5, 6, 7, 8 y 8. Todos estos ítems preguntan acerca de acciones de la empresa en materia de seguridad, sobre interés o prioridad de la seguridad para la empresa. Todos ellos tienen en común que son acciones o intenciones realizadas directamente por la empresa como organización. El segundo factor agrupa los ítems 2 y 3. El ítem dos cuestiona acerca de la existencia de canales de comunicación adecuados en la empresa para informar de aspectos de seguridad. El tercer ítem pregunta si en la empresa está primero la rapidez que la seguridad. Los dos ítems presentan en común que en este caso se mide alguna acción que realiza directamente el trabajador. La empresa facilita, pone los medios, coacciona, etc pero no trabaja directamente. Son los propios trabajadores los que informan y los que trabajan bajo presión de tiempo sin poder atender a los aspectos de seguridad. Esta estructura introduce un nuevo factor diferente del primero por una leve matización teórica. Nos servirá en este caso más como modelo competitivo de comparación en el ajuste.

3) Por último se plantea una segunda estructura bifactorial. En este caso el primer factor incluiría los ítems 1, 3, 5, 6, y 8. El segundo factor incluye teóricamente los ítems 2, 4, y 7. Los contenidos de los ítems del segundo factor definen un componente de acciones informativas de seguridad, ya que todos preguntan acerca del flujo de información que se produce sobre aspectos de seguridad, bien en sentido descendente, donde la empresa informa sobre seguridad o en sentido ascendente, donde la empresa especifica unos canales de comunicación vertical claros para que el trabajador pueda informar a los superiores sobre la seguridad. El primer factor recogería el resto de contenidos, bastante dispares, y que podrían resumirse en ítems globales sobre prioridad percibida de la seguridad para la empresa y de acciones no informativas que emprende la dirección -sistemas de recompensas por conducta segura o inspecciones de seguridad-.

En este caso las tres estructuras comprobadas mediante análisis factorial confirmatorio estaban motivadas por agrupación teórica. El Lagrangian Multiplier test y el Wald test no proponían ningún cambio estadístico que resultara en un decremento en la ji-cuadrado y, como consecuencia, en un mejor ajuste.

Los resultados de ajuste global de los modelos pueden verse en la tabla 3.13. Aparecen las tres estructuras hipotetizadas, por orden, y para cada una de ellas se listan los índices calculados para cada uno de los dos métodos de estimación empleados, mínimos cuadrados generalizados y máxima verosimilitud.

Tabla 3.13. Índices de ajuste global de las tres estructuras hipotetizadas para la escala de política organizacional de seguridad. Los índices son: *chi*=chi-cuadrado; *g*=grados de libertad del modelo; *p*=probabilidad; *F*=función de ajuste a minimizar; *LHR*=razón de verosimilitud escalada; *CAK*=Akaike information criterion; *CSK*=Schwarz information criterion; *NFI*=normed fit index; *NNFI*=non-normed fit index; *CFI*=comparative fit index. Métodos de estimación: *GLS*=mínimos cuadrados generalizados; *ML*=máxima verosimilitud.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
UN	GLS	23.8	20	=.248	.24	.885	1.19	.566	.986	.977	.995	.996
FACTOR	ML	27.3	20	=.126	.27	.869	1.36	.581	1.02	.836	.926	.947
BIFACTO-	GLS	21.7	19	=.298	.22	.895	1.14	.564	1.01	.979	.996	.997
RIAL (1)	ML	26.6	19	=.112	.27	.872	1.40	.615	1.06	.840	.918	.945
BIFACTO-	GLS	23.4	19	=.218	.24	.887	1.23	.582	1.02	.977	.993	.996
RIAL (2)	ML	26.8	19	=.108	.27	.871	1.41	.617	1.06	.838	.916	.943

Las tres estructuras planteadas muestran un buen ajuste de ji-cuadrado tanto mediante GLS como mediante ML. Los valores del resto de índices, menos potentes, concuerdan en mostrar que las tres estructuras son buenas representaciones de los datos, salvo el índice NFI que no alcanza el valor de 0.9. Los resultados de ajuste global son consistentemente mejores para GLS que para ML.

Si nos centramos en los resultados para mínimos cuadrados generalizados y comparamos las tres estructuras observaremos la igualdad a la hora de explicar los datos originales. Así, la prueba ji-cuadrado resulta ser para la estructura unifactorial significativa con probabilidad 0.248, la probabilidad sube al 0.298 para la primera estructura bifactorial, y baja algo para la segunda estructura de dos factores (0.218). Ahora hay que determinar si esos cambios son significativos. Si escogemos ahora un índice que represente parsimonia, como la prueba ji-cuadrado partida por los grados de libertad del modelo volvemos a apreciar la igualdad entre los modelos hipotetizados, todos muy por debajo del mínimo más restrictivo exigido. Las diferencias siguen siendo muy ligeramente positivas en favor de la primera estructura bifactorial. Si ahora escogemos un índice representativo de los que comparan el ajuste del modelo frente a un modelo nulo, el CFI, las diferencias se acortan todavía más, acercándose los tres modelos al valor teóricamente superior de 1, indicativo de un ajuste perfecto. Insistimos, diferencias apenas apreciables. Exactamente el mismo esquema aparece para la estimación máximo versosímil en cuanto al ajuste global, igualdad máxima entre los tres modelos hipotetizados en todos los índices

Se ha realizado, por tanto una prueba de las diferencias entre las ji-cuadrado de los tres modelos en los dos métodos de estimación.

Si se compara el modelo unifactorial frente al bifactorial (1) se obtiene la siguiente prueba para la estimación mediante GLS:

$\Delta\chi^2=2.1$ , con  $\Delta gl=1$  que resulta en una probabilidad  $p>0.1$

indicando que no hay diferencias estadísticas a la hora de representar los datos entre las dos estructuras teóricas. El modelo bifactorial no explica mejor los datos que el unifactorial.

La misma prueba de las diferencias entre ji-cuadrado calculada para los resultados de ML:

$\Delta\chi^2=0.7$ , con  $\Delta gl=1$ ,  $p>0.1$ .

indicando nuevamente que no existen diferencias estadísticamente significativas entre ambas representaciones de los datos.

Si calculamos ahora las diferencias de ji-cuadrado entre la estructura unifactorial y la estructura bifactorial (2) obtenemos, para GLS, los siguientes resultados:

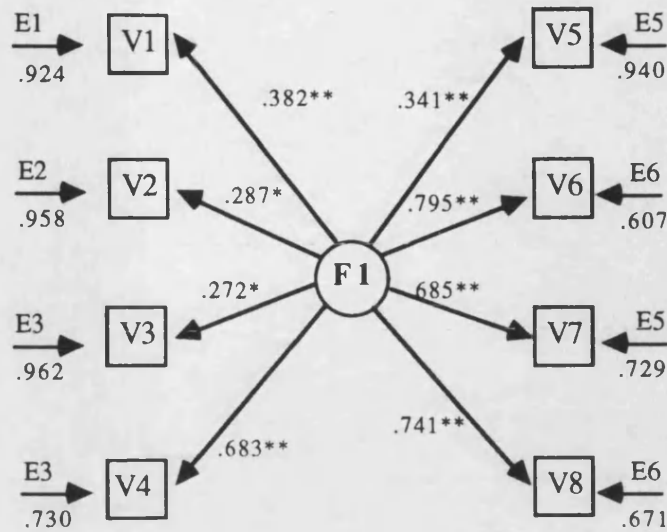
$\Delta\chi^2=0.4$ , con  $\Delta gl=1$  que no es estadísticamente significativo ( $p>0.1$ ),

y para máxima verosimilitud

$\Delta\chi^2=0.5$ , con  $\Delta gl=1$  que tampoco resulta significativo ( $p>0.1$ ).

A la vista de las pruebas de diferencias entre las ji-cuadrado, y teniendo en cuenta el principio de parsimonia científica, el modelo unifactorial se escoge como mejor representación de los datos, dado que otros modelos más complejos no ofrecen ganancias significativas en cuanto a la explicación de los datos.

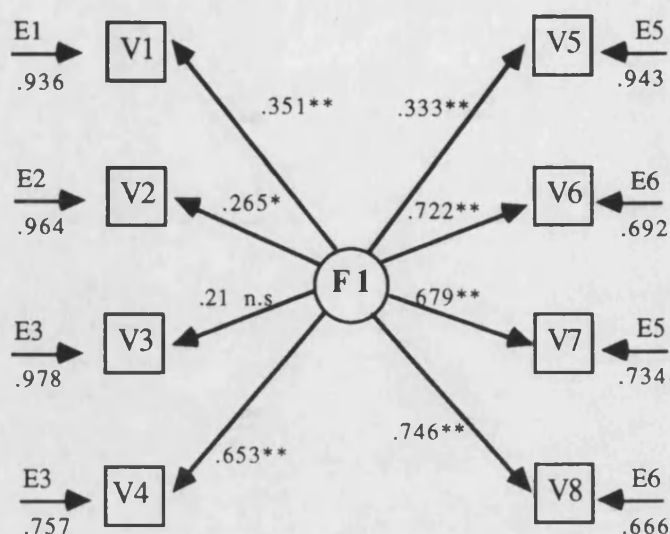
La gráfica 3.4. presenta los resultados, para el modelo de un solo factor, de la estimación mínimo cuadrática, incluyendo los datos de ajuste analítico.



Gráfica 3.4. Estructura factorial estimada mediante mínimos cuadrados generalizados para la escala de política organizacional de seguridad. Ajuste analítico: \*\*= $p < .01$ ; \*= $p < .05$ .

Las saturaciones factoriales son todas estadísticamente significativas, estadísticamente diferentes de cero, bien al nivel de confianza del 99% o del 95%. No obstante, salvo los ítems 4, 6, 7 y 8, los valores que alcanzan son bastante bajos. El factor de política de seguridad que se hipotetiza como subyacente a la escala no consigue explicar satisfactoriamente el resto de ítems.

Para la estimación máximo verosímil también se muestran, en la gráfica 3.5, las saturaciones factoriales y su significación estadística.



Gráfica 3.5. Estructura factorial estimada mediante máxima verosimilitud para la escala de política organizacional de seguridad. Ajuste analítico: \*\*= $p < .01$ ; \*= $p < .05$ ; n.s.= no significativo.

Las saturaciones factoriales estimadas mediante máxima verosimilitud son muy similares a las encontradas mediante mínimos cuadrados generalizados. En general presentan valores ligeramente inferiores. En términos de diferencias estadísticas, mediante ML se pierde la significatividad del ítem dos, que pasa a ser no significativo. Como en GLS, sólo los ítems 4, 6, 7, y 8 presentan saturaciones elevadas con el factor.

El ajuste analítico no parece ser demasiado adecuado, aunque se alcanzan valores de significatividad estadística, cabe pensar que no hay un solo factor subyaciendo a algunos ítems de la escala, especialmente los ítems 2 y 3. Las estructuras bifactoriales propuestas no ofrecieron cambios significativos frente a la unifactorial, pero sí podrían aportar alguna mejora en el ajuste analítico.

En la siguiente tabla se ofrecen los ajustes analíticos que se obtienen mediante las correspondientes pruebas z. Estas pruebas se han efectuado para cada saturación factorial del ítem en los tres modelos probados por la estimación mínimo cuadrática. En la estructura unifactorial, los resultados son los que ya aparecen en las gráficas 3.4. y 3.5, que ya han sido comentados, con saturaciones bajas para los ítems número 1, 5, y, especialmente 2 y 3. Se trataba de observar los cambios que otras estructuras podían comportar a nivel de cargas factoriales, y específicamente si se podían obtener mejoras sustanciales en la explicación de los cuatro ítems con cargas bajas.



Tabla 3.14. Saturaciones factoriales de cada ítem en las tres estructuras factoriales puestas a prueba. Estimación por mínimos cuadrados generalizados. \*\*= $p < .01$ ; \*= $p < .05$ , y n.s.=no significativo.

ITEM	Estructura unifactorial	Estructura bifactorial 1	Estructura bifactorial 2
1	.382**	.389**	.383**
2	.287*	.488*	.299*
3	.272*	.452*	.298*
4	.683**	.688**	.694**
5	.341**	.339**	.339**
6	.795**	.793**	.806**
7	.685**	.684**	.702**
8	.741**	.742**	.746**

Si se observa de modo global la tabla 3.14 puede apreciarse cómo a pesar de cambiarse de estructura factorial los efectos en las saturaciones factoriales son mínimos, de forma que no se produce ni un solo cambio en la significatividad estadística. El único cambio que parece apreciable es el producido en los ítems dos para la estructura bifactorial 1, comparando con la estructura unifactorial y la bifactorial 2. Esta ganancia que se produce en la saturación factorial no consigue, sin embargo, que estos ítems pasen de una significatividad de 0.05 a 0.01. Por otro lado, este aumento en la carga factorial es absolutamente previsible, dado que en la estructura bifactorial 1, son los ítems 2 y 3 los que conforman el factor 2, por lo que es muy fácil que tengan mayor homogeneidad con éste que al estar integrados en un factor de ocho ítems con el que su homogeneidad no era elevada.

Tabla 3.15. Saturaciones factoriales de cada ítem en las tres estructuras factoriales comprobadas. Estimación por máxima verosimilitud. \*\*= $p < .01$ ; \*= $p < .05$ , y n.s.=no significativo.

ITEM	Estructura unifactorial	Estructura bifactorial 1	Estructura bifactorial 2
1	.351**	.352**	.357**
2	.265*	.411*	.274*
3	.21n.s.	.32n.s.	.224*
4	.653**	.655**	.666**
5	.333**	.333**	.334**
6	.722**	.720**	.735**
7	.679**	.679**	.701**
8	.746**	.747**	.747**

También se ha querido presentar una tabla que nos ofreciera los valores de ajuste analítico de los tres modelos hipotetizados, esta vez para la estimación máximo verosímil. Estos datos son, como es habitual, muy similares a los hallados para la estimación por mínimos cuadrados generalizados, aunque algo más bajos. Por lo demás, en el ajuste analítico no se aprecian cambios significativos a través de los modelos. Nuevamente, ni en aquellos ítems con más baja carga



factorial en la estructura unifactorial (2 y 3) se consiguen cambios relevantes.

Así, los resultados recogidos para la validez factorial de la escala tanto a un nivel de ajuste global como analítico permiten afirmar para esta muestra que la escala es razonablemente unifactorial. Otras estructuras más complejas no ofrecen ninguna ganancia y, por razones de parsimonia, se escoge el modelo de un factor para representar a los datos.

Los dos métodos de análisis presentados hasta ahora se basan en el supuesto de multinormalidad de las variables. La siguiente cuestión es si cumplen el supuesto razonablemente y, si no es así, realizar la estimación, arbitraria o elíptica, que se ajuste más razonablemente a los datos.

La normalidad univariada es una condición necesaria para la normalidad multivariable. Se ha realizado una prueba de bondad de ajuste para cada variable, con objeto de comprobar si se alejaban significativamente de la normalidad. Los resultados de la prueba se ofrecen en la tabla 3.16. No puede mantenerse para ninguna de las variables la distribución normal, lo que no es de extrañar dada la escala de respuesta de los ítems, de dos, tres o cuatro anclajes. Difícilmente escalas de medida de estas características pueden soportar una prueba de normalidad.

*Tabla 3.16. Prueba de Kolmogorov-Smirnov de una muestra para comprobar el ajuste de los ítems de la escala de formación-información a la distribución normal univariada.*

ITEM	Casos	Diferencia máxima	Probabilidad
1	112	0.565	0.000
2	122	0.500	0.000
3	120	0.500	0.000
4	123	0.500	0.000
5	115	0.500	0.000
6	123	0.695	0.000
7	120	0.500	0.000
8	117	0.500	0.000

A continuación se presentan los datos de asimetría y apuntamiento (curtosis) de cada variable. En asimetría, sólo el ítem cuatro está cercano a cero. Todas las variables muestran algún grado de curtosis.

Tabla 3.17. Datos de asimetría y curtosis para los ocho ítems de la escala de política organizacional.

	Asimetría	Curtosis
Item 1	-1.0788	-0.8361
Item 2	-0.2236	-1.950
Item 3	-0.4777	-1.7718
Item 4	0.1011	-1.9898
Item 5	1.7409	1.0308
Item 6	-0.3806	-0.5971
Item 7	0.5669	-1.6786
Item 8	0.5813	-0.6808

Hasta ahora hemos presentado los valores descriptivos de asimetría y apuntamiento univariados y una prueba de bondad de ajuste, también univariada. No se puede mantener, a la vista de los resultados, la normalidad de ningún ítem de la escala de política, lo que resulta necesario para mantener la multinormalidad de la escala. Analizar los resultados de curtosis multivariada no sería necesario en este caso, puesto que no se cumple una condición previa. En cualquier caso, el coeficiente de Mardia para la curtosis multivariada ofrece un valor claramente superior al esperado mediante la normalidad multivariada de los ítems.

COEFICIENTE DE MARDIA MULTIVARIADO= -4.6266 Coef. esperado bajo normal multivariada= -1.8197
---

En función de los resultados se pensó que la estimación elíptica no era la más adecuada ya que las variables presentaban asimetría, y se optó por realizar estimaciones arbitrarias. En la tabla 3.18 se presentan los datos de ajuste global de los tres métodos de estimación (máxima verosimilitud, mínimos cuadrados generalizados y estimación arbitraria) para la estructura unifactorial retenida. Los datos coinciden en los métodos de estimación máximo verosímil y mínimo cuadrático con los ya presentados. Se han colocado aquí por comodidad del lector, que así no tiene que volver atrás para comparar ajustes.

Tabla 3.18. Índices de ajuste global para la escala de política organizacional de seguridad. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; CAK=Akaike information criterion; CSK=Schwarz information criterion; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: GLS=mínimos cuadrados generalizados; AGLS= estimación arbitraria; ML=máxima verosimilitud.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
MODELO	GLS	23.8	20	=.248	.24	.885	1.19	.566	.986	.977	.995	.996
DE UN	AGLS	34.2	20	=.024	.34	.839	1.71	.672	1.09	.811	.870	.907
FACTOR	ML	27.3	20	=.126	.27	.869	1.36	.581	1.02	.836	.926	.947

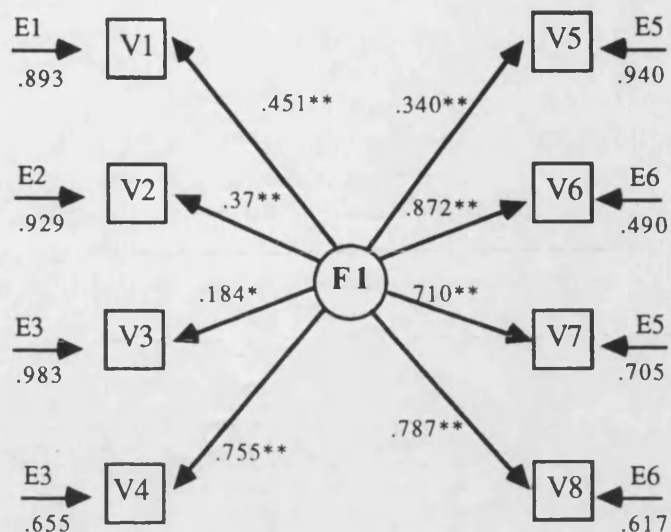
Como puede verse, se ha producido un cambio sustancial en el ajuste global medido por ji-cuadrado, no así para el resto de índices que muestran cambios pero no tan apreciables. Se ha pasado de una estructura no significativa a una significativa, indicando el ajuste por chi-cuadrado que un factor no da cuenta adecuadamente de los datos observados. Este dato de falta de ajuste no es consistente a través de todos los índices. Así, chi-cuadrado partido por los grados de libertad, como índice de parsimonia, muestra un ajuste muy adecuado y el índice CFI se encuentra por encima de .9.

Además, cabe recordar que cuando se utiliza la estimación arbitraria el programa EQS ofrece otros dos índices de ajuste. Estos dos índices son el GFI y el AGFI. Sus valores son alentadores, 0.999 para el GFI y 0.998 para el AGFI, indicando un altísimo grado de ajuste del modelo a los datos.

La conclusión global es que aún cuando se ha perdido capacidad explicativa la estructura unifactorial puede seguir manteniéndose también para la estimación arbitraria.

Falta saber qué cambios se han producido en la estimación de las cargas factoriales de los items con el factor latente, y reevaluar en base a los nuevos errores estándar el ajuste analítico de las saturaciones.

La gráfica 3.6 ofrece las cargas factoriales para el modelo unifactorial para la estimación arbitraria, además del ajuste analítico.



Gráfica 3.6. Estructura unifactorial probada por el método de estimación arbitraria para la escala de política organizacional de seguridad. Ajuste analítico: \*\*= $p < .01$ ; \*= $p < .05$ ; n.s.= no significativo.

Se han producido algunos cambios, si bien los resultados son consistentes con los hallados mediante los dos métodos de estimación basados en la distribución multinormal. Todos los ítems han ganado algo en el valor de su carga factorial, a excepción del ítem 3 que baja ligeramente su carga factorial. No obstante esta saturación es estadísticamente significativa a pesar de ser más baja que la correspondiente en GLS y ML. Si volvemos al ajuste analítico por máxima verosimilitud observaremos que allí con un índice mayor la saturación no resultaba estadísticamente significativa. Este resultado indica que el error estándar calculado para este ítem no era demasiado fiable (Bentler, 1988; Tanaka, 1984). Los ítems números 4, 5, 6, y 7 han permanecido prácticamente iguales, con ligeras modificaciones. Los ítems 1 y 2 son los únicos que han contado con una modificación sustancial en torno al 0.1 de ganancia frente a la estimación mínimo cuadrática y máximo verosímil.

En general, parece que se han producido algunos cambios en el ajuste global y analítico de la escala, pero que no alteran las conclusiones que ya se derivaban de los análisis realizados con los dos tipos de estimación basados en el supuesto de distribución normal multivariada.

### 3.3. Fiabilidad y homogeneidad de la escala y los ítems

Se ha estudiado la fiabilidad de la escala de política organizacional de seguridad a través de los habituales coeficientes de consistencia interna. El coeficiente alfa para la escala total se sitúa en torno al 0.7, lo que no puede verse como un resultado demasiado satisfactorio. El resto de coeficientes se sitúan en torno a estos valores. La consistencia interna de la escala de política se presenta en la tabla 3.19.

Estos resultados de consistencia ya eran previsibles desde el análisis factorial confirmatorio. Un solo factor difícilmente explicaba allí los ítems 1, 2, 3, y 5. El resto de ítems sí eran homogéneos con el factor hipotetizado indicando que podían estar midiendo lo mismo. Si observamos los coeficientes alfas para la parte par e impar se corrobora esta impresión de que algunos ítems muy poco homogéneos con el resto están bajando la consistencia interna de la escala. Efectivamente el coeficiente alfa para los ítems impares ofrece un resultado muy bajo de consistencia interna (0.367), precisamente porque incluye tres ítems que en el análisis factorial se mostraban poco relacionados con el factor, los ítems 1, 3 y 5. No obstante estos resultados son relativamente comunes para un alfa calculado sólo en base a cuatro ítems. Sólo el ítem 7, también impar, presentaba alta saturación en el factor. El coeficiente alfa para los ítems pares, por contra, incluye un sólo ítem poco relacionado con el factor, y su valor es más alto (0.609), contando lógicamente con que tiene un menor número de ítems que el total.

Tabla 3.19. Coeficientes de consistencia interna de la escala de política organizacional de seguridad.

<i>Consistencia interna</i>	
Correlación dos mitades	0.600
Coeficiente Spearman-Brown	0.750
Coeficiente Guttman-Rulon	0.718
Coeficiente alpha total	0.702
Alpha ítems impares	0.367
Alpha ítems pares	0.609

Los estadísticos descriptivos muestran este mismo patrón de relaciones, y pueden consultarse en la tabla 3.20. A través de los diversos coeficientes de homogeneidad, y fiabilidad para los ítems se muestra cómo los ítems 1, 2, 3, y 5 son muy poco consistentes con el total de la escala, provocando los desequilibrios en los análisis realizados hasta ahora.

### Capítulo 3. Clima organizacional hacia la seguridad.

Tabla 3.20. Media, desviación típica, coeficiente de homogeneidad (correlación ítem-total), fiabilidad, homogeneidad excluyendo el ítem y coeficiente alpha de la escala si se excluye el ítem, para cada ítem de la escala de política organizacional de seguridad.

ITEM	Med.	D. T.	Homogen.	Fiabilidad	Homogen. sin el ítem	Alpha sin el ítem
1	0.737	0.44	0.441	0.194	0.270	0.697
2	0.556	0.55	0.395	0.196	0.194	0.714
3	0.616	0.61	0.416	0.202	0.221	0.708
4	0.475	0.47	0.655	0.327	0.503	0.649
5	0.172	0.17	0.423	0.159	0.276	0.695
6	1.465	1.46	0.818	0.671	0.629	0.610
7	0.364	0.36	0.704	0.339	0.573	0.634
8	0.313	0.31	0.710	0.234	0.627	0.646

Los ítems 4, 6, 7, y 8, son los ítems que sí son homogéneos con la escala, y también presentan mayores fiabilidades. Habría una pérdida considerable en el valor alfa de la escala (última columna) si se eliminaran estos ítems que son los que más contribuyen a mantener los niveles de consistencia interna. Por contra, si se eliminaran los ítems poco homogéneos no habría pérdida, o incluso se produciría una ligera ganancia.

Las tres siguientes tablas presentan los resultados de consistencia interna de los ítems de la escala calculados a través del análisis factorial confirmatorio. Se presenta una tabla para cada tipo de estimación: mínimos cuadrados generalizados, máxima verosimilitud, y métodos de estimación arbitrarios. Dentro de cada tabla se presentan dos fórmulas. La primera columna ( $\lambda^2$ ) presenta la saturación factorial al cuadrado, que es una medida del porcentaje de varianza de la variable que se puede explicar desde el factor. La segunda fórmula ( $\rho_i$ ) cumple básicamente la misma función, pero toma en consideración la varianza de error de cada variable para calcular el coeficiente.

Tabla 3.21. Índices de fiabilidad calculados a partir de las saturaciones factoriales estimadas para la estructura unifactorial mínimo cuadrática. La primera columna de datos corresponde a las saturaciones elevadas al cuadrado propuestas por Long (1983) como medida de la fiabilidad, la segunda columna de datos se corresponde con la fórmula para la fiabilidad presentada por Bagozzi (1981).

	$\lambda^2$	$\rho_i$
Item 1	0.145	0.136
Item 2	0.082	0.079
Item 3	0.073	0.070
Item 4	0.466	0.389
Item 5	0.116	0.109
Item 6	0.632	0.477
Item 7	0.469	0.391
Item 8	0.549	0.450

Tabla 3.22. Índices de fiabilidad calculados a partir de las saturaciones factoriales estimadas para la estructura unifactorial máximo verosímil. La primera columna de datos corresponde a las saturaciones elevadas al cuadrado propuestas por Long (1983) como medida de la fiabilidad, la segunda columna de datos se corresponde con la fórmula para la fiabilidad presentada por Bagozzi (1981).

	$\lambda^2$	$\rho_i$
Item 1	0.123	0.116
Item 2	0.070	0.067
Item 3	0.043	0.042
Item 4	0.426	0.360
Item 5	0.110	0.105
Item 6	0.521	0.429
Item 7	0.469	0.385
Item 8	0.556	0.455

Tabla 3.23. Índices de fiabilidad calculados a partir de las saturaciones factoriales para el modelo unifactorial estimado mediante estimación arbitraria. La primera columna de datos corresponde a las saturaciones elevadas al cuadrado propuestas por Long (1983) como medida de la fiabilidad, la segunda columna de datos se corresponde con la fórmula para la fiabilidad presentada por Bagozzi (1981).

	$\lambda^2$	$\rho_i$
Item 1	0.203	0.185
Item 2	0.137	0.129
Item 3	0.033	0.033
Item 4	0.570	0.465
Item 5	0.115	0.109
Item 6	0.760	0.608
Item 7	0.504	0.416
Item 8	0.619	0.500

Los resultados de ambas fórmulas son coincidentes con los ya expuestos y presentan el mismo panorama que ya se ha recogido mediante los coeficientes de homogeneidad y fiabilidad de los items de la escala. Claramente los items 4, 6, 7, y 8 son suficientemente consistentes, y se pueden explicar razonablemente bien desde el factor latente. El resto, los items 1, 2, 3, y 5, no son consistentes y son muy pobremente explicados desde el factor latente.

Los resultados para los items apuntan en la dirección de que sólo cuatro items de la escala están midiendo claramente y en esta muestra el factor política organizacional de seguridad.



### **3.4. Validez criterial de los items**

Las hipótesis de relaciones criterioales son similares a las que se plantearon en el caso de la escala de formación-información. En primer lugar, y al respecto de la relación con los otros componentes de clima -formación-información y estructura de seguridad-, se espera coeficientes de correlación positivos y elevados. Se espera una relación mayor con formación-información dado que es una de las principales acciones de seguridad que puede llevar a cabo la empresa, y muestra por tanto la política de seguridad de la empresa. La relación con estructura puede ser algo menor dado que la existencia de encargados y comités de seguridad está en parte legislada, de forma que puede no concordar con la política de seguridad que exista en la empresa, sino estar determinada desde el exterior. En cuanto al resto de criterios, se espera una relación positiva con respuesta de los superiores. De hecho la mayoría de las acciones emprendidas por los supervisores y mandos intermedios es previsible que estén determinadas por la política y cultura imperante en la empresa, siendo éstos los principales trasmisores de ésta a los trabajadores. Se vislumbra, pues, como uno de los criterios más importantes. También se esperan relaciones positivas pero más débiles sobre los trabajadores (compañeros y actitud y conducta del trabajador), dado que son relaciones en principio mediatizadas por el efecto más directo de la respuesta de los supervisores y mandos directos del trabajador. No debe existir, por hipótesis, más que una débil relación con el nivel de riesgo basal y, si hubiera, una débil relación negativa con riesgo real que puede estar muy mediatizada por otras variables. No se espera relación con los accidentes y con la tensión.

De especial interés, además de las tendencias generales que presenten los items y que deben ser concordantes con las hipótesis, resulta comprobar las relaciones de los items 1, 2, 3, y 5 con los criterios. Efectivamente, estos cuatro items de la escala de política de seguridad no son homogéneos con el total de la escala, de forma que no puede mantenerse desde un punto de vista psicométrico que midan bien, para esta muestra, la política de seguridad. Resulta, por tanto, relevante ver con qué otras variables se encuentran más relacionados.

Los datos de correlación entre los ocho items de la escala de política y los criterios se presentan en la siguiente tabla, junto con los datos de significatividad estadística basados en tablas (Welkowitz, Ewen y Cohen, 1981).



Tabla 3.24. Correlaciones de los ítems de la escala de política organizacional de seguridad con los principales criterios. ES=estructura de seguridad; F=formación-información; RS=respuesta de los superiores en seguridad; RC=respuesta de los compañeros en seguridad; L=actitud hacia la seguridad; CHS=conducta hacia la seguridad; RB=riesgo basal; RR=riesgo real; T=tensión psicológica; y AC= sinestralidad. Símbolos de significatividad estadística de las correlaciones: n.s.= no significativo; #= marginalmente significativo; \*= probabilidad <0.05; y \*\*= probabilidad <0.01.

	ES	F	RS	RC	L	CHS	RB	RR	T	AC
1	.061ns	.286 **	.305 **	.280 **	-.101ns	.076 ns	.069 ns	.209 *	.128 ns	.027 ns
2	.166 #	.258 **	.166 #	.251 **	-.031ns	.002 ns	.170 #	.028 ns	-.112ns	-.089ns
3	-.013ns	.106 ns	.333 **	.331 **	-.241**	.455 **	-.192 *	-.411**	-.390**	-0.178#
4	.442 **	.428 **	.401 **	.273 **	-.163 #	-.010 ns	-.234 *	-.080ns	.046 ns	-.047ns
5	.233 **	.251 **	.398 **	.182 *	-.168 #	.156 #	-.005 ns	-.126ns	-.116ns	0.039ns
6	.254 **	.518 **	.753 **	.515 **	-.407**	.400 **	-.184 *	-.211 *	-.338**	-0.168#
7	.357 **	.449 **	.518 **	.284 **	-.105ns	.193 *	-.084ns	.011 ns	-.182 *	0.030ns
8	.370 **	.495 **	.538 **	.403 **	-.168 #	.397 **	-.121ns	-.143 ns	-.171 #	-0.165#

En cuanto a la validez criterial de los ítems con los otros dos constructos de clima, los resultados avalan, en general, las hipótesis de relación positiva y elevada. Estos resultados son ciertos, sobre todo en lo que a los ítems 4, 6, 7, y 8 se refiere, justamente los que mejor representaban el constructo de política de seguridad.

Las correlaciones de los ítems con el criterio respuesta de los superiores son conforme a hipótesis, especialmente los cuatro ítems ya mencionados. Una escala formada por los ítems 4, 5, 6, y 8 resultaría, previsiblemente, un adecuado predictor de la respuesta hacia los superiores, tal y como se hipotetizaba.

Las correlaciones empiezan a bajar, conforme se planteaba teóricamente, para las variables que representan la actitud y conducta de los trabajadores y sus compañeros. No obstante existen relaciones significativas, aunque menos elevadas que para los anteriores criterios, más cercanos por hipótesis.

Es prácticamente nula la relación de casi todos los ítems de la escala con el nivel de riesgo basal, lo que ya era previsible teóricamente dado que el nivel de riesgos basales se plantea como una variable del propio proceso de trabajo, independiente de las medidas que se tomen para rebajarlo. El mismo esquema se reproduce para la tensión laboral.

No hay relaciones fuertes entre los ítems de política y la escala de riesgo real, como cabía esperar por hipótesis. A resaltar, sin embargo, que el ítem 3 sí correlaciona de forma elevada con el riesgo real. Si recordamos el contenido del ítem planteaba si la rapidez era una prioridad frente a la seguridad. Creemos que este ítem puede estar más relacionado con el nivel de riesgos que con la política de seguridad. Es cierto que la empresa puede primar la rapidez, y esto puede ser una medida del interés que tienen en la seguridad. Muchos

procesos de trabajo implican, necesariamente, cierto nivel de rapidez que puede estar relativamente al margen de la actividad de seguridad de la empresa, o al menos el trabajador no ver esta presión de tiempo como un aspecto de seguridad planteado directamente por la empresa. En resumen, su contenido, como mínimo no es claro, pudiendo estar implicado en diversas variables como las de riesgo (basal o real), respuesta de los superiores o de compañeros y conducta del propio trabajador. Efectivamente, este ítem presenta correlaciones con todos estos criterios y su intensidad es muy similar.

Los cuatro ítems que mejor definen el constructo de política son, como hemos visto a través de los diversos análisis, el 4, 6, 7, y 8. Estos ítems son de contenido general, definiendo de forma global la 'cultura' imperante en la empresa. Las respuestas afirmativas a estos ítems se corresponden con una percepción por parte de los trabajadores de que la empresa muestra interés en seguridad y hace cosas para fomentarla.

Cuestión aparte son los ítems con mal funcionamiento en la escala de política. Queda observar qué otro constructo de los tenidos en cuenta presenta una mayor homogeneidad con estos ítems, o si son hasta cierto punto irrelevantes dentro del marco teórico planteado en este trabajo. La tabla de correlaciones con los criterios muestra para el ítem tres el comportamiento que ya se ha descrito, bastante relacionado con constructos muy diversos, indicando que su contenido no es fácilmente adscribible a ninguno de ellos y puede estar midiendo diversas cosas. El ítem se relaciona con los criterios conforme a hipótesis, aunque con menos intensidad en la relación que los ítems que mejor representan el factor de política. Por su parte los ítems 2 y 5, son demasiado específicos. El primero de ellos (2) plantea la posibilidad de informar a la empresa sobre deficiencias en la seguridad y el segundo (5) sobre la existencia de recompensas e incentivos por trabajo seguro. Las empresas de nuestra muestra no se caracterizan por acciones tan decididas en materia de seguridad, haciendo descansar las acciones de seguridad básicamente en normativas y ocasionalmente en formación o información. De forma que estos ítems no consiguen discriminar entre empresas por su nivel de interés en seguridad, ni tienen efecto, por no llevarse a cabo, en la respuesta de los superiores o en la conducta de los trabajadores.

### **3.5. Conclusiones para la escala de política.**

Al igual que se ha hecho para la escala de formación, se trata aquí no de recordar todos los datos ya comentados junto a los análisis concretos sino de extraer las conclusiones globales del comportamiento de la escala. La escala en su conjunto presenta un

adecuado comportamiento medido por los parámetros psicométricos. Presenta una adecuada consistencia interna ( $\alpha=0.702$ ), máxime dado el número de ítems. La estructura factorial unifactorial es parsimoniosa y sostenible desde metodología confirmatoria y los ítems se comportan en la validez criterial conforme a las hipótesis planteadas.

Si bajamos de este nivel tan global a un estudio más particular de los ítems ya se ha visto en los análisis que existen fuertes problemas. Cuatro ítems son consistentes, fiables y factorial y criterialmente muy adecuados para medir el constructo política de seguridad, los otros cuatro no. Los ítems que, en esta muestra, sí muestran un muy adecuado comportamiento psicométrico son los de contenido más general que preguntan acerca del interés de la empresa en seguridad, existencia de inspecciones de seguridad y comunicación explícita de la normativa y políticas de seguridad de la empresa. Los cuatro ítems restantes, inadecuados desde un punto de vista psicométrico, rastrean contenidos más específicos que han resultado poco homogéneos y discriminativos. Si ahora recordamos sus contenidos brevemente:

Ítem 1. ¿Considera usted que los aspectos de seguridad son una prioridad dentro de los objetivos de su empresa?

Ítem 2. ¿Conoce Ud. algún procedimiento en su empresa para informar a la Dirección de las condiciones inseguras que pudieran existir en su área de trabajo?

Ítem 3. En el desempeño de su trabajo ¿está primero la rapidez que la seguridad?

Ítem 5. ¿Su empresa tiene algún sistema para recompensar, incentivar o premiar el trabajar en forma segura?

Los ítems dos, tres y cinco rastrean contenidos específicos. Así, por ejemplo, la existencia de incentivos o recompensas por trabajo seguro es inusual en nuestro contexto laboral y el ítem por tanto presenta poca variabilidad. La preponderancia de la rapidez frente a la seguridad -ítem 3- es un contenido muchas veces tenido en cuenta en la literatura, pero quizá no se ha preguntado aquí en la forma más idónea. Muchos trabajos implican intrínsecamente un proceso rápido, como por ejemplo algunas cadenas de montaje. Podría reformularse la pregunta para intentar que fuera más discriminativa de la siguiente forma: ¿se le hace trabajar más aprisa del ritmo habitual descuidando por ello los aspectos de seguridad?, e intentando otro tipo de escala de respuesta con más anclajes, del tipo de nunca a siempre. El ítem uno, sin embargo, no es de contenido específico. De hecho es muy similar al ítem seis que en lugar de la palabra prioridad emplea la palabra interés. Varias hipótesis caben para explicar este mal comportamiento, desde la de falta de comprensión, a la estrecha escala de respuesta. Efectivamente, mientras la escala de respuesta del ítem seis es de cuatro puntos con contenidos fácilmente discriminables para los

sujetos, donde se puede situar las empresas en una gradación relativamente adecuada, el ítem uno fuerza excesivamente a una respuesta sí/no que puede no ajustarse a la realidad de las empresas. No obstante, dado el buen comportamiento del ítem seis, nos parece que este ítem uno puede ser innecesario.

En resumen, y para esta muestra, se propone una reducción drástica de la escala de ocho a cuatro ítems, los de mejor comportamiento psicométrico. Los contenidos de los cuatro ítems restantes, especialmente los más específicos, 2, 3, y 5, puede utilizarse en análisis diferenciales, dado que nos parecen de indudable interés. Pudieran existir diferencias significativas en algunas variables clave del modelo entre las empresas con o sin incentivos por trabajo seguro, o entre aquellas que priman la seguridad frente a la rapidez y las que priman la rapidez.

### **3.6. Presentación de una escala de política depurada**

Para asegurarnos del buen funcionamiento psicométrico de la escala de política reducida a cuatro ítems se van a efectuar los correspondientes análisis de estructura factorial y de fiabilidad de los ítems y la escala. No se añadirán los análisis de validez criterial dado que son redundantes con los presentados anteriormente, ya que eran ítem a ítem.

Se ha probado una estructura unifactorial, ya encontrada para la versión original de la escala de política. Los resultados indican que la estructura es un excelente ajuste: ji-cuadrado=0.016 para dos grados de libertad ( $p=0.991$ ). Por su parte las saturaciones factoriales han sido:

- Relación del factor con el ítem 4, 0.657, lo que indica una fiabilidad (al multiplicarlo por cien resulta el porcentaje de varianza explicada) elevada, 0.431.

- La relación entre el ítem 5 y el factor es menos adecuada, 0.329, pero es estadísticamente significativa ( $p<.01$ ),. El porcentaje de varianza explicado es del 10.8%. A pesar de la baja fiabilidad es un contenido teóricamente relevante.

- La saturación del ítem seis vuelve a ser elevada, 0.681, que indica una elevada fiabilidad para medir el factor, .463.

- El último de los ítems es el más homogéneo con el factor, con una saturación de 0.775. La fiabilidad alcanza un valor de 0.60.

En cuanto a análisis clásicos de fiabilidad de la escala y de los ítems, pueden consultarse los datos de las tablas 3.25 y 3.26.

Tabla 3.25 Coeficientes de consistencia interna de la escala de política organizacional de seguridad.

<b>Consistencia interna</b>	
Correlación dos mitades	.526
Coeficiente Spearman-Brown	.689
Coeficiente Guttman-Rulon	.598
Coeficiente alpha total	.641
Alpha items impares	.543
Alpha items pares	.426

Aunque el alfa para la escla no es muy elevado, 0.641, sí puede considerarse alto para una escala de cuatro items, ya que es un coeficiente afectado que aumenta a medida que aumentan el número de items.

Tabla 3.26. Media, desviación típica, coeficiente de homogeneidad (correlación ítem-total), fiabilidad, homogeneidad excluyendo el ítem y coeficiente alpha de la escala si se excluye el ítem, para cada ítem de la escala de política organizacional de seguridad.

ITEM	Med.	D. T.	Homogen.	Fiabilidad	Homogen. sin el ítem	Alpha sin el ítem
1	0.464	0.499	.725	.362	.494	.525
2	0.173	0.378	.525	.199	.304	.645
3	1.482	0.806	.855	.689	.517	.577
4	0.327	0.333	.723	.241	.585	.532

Como puede verse en los datos de homogeneidad y fiabilidad de los items de la tabla 3. 26, todos los items son bastante homogéneos con el total de la escala.

A la vista de los datos de estructura factorial y de fiabilidad de la escala, puede concluirse que la versión reducida de la escala de política tiene un comportamiento psicométrico adecuado. Unido esto, a los problemas de validez de contenido detectadas, hace que tengamos más confianza en la validez de la escala reducida que en la original para medir la variable política organizacional sin confusiones con otras variables.

#### 4. Estructura de seguridad

Por ley, en las empresas es necesario algún responsable de seguridad. Según el tamaño de las empresas esta exigencia puede concretarse en la existencia de un único encargado, en la de un comité o en un departamento. Los contenidos de la escala son la existencia de encargados, departamento o comités de seguridad, el conocimiento que de esos representantes tienen los trabajadores (lo 'visibles' que son para ellos) y el conocimiento que tienen de sus funciones.

##### 4.1. Medida

###### ESCALA:

1. ¿Existen en su empresa comités de seguridad e higiene en el trabajo? ..... Si No
2. ¿Existe en su empresa un departamento o un encargado de seguridad e higiene en el trabajo? Si ..... No
3. ¿Conoce Ud. los representantes de los comités de seguridad e higiene en el trabajo de su Empresa? ..... Si No
4. ¿Conoce Ud. las funciones de los comités de seguridad e higiene en el trabajo? Si No

##### 4.2. Estructura factorial confirmatoria

Se han planteado dos estructuras factoriales teóricas para la escala. Dos modelos competitivos para representar los datos, uno unifactorial que nos servirá como medida de 'nivel basal' de explicación, y una estructura de dos factores correlacionados. Por un lado los dos primeros items definirían el factor *existencia de estructura* de seguridad. Este factor incluye dos items que solo preguntan si en la empresa hay algún tipo de estructura formal de seguridad, bien encargado, comité o departamento. El segundo factor viene definido por los contenidos de los items tres y cuatro, y lo hemos denominado *conocimiento de la estructura*. Estos dos items cuestionan acerca del conocimiento que el trabajador tiene de las personas encargadas de la seguridad en su empresa y del conocimiento de las funciones que deben desempeñar.

Los ajustes globales a las dos estructuras hipotetizadas se presentan a continuación y nos permitirán decidir entre la estructura que mejor explique las covariaciones entre los cuatro items.



Tabla 3.27. Índices de ajuste global para las dos estructuras hipotetizadas a los items de la escala de estructura de seguridad. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; CAK=Akaike information criterion; CSK=Schwarz information criterion; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: GLS=mínimos cuadrados generalizados; y ML=máxima verosimilitud.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
UN	GLS	10.5	2	.005	.09	.95	5.27	.246	.445	.988	.971	.990
FACTOR	ML	13.7	2	.001	.12	.94	6.86	.276	.475	.971	.751	.917
DOS	GLS	1.13	1	.285	.01	.99	1.13	.177	.400	.999	.999	1.00
FACTORES	ML	1.15	1	.281	.01	.99	1.15	.177	.400	.992	.993	.999

Comentando en profundidad los índices, tenemos primeramente el ajuste mediante ji-cuadrado, que resulta en una probabilidad significativa para la estructura unifactorial y no significativa para el bifactorial, tanto con mínimos cuadrados como para máxima verosimilitud. Dado que la significatividad de la ji-cuadrado indica diferencias entre el modelo hipotetizado y los datos hay que concluir que el modelo bifactorial se ajusta mejor a la muestra bajo estudio. También, son menores las funciones de ajuste del bifactorial que las del unifactorial, indicando una mejor minimización del error. LHR ofrece mayores valores de ajuste para el bifactorial. La ji-cuadrado partida por los grados de libertad, a pesar de penalizar por ser menos parsimonioso al de dos factores, ofrece resultados inequívocos. Muy adecuado ajuste para el de dos factores e inadecuado para el bifactorial, por encima incluso del valor máximo aceptable de 5. CAK y CSK ofrecen los mismos resultados y sus valores son menores (mejor ajuste) para el bifactorial. Los índices NFI, NNFI y CFI obtienen idénticas conclusiones, si bien hay que destacar que el diseño unifactorial es plausible por todos estos índices si tomamos como valor de ajuste el indicado por Bentler y Bonnett (1980) de .90. A nuestro entender estos índices no nos permitirían discriminar el mejor modelo.

Por otro lado, puede generarse una prueba de las diferencias entre las ji-cuadrado de los dos modelos hipotetizados. Así, para la estimación mínimo cuadrática ofrece los siguientes resultados:

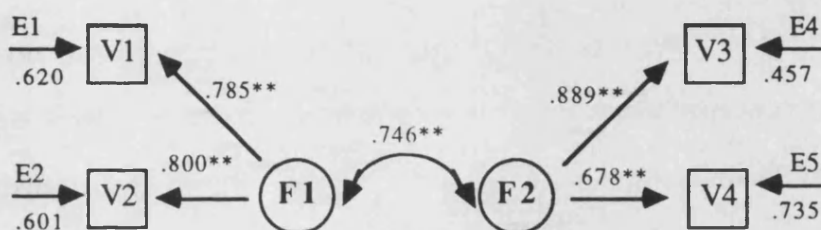
$\Delta\chi^2=9.419$ , con  $\Delta gl=1$  que resulta estadísticamente significativo ( $p<.005$ ), con una menor chi-cuadrado para el modelo bifactorial,

y para máxima verosimilitud

$\Delta\chi^2=12.568$ , con  $\Delta gl=1$  que resulta estadísticamente significativo ( $p<.005$ ) también indicando un ajuste estadísticamente mejor para el modelo bifactorial.

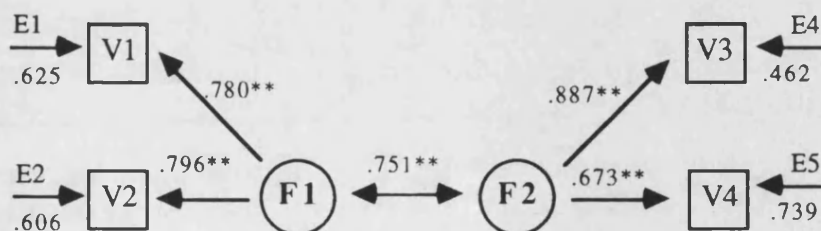
En cuanto a los resultados que ofrecen los índices analíticos, pueden consultarse en las gráficas que ofrecen los resultados del modelo bifactorial estimado por mínimos cuadrados y por máxima verosimilitud que se ofrecen a continuación. Todos los test z realizados para las cargas factoriales y la correlación son significativos estadísticamente a una  $p < 0.001$ , indicando que los valores son significativamente distintos de cero.

Gráfica 3.7. Estructura bifactorial confirmatoria estimada mediante mínimos cuadrados generalizados para la escala de estructura de seguridad. (\*\*= $p < 0.001$ ).



En cuanto a los resultados del LM test y de W test, son los mismos tanto para la estimación mínimos cuadrática, como para la máximo verosímil. No se obtendrían ganancias significativas si se hipotetizara alguna otra relación de un factor con los items adscritos al otro (LMtest), por lo que la estructura es simple, con todos los items cargando en un solo factor. Por otro lado, el W test ofrece como resultado que no habría una pérdida significativa de explicación en el modelo si se hipotetizara que el error de ítem 3 es cero. Este dato indica que la predicción que el factor 2 realiza de la varianza del ítem 3 es especialmente adecuada.

Gráfica 3.8. Estructura bifactorial confirmatoria estimada mediante mínimos cuadrados generalizados para la escala de estructura de seguridad. (\*\*= $p < 0.001$ ).



Al igual que se realizó para las anteriores escalas presentadas, queda por comprobar si puede mantenerse el supuesto de normalidad multivariable. Este supuesto está a la base del uso de la estimación máximo verosímil y mínimo cuadrática. Se han realizado pruebas de bondad de ajuste -prueba de Kolmogorov-Smirnov- para cada variable al objeto de probar la normalidad univariada, condición necesaria para la normalidad multivariable. Los resultados de esta prueba de ajuste a la normalidad se presentan en la siguiente tabla.



Tabla 3.28. Prueba de Kolmogorov-Smirnov de una muestra para comprobar el ajuste de los ítems de la escala de formación-información a la distribución normal univariada.

ITEM	Casos	Diferencia máxima	Probabilidad 2 colas
1	120	0.500	0.000
2	118	0.500	0.000
3	120	0.500	0.000
4	121	0.500	0.000

Ninguna prueba de normalidad a ofrecido apoyo a la hipótesis de normalidad multivariable. En cualquier caso, hay que recordar que los cuatro ítems de la escala se contestan mediante una escala dicotómica, lo que hace difícil que ajuste una prueba de normalidad.

Se presentan, además, los estadísticos descriptivos de asimetría y apuntamiento para cada uno de los ítems de la escala de estructura de seguridad. Ninguno de los coeficientes (tabla 3.29) ofrece valores que puedan mantener la normalidad.

Tabla 3.29. Datos de asimetría y curtosis para los ocho ítems de la escala de política organizacional.

	Asimetría	Curtosis
Item 1	0.2996	-1.9102
Item 2	0.6205	-1.615
Item 3	0.9923	-1.0154
Item 4	0.8922	-1.2039

Por su parte, aún cuando la normalidad multivariada es condición necesaria para la normalidad multivariable el programa EQS ofrece los datos de apuntamiento multivariado, junto con el esperado bajo la normalidad multivariable.

COEFICIENTE DE MARDIA MULTIVARIADO= 4.123
Coef. esperado para la normal multivariada= 3.0923

Puede verse cómo el coeficiente de mardia para los cuatro ítems ofrece un valor cercano a 4, mientras que lo esperado bajo la normalidad multivariable ofrece un valor de 3. Son valores relativamente cercanos. Desgraciadamente no puede probarse su significatividad estadística.

Ante los datos de asimetría y curtosis se ha optado por calcular el análisis factorial confirmatorio retenido por su mejor ajuste por los métodos convencionales de máxima verosimilitud y mínimos cuadrados mediante estimación arbitraria. Permitirá comparar los datos de ajuste global y analítico a través de métodos de estimación.

### Capítulo 3. Clima organizacional hacia la seguridad.

Tabla 3.30. Índices de ajuste global para la estructura de dos factores a los ítems de la escala de estructura de seguridad. Los índices son: *chi*=chi-cuadrado; *g*=grados de libertad del modelo; *p*=probabilidad; *F*=función de ajuste a minimizar; *LHR*=razón de verosimilitud escalada; *CAK*=Akaike information criterion; *CSK*=Schwarz information criterion; *NFI*=normed fit index; *NNFI*=non-normed fit index; *CFI*=comparative fit index. Métodos de estimación: *GLS*=mínimos cuadrados generalizados; *ML*=máxima verosimilitud; *AGLS*= estimación arbitraria.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
MODELO	GLS	1.13	1	.285	.01	.99	1.13	.177	.400	.999	.999	1.00
DE DOS	ML	1.15	1	.281	.01	.99	1.15	.177	.400	.992	.993	.999
FACTORES	AGLS	0.68	1	.408	.00	.99	.684	.173	.396	.992	1.02	1.00

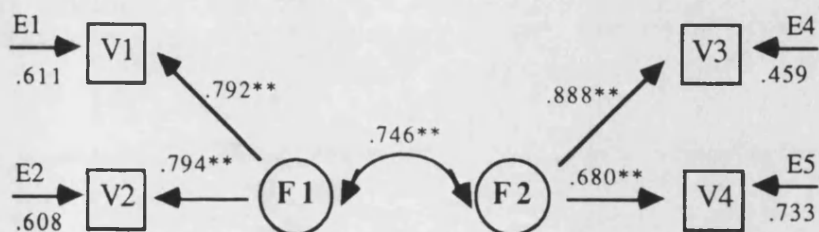
La prueba de chi-cuadrado ha sufrido, como puede verse en la tabla 3.28, cambios en la significatividad. Se ha pasado de probabilidades de 0.28 a probabilidades de 0.4. En cualquier caso, todos los índices de ajuste global muestran un muy adecuado grado de ajuste de la estructura bifactorial hipotetizada a los datos observados.

INDICE DE BONDAD DE AJUSTE, GFI= .999  
 Índice de bondad de ajuste ajustado, AGFI= .995

Los dos índices de ajuste global que ofrece el programa EQS para la estimación arbitraria, GFI y AGFI, ofrecen valores elevados que indican una elevadísima proporción de varianza explicada por la estructura factorial propuesta.

Los datos de ajuste analítico asociados a los coeficientes de la estructura bifactorial se muestran en la gráfica 3.9. Se espera, conforme hemos comprobado en las anteriores escalas que los coeficientes obtenidos mediante estimación arbitraria sean consistentes con los de los otros métodos, a diferencia de los cambios que produce en la significatividad de la prueba de chi-cuadrado.

Gráfica 3.9. Estructura bifactorial confirmatoria por estimación arbitraria para la escala de estructura de seguridad. (\*\*= $p < 0.001$ ).



Todas las saturaciones factoriales, así como la correlación entre los dos factores, son altamente significativas. Los resultados de la estimación arbitraria, como antes los de máxima verosimilitud y mínimos cuadrados generalizados muestran una escala consistente y con factores altamente relacionados con sus indicadores (items).

Tabla 3.31. Saturaciones factoriales estimadas mediante mínimos cuadrados generalizados, máxima verosimilitud y estimación arbitraria.

	GLS	ML	AGLS
Item 1	.785	.780	.792
Item 2	.80	.796	.794
Item 3	.889	.887	.888
Item 4	.678	.673	.680
Relación factores	.746	.751	.746

La tabla 3.29 pone de manifiesto la altísima consistencia y mínimas diferencias entre métodos de estimación que se dan en el cálculo de los coeficientes. Los resultados en este sentido concuerdan con los ya expuestos por Tanaka (1984) a nivel estadístico. Efectivamente, las diferencias entre los tres métodos son de centésimas, no cambiando tampoco su significatividad estadística.

#### 4.3. Fiabilidad y homogeneidad de la escala y los items

La fiabilidad de la escala se ha medido mediante los coeficientes de consistencia interna que aparecen en la tabla 3. 32.

Tabla 3.32. Coeficientes de fiabilidad, consistencia interna, de la escala de estructura de seguridad.

<i>Consistencia interna</i>	
Correlación dos mitades	0.756
Coefficiente Spearman-Brown	0.861
Coefficiente Guttman-Rulon	0.860
Coefficiente alpha total	0.806
Alpha items impares	0.674
Alpha items pares	0.541

Los coeficientes de consistencia interna son adecuados para el contexto de medición, mostrando buenas covariaciones entre los cuatro items de la escala. Así, el coeficiente alfa para el total alcanza el valor de 0.806.

Tabla 3.33. Media, desviación típica, coeficiente de homogeneidad (correlación ítem-total), fiabilidad, homogeneidad excluyendo el ítem y coeficiente alpha que quedaría en la escala si se excluyera cada ítem, para la escala de estructura de seguridad.

ITEM	Med.	D. T.	Homogen.	Fiabilidad	Homogen. sin el ítem	Alpha sin el ítem
1	0.426	0.494	0.813	0.402	0.638	0.749
2	0.352	0.478	0.801	0.382	0.626	0.754
3	0.278	0.448	0.823	0.369	0.678	0.731
4	0.296	0.457	0.744	0.340	0.548	0.790

Los resultados de homogeneidad de la escala son adecuados, elevados si se considera la correlación del ítem con el total sin excluir el ítem y apreciablemente más bajo, pero aún elevado, si se considera la correlación excluido el ítem, lo cual es lógico teniendo en cuenta que la escala la forman únicamente cuatro ítems. Los resultados de fiabilidad para los ítems son altos teniendo en cuenta que son ítems dicotómicos. Por último, no es de extrañar que al alpha sin el ítem (coeficiente alpha que tendría la escala si se eliminara el ítem bajo estudio) sean más bajos. La escala que quedaría sería de tan solo tres ítems.

Se puede nuevamente calcular una medida de la fiabilidad de los ítems desde el análisis factorial confirmatorio como el cuadrado de la saturación, que indica la proporción de varianza del ítem explicada por el factor.

Tabla 3.34. Saturaciones factoriales estimadas mediante mínimos cuadrados generalizados y máxima verosimilitud elevados al cuadrado de los cuatro ítems de la escala de estructura de seguridad.

	GLS	ML	AGLS
Item 1	0.616	0.608	0.627
Item 2	0.640	0.633	0.630
Item 3	0.790	0.786	0.788
Item 4	0.459	0.452	0.462

Todos los ítems presentan adecuados niveles de fiabilidad a través de todos los métodos de estimación. La proporción de varianza de cada ítem explicada por los factores latentes hace pensar que los ítems son consistentes con los componentes que mide la escala de estructura de seguridad.

Si aplicamos la fórmula de Bagozzi (1981) para la fiabilidad de los ítems obtenemos los siguientes resultados -ver tabla 3.35 para la estimación mínimo cuadrática, tabla 3.36 para máxima verosimilitud, y tabla 3.37 para la estimación arbitraria-

Tabla 3.35. Índices de fiabilidad de los ítems calculados desde los resultados del análisis factorial confirmatorio de la escala de estructura de seguridad por la fórmula propuesta por Bagozzi (1981). Estimación mínimo cuadrática.

	Factor I	Factor II
Item 1	0.498	-
Item 2	0.515	-
Item 3	-	0.633
Item 4	-	0.384

Tabla 3.36. Índices de fiabilidad de los ítems calculados desde los resultados del análisis factorial confirmatorio de la escala de estructura de seguridad por la fórmula propuesta por Bagozzi (1981). Estimación máximo verosímil.

	Factor I	Factor II
Item 1	0.493	-
Item 2	0.511	-
Item 3	-	0.630
Item 4	-	0.379

Tabla 3.37. Índices de fiabilidad de los ítems calculados desde los resultados del análisis factorial confirmatorio de la escala de estructura de seguridad por la fórmula propuesta por Bagozzi (1981). Estimación arbitraria.

	Factor I	Factor II
Item 1	0.506	-
Item 2	0.509	-
Item 3	-	0.632
Item 4	-	0.386

Los cálculos mediante la fórmula de Bagozzi baja ligeramente los resultados de fiabilidad para los cuatro ítems. Sin embargo, los resultados son consistentes a través de los diferentes métodos, desde los clásicos hasta los basados en los resultados del análisis factorial tanto en máxima verosimilitud como en mínimos cuadrados generalizados, o estimación arbitraria y muestran una adecuada fiabilidad de la escala y los ítems que la componen.

#### 4.4. Validez de los items

En cuanto a la validez, las hipótesis de relación con los criterios son similares a las ya planteadas para formación-información y política, las otras dos variables del clima, pero con matizaciones. Así, se espera que las relaciones de esta variable sean, en principio, menores con todos los criterios. Esto debido a que la escala mide básicamente existencia de la estructura de seguridad y no si su funcionamiento es adecuado o no, no si los trabajadores perciben efectos positivos en la seguridad por la intervención de la estructura de seguridad de la empresa. Por estar la estructura en cierta medida legalmente establecida puede que la aparición de encargados no afecte excesivamente el clima de la empresa. Podría ocurrir que, aún existiendo, no fueran activos y sólo se plantearan para cumplir con la legislación. No obstante, se esperan las siguientes relaciones con los criterios: relación positiva con formación-información y con política por ser partes de la variable clima; relación positiva con respuesta de los superiores; relaciones positivas pero más débiles sobre los trabajadores (compañeros y actitud y conducta del trabajador); no se espera relación con riesgo basal y, si hubiera, una débil relación negativa con riesgo real que debe estar muy mediatizada por otras variables; se espera poca relación con los accidentes y con la tensión.

Tabla 3.38. Correlaciones de los items de la escala de estructura de seguridad con los principales criterios. F=formación-información; PO=política organizacional de seguridad; RS=respuesta de los superiores en seguridad; RC=respuesta de los compañeros en seguridad; L=actitud hacia la seguridad; CHS=conducta hacia la seguridad; RB=riesgo basal; RR=riesgo real; T=tensión psicológica; y AC= siniestralidad. Símbolos de significación estadística de las correlaciones: n.s.= no significativo; #= marginalmente significativo; \*= probabilidad <0.05; y \*\*= probabilidad <0.01, extraídos de tablas (Welkowitz, Ewen y Cohen, 1981).

	F	PO	RS	RC	L	CHS	RB	RR	T	AC
1	.169 #	.309**	.204*	.145 ns	.089 ns	-.040ns	.118 ns	-.047ns	-.063ns	.002 ns
2	.114 ns	.328**	.209 *	.164 #	.083 ns	-.026ns	.086 ns	-.028ns	.037 ns	-.053ns
3	.126 ns	.275**	.170 #	.196 *	.012 ns	-.030ns	.002 ns	.156 ns	.100 ns	-.059ns
4	.163 #	.400**	.185 #	.074 ns	-.011ns	-.005ns	.167 #	.035 ns	-.039ns	.080 ns

Las relaciones son positivas con formación-información y con política, pero muy débiles con las primeras. Esto podría indicar que la estructura de seguridad no está, en nuestra muestra, muy implicada en la formación de seguridad. Sin embargo, y observando detenidamente esta relación, cabe recordar que los items 4, 5 y 6 de la escala de formación-información tenían una mayor relación con la variable estructura de seguridad en el estudio de validez criterial. Estos tres items son precisamente los que componen el factor denominado información de la escala de formación-información. Las correlaciones de estos tres items de información con la variable estructura eran relativamente elevadas: 0.324, 0.522, y 0.499 respectivamente. Una hipótesis para la correlación de los items de estructura con la variable formación es que al estar ésta formada por

seis items pero siendo precisamente los tres primeros de seis alternativas, mientras los tres últimos sólo de dos, son precisamente los tres primeros los que más pesan en la escala, modificando las correlaciones más en su sentido que en el del factor información. La correlación con clima, por su parte, es positiva y relativamente elevada tal y como se hipotetizaba.

Las relaciones con respuesta de los superiores y de los compañeros son positivas, pero débiles, mostrando que la relación de la variable estructura de seguridad con estos dos criterios es menor de la esperada. La relación con la actitud y la conducta hacia la seguridad del trabajador es inexistente, no mostrándose la estructura como capaz de producir cambios en esas variables. Tampoco existe relación con las variables accidentabilidad ni tensión.

La debilidad o carencia de relación de los items, y por tanto de la escala total, con los criterios evidencia que la estructura de seguridad tal y como está definida presenta poco efecto sobre la seguridad de la empresa, respuesta de los superiores, de los compañeros, niveles de riesgo, conductas seguras, etc.

#### **4.5. Conclusiones para la escala de estructura de seguridad.**

A lo largo de los análisis presentados pueden entreverse, a nuestro entender, dos conclusiones psicométricas para la escala. La primera, que los items son internamente consistentes, homogéneos. La segunda, que la validez predictiva (criterial) de la escala no es adecuada. En apoyo de la primera conclusión pueden aducirse los resultados de consistencia interna de la escala y de los items, elevados en ambos casos, y los resultados del análisis factorial confirmatorio, que mostraban con un muy adecuado ajuste que una estructura bifactorial fuertemente oblicua explicaba adecuadamente la varianza de los items. Sin embargo, la validez criterial ha resultado hasta cierto punto decepcionante. En todo caso la interpretación de la validez de la variable estructura de seguridad es cualitativamente distante a la de otras variables debido a que esta variable no se refiere a un constructo psicológico propiamente sino más bien a la mera constitución y conocimiento de la presencia o ausencia de determinadas estructuras. En este sentido propiamente no es una variable a validar desde otras variables del modelo, si bien, obviamente también está sujeta a una determinada posición y relaciones en la red nomológica de constructos. Los items sólo se muestran relacionados significativamente conforme a hipótesis con otro constructo de clima, la política organizacional de seguridad. No hay correlación significativa con el otro constructo de clima, y moderada a baja con la respuesta de los superiores. Tampoco hay relaciones apreciables con el resto de criterios. Estos resultados auguran un poco peso de la variable en la cadena causal de

explicación de los accidentes, señalando una relativa independencia entre la estructura de seguridad de la empresa y otros aspectos relevantes del funcionamiento de la empresa en materia de seguridad. La estructura de seguridad en esta muestra puede no ser vista como integrada en la 'cultura' y/o 'política' de seguridad imperante en la empresa. Obviamente, estas conclusiones son siempre entendiendo la variable estructura tal y como aquí ha sido operacionalizada, con los contenidos concretos de los cuatro items presentados. Los items son adecuados para lo que tratan de medir, constatar si existe o no estructura de seguridad y si es o no conocida, pero se muestran esos contenidos como poco relacionados, por lo que quizá se pueda ampliar sus contenidos para mejorar el comportamiento de la variable y recoger aspectos que pueden resultar relevantes. En los contenidos de la variable ahora definida no se pregunta qué actividad desarrollan los encargados o comités de seguridad. Quizá la escala podría mejorarse incluyendo items relativos a la percepción de los sujetos acerca de la actividad y utilidad real de las tareas desarrolladas por los encargados de seguridad. En definitiva, items que traten de medir si para los trabajadores es relevante esta función en la consecución de los fines empresariales en materia de seguridad o, por el contrario, se encuentra desligado de ellos.





## 4. Medición de los Riesgos laborales.

### 1. Introducción

La Línea de Investigación de Seguridad Laboral de la Universitat de València viene trabajando en una doble conceptualización de los riesgos laborales. Se distingue entre un riesgo basal en el puesto de trabajo y un riesgo real.

Riesgo Basal. Puede definirse como el nivel de riesgo de accidente presente en el puesto de trabajo por la misma naturaleza de la ejecución de la tarea. Un montador de grúas en el sector de la construcción está sometido a un riesgo intrínseco en su puesto mucho mayor al que se ve sometido un profesor universitario. El riesgo de sufrir descargas eléctricas es elevado para un electricista pero bajo para un representante.

Riesgo real. Es la probabilidad de accidentes que permanece tras tomar las medidas de seguridad oportunas. Es la probabilidad real de accidentarse dadas las acciones emprendidas a diversos niveles de la organización y por el propio trabajador para rebajar este riesgo o al menos sus efectos. Por el contrario, si se realizan conductas inseguras el riesgo real puede incrementarse aún con riesgo basal bajo.

Efectivamente, y al menos teóricamente, ambos conceptos son discernibles y tienen diferentes implicaciones. El riesgo basal en el puesto de trabajo determinará en muchos casos la probabilidad de accidentarse. Aun cuando el riesgo basal parece ser un predictor directo de la ocurrencia de accidentes, cabe pensar más en el concepto de riesgo para afinar esta primera impresión. Es evidente que un montador de grúas tienen alta probabilidad de sufrir accidentes por el riesgo basal a que se ve sometido. No obstante, este riesgo puede amortiguarse casi en su totalidad mediante el constante y correcto uso de un sencillo equipo de protección, un cinturón de seguridad y un casco. Es bien conocido que por múltiples factores una gran cantidad de montadores no hacen uso de estos elementos de seguridad. Dos montadores, uno que sí hiciera uso del equipo de protección y otro que no, estarían sometidos al mismo nivel de riesgo basal pero a muy diferente nivel de riesgo real. Así, el riesgo basal puede considerarse un predictor directo de la probabilidad de accidentarse sólo cuando no se toman medidas de seguridad. Sin embargo el que sí es un predictor directo siempre de los accidentes es el riesgo real.

Una vez vista la distinción a nivel de contenido e implicaciones de ambos tipos de riesgo, cabe hablar de la medida de los mismos. Nuestro trabajo, por su formato de encuesta a trabajadores, está basado en las percepciones de esos mismos trabajadores sobre los aspectos que se cuestionan, y depende en gran medida de su visión particular y de su interpretación cognitiva de las preguntas. En cuanto a las variables tomadas una a una, hay que tener en cuenta que no disponemos para cada puesto de trabajo de los representados en la muestra de una medida externa al propio trabajador del nivel de riesgos. Sería muy adecuado tener para cada puesto una medida, por ejemplo, extraída de la evaluación de un experto en riesgos del nivel de riesgo basal y de riesgo real. Desgraciadamente esto no es posible. Así, contamos con una medida de riesgos percibidos por el trabajador y esto no hay que olvidarlo. En la medida en que la percepción sea realista y ajustada, las conclusiones serán válidas. Por otro lado, y en cuanto a la distinción entre nivel de riesgo basal y nivel de riesgo real como variables diferenciables, es teóricamente clara pero puede fácilmente ocurrir que la distinción teórica no se corresponda con la percepción de los trabajadores que no distinguen entre ambos conceptos. Como se ve, ciertos problemas aquejan la medición de riesgos, y lo que es peor es difícil medir su impacto en nuestros instrumentos de medida. No obstante, este problema es común a la mayoría de investigaciones que incluyen la medida de los riesgos como variable relevante.

## 2. Riesgo Basal

### 2.1. Medida

Para la medición de la variable riesgo basal se dispone de una escala de diecisiete items dicotómicos. La mayor puntuación de la escala se corresponde con un mayor nivel de riesgos en el puesto de trabajo.

#### ESCALA:

1. Trabajo con materiales tóxicos o peligrosos.....	SI	NO
2. Trabajo en lugares peligrosos o tóxicos. ....	SI	NO
3. Manejo máquinas o herramientas cortantes, punzantes o peligrosas. ....	SI	NO
4. En mi trabajo puedo sufrir atrapamientos, descargas eléctricas o vibraciones. ....	SI	NO
5. En mi trabajo puedo sufrir quemaduras, golpes o cortes. ....	SI	NO
6. En mi trabajo estoy expuesto a derrumbamientos, explosiones o infecciones. ....	SI	NO
7. En mi trabajo puedo inhalar gases o polvo o sufrir algún tipo de envenenamiento .....	NO	SI
8. A causa de mi trabajo puedo sufrir un accidente de tráfico o violencia de otras personas ..... .....	NO	SI
9. Trabajo expuesto a ruido, humo, temperaturas extremas o bajo fuertes presiones de tiempo. .... .....	NO	SI
10. En mi trabajo tengo que levantar, cargar o descargar objetos pesados, o puedo ser golpeado o arrollado por alguna máquina o herramienta. ....	SI	NO
11. En mi trabajo puedo sufrir daños en los ojos o en la cara. ....	SI	NO
12. Puedo sufrir un accidente en las manos o en los brazos. ....	SI	NO
13. En mi trabajo puedo padecer daños en las piernas, en los pies o en el tronco. ....	SI	NO
14. Estoy expuesto frecuentemente a sufrir daños físicos importantes en mi puesto de trabajo. .... .....	NO	SI
15. La avería de algún vehículo o máquina o el error de otros podría llevar a que yo sufriera un accidente laboral.....	SI	NO

16. En mi trabajo deben tomarse estrictas medidas de seguridad para evitar tener accidentes..... SI  
..... NO
17. En su trabajo está expuesto fácilmente a accidentes de trabajo donde podría perder la vida o quedar inválido..... SI NO

En el contenido de los items se ha intentado rastrear los tipos de riesgo más comunes, dado que es una escala de uso general. El hecho de que tenga que usarse en un amplio rango de empresas - químicas, construcción, cadenas de montaje, servicios- supone una dificultad adicional; están sometidas a muy diferentes tipos de riesgos. Para superar este problema y conseguir una medida global de riesgos sin demasiados items sin contestar se ha optado por introducir riesgos diferentes dentro del contenido de los items. Esta opción genera un cuestionario muy global de riesgos, lo que puede ofrecer problemas en el estudio de la validez factorial del cuestionario por la mezcla de contenidos dispares.

## **2.2. Estructura factorial confirmatoria**

Para el estudio de la validez factorial de los items se han planteado diversos análisis factoriales confirmatorios, siguiendo la misma lógica de modelos anidados ya comentada y aplicada en capítulos anteriores. Aquí se han planteado tres modelos básicos para representar a los datos, dos que nos sirven como modelos basales competitivos, y finalmente un tercero que es el modelo sustantivo, teóricamente planteado. Las tres estructuras propuestas inicialmente se completan con una última estructura calculada a partir de las sugerencias de los índices de ajuste analítico:

1) Estructura unifactorial. Esta estructura nos permite tener un modelo basal, a partir del cual evaluar las ganancias de otros modelos de más factores, sin los inconvenientes que presenta el simple uso de un modelo nulo, de completa independencia de las medidas. Además de como simple modelo basal, en este caso la estructura de un solo factor se contempla como hipótesis teórica razonable dado que los items incluyen dentro de sus contenidos particulares riesgos dispares muy relacionados con contenidos en el resto de la escala.

2) Estructura bifactorial. Construida por oposición entre los items que preguntan por riesgos tóxicos, muy distintivos y que suelen aparecer sobre todo en empresas químicas, frente al resto de riesgos (no tóxicos) mucho más comunes en amplios rangos de empresas, como atrapamientos, cortes, desgarros, etc. Así, el primer factor quedaría compuesto por los items 3 a 6 y 8 a 17. Se denominaría riesgos no tóxicos e incluiría manejo de máquinas peligrosas, exposición a derrumbamientos, cargas y descargas de objetos pesados, temperaturas extremas, presión de tiempo, y

también incluye los ítems de consecuencias y gravedad de los riesgos. Incluye todo lo que claramente no son riesgos tóxicos, químicos. El factor dos, riesgos tóxicos, incluye los ítems 1, 2 y 7. Los contenidos de estos tres ítems presentan un fuerte contenido de riesgos por toxicidad o envenenamiento.

3) Estructura trifactorial. El primer factor estaría constituido por los ítems 3, 5, 10 a 14, 16 y 17, y se denominaría riesgos manipulativos. Incluye riesgos por manejo de herramientas o maquinaria, vehículos, etc; carga o descarga de objetos pesados, etc. Frente a los otros factores presenta una mayor implicación del sujeto. El segundo factor, idéntico al segundo de la estructura factorial, está compuesto por los ítems 1, 2 y 7, y se ha etiquetado como riesgos tóxicos. El último factor, riesgos pasivos, agrupa los restantes ítems - 4, 6, 8, 9 y 15-. Los contenidos hacen referencia a peligros de accidente en el puesto de trabajo pero básicamente producidos por terceros, en que el trabajador sufre de forma pasiva las consecuencias. No obstante, debemos decir nuevamente que los diferentes tipos de riesgo no estaban claramente delimitados, y aunque nosotros pensamos que esta estructura puede ser plausible, descansa en nuestras impresiones sobre qué contenido es más relevante para cada ítem. De hecho, la mayoría de los ítems presentan diversos contenidos juntos.

4) Estructura trifactorial modificada. Como el ajuste trifactorial no era del todo adecuado, y teniendo en cuenta las modificaciones que proponían el W test y la LM test, se han planteado modificaciones de la estructura factorial. No cabe aquí hablar de qué ítems saturan en cada factor, dado que surgen diferentes estructuras según el tipo de estimación empleado. Se verán más adelante qué cambios se han producido y qué sentido teórico pueden tener.

Se van a presentar una a una todas las estructuras factoriales probadas, en sus datos de ajuste global y analítico. La especial complejidad de esta escala así lo recomienda.

En la tabla 4.1 se presentan los datos de ajuste global para la estructura unifactorial. Puede verse que tanto para la estimación por mínimos cuadrados generalizados (GLS), como para la máximo verosímil (ML) el ajuste del modelo a los datos es inadecuado.

En la estimación mediante mínimos cuadrados la prueba de ji-cuadrado es estadísticamente significativa, indicando ajuste inadecuado y la razón de verosimilitud escalada -LHR-, que oscila de cero a uno, indicando un ajuste perfecto, se sitúa en un pobre 0.464. Sin embargo, ji-cuadrado partido por los grados de libertad muestra una estructura altamente parsimoniosa, debido al elevado número de grados de libertad, a la enorme reducción de la complejidad que supone pasar de diecisiete medidas a una, el factor propuesto. Además, los índices de ajuste con un modelo nulo son todos

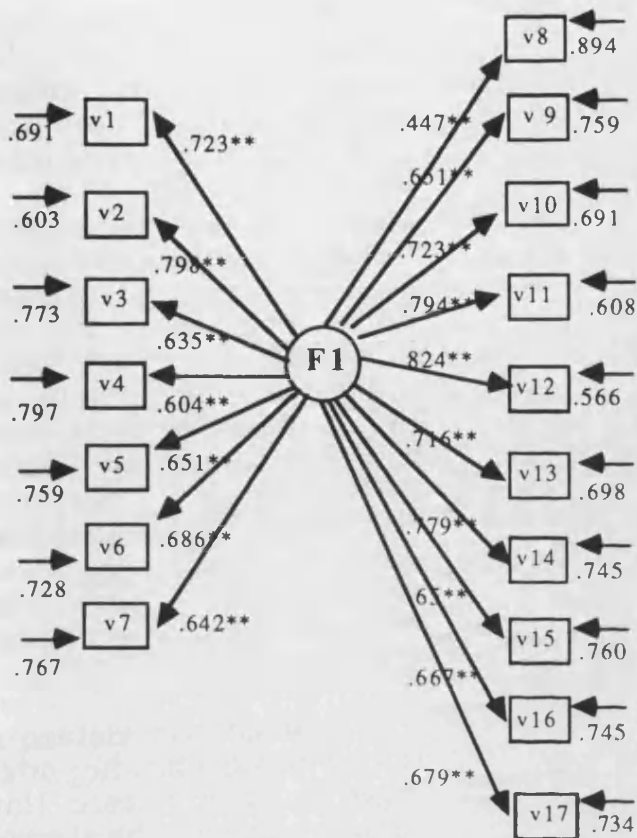
satisfactorios, llegando incluso al 0.99. Índices contradictorios, teniendo en cuenta que no puede despreciarse el ajuste mediante ji-cuadrado; con 107 personas como N muestral no puede pensarse que sea el tamaño de la muestra el causante de un mal ajuste.

No hay dudas en lo que se refiere a la estimación máximo verosímil. Empezando en la ji-cuadrado y acabando en el índice de ajuste comparativo (CFI), todos los índices muestran inadecuado ajuste, a excepción de la chi-cuadrado partido los grados de libertad, y aún este, sólo si consideramos el criterio menos restrictivo para considerarlo indicativo de un modelo parsimonioso.

*Tabla 4.1. Índices de ajuste global para la estructura unifactorial. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; CAK=Akaike information criterion; CSK=Schwarz information criterion; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: GLS=mínimos cuadrados generalizados; ML=máxima verosimilitud. Tamaño de la muestra= 107.*

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
UN	GLS	162.47	119	.005	1.53	.464	1.365	2.16	3.01	.971	.991	.992
FACT.	ML	359.53	119	<.001	3.39	.183	3.021	4.02	4.87	.586	.625	.672

Sin embargo, y pese a parecer que el ajuste monofactorial es, como mínimo muy mejorable, conviene observar las saturaciones factoriales y las pruebas de significatividad estadística asociadas. Así, en las gráficas 4.1 y 4.2 pueden verse estos datos para los dos métodos de estimación, mínimos cuadrados y máxima verosimilitud, respectivamente.



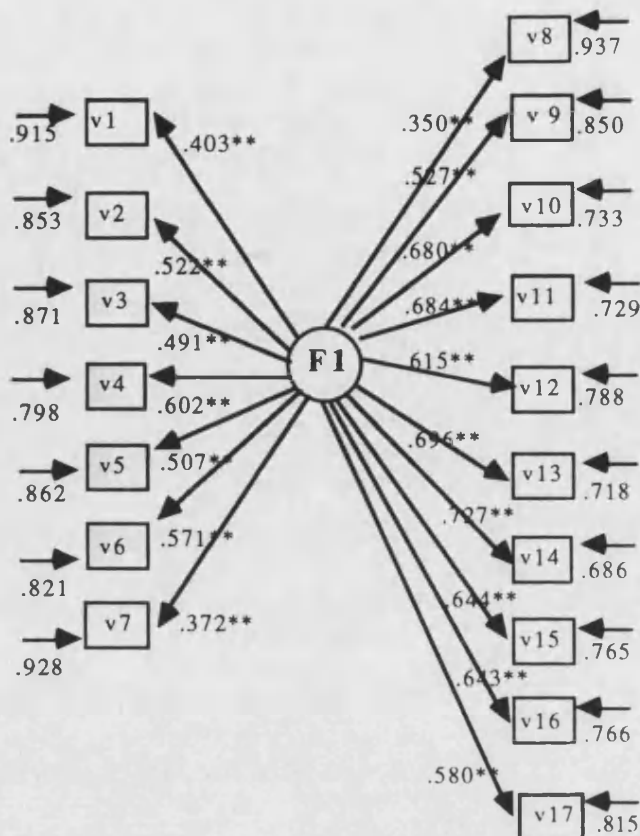
Gráfica 4.1. Modelo de un factor estimado mediante mínimos cuadrados generalizados.

La saturación al cuadrado nos indica el porcentaje de varianza del ítem explicado por el factor. El ítem ocho es pobremente explicado por el factor, tan sólo un 19.98%, pero esto no es lo común. Por simplificar, y a la vista de los datos de la gráfica, la mayoría de las saturaciones factoriales se sitúan en torno al 0.75, lo que significa una apreciable explicación de la varianza del ítem de en torno al 56%, con una máxima explicación del 67.89% del ítem 12. Otra serie importante de ítems presentan saturaciones en torno al 0.65, que representa un porcentaje de explicación en torno al 43%, nada despreciable.

Ahora queda observar, de la misma forma, los resultados del análisis factorial confirmatorio de un factor para la estimación máximo verosímil, gráfica 4.2. Dados los datos de ajuste global puede preverse que los porcentajes de varianza explicada serán sensiblemente inferiores, pero resulta relevante, por un lado, ver qué cambios se producen con respecto a la estimación mínimo cuadrática, y por otro, si con todo un factor da cuenta de parte importante de la variabilidad de los datos.

Todas las saturaciones factoriales son estadísticamente significativas ( $p < .001$ ). Este dato en sí no es excesivamente importante, dado que únicamente indica que la saturación no igual a cero en la población. Esto es, indica simplemente que existe relación entre el ítem y el factor, una relación que no puede deberse al azar. En un análisis factorial confirmatorio buscamos algo más que esto; buscamos que el factor que se asume como subyacente al ítem sea capaz de explicar una parte importante de la varianza observada en el ítem. Bien, pues esto también se cumple en la mayor parte de las saturaciones factoriales; no sólo son significativas, sino que son elevadas. Sabemos que la





Gráfica 4.2. Modelo de un factor estimado mediante máxima verosimilitud.

Así, por un lado podemos comparar las saturaciones de ambos métodos de estimación. Son iguales en el sentido de la relación y en la significatividad estadística. No hay cambios bruscos, pero ocurre lo mismo que ya hemos apreciado en otras escalas en capítulos anteriores: el método de máxima verosimilitud ofrece estimaciones más bajas que mínimos cuadrados generalizados. Estas diferencias se sitúan en torno a 0.1, a favor de la estimación mínimo cuadrática.

Echemos un vistazo a la varianza explicada por el único factor propuesto. Hay saturaciones bastante diferentes, pero con ánimo de resumir, podemos considerar que los items 4, 6, 10, 11, 12, 13, 14, 15, 16, y 17 puede

considerarse como en torno a la saturación del 0.65; esto significa que la explicación de la varianza se sitúa en torno al 43%. Los items 2, 5 y 9 presentan saturaciones en torno al 0.5, que resulta en un porcentaje de explicación del 25%. Por último, en la parte más baja de explicación el factor apenas explica un 13% de los items 7 y 8. Aunque los porcentajes de varianza explicada bajan considerablemente, 10 de los 17 items de la escala se verían relativamente bien explicados por un único factor de primer orden. Esta cantidad de varianza explicada puede considerarse insuficiente, pero hay que recordar que una estructura unifactorial supone una reducción drástica de la complejidad de diecisiete items y es, por tanto, altamente parsimoniosa. Estos resultados van a ser especialmente relevantes porque cualquier estructura más compleja deberá evaluarse a la luz de éstos, y no sólo del ajuste global, desde nuestro punto de vista.

La segunda estructura factorial hipotetizada era bifactorial, como hemos visto al principio. Se probó esta estructura pensando en que quizá los riesgos tóxicos, de envenenamiento con gases, etc era más específico que el resto de riesgos, más mecánicos y por

manipulación de máquinas. También se planteó pensando en que su contenido en los ítems estaba más aislado, menos mezclado con otros contenidos. Se esperaba que formara, frente al resto de ítems de la escala, un factor claramente aislado.

Los datos de ajuste global para la estimación por mínimos cuadrados generalizados y máxima verosimilitud se presentan en la tabla 4.2. En cualquier caso, y mirando globalmente los resultados, no puede decirse que haya una ganancia clara con respecto al modelo de un solo factor. Así, el ajuste mediante ji-cuadrado es, para GLS, significativo a un nivel del 0.009, mientras en el modelo unifactorial era de 0.005. Por lo que respecta a ML, tanto en el modelo unifactorial como en el bifactorial la significatividad es más elevada (<0.001). Se reproducen los resultados también en el resto de índices. Para mínimos cuadrados generalizados, los índices de comparación con un modelo nulo indican ajuste adecuado, con valores por encima de .97; mientras chi-cuadrado partido los grados de libertad nos informa de un modelo altamente parsimonioso, como ya ocurría en el modelo monofactorial. La prueba de ji-cuadrado y la razón de verosimilitud escalada indican ajuste inadecuado. En la estimación mediante máxima verosimilitud todas las pruebas de ajuste global, a excepción de la ji-cuadrado partido los grados de libertad indican mal ajuste del modelo a los datos.

Tabla 4.2. Índices de ajuste global para la estructura bifactorial. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; CAK=Akaike information criterion; CSK=Schwarz information criterion; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: GLS=mínimos cuadrados generalizados; ML=máxima verosimilitud.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
BIFAC-	GLS	157.04	118	.0094	1.48	.476	1.33	2.13	3.01	.972	.992	.993
TORIAL	ML	285.02	118	<.001	2.68	.260	2.415	3.34	4.21	.672	.737	.772

Queda por ver si mediante una estructura más compleja se ha ganado algo en la explicación de los datos, y si esa ganancia es estadísticamente significativa. Para comprobar qué estructura representa mejor los datos puede realizarse una prueba de las diferencias entre las dos ji-cuadrados. Realizada la prueba para los ji-cuadrado calculadas mediante mínimos cuadrados generalizados los resultados son:

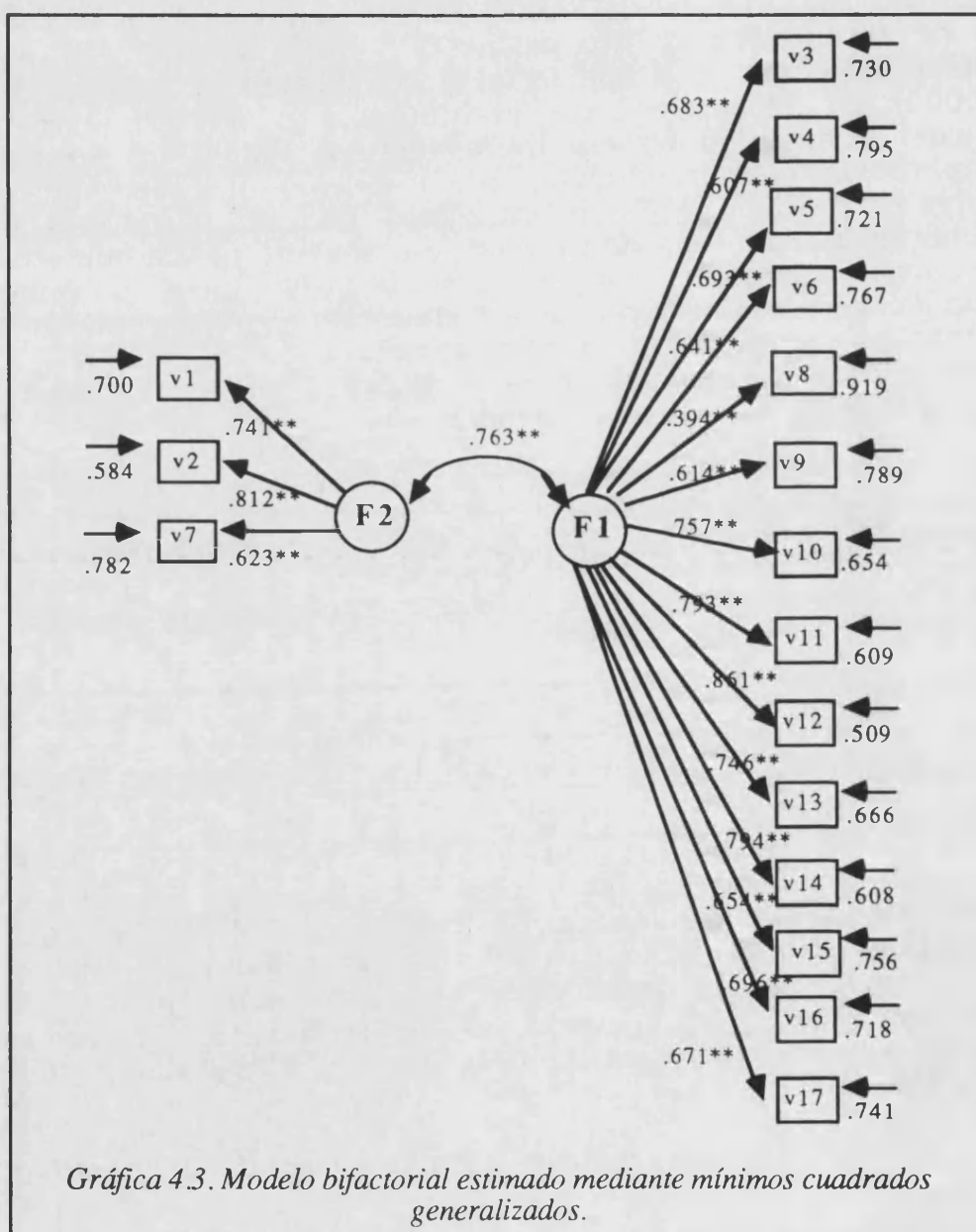
$\Delta\chi^2 = 5.43$ , con  $\Delta gl = 1$  que resulta estadísticamente significativa ( $p < .05$ ),

y para máxima verosimilitud:

$\Delta\chi^2 = 74.51$ , con  $\Delta gl = 1$  que también resulta significativa ( $p < .005$ ).

Puede afirmarse, por tanto, que en la explicación de los datos, el modelo de dos factores es estadísticamente mejor que el de un factor.

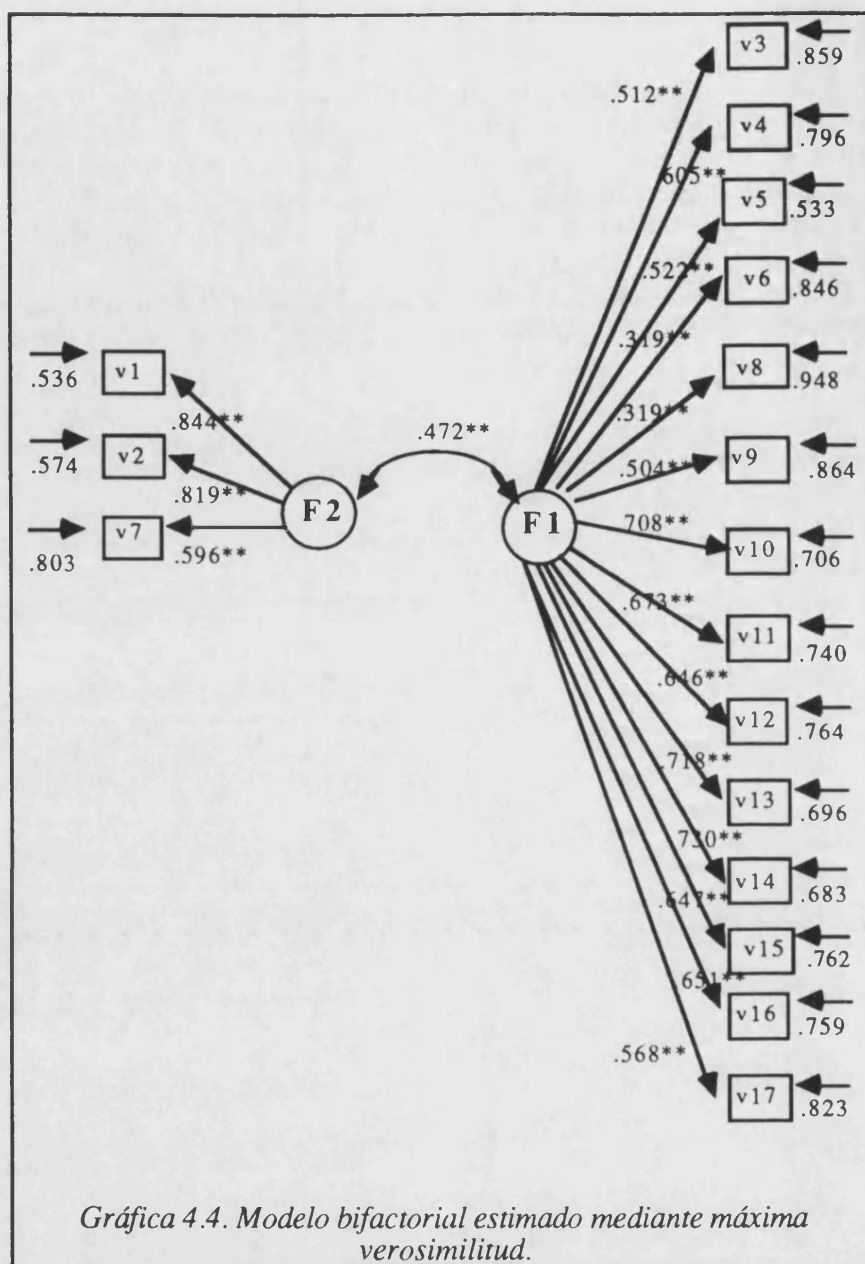
En la gráfica 4.3 se presentan las saturaciones factoriales y los datos de ajuste analítico para la estructura bifactorial en estimación mínimo cuadrática.



La pregunta aquí es: ¿ganamos algo en la explicación de cada uno de los ítems desde los nuevos factores hipotetizados?. Y esta

pregunta se debe contestar desde la comparación de las saturaciones de ambos modelos (unifactorial y bifactorial) estimados por los dos métodos de estimación. La ganancia producida en las saturaciones factoriales se encuentra en el margen de las centésimas. Y aún estas pequeñas ganancias no son en todos los ítems. La explicación de los ítems 7, 8 y 9 decrece en la misma medida. Estas pequeñas diferencias pueden, en conjunto, producir un cierta mejora en el modelo global, que llegue incluso a ser estadísticamente significativa, como hemos visto. Tienen, sin embargo, muy poca importancia práctica.

Comparemos ahora las saturaciones factoriales del modelo de dos factores frente al de un factor en estimación máximo verosímil. La ganancia global que se producía era en este caso mayor que en el de estimación mínimo cuadrática, es de esperar que se produzca un cambio paralelo en el ajuste analítico, gráfica 4.4.



Gráfica 4.4. Modelo bifactorial estimado mediante máxima verosimilitud.

Efectivamente se ha producido un importante cambio en las saturaciones factoriales, pero localizado. Localizado en los tres ítems que forman el factor 2, riesgo tóxico. El ítem 1 ha pasado de una saturación de 0.403 a 0.844; el ítem 2 de 0.522 a 0.819; y el ítem 7 de 0.372 a 0.596. En el resto de ítems las ganancias desde un máximo de 0.03 hasta un mínimo de 0.003. Excepto las ganancias que se producen en el factor dos, los demás ítems tienen un grado de predicción muy similar al conseguido mediante la estructura unifactorial.

La tercera estructura hipotetizada era trifactorial, con un factor uno de riesgos manipulativos, un segundo factor de riesgos tóxicos, y

un tercero de riesgos pasivos. Los datos de ajuste global pueden consultarse en la tabla 4.3. Nuevamente el ajuste no es claramente satisfactorio.

Los datos de ajuste para la estimación por mínimos cuadrados generalizados muestran una ligera mejora con respecto al modelo bifactorial. No obstante esta mejora, la prueba de ji-cuadrado resulta significativa, y la razón de máxima verosimilitud apenas sobrepasa el .5. El CAK y el CSK apenas muestran diferencias, 2.06 por 2.13 y 2.98 por 3.01, respectivamente. Por su parte, y con ligeros cambios, la ji-cuadrado partido los grados de libertad y el NFI, NNFI y CFI indican ajuste satisfactorio.

Muy similar son, también, los datos de ajuste global para la estimación máximo verosímil de la estructura trifactorial frente a los calculados para la bifactorial. Nuevamente sólo la ji-cuadrado partido los grados de libertad da algún apoyo a la plausibilidad del modelo.

*Tabla 4.3. Índices de ajuste global para la estructura trifactorial. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; CAK=Akaike information criterion; CSK=Schwarz information criterion; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: GLS=mínimos cuadrados generalizados; ML=máxima verosimilitud.*

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
TRIFAC	GLS	145.28	116	.034	1.37	.504	1.252	2.06	2.98	.974	.994	.995
TORIAL	ML	257.94	116	<.001	2.43	.296	2.223	3.08	4.04	.703	.773	.806

Se han realizado pruebas de las diferencias de ji-cuadrado para comparar si hay ganancia del modelo de tres factores frente al de dos. Para la estimación mínimo cuadrática tenemos:

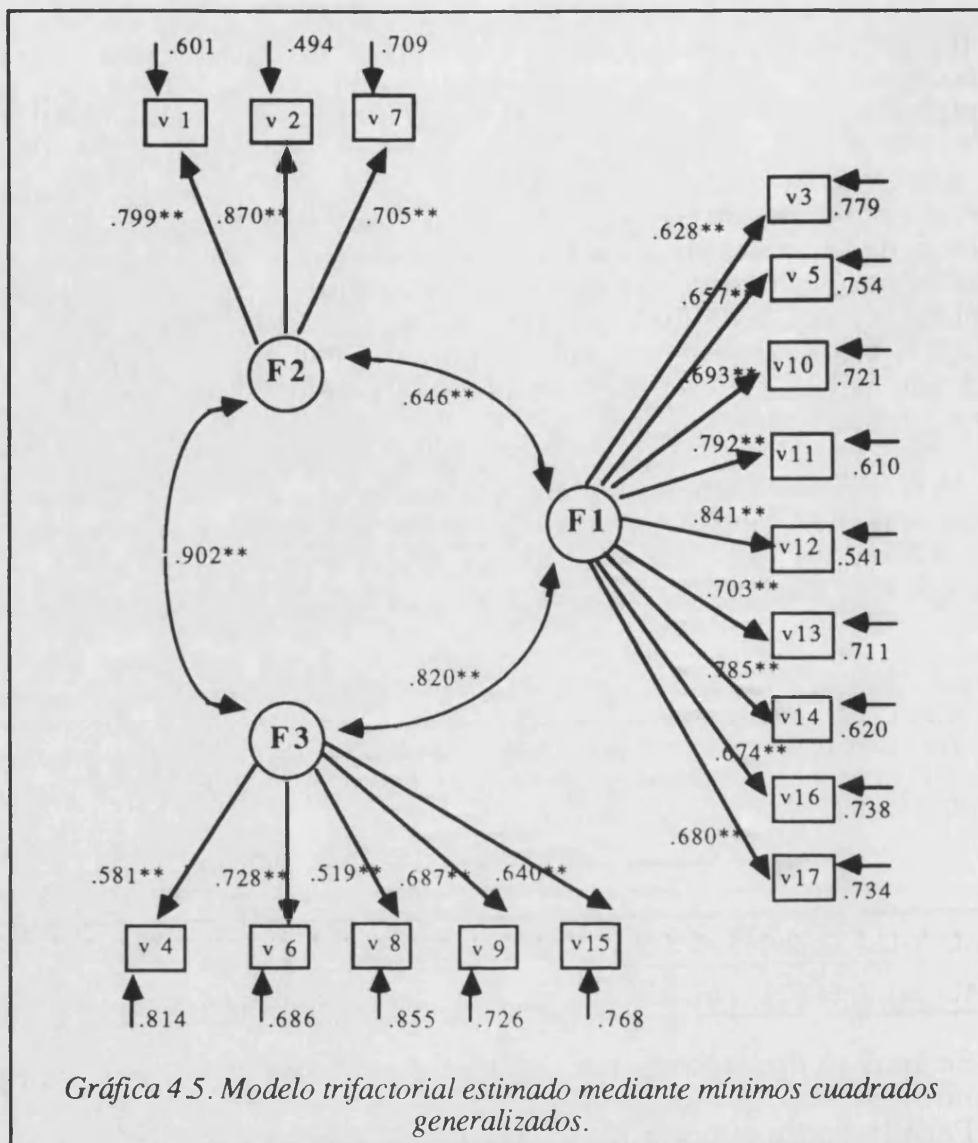
$\Delta\chi^2=11.76$ , con  $\Delta g=2$  que resulta estadísticamente significativa ( $p<.005$ ),

y para máxima verosimilitud:

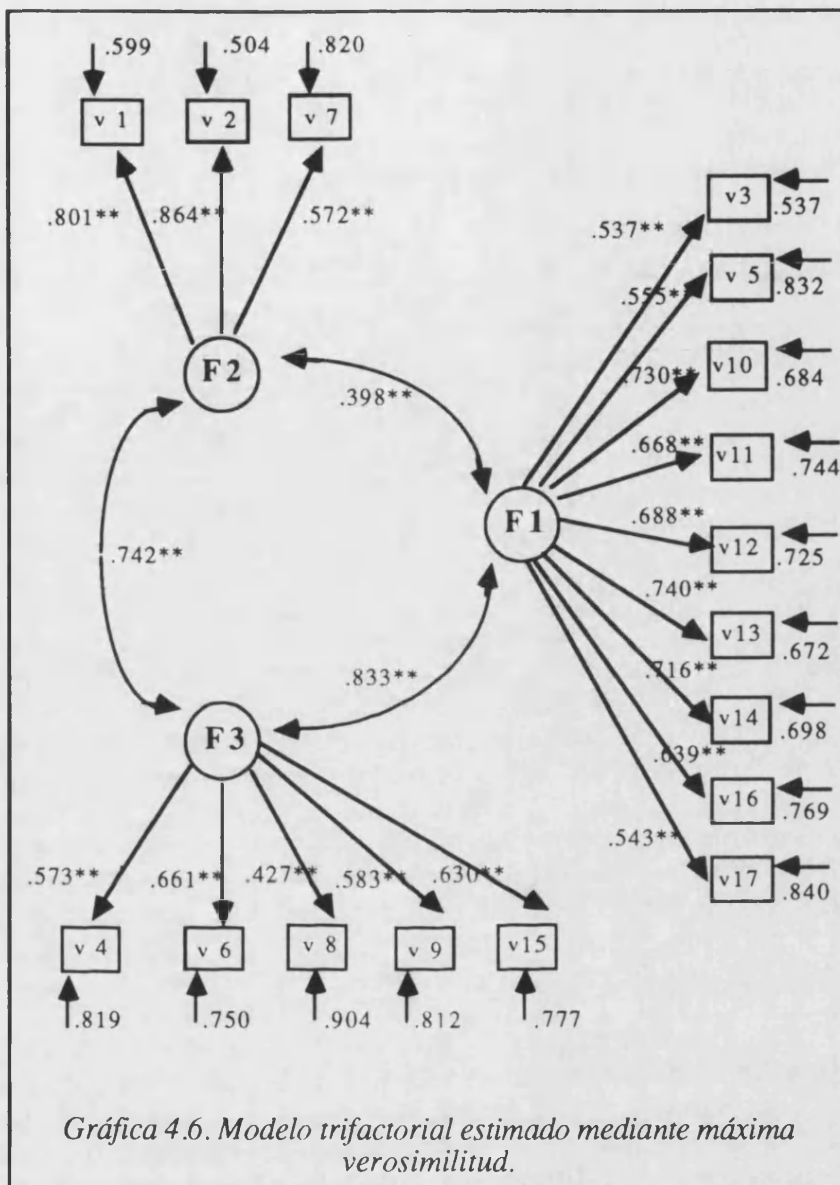
$\Delta\chi^2=27.08$ , con  $\Delta g=2$  que también resulta significativa ( $p<.005$ ).

La prueba aplicada a ambos tipos de estimación arroja los mismos resultados. Existe una ganancia estadísticamente significativa al utilizar una estructura trifactorial para reproducir los datos observados. Ahora hay que observar las ganancias en cuanto al ajuste analítico.





Las saturaciones factoriales presentan, en general, pocas diferencias cuantitativas con respecto al modelo bifactorial. Algunos items ganan hasta 0.1 (v. gr. 6, 7, 8), otros algo menos (v. gr. 1, 2), otros se quedan igual (v.gr. 11, 12), y algunos pierden algo (v. gr. 4, 5). Además, no aparece claramente ningún patrón en estas alzas y bajas. Lo que sí es importante resaltar es que se consigue que cada factor sea capaz de explicar razonablemente bien una parte de la varianza de los items adscritos a ellos. Sólo las saturaciones para los items 4 y 8, especialmente el ocho, pueden considerarse bajas.



En cuanto a las saturaciones factoriales estimadas mediante máxima verosimilitud, puede hacerse un análisis paralelo al realizado para mínimos cuadrados. Aquí las diferencias son todavía más cortas. Únicamente se ha mejorado claramente el ítem 8, que aún así continúa siendo el más bajo. La proporción de varianza que los factores explican de los ítems es, lógicamente, bastante menor al ser las saturaciones factoriales algo más bajas, pero los resultados se mantienen con cierta claridad entre métodos. Se aprecian diferencias en las correlaciones entre factores, más altas en la estimación mínimo cuadrática.

El LM test y el W test se aplicaron a la estructura trifactorial en los dos métodos de estimación, para conocer los cambios que podrían contribuir a un mejor ajuste del modelo.

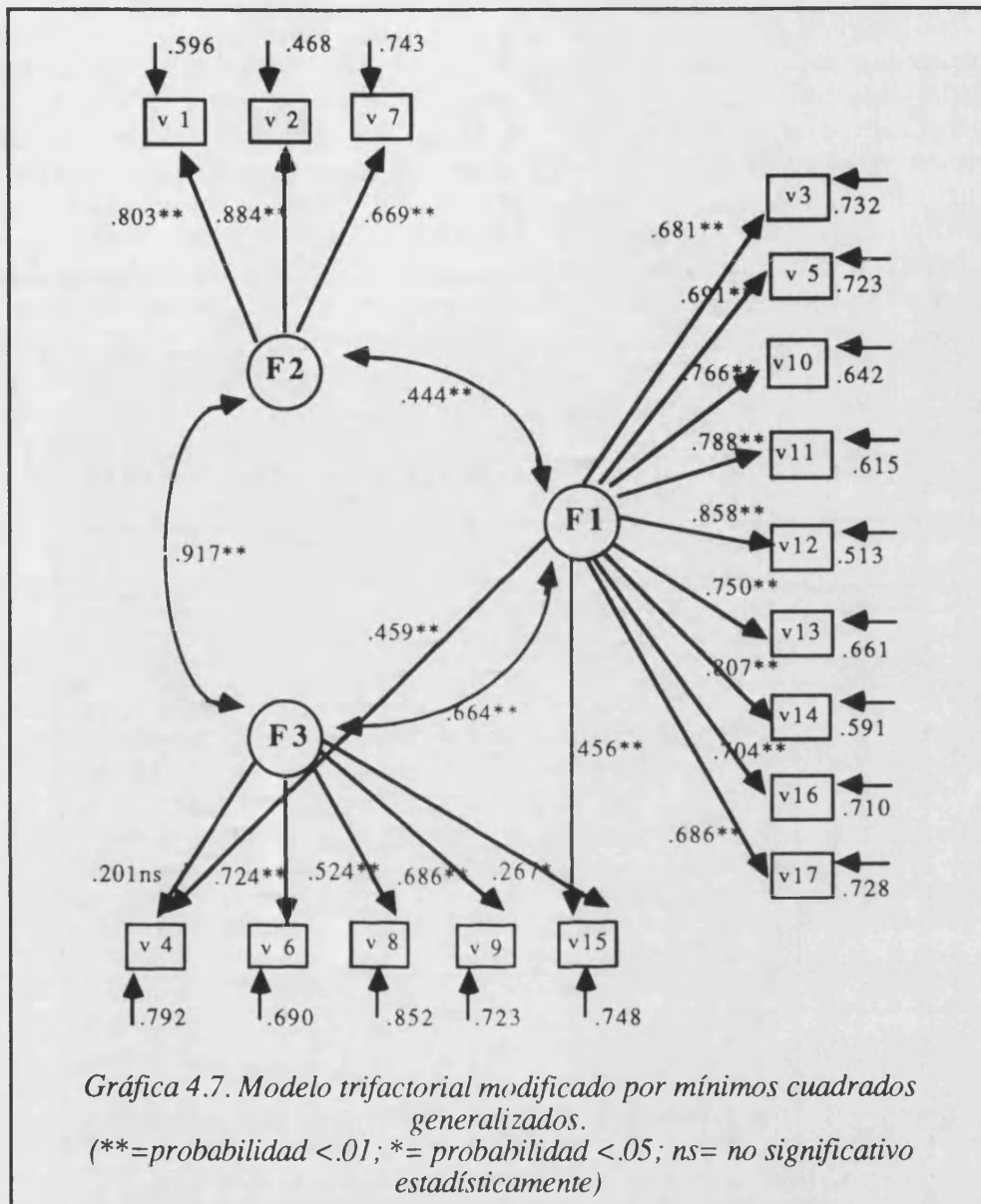


Tabla 4.4. Índices de ajuste global para la estructura trifactorial modificado. Los índices son: *chi*=chi-cuadrado; *g*=grados de libertad del modelo; *p*=probabilidad; *F*=función de ajuste a minimizar; *LHR*=razón de verosimilitud escalada; *CAK*=Akaike information criterion; *CSK*=Schwarz information criterion; *NFI*=normed fit index; *NNFI*=non-normed fit index; *CFI*=comparative fit index. Métodos de estimación: *GLS*=mínimos cuadrados generalizados; *ML*=máxima verosimilitud.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
3 FACTO	GLS	131.02	114	.131	1.23	.539	1.149	1.96	2.93	.977	.996	.997
RES (2)	ML	132.03	106	.044	1.24	.536	1.245	1.97	2.94	.848	.954	.964

El ajuste conseguido mediante mínimos cuadrados generalizados es satisfactorio. Ahora todos los índices, salvo LHR que no es claro, muestran una estructura que predice razonablemente los datos y parsimoniosa. Los índices CAK y CSK han bajado, los índices de comparación con un modelo nulo siguen muy altos, y lo que es más importante, la prueba de ji-cuadrado ha dejado de ser significativa. Los datos de ajuste global también muestran una evidente mejora para el caso de estimación máximo verosímil. A pesar de todo, la prueba de chi-cuadrado todavía resulta significativa, si bien a punto de superar el .05, y el NFI no alcanza el valor de .9 propuesto como indicativo de buen ajuste. Aquí cabe recordar que éste último puede ser anormalmente bajo en comparación con el NNFI y CFI por ser la muestra relativamente pequeña (Anderson y Gerbing, 1984). No cabe aquí realizar pruebas de diferencias entre las ji-cuadrado, dado que los modelos se plantean por sugerencias del LM test y el W test, que incluyen una prueba de diferencias de las chi-cuadrado para evaluar la significatividad estadística de los cambios.

Aunque hemos presentado juntos los datos de ajuste global para los dos tipos de estimación, los modelos a los que dan lugar son bien diferentes, como puede comprobarse fácilmente en la tabla 4.4 observando la diferencia entre los grados de libertad de uno y otro modelo. Veamos primero las estimaciones y el ajuste analítico para la estimación mínimo cuadrática. El LM test sólo proponía aquí dos cambios, referidos siempre a ítems adscritos originalmente al factor 3. Parecía que los ítems 4 y 15 podían ser explicados, además de por el factor 3, por el factor 1. Revisando el contenido de los ítems, y considerando la diferencia teórica entre ambos factores (factor 1, riesgos manipulativos; factor 3, riesgos pasivos) se pensó que era razonable poner en práctica estas hipótesis. Efectivamente, el ítem 4 es de contenido muy general, preguntando si en el trabajo se pueden sufrir atrapamientos, etc, que no tienen porqué ser únicamente pasivos. Lo mismo ocurre, aunque menos claramente para el ítem 15, tiene una parte de contenido que puede hacer referencia a riesgo 'activo'. Una vez calculado el modelo surgido en base a las propuestas del LM test, las estimaciones quedaron como pueden consultarse en la gráfica 4.7.



Aunque se han producido los dos cambios mencionados, lo demás permanece igual, se ha ganado ajuste del modelo y también, como consecuencia, se ha subido algo la cantidad de varianza explicada por los factores. El ítem 4 ha pasado a ser explicado por el factor 1, mientras que el ítem 15 se explica por los factores 1 y 3. No obstante, la estructura trifactorial teórica se mantiene bastante bien. Esta solución se ha alcanzado en una sola iteración, y tanto el LM test como el W test no proponen más cambios.

La situación es muy diferente para la estimación máximo verosímil. Mientras sólo se necesitaba una iteración para llegar a una solución razonable en GLS, se necesitaron hasta 12 modelos para llegar a una solución no tan buena en máxima verosimilitud. Puede decirse que se realizó un verdadero algoritmo exploratorio, en lugar

de realizar propiamente un análisis confirmatorio. No obstante, siempre las indicaciones del LM test y del W test se analizaron bajo el prisma de su substantividad, no empeñándonos en cambios sin sentido. No vamos a presentar el diagrama correspondiente al último modelo prouesto. Los cambios son tan relevantes que parece más adecuado tratarlos uno a uno, y los datos concretos de las estimaciones no ayudarían tampoco demasiado, dado que la explicación por las saturaciones no se han visto sustancialmente incrementadas, sino más bien cambiadas de lugar. Vamos, por contra, a ir listando los cambios ítem a ítem, para comentar exhaustivamente qué sentido teórico presentan los cambios y qué consecuencias prácticas pueden derivarse.

ITEM	CAMBIOS PRODUCIDOS Y SENTIDO TEORICO
Item 1	No se producen cambios en este ítem, sigue adscrito únicamente al factor 2, riesgos tóxicos.
Item 2	No se producen cambios, sigue adscrito al factor dos, riesgos tóxicos.
Item 3	Pasa a estar explicado por los factores 1 y 3. Este ítem se asociaba inicialmente sólo al factor 1. Efectivamente, el contenido del ítem es "manejo máquinas cortantes o peligrosas". Aunque básicamente es un riesgo por manipulación, también en circunstancias el riesgo puede ser pasivo, puede recibirlo el sujeto sin su intervención.
Item 4	Pasa a ser explicado por el factor 1, mientras se hipotetizó en un principio asociado al factor 3. Este cambio ya ocurrió en la estimación por mínimos cuadrados generalizados.
Item 5	Además de estar explicado por el factor 1, como se hipotetizó en un principio, pasa ahora a ser explicado significativamente también por el factor 3. Efectivamente el contenido es general, rastreando la posibilidad de sufrir quemaduras, golpes etc, ya sea por manipulación o recibidos 'pasivamente'.
Item 6	El ítem seis dejaría de formar parte del factor 3 y pasaría a estar explicado por el factor 2. Este cambio puede deberse a que el ítem incluye como riesgo las explosiones o infecciones, típicas de industrias que trabajan con sustancias tóxicas.
Item 7	Sigue siendo explicado, como se hipotetizó, por el factor 2 de riesgos tóxicos.
Item 8	Además de explicarse por el factor 3, como se hipotetizaba, viene ahora explicado en una parte significativa por el factor 2. A la vista del contenido no hay una hipótesis clara del porqué de este comportamiento.

#### Capítulo 4. Medición de los riesgos laborales.

Item 9	Los factores 1 y 2 pasan a explicar ahora este ítem, en lugar del factor 3 hipotetizado en un principio. Los contenidos a que se refiere el ítem pueden considerarse pasivos porque el sujeto los recibe del ambiente, y por ello se hipotetizó como relacionado con el factor 3. Ahora bien, las fuertes presiones de tiempo y el ruido pueden llevar a riesgos manipulativos y el humo puede considerarse un riesgo que tienda a agruparse con los tóxicos.
Item 10	El ítem diez permanece igual, sigue significativamente relacionado con el factor 1.
Item 11	Se relaciona, como se hipotetizó, con el factor 1; pero existe una relación estadísticamente significativa con el factor 2. Se refiere al riesgo de sufrir daños en cara u ojos. Es, por tanto, un ítem de consecuencias de los riesgos y los daños en ojos y cara son típicos de riesgos tóxicos.
Item 12	Además de saturar en el factor 1, como se planteaba por hipótesis, satura en el factor 3. Por ser, al igual que el ítem 11, un ítem de consecuencia o daños del riesgo, esta vez en manos y brazos puede estar relacionado con los tres factores. Aunque los daños de este tipo son típicos de procesos con riesgos manipulativos, nada dice que no se puedan sufrir este mismo tipo de daños por riesgos pasivos.
Item 13	Permanece igual, explicado por el factor 1.
Item 14	El factor 1 explica la mayor parte de la varianza del ítem pero el factor 3 resulta también un predictor estadísticamente significativo. Esto último no se había planteado por hipótesis. Claramente es un ítem general, que cuestiona por el riesgo de daños físico en el puesto, ya sea de forma activa o pasiva. Sólo parece estar más alejado del factor de riesgos tóxicos.
Item 15	Se hipotetizaba como un ítem que saturaría en el factor 3 de riesgos pasivos pero ahora, además satura significativamente en el factor 1. Observando el contenido parece claro que es un ítem de riesgo pasivo, pero como se han producido tantos cambios en ítems que ahora cargan en los factores ya no puede decirse que los factores sean exactamente como se han etiquetado.
Item 16	Se planteaba como un ítem perteneciente al factor 1. No obstante su contenido general haría posible incluirlo en todos los factores. "En mi trabajo deben tomarse estrictas medidas de seguridad" no evalúa qué tipo de riesgos hacen tomar estas medidas, si son riesgos por manipulación, pasivos o tóxicos. Así, el factor 3 es capaz de explicar significativamente una parte de la varianza del ítem.
Item 17	Pasa a estar explicado por el factor 3, frente a la hipótesis que lo asociaba al factor 1. Nuevamente vemos un contenido general, "estoy expuesto a accidentes graves", que podría saturar en cualquier factor.

Además de los datos sobre las saturaciones, para evaluar el problema totalmente cabe observar las correlaciones entre los factores. El factor 1 y el 2 correlacionan 0.420; los factores 2 y 3 0.420; y los factores 1 y 3 0.681. El factor dos es el que presenta menos 'contenidos comunes', es el de estructura más simple y por ello las correlaciones son algo menores. Por el contrario los factores 1 y 3 comparten los contenidos de muchos items y correlacionan más alto. Aún con todo la correlación no es tan alta como la que puede verse para la estimación por mínimos cuadrados generalizados, 0.917.

La vista de esta tabla con los cambios producidos en los items y las correlaciones entre los factores ofrece un panorama complejo. Sólo cuatro items han permanecido sin cambios, no se ha añadido ningún factor en su explicación o cambiado de factor en el que saturar con respecto a la estructura trifactorial simple. De estos cuatro items, tres son los referidos a riesgos tóxicos (1, 2, y 7), que parecen suficientemente delimitados. El resto de items, trece, han cambiado en algo. Los cambios básicamente son para pasar a ser explicados por más de un factor, pero otro cambio muy sintomático es la confusión que aparece entre los factores 1 y 3. Visto con una óptica amplia, ya no puede decirse que tengamos un factor de riesgos manipulativos y uno de riesgos pasivos. Los contenidos de uno y otro están confundidos.

No asumimos sin crítica la idoneidad de la estimación basada en el supuesto de normalidad multivariable, confiando en que el efecto de su incumplimiento sea leve. Por el contrario, ya en capítulos anteriores hemos probado métodos de estimación no basados en este restrictivo supuesto. Los items de un cuestionario difícilmente cumplen con la normalidad univariada y más difícilmente con la multivariada. Nos dispusimos a utilizar el método arbitrario aquí también caso de que las pruebas de normalidad hicieran aconsejable su uso. Se han realizado pruebas de Kolmogorov-Smirnov para cada una de variables, al objeto de probar el ajuste a una distribución normal, condición necesaria para la normalidad multivariable (tabla 4.5).

#### Capítulo 4. Medición de los riesgos laborales.

Tabla 4.5. Prueba de Kolmogorov-Smirnov de una muestra para comprobar el ajuste de los ítems de la escala de riesgo basal a la distribución normal univariada.

ITEM	N muestral	Diferencia máxima	Probabilidad
1	117	.500	.000
2	116	.500	.000
3	118	.500	.000
4	117	.500	.000
5	118	.562	.000
6	117	.500	.000
7	116	.500	.000
8	117	.500	.000
9	120	.500	.000
10	120	.500	.000
11	120	.500	.000
12	119	.572	.000
13	118	.500	.000
14	117	.500	.000
15	119	.500	.000
16	119	.500	.000
17	120	.500	.000

Los resultados de la prueba de bondad de ajuste apuntan a que las variables no cumplen una ley normal. Esto era previsible dado que la escala de respuesta de los ítems es dicotómica y difícilmente pueden distribuirse normalmente. Sin embargo, aún cuando habría sido interesante probar el ajuste de la estructura de tres factores por métodos arbitrarios, el número de parámetros a estimar y el tamaño de la muestra han hecho imposible que el programa encontrara estimaciones con garantías, y por ello no ha llevado a cabo el análisis.

Otra cuestión, al margen del tipo de estimación, es la misma estructura factorial del cuestionario. Ante la dificultad de obtener una representación de los datos factorialmente simple y a la vista que las correlaciones elevadas entre los factores, tanto en GLS como en ML, se pensó en la idoneidad de someter a los datos a una análisis factorial de segundo orden. Los tres factores de primer orden vendrían explicados por un factor de orden superior que representaría el nivel de riesgo basal. La importancia de este modelo radica básicamente en probar la idoneidad o plausibilidad de utilizar los ítems de la escala de riesgo basal como una variable unitaria, dadas también las altas saturaciones factoriales encontradas para el unifactorial probado con anterioridad. Así, aquí lo único importante es obtener una medida de ajuste global. Comentamos esto porque para poder calcular mediante análisis factorial confirmatorio de segundo orden esta estructura es necesario realizar algunas constricciones poco realistas. Así, si el factor de segundo orden presenta únicamente dos indicadores, factores de primer orden, es



necesario realizar una restricción de igualdad de las saturaciones entre los factores de primer orden y el de segundo (Bentler, 1989). Si tenemos, como es nuestro caso, tres factores de primer orden y uno de segundo también es necesario realizar la restricción de igualdad de saturaciones para todos los factores de primer orden en el factor de segundo orden (Rindskopf y Rose, 1988). Hay que entender que estas restricciones se hacen para conseguir que el modelo esté identificado y pueda, al menos, ser calculado. El programa EQS prueba la plausibilidad de cada una de las restricciones, pero aunque sean poso razonables no podemos liberar de ellas al modelo. Así, el programa nos indicó que todas las restricciones realizadas, tanto en mínimos cuadrados como en máxima verosimilitud eran poco razonables. Esto invalida el valor concreto de las saturaciones factoriales de los factores de primer orden en los de segundo, pero no el ajuste global del modelo. Es en este sentido que vamos a utilizar este ajuste global para dar cierto apoyo al modelo de un factor de segundo orden, pero siempre con las comentadas matizaciones y reservas.

Tabla 4.6. Índices de ajuste global de la estructura de segundo orden hipotetizada para la escala de riesgo basal. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: GLS=mínimos cuadrados generalizados; ML=máxima verosimilitud.

		chi	g	p	F	LHR	chi/g	NFI	NNFI	CFI
2°	GLS	139.68	116	.066	1.31	.517	1.20	.975	.995	.996
ORDEN	ML	137.31	106	.022	1.29	.524	1.29	.842	.945	.957

A la vista de los resultados de ajuste global, presentados en la tabla 4.6, parece razonable pensar que un factor de riesgo basal subyace a los tres factores de primer orden hipotetizados por la escala. En este sentido parece también razonable utilizar la escala como medida de este factor de riesgo basal. Esta hipótesis también queda respaldada por las elevadas interrelaciones halladas entre los factores de primer orden a través de los diferentes métodos de estimación.

### 2.3. Fiabilidad y homogeneidad de la escala de riesgo basal y sus ítems

En la tabla 4.7 pueden consultarse diversos coeficientes de consistencia interna de la escala de riesgo basal. El coeficiente alfa de Cronbach muestra una escala altamente homogénea, 0.890.

Tabla 4.7. Coeficientes de consistencia interna de la escala de riesgo basal.

<i>Consistencia interna</i>	
Correlación dos mitades	0.823
Coeficiente Spearman-Brown	0.903
Coeficiente Guttman-Rulon	0.901
Coeficiente alpha total	0.890
Alpha items impares	0.793
Alpha items pares	0.805

Por otro lado, se pueden calcular diversos índices de homogeneidad y consistencia de los ítems, todos ellos se ofrecen en la siguiente tabla. Los índices de homogeneidad corregidos están como mínimo cercanos al .5, salvo los ítems 1, 7 y 8 que son algo más bajos. Por su parte los índices de fiabilidad son adecuados para ítems dicotómicos como los de esta escala. La consistencia interna de la escala, medida mediante el coeficiente alpha, no cambiaría sustancialmente en ningún caso si se eliminara alguno de los ítems.

Tabla 4.8. Media, desviación típica, coeficiente de homogeneidad (correlación ítem-total), fiabilidad, homogeneidad excluyendo el ítem, y coeficiente alpha que quedaría en la escala si se excluyera cada ítem, para los ítems de la escala de riesgo basal.

ITEM	Media	D. T.	Homogeneidad	Fiabilidad	Homogen. sin el ítem	Alpha sin el ítem
1	.178	.382	.494	.189	.429	.887
2	.196	.397	.601	.239	.544	.884
3	.514	.500	.525	.263	.443	.888
4	.308	.462	.627	.290	.562	.883
5	.720	.449	.537	.241	.465	.886
6	.187	.390	.619	.241	.564	.883
7	.243	.429	.454	.195	.379	.889
8	.393	.488	.427	.208	.337	.891
9	.411	.492	.597	.294	.524	.884
10	.617	.486	.670	.326	.607	.881
11	.411	.492	.701	.345	.642	.880
12	.720	.449	.634	.285	.572	.883
13	.636	.481	.699	.336	.641	.880
14	.421	.494	.725	.358	.670	.879
15	.477	.499	.667	.333	.603	.881
16	.467	.499	.661	.330	.596	.882
17	.364	.481	.591	.285	.519	.885

#### 2.4. Validez criterial de los ítems

El riesgo basal, ya se ha comentado, es un constructo que trata de medir el nivel de riesgos intrínsecos al puesto de trabajo; son riesgos propios de la tarea que se realiza. De forma que las relaciones esperadas con las variables de clima, de respuesta de los superiores y de respuesta de los compañeros son nulas o muy bajas. El riesgo



puede, no obstante, tener cierta relación con la actitud del trabajador ante la seguridad ya que esta variable, tomada de Leather, supone una medida de locus de control de los accidentes, y si el nivel de riesgos es muy elevado esto puede afectar al locus del sujeto, que se hace más externo. Puede aparecer una ligera relación negativa entre actitud y riesgo basal, a mayor riesgo basal menor puntuación en la escala de Leather (mayor locus de control externo). El riesgo basal puede estar relacionado con conducta hacia la seguridad, ya que los trabajadores sometidos a mayores riesgos pueden verse impelidos a realizar una conducta más segura. Un alto nivel de riesgos en el lugar de trabajo puede estar relacionado con una mayor tensión laboral. El criterio más importante es, no obstante, el riesgo real. Debe haber una importante relación entre riesgo real y basal. Ambos tipos de riesgo van hasta cierto punto unidos, aunque deben poderse discriminar, lo que será probado también. Por último pueden aparecer relaciones con accidentabilidad, aunque se espera que éstas sean menores que las que aparezcan con riesgo real que intenta medir la probabilidad real de tener accidentes laborales.

Tabla 4.9. Correlaciones de los items de la escala de riesgo basal con los principales criterios. ES=estructura de seguridad; PO=política organizacional de seguridad; F=formación-información; RS=respuesta de los superiores en seguridad; RC=respuesta de los compañeros en seguridad; L=actitud hacia la seguridad; CHS=conducta hacia la seguridad; RR=riesgo real; T=tensión psicológica; y AC= siniestralidad. Símbolos de significatividad estadística de las correlaciones: n.s.= no significativo; #= marginalmente significativo; \*= probabilidad <0.05; y \*\*= probabilidad <0.01, extraídos de tablas (Welkowitz, Ewen y Cohen, 1981).

	ES	PO	F	RS	RC	L	CHS	RR	T	AC
1	-.012ns	-.00ns	-.04ns.	-.01ns	.016ns	-.05ns	.081ns	.175ns	.081ns	.044ns
2	-.081ns	.058ns	-.012ns	.084ns	.073ns	.033ns	.092ns	.247**	-.02ns	.185*
3	-.022ns	.033ns	.124ns	.103ns	.081ns	.094ns	-.01ns	.221*	-.01ns	.195*
4	.119ns	-.213*	-.17ns	-.132ns	-.10ns	.087ns	-.11ns	.299**	.023ns	.226*
5	-.125ns	-.09ns	.011ns	.025ns	-.05ns	.049ns	.011ns	.162#	-.06ns	.045ns
6	.092ns	-.006ns	-.028ns	.056ns	.064ns	.015ns	-.07ns	.385**	.154#	.172#
7	-.064ns	-.134ns	-.144ns	-.02ns	-.08ns	-.11ns	.027ns	.093ns	.005ns	.089ns
8	-.037ns	-.096ns	-.068ns	-.071ns	-.04ns	-.243**	-.09ns	.136ns	.158#	.179#
9	.005ns	-.176#	-.185#	-.129ns	-.09ns	.028ns	-.14ns	.225*	.159#	.224*
10	.007ns	-.210*	-.061ns	-.184#	-.19*	.135ns	-.26**	.341**	.141#	.145ns
11	.039ns	-.156#	-.06ns	-.077ns	-.07ns	.037ns	-.06ns	.361**	-.10ns	.206*
12	.018ns	-.058ns	.045ns	.005ns	-.03ns	.033ns	-.08ns	.241**	-.05ns	.151#
13	.014ns	-.112ns	.011ns	-.029ns	-.04ns	.069ns	-.11ns	.293**	-.01ns	.202*
14	.006ns	-.206*	-.09ns	-.166#	-.06ns	.099ns	-.20*	.547**	.157#	.277**
15	.013ns	-.117ns	-.05ns	-.142ns	-.07ns	.103ns	-.10ns	.371**	.132ns	.152#
16	.100ns	.100ns	.115ns	-.047ns	.07ns	.081ns	-.13ns	.378**	.175#	.254**
17	.059ns	.059ns	.096ns	-.200ns	-.09ns	.085ns	-.16#	.503**	.155#	.155#

En la tabla 4.9 pueden consultarse los valores de correlación entre los items de la escala y los criterios considerados. Tal y como se hipotetizaba, las relaciones de los items de riesgo basal y los tres constructos de clima (estructura de seguridad, formación y política) son nulas en general. El mismo patrón se produce en las relaciones con respuesta de los superiores y respuesta de los compañeros.

Habíamos considerado la posibilidad de que existiera una relación entre nivel de riesgos y evaluación de locus de control externo de los accidentes, actitud negativa de seguridad. No ha ocurrido así, parece que la evaluación del locus de control de los accidentes, medido mediante la escala de Leather no se ve afectado por el nivel de riesgo basal. Únicamente el ítem 8 se encuentra relacionado negativa y significativamente con la actitud. Esto es probablemente debido a la formulación de la pregunta que habla del trabajo como *causa* de un posible accidente. Tampoco parece afectar el nivel de riesgo basal a la conducta segura o insegura que practica el trabajador, únicamente los ítems 10 y 14 alcanzan una relación estadísticamente significativa, y el 17 marginalmente significativa. Las correlaciones de riesgo basal con tensión tampoco alcanzan significatividad estadística; no obstante, sí existen un buen número de relaciones marginalmente significativas ( $p=0.1$ ) en la dirección prevista de mayor tensión cuanto mayor nivel de riesgos. La variedad de causas posibles de la tensión explique quizá esta falta de relación.

Entrando ahora con los criterios más importantes, cabe destacar el buen comportamiento general de los ítems de riesgo basal. Casi todos ellos están significativamente relacionados y fuertemente en algunos casos con riesgo real. Las relaciones con accidentabilidad son, en bastantes casos, estadísticamente significativas, aunque más moderadas que con riesgo real, como era esperado por hipótesis.

## **2.5. Conclusiones sobre la escala de riesgo basal.**

Una vez analizada la estructura factorial, vistas las altas correlaciones entre los factores de primer orden, la confusión de contenidos en los ítems y el alto grado de explicación que consigue el factor de segundo orden sobre los de primero, lo más razonable parece ser utilizar la escala como indicador de un solo factor de riesgo basal, que es la idea con que se construyó. Por otro lado, los resultados de fiabilidad de los ítems, de fiabilidad de la escala y de validez criterial son adecuados.

Una cuestión metodológica a destacar es la diferencia entre métodos de estimación en análisis factorial confirmatorio. Siempre habíamos encontrado en nuestra experiencia ciertas diferencias entre la máxima verosimilitud y mínimos cuadrados. En concreto, la estimación máximo verosímil ofrece generalmente unas estimaciones de las relaciones ligeramente inferiores, redundando ésto en un ligeramente menor ajuste global, apreciable en los índices. No obstante, ésto no había provocado hasta aquí fuertes diferencias en las estructuras que subyacían a ambos métodos. Sin embargo, para esta escala las diferencias entre métodos han sido fuertes en sus efectos, de forma que una estructura trifactorial aparece relativamente clara con mínimos cuadrados generalizados, mientras

que los menores valores de explicación de máxima verosimilitud provocan multitud de cambios para encontrar un ajuste razonable. De este modo, las aparentes despreciables diferencias entre métodos, que se despachan siempre con un 'asintóticamente son iguales', pueden provocar modelos sustancialmente diferentes.

### 3. Riesgo Real

#### 3.1. Medida

El cuestionario de riesgo real está compuesta por siete items con escala de respuesta tipo Likert de cinco puntos. Una alta puntuación en la escala se corresponde con un elevado nivel de riesgo aún después de tomar todas las medidas para controlarlo.

##### ESCALA:

1. Dada su experiencia y las medidas de seguridad que Ud. toma ¿Qué riesgo real tiene Ud. de sufrir un accidente laboral?  
0      1      2      3      4      5  
Ninguno Muy Poco    Alguno    Bastante    Mucho    Extremado
2. Considerando su formación, su experiencia y el modo en que Ud. trabaja realmente ¿Qué riesgo real existe de que Ud. sufra un accidente mortal o muy grave?  
0      1      2      3      4      5  
Ninguno Muy Poco    Alguno    Bastante    Mucho    Extremado
3. En su trabajo, ¿Qué riesgo cree Ud. que tiene realmente de sufrir daño o lesiones en sus ojos o cara?  
0      1      2      3      4      5  
Ninguno Muy Poco    Alguno    Bastante    Mucho    Extremado
4. ¿Qué riesgo cree Ud. que tiene realmente de sufrir un accidente laboral que afecte a sus manos o brazos?  
0      1      2      3      4      5  
Ninguno Muy Poco    Alguno    Bastante    Mucho    Extremado
5. Teniendo en cuenta sus hábitos de trabajo ¿en qué medida cree Ud. que está protegido realmente de un posible accidente?  
5      4      3      2      1      0  
Nada    Poco    Algo    Bastante    Muy    Extremadamente
6. Realmente, ¿Qué riesgo cree Ud. que tiene de padecer una accidente leve?  
0      1      2      3      4      5  
Ninguno Muy Poco    Alguno    Bastante    Mucho    Extremado
7. Considerando como trabaja Ud. y las condiciones de su lugar de trabajo, ¿A cuanto "peligro de accidente" cree Ud. que está expuesto realmente?  
0      1      2      3      4      5  
Ninguno Muy Poco    Alguno    Bastante    Mucho    Extremado

#### 3.2. Estructura factorial confirmatoria

La estructura hipotetizada para la escala es unifactorial. La escala mide un único factor de riesgo real. En este caso no se han probado estructuras alternativas, dado que no habían razones teóricas para plantearlas. Los contenidos son, al menos teóricamente, homogéneos. En la tabla 4.10 se presentan los índices de ajuste global para la escala de riesgo real mediante los métodos de estimación basados en

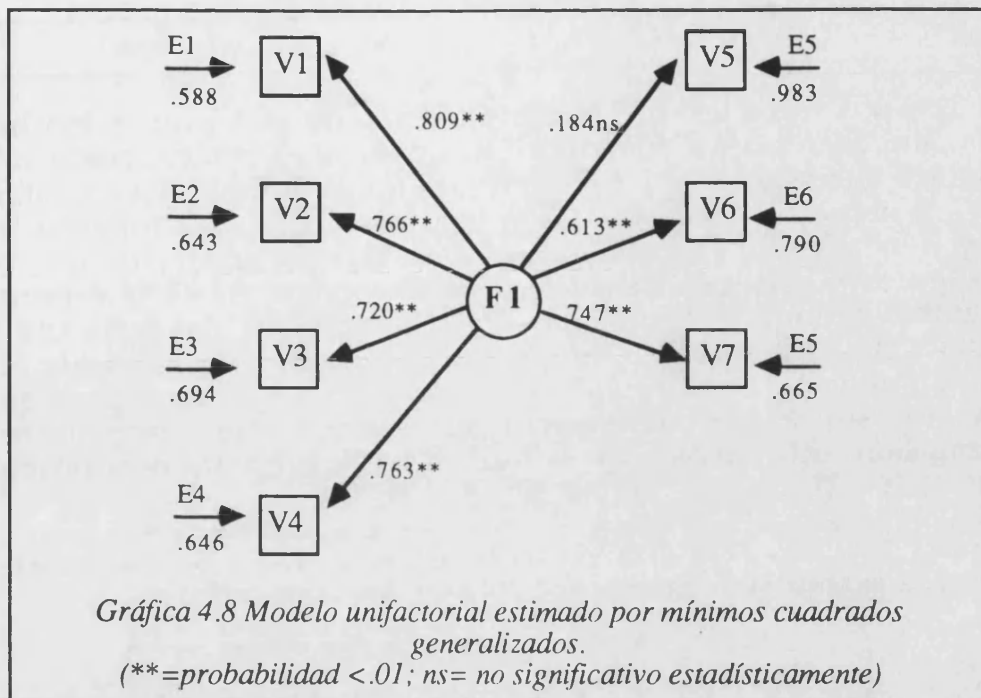
el supuesto de normalidad multivariable, mínimos cuadrados generalizados y máxima verosimilitud.

*Tabla 4.10. Índices de ajuste global para la estructura unifactorial hipotetizada para la escala de riesgo real. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; CAK=Akaike information criterion; CSK=Schwarz information criterion; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: GLS=mínimos cuadrados generalizados; ML=máxima verosimilitud.*

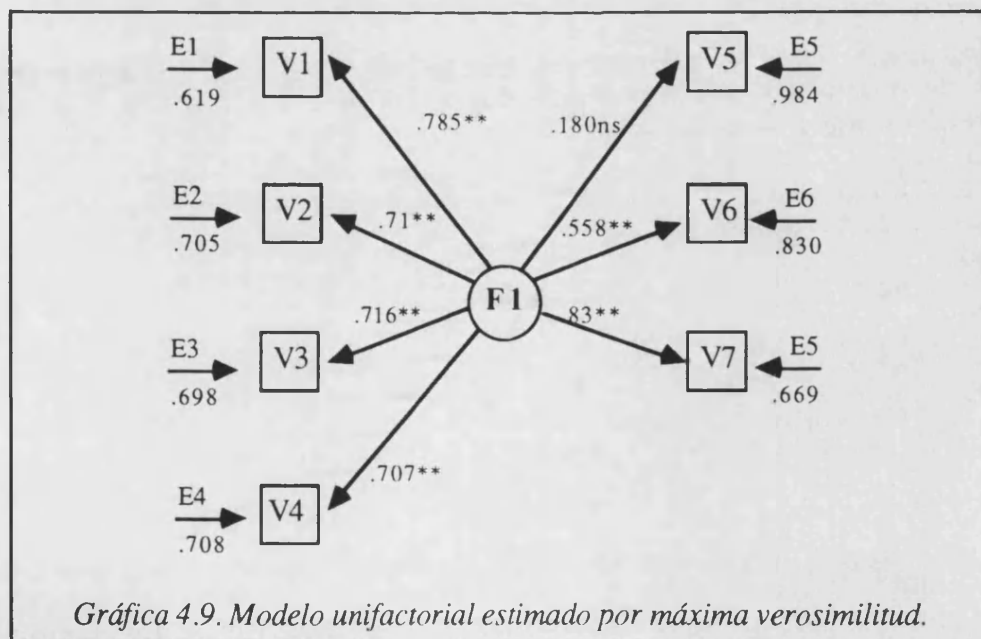
		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
UN	GLS	23.67	14	.0501	.202	.903	1.69	.439	.768	.986	.991	.994
FACTOR	ML	30.13	14	.0073	.257	.879	2.15	.494	.823	.898	.912	.941

Para el método mínimo cuadrático, el ajuste de la prueba ji-cuadrado muestra que el modelo ajusta a los datos observados ( $p=.0501$ ). Por su parte, LHR ofrece un ajuste muy satisfactorio por su cercanía a 1, ajuste perfecto. La prueba de chi-cuadrado partido los grados de libertad muestra que la solución de un factor es adecuada para explicar las relaciones entre las variables y altamente parsimoniosa. Por su parte CAK y CSK son bajos, indicando también buen ajuste. Por último, los índices de comparación con un modelo nulo están en valores elevados, muy por encima de .9.

En máxima verosimilitud los resultados son similares, pero con todos los índices mostrando un ajuste ligeramente menor. Así la prueba de ji-cuadrado gana cerca de 7 puntos que la lleva a ser estadísticamente significativa. No obstante, LHR es muy elevado, cercano al uno. Por su parte, chi-cuadrado partido los grados de libertad muestra una estructura altamente parsimoniosa y con buen ajuste a los datos. CAK y CSK son muy similares a los calculados para mínimos cuadrados generalizados. NNFI y CFI muestra un buen ajuste, dado que están por encima de .9. NFI está ligeramente por debajo de .9, pero hay que recordar que en muestras relativamente bajas, como la que ahora tenemos, puede ser anormalmente bajo. Como conclusión general, también por máxima verosimilitud puede considerarse que el modelo unifactorial ajusta razonablemente a los datos.



En la gráfica 4.8 puede consultarse el diagrama causal para la estimación mínimo cuadrática. Todos los ítems pueden ser explicados significativamente desde el factor de riesgo real, a excepción del ítem 5, cuya saturación es claramente deficiente. El mismo patrón de relaciones se presenta en la gráfica 4.9, esta vez para la estimación máximo verosímil. Los valores para las saturaciones factoriales son muy similares, aunque en general ligerísimamente menores. Nuevamente sólo el ítem 5 presenta una saturación factorial muy baja.



(\*\*=probabilidad <.01; ns= no significativo estadísticamente)

Los dos tipos de estimación presentados se basan, como hemos visto, en el supuesto de distribución normal multivariada de las variables observables. Si este supuesto no se cumple, y aún sabiendo que los métodos son robustos a la violación de la asunción, sería razonable calcular estimación mediante método arbitrario, que no se basa en este supuesto restrictivo. Para conocer hasta qué punto es razonable asumir la normalidad multivariable, se han realizado siete pruebas de Kolmogorov-Smirnov sobre los ítems de la escala. Dado que la condición de normalidad univariada es necesaria para que se cumpla el supuesto de normalidad multivariable puede ofrecernos una información muy valiosa. En la tabla 4.11 se presenta esta prueba de ajuste a la distribución normal de las variables.

Tabla 4.11. Prueba de Kolmogorov-Smirnov de una muestra para comprobar el ajuste de los ítems de la escala de riesgo real a la distribución normal univariada.

ITEM	N muestral	Diferencia máxima	Probabilidad
1	124	.801	.000
2	121	.552	.000
3	122	.579	.000
4	122	.751	.000
5	122	.879	.000
6	122	.800	.000
7	124	.801	.000

Los resultados de la prueba de bondad de ajuste apuntan a que las variables no cumplen una ley normal una a una, con lo que no pueden cumplir una distribución normal multivariable.

Además de esta prueba de ajuste, el programa EQS ofrece los datos de curtosis y asimetría para las variables observables, datos que pueden verse en la tabla 4.12.

Tabla 4.12. Datos de asimetría y curtosis para los ítems de la escala de riesgo real.

	Asimetría	Curtosis
Item 1	0.8013	0.7921
Item 2	1.1052	1.0967
Item 3	0.7718	-0.1576
Item 4	0.1792	-0.6665
Item 5	0.1387	-0.2920
Item 6	0.3826	-0.6022
Item 7	0.5889	-0.2313

El coeficiente de Mardia multivariado, que ofrece el programa EQS, puede también utilizarse como una medida de la desviación de la normalidad multivariable. Junto con este coeficiente se ofrece el



valor esperado del mismo si las variables cumplieren una ley multinomial. Si se alejan mucho ambos valores se tiene un dato a favor de la inadecuación del supuesto.

COEFICIENTE DE MARDIA MULTIVARIADO= 16.0714  
 Coeficiente esperado normal multivariada= 7.7764

A la vista de los resultados descriptivos en torno a la normalidad de las variables, y aún sabiendo que los métodos antes empleados son robustos a violaciones de los supuestos, parece adecuada la aplicación de métodos arbitrarios. Los resultados de ajuste global pueden verse en la tabla 4.13.

*Tabla 4.13. Índices de ajuste global para la estructura unifactorial hipotetizada para la escala de riesgo real estimada por método arbitrario. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; CAK=Akaike information criterion; CSK=Schwarz information criterion; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index.*

	chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
1 FACTOR AGLS	37.989	14	<.001	.32	.849	2.71	.562	.891	.588	.494	.663

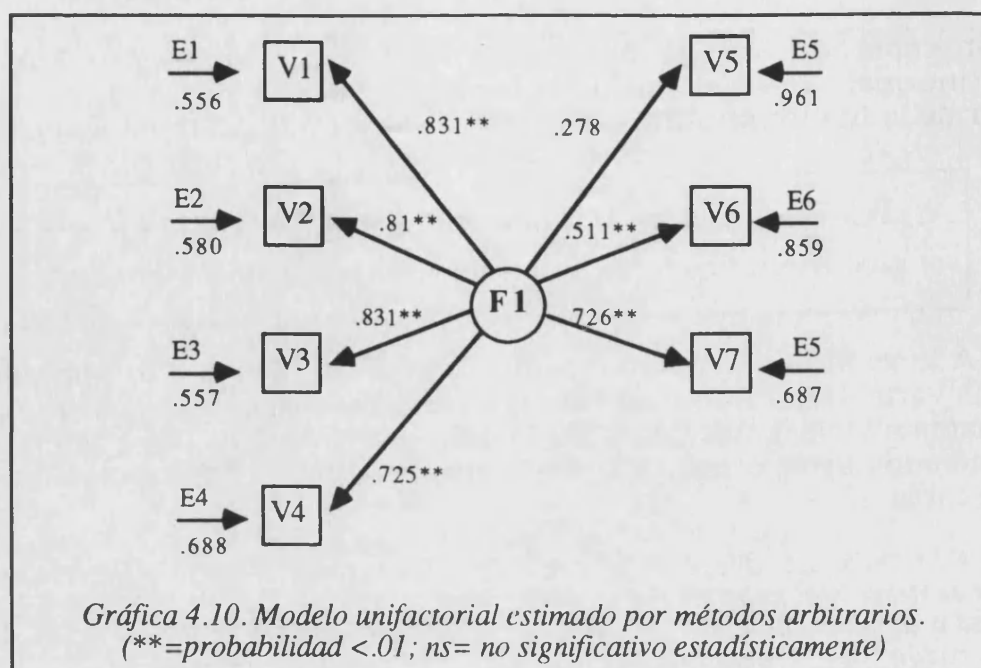
INDICE DE BONDAD DE AJUSTE, GFI= .902  
 Índice de bondad de ajuste ajustado, AGFI= .804

El ajuste global no es tan claro como era en mínimos cuadrados o máxima verosimilitud. Sólo LHR, chi/gl, GFI y AGFI, CAK y CSK muestran un ajuste adecuado del modelo; mientras que la prueba de chi-cuadrado, y los índices de ajuste que comparan con un modelo nulo muestran un inadecuado ajuste a los datos.

Hay que recordar aquí que, según Tanaka (1986), mientras las estimaciones de los parámetros son muy similares por GLS, ML y métodos arbitrarios, los niveles de probabilidad asociados con las distintas pruebas de ajuste global o analítico sí sufren cambios considerables.

En cuanto al ajuste analítico, en la gráfica 4.10 pueden verse las estimaciones de los parámetros junto con las pruebas z de contraste de hipótesis sobre las saturaciones factoriales.





Los resultados de las estimaciones de los parámetros son muy similares a las encontradas por los otros dos métodos, y no alteran las conclusiones de los contrastes de hipótesis. Nuevamente es el ítem 5 el que presenta graves problemas de explicación desde el factor de riesgo real.

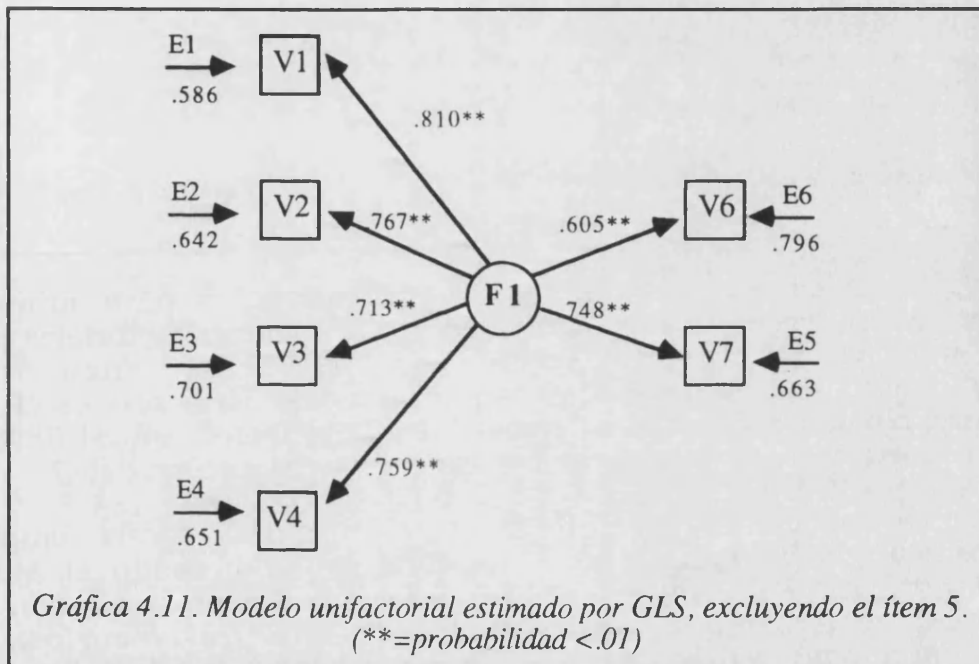
A través de métodos de estimación, la conclusión más relevante que puede extraerse es la inadecuación del ítem 5 para medir el constructo que trata de medir la escala. Este ítem no es fiable para medir el riesgo real. Parecía adecuado, por tanto, eliminar este ítem y recalculer los análisis factoriales confirmatorios.

Los resultados de ajuste global, para los métodos de estimación de máxima verosimilitud y mínimos cuadrados generalizados, de la escala de riesgo real excluyendo el ítem 5 pueden consultarse en la tabla 4.14.

Tabla 4.14. Índices de ajuste global para la estructura unifactorial hipotetizada para la escala de riesgo real excluyendo el ítem 5. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; CAK=Akaike information criterion; CSK=Schwarz information criterion; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: GLS=mínimos cuadrados generalizados; ML=máxima verosimilitud.

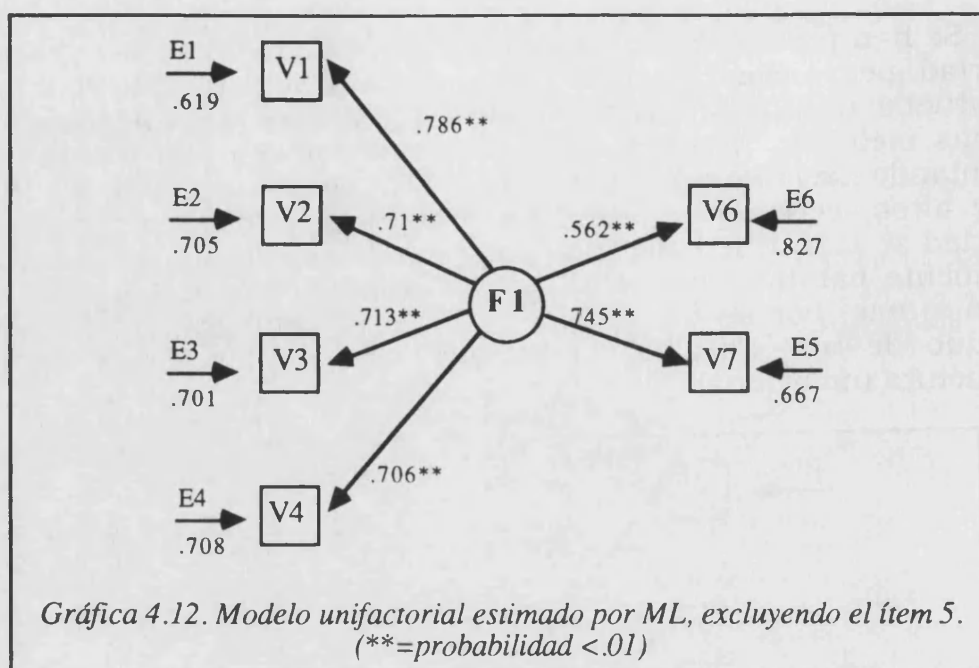
		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
UN	GLS	20.1	9	.017	.17	.917	2.23	.375	.657	.987	.988	.993
FACTOR	ML	25.94	9	.002	.22	.895	2.88	.425	.706	.910	.897	.938

Se han producido, como es lógico, una pérdida en los grados de libertad, por la eliminación de una variable observable. De esta forma, la prueba de chi-cuadrado alcanza resultados significativos para ambos métodos. No obstante, los demás índices son coincidentes, apuntando hacia un adecuado ajuste. LHR se encuentra en resultados muy altos, cercanos al uno. Chi-cuadrado partido los grados de libertad se mantiene bajo, indicando que la estructura unifactorial es altamente parsimoniosa. CAK y CSK también son bajos, indicando buen ajuste. Por su parte, NFI, NNFI, y CFI son coincidentes en el sentido de una adecuada representación de los datos desde la estructura unifactorial.



Gráfica 4.11. Modelo unifactorial estimado por GLS, excluyendo el ítem 5. (\*\*=probabilidad <.01)

Los datos de estimación de los parámetros para GLS pueden consultarse en la gráfica 4.11. Cabe destacar en los resultados de las saturaciones factoriales que, al eliminarse el ítem 5, la explicación de los ítems desde el factor es ahora plenamente adecuada. Algunos ítems elevan ligeramente sus saturaciones y otros lo bajan, también ligeramente. Pero en resumen, cambios apenas relevantes en cuanto a proporción de varianza explicada.



En niveles muy similares a los presentados para mínimos cuadrados generalizados se mantienen las saturaciones factoriales para máxima verosimilitud. Todos los ítems se encuentran significativamente relacionados con el factor. El ítem seis es el que viene explicado más deficientemente desde el factor, un 31.58% de la varianza. El resto de saturaciones se sitúan por encima del .7

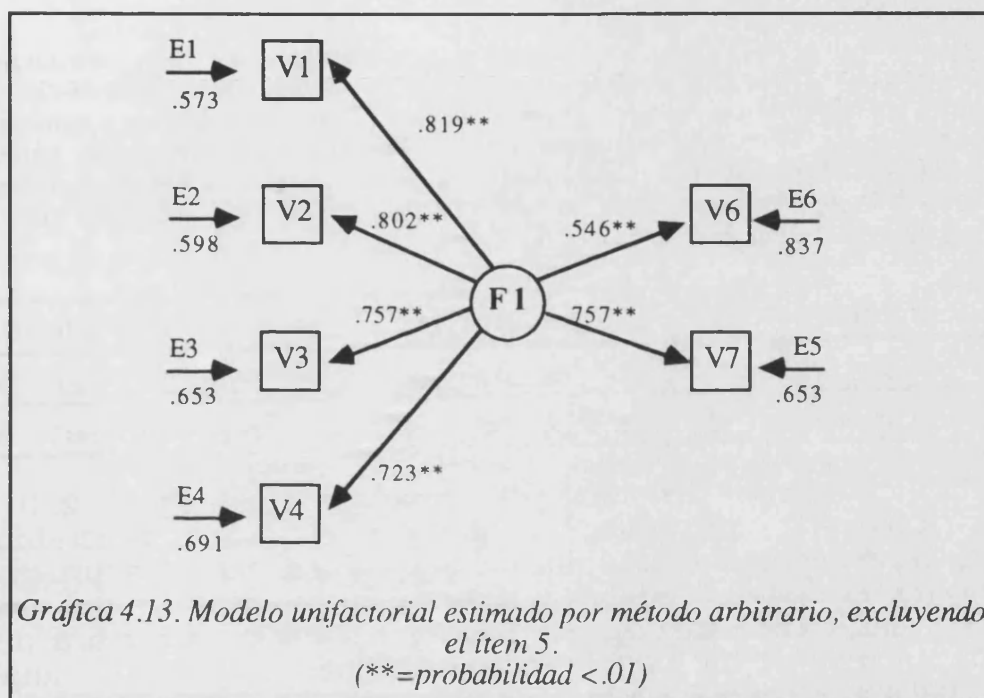
Al igual que cuando se probó la estructura de la escala completa ahora, con la versión de seis, también se ha probado el ajuste mediante métodos arbitrarios. Las variables, excepto una, son las mismas, de forma que los resultados de asimetría y curtosis se mantienen y no se hace necesario repetirlos aquí. En la tabla 4.15 pueden consultarse los resultados de ajuste global.

Tabla 4.15. Índices de ajuste global para la estructura unifactorial hipotetizada para la escala de riesgo real (excluido el ítem 5) estimada por método arbitrario. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; CAK=Akaike information criterion; CSK=Schwarz information criterion; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: AGLS, método arbitrario.

	chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
1 FACTOR AGLS	27.41	9	.001	.23	.889	3.045	.437	.719	.607	.439	.663

INDICE DE BONDAD DE AJUSTE, GFI= .902  
Indice de bondad de ajuste ajustado, AGFI= .770

Los índices para la estimación por métodos arbitrarios se mantiene en los mismos términos que los encontrados para la escala de siete ítems en este tipo de estimación. Chi-cuadrado, y los índices de comparación con un modelo nulo muestran ajuste inadecuado; el resto de índices apuntan hacia un buen ajuste y una representación parsimoniosa. Ante esta situación, es especialmente importante estudiar las relaciones analíticas, los parámetros a estimar.



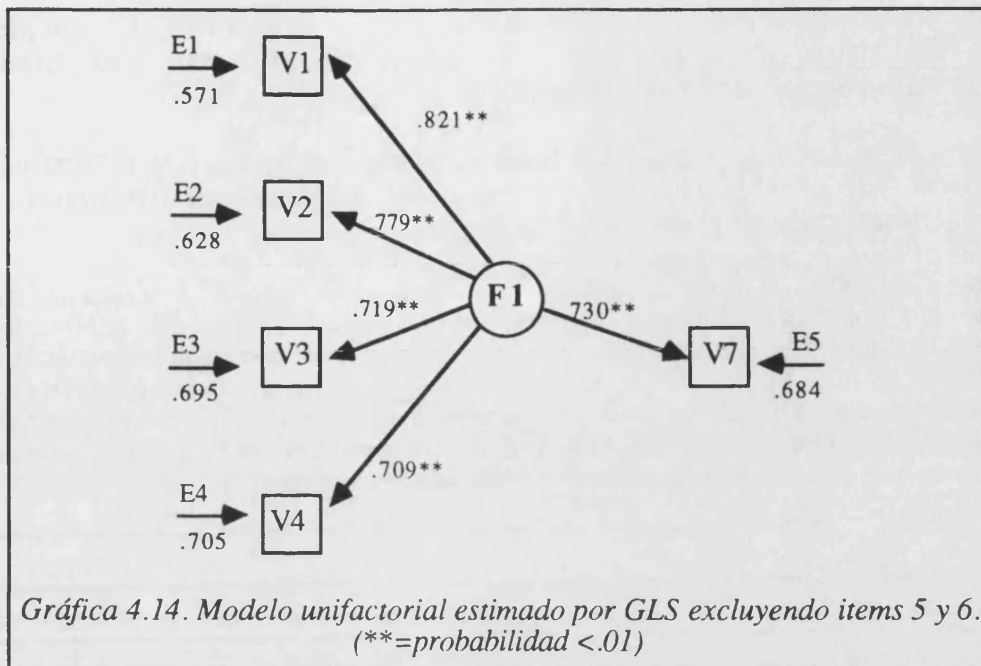
Como puede verse en la gráfica 4.13 todos los ítems se muestran significativa y altamente relacionados con el factor de riesgo real. Los valores de las saturaciones y de las varianzas no explicadas, errores, son muy similares a los hallados para los métodos basados en la distribución multivariable. La saturación más baja se corresponde con el ítem 6, al igual que ocurría en los otros tipos de estimación. Para este ítem el porcentaje de varianza explicada, la fiabilidad, se sitúa en un 29.8%. Con el ítem cinco, los resultados del análisis factorial confirmatorio señalaban claramente que no era adecuado como indicador del riesgo real. Ahora, con el ítem 6 cabe preguntarse si es un ítem teóricamente importante o no; esto es, si se perdería en la validez de contenido de la escala por eliminar el ítem o su contenido se recoge suficientemente bien con el resto de ítems. El ítem seis recoge la probabilidad de accidente leve, añade la noción de gravedad. No obstante éste contenido ya se recoge directamente o indirectamente en el resto de la escala, que por otro lado es bastante homogénea. Parece pues que a la vista de su deficiente funcionamiento es razonable eliminarlo para reducir aún más la escala. No obstante esta decisión no es definitiva, y debe corroborarse mediante los análisis de validez criterial.

Se han vuelto a estimar tres análisis factoriales confirmatorios, uno para estimación mínimo cuadrática, otro para máximo verosímil y un último para métodos arbitrarios. Son los análisis factoriales resultantes de hipotetizar una estructura unifactorial a los cinco items que restan del cuestionario de riesgo real una vez eliminados los items 5 y 6. En la siguiente tabla pueden consultarse los datos de ajuste global para los dos métodos de estimación basados en el supuesto de distribución normal multivariable.

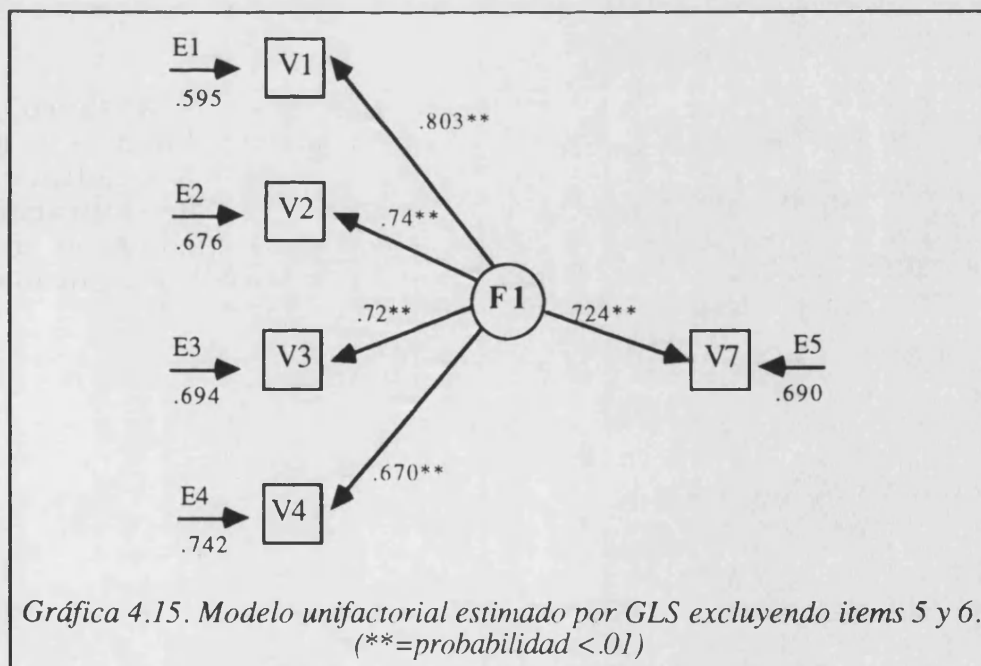
Tabla 4.16. Índices de ajuste global para la estructura unifactorial hipotetizada para la escala de riesgo real excluyendo los items 5 y 6. Los índices son: *chi*=chi-cuadrado; *g*=grados de libertad del modelo; *p*=probabilidad; *F*=función de ajuste a minimizar; *LHR*=razón de verosimilitud escalada; *CAK*=Akaike information criterion; *CSK*=Schwarz information criterion; *NFI*=normed fit index; *NNFI*=non-normed fit index; *CFI*=comparative fit index. Métodos de estimación: *GLS*=mínimos cuadrados generalizados; *ML*=máxima verosimilitud.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
UN	GLS	12.459	5	.0290	.106	.948	2.491	.275	.510	.991	.989	.994
FACTOR	ML	12.186	5	.0323	.104	.949	2.437	.273	.508	.949	.937	.969

El ajuste global para la escala formada por los items 1, 2, 3, 4, y 7 de riesgo real son, tanto para mínimos cuadrados generalizados como para máxima verosimilitud, aceptables. Salvo la prueba ji-cuadrado, el resto de índices muestran un buen ajuste del modelo unifactorial a los datos, y alta parsimonia de la estructura de un factor. Incluso se ha ganado en ajuste global frente a los análisis factoriales realizados para la escala con seis items. Es relevante, por otro lado, comprobar el grado de ajuste analítico, y la proporción de varianza de los items que puede predecirse desde los factores. El objetivo es generar una escala homogénea, que mida el rasgo riesgo real.



Como puede verse en la gráfica 4.14. las saturaciones son, ahora sí, elevadas en todos los items. Se consiguen, como mínimo, porcentajes de varianza explicada como mínimo del 50%, y hasta del 67%. Parece que de esta forma, reduciendo el número de items, se garantiza la homogeneidad de la escala sin perder validez de contenido.



Otro tanto ocurre con el método de máxima verosimilitud (gráfica 4.15). Con la eliminación de los dos items, la capacidad



predictiva del factor sobre los ítems se ve aumentada. La escala es, por lo tanto, más homogénea. Se pasa siempre del 50% de explicación de la varianza en cada ítem.

Para la escala que queda tras eliminar los ítems 5 y 6 también se ha probado la estimación mediante métodos arbitrarios. Los resultados pueden verse a continuación.

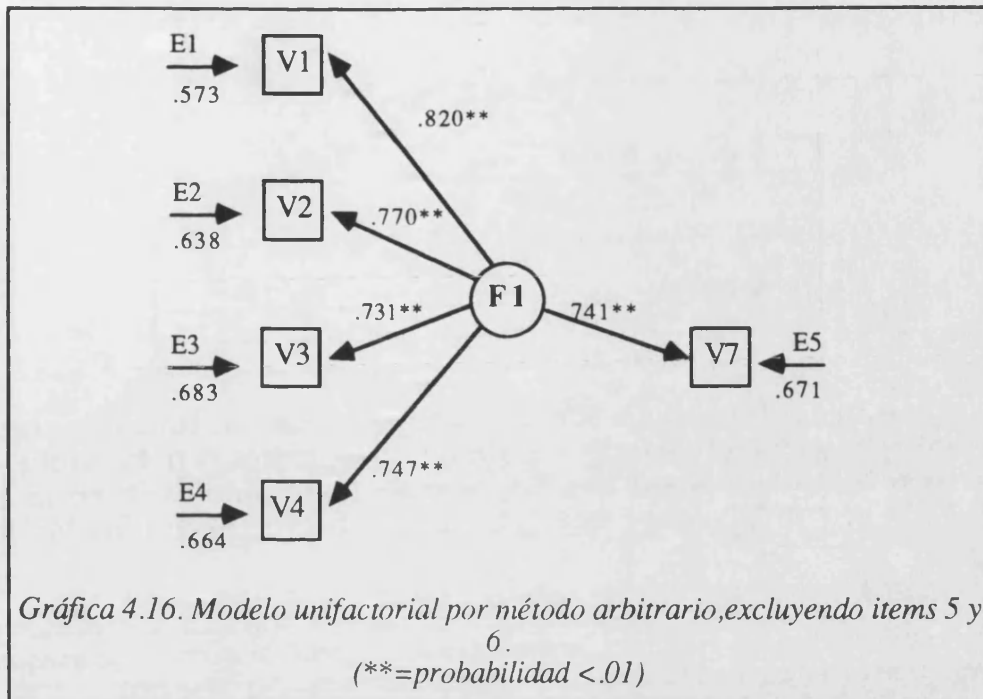
Tabla 4.17. Índices de ajuste global para la estructura unifactorial hipotetizada para la escala de riesgo real (excluidos los ítems 5 y 6) estimada por método arbitrario. Los índices son: *chi*=chi-cuadrado; *g*=grados de libertad del modelo; *p*=probabilidad; *F*=función de ajuste a minimizar; *LHR*=razón de verosimilitud escalada; *CAK*=Akaike information criterion; *CSK*=Schwarz information criterion; *NFI*=normed fit index; *NNFI*=non-normed fit index; *CFI*=comparative fit index. Métodos de estimación: *GLS*=mínimos cuadrados generalizados; *ML*=máxima verosimilitud.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
1 FACTOR	AGLS	16.71	5	.005	.14	.931	3.34	.346	.628	.716	.521	.761

INDICE DE BONDAD DE AJUSTE, GFI= .928  
 Índice de bondad de ajuste ajustado, AGFI= .783

Los índices son contradictorios. Mientras chi-cuadrado, NFI, NNFI y CFI indican ajuste inadecuado, LHR, CAK, CSK y chi/g GFI y AGFI indican un adecuado ajuste a los datos y una estructura parsimoniosa.

Relevante es, también, observar los datos de ajuste analítico, que se ofrecen en el siguiente diagrama causal. Los resultados son muy similares a los ya encontrados mediante mínimos cuadrados y máxima verosimilitud. Para valores tan altos de las saturaciones factoriales, la evaluación de la significatividad estadística no se ve, lógicamente, alterada. El porcentaje mínimo de varianza explicada de cada ítem es aquí del 53.43%.



### 3.3. Fiabilidad y homogeneidad de la escala y los ítems

Los coeficientes de fiabilidad de la escala pueden consultarse en la tabla adjunta. El alfa de Cronbach se sitúa en un .82, que puede considerarse muy adecuado. Los resultados para los ítems impares y pares son más bajos, lo que es de esperar dado que se han calculado en base a pocos ítems, 4 y 3 respectivamente.

Tabla 4.18. Coeficientes de consistencia interna de la escala de riesgo real.

<i>Consistencia interna</i>	
Correlación dos mitades	0.746
Coefficiente Spearman-Brown	0.855
Coefficiente Guttman-Rulon	0.854
Coefficiente alpha total	0.822
Alpha ítems impares	0.685
Alpha ítems pares	0.675

También se ha calculado la consistencia interna de la escala una vez excluidos los ítems 5 y 6. Era de esperar que se ganara algo de consistencia, como así ha sucedido para todos los coeficientes. El alfa para los ítems pares es mucho más bajo que el resto, por calcularse en base a dos ítems.



Tabla 4.19. Coeficientes de consistencia interna de la escala de riesgo real, una vez eliminados los ítems 5 y 6.

<b>Consistencia interna</b>	
Correlación dos mitades	.788
Coefficiente Spearman-Brown	.881
Coefficiente Guttman-Rulon	.849
Coefficiente alpha total	.845
Alpha ítems impares	.775
Alpha ítems pares	.540

Se han calculado, como viene siendo habitual, algunos estadísticos descriptivos de los ítems. Se ofrecen en la tabla 4.20 para la escala compuesta por los siete ítems originales, y en la tabla 4.21 para la escala compuesta por cinco, una vez excluidos los ítems 5 y 6.

Tabla 4.20 Media, desviación típica, coeficiente de homogeneidad (correlación ítem-total), fiabilidad, homogeneidad excluyendo el ítem, y coeficiente alpha que quedaría en la escala si se excluyera cada ítem, para los ítems de la escala de riesgo real.

ITEM	Media	D. T.	Homogeneidad	Fiabilidad	Homogen. sin el ítem	Alpha sin el ítem
1	1.947	1.067	.775	.827	.688	.782
2	1.132	1.030	.726	.748	.628	.791
3	1.439	1.278	.780	.996	.671	.782
4	2.333	1.368	.773	1.058	.653	.785
5	2.763	1.149	.342	.394	.159	.852
6	2.430	1.284	.666	.855	.519	.806
7	2.053	1.130	.775	.876	.681	.782

Efectivamente, los ítems 5 y 6 tal y como habíamos encontrado en los diversos análisis factoriales llevados a cabo sobre la escala son los ítems menos homogéneos con el total de la escala, y por ende con el factor que mide ésta. Este resultado es especialmente acusado en el ítem 5. Se conseguiría una ganancia considerable en la consistencia interna, medida por alfa, si se eliminara.

Se han realizado los mismos análisis de ítems para los cinco que quedan tras eliminar de la escala de riesgo real los ítems 5 y 6. Todos los índices, tanto de homogeneidad, como de fiabilidad han mostrado mejoras apreciables.

Tabla 4.21. Media, desviación típica, coeficiente de homogeneidad (correlación ítem-total), fiabilidad, homogeneidad excluyendo el ítem, y coeficiente alpha que quedaría en la escala si se excluyera cada ítem, para los ítems de la escala de riesgo real una vez eliminados los ítems 5 y 6.

ITEM	Media	D. T.	Homogeneidad	Fiabilidad	Homogen. sin el ítem	Alpha sin el ítem
1	1.941	1.056	.819	.864	.717	.798
2	1.109	1.027	.760	.780	.636	.818
3	1.420	1.267	.801	1.014	.659	.811
4	2.336	1.355	.780	1.057	.613	.828
7	2.050	1.129	.790	.892	.664	.810

A partir de los datos de los análisis factoriales confirmatorios se han calculado los índices de fiabilidad de los ítems, como la saturación factorial elevada al cuadrado. Este cuadrado es un coeficiente de determinación, que indica la proporción de varianza del ítem que puede explicarse desde el factor. Desde las estimaciones para la escala total, los resultados de los tres métodos de estimación son coincidentes. El ítem 5 presenta una bajísima fiabilidad, mientras el ítem 6 no alcanza una elevada proporción de varianza explicada.

Tabla 4.22. Saturaciones factoriales para la escala de riesgo real estimadas mediante mínimos cuadrados generalizados, máxima verosimilitud y estimación arbitraria elevados al cuadrado, como medida de la fiabilidad de los ítems. Los resultados indican la proporción de la varianza de cada ítem explicada por los factores latentes.

	GLS	ML	AGLS
Item 1	.654	.616	.69
Item 2	.586	.505	.656
Item 3	.518	.512	.69
Item 4	.582	.499	.525
Item 5	.033	.03	.07
Item 6	.375	.311	.261
Item 7	.558	.688	.527

El mismo cálculo de la fiabilidad se ha realizado para los análisis factoriales de la escala eliminado el ítem 5 (tabla 4.23). Los resultados son muy similares a los encontrados para la escala total, el ítem 6 es el único pobremente explicado, aunque sí significativamente ( $p < 0.01$ ).

Tabla 4.23. Saturaciones factoriales para la escala de riesgo real (sin ítem 5) estimadas mediante mínimos cuadrados generalizados, máxima verosimilitud y estimación arbitraria elevados al cuadrado, como medida de la fiabilidad de los ítems. Los resultados indican la proporción de la varianza de cada ítem explicada por los factores latentes.

	GLS	ML	AGLS
Item 1	.656	.617	.69
Item 2	.588	.504	.656
Item 3	.508	.508	.69
Item 4	.576	.498	.525
Item 6	.366	.315	.298
Item 7	.559	.555	.573

Lógicamente, cuando se elimina el ítem seis y se repiten los cálculos de fiabilidad se encuentra que todos los ítems son fiables (tabla 4.24). No obstante, hay que reseñar que no produce, para los ítems que quedan en la escala, una ganancia relevante en la proporción de varianza explicada de éstos.

Tabla 4.24. Saturaciones factoriales para la escala de riesgo real (sin ítems 5 y 6) estimadas mediante mínimos cuadrados generalizados, máxima verosimilitud y estimación arbitraria elevados al cuadrado, como medida de la fiabilidad de los ítems. Los resultados indican la proporción de la varianza de cada ítem explicada por los factores latentes.

	GLS	ML	AGLS
Item 1	.674	.644	.672
Item 2	.606	.547	.592
Item 3	.516	.518	.534
Item 4	.502	.448	.558
Item 7	.532	.524	.549

### 3.4. Validez criterial de los ítems

Para la validación criterial de los ítems de la escala de riesgo real se han utilizado los mismos criterios externos que se vienen usando hasta aquí. Las variables con las que, por hipótesis, han de encontrarse relacionados los ítems de riesgo real son: conducta hacia la seguridad, riesgo basal y accidentabilidad. Las relaciones que se esperan con el resto de variables son nulas o muy moderadas, en todo caso mucho menores que las que se encuentren con las tres nombradas anteriormente.

Tabla 4.25. Correlaciones de los ítems de la escala de riesgo real con los principales criterios. ES=estructura de seguridad; PO=política organizacional de seguridad; F=formación-información; RS=respuesta de los superiores en seguridad; RC=respuesta de los compañeros en seguridad; L=actitud hacia la seguridad; CHS=conducta hacia la seguridad; RB=riesgo basal; T=tensión psicológica; y AC= siniestralidad. Símbolos de significación estadística de las correlaciones: ns= no significativo; #= marginalmente significativo; \*= probabilidad <0.05; y \*\*= probabilidad <0.01, extraídos de tablas de significatividad de las correlaciones (Welkowitz, Ewen y Cohen, 1981).

	ES	PO	F	RS	RC	L	CHS	RB	T	AC
1	.004ns	-.04ns	.068ns	-.15#	-.03ns	.056ns	-.32**	.37**	.23*	.278**
2	.192*	.122ns	.000ns	-.05ns	.152#	.045ns	-.16#	.33**	.133ns	.031ns
3	.06ns	-.08ns	-.05ns	-.12ns	.004ns	.192*	-.28**	.48**	.177#	.256**
4	-.02ns	-.14ns	-.01ns	-.21*	-.09ns	.252**	-.30**	.38**	.222*	.196*
5	.062ns	-.40**	-.28**	-.48**	-.27**	.381**	-.44**	.029ns	.246*	-.028ns
6	-.133ns	-.09ns	-.04ns	-.07ns	-.06ns	.293**	-.25**	.29**	.181#	.265**
7	.031ns	-.13ns	-.13ns	-.22*	-.16#	.219*	-.22*	.43**	.233*	.208ns

Casi todos los ítems correlacionan negativa y estadísticamente con la conducta hacia la seguridad, especialmente el ítem 5. Si se observa el contenido de este ítem puede apreciarse en él un fuerte contenido de conducta de seguridad, de forma que su contenido puede estar confundido con el que puede recoger esta variable. Todos los ítems, excepto el cinco, correlacionan según hipótesis y bastante fuertemente con riesgo basal, el principal criterio. Las relaciones con accidentabilidad no son tan claras, pero los ítems 1, 3, 4 y 6 muestran una adecuada relación con la variable.

Las relaciones con el resto de criterios son de acuerdo a hipótesis, es decir, bajas o nulas. Sólo cambia este patrón para la relación con actitud hacia la seguridad. No obstante, cuando veamos la medida de actitud se entenderá mejor este resultado. Efectivamente, la escala de Leather incluye ítems que miden la frecuencia de riesgos laborales. Además de estas relaciones, podemos ver en la tabla 4.25 cómo el ítem 5 muestra importantes relaciones con estructura de seguridad, formación, respuesta de los superiores, respuesta de los compañeros, y actitud, con los que se esperaba una relación baja o nula, mientras no se correlaciona adecuadamente con riesgo basal, y accidentabilidad, principales criterios. Tal y como hemos visto el contenido del ítem cinco puede tener algo en común con conducta de seguridad, ya que pregunta básicamente sobre hábitos de trabajo seguros y, por ello, parece mostrar un patrón de relaciones más acorde a la medida de conducta hacia la seguridad.

### **3.5. Conclusiones sobre la escala de riesgo real.**

La discusión metodológica tiene dos vertientes principales. Por un lado la vertiente de la homogeneidad de los ítems de la escala con el factor de riesgo real, ya sea estudiada ésta por modelos causales o por métodos tradicionales. Por otro lado, el estudio de validez criterial de los ítems. Ambas discusiones deben ponerse en común para escoger qué ítems deben eliminarse, si es que hay que eliminar alguno. Los resultados de homogeneidad ofrecen en general un muy buen panorama, con un único ítem inadecuado (el 5) y un ítem menos homogéneo (el 6). La validez criterial de los ítems de la escala es menos clara, parece claro que el ítem seis tiene una importancia en este sentido por su importante relación con la variable accidentabilidad. Es en este sentido que valorando ambos aspectos creemos que para dar un uso lo más adecuado posible metodológicamente, habría que eliminar únicamente el ítem cinco y conservar el resto de la escala.

#### **4. Validez convergente-discriminante de las escalas de riesgo basal y real**

Cuando se distingue entre riesgo real y basal se está haciendo una suposición teórica, una diferenciación clara para el investigador, pero que puede no apreciarse por parte del trabajador que contesta el cuestionario. Podría ocurrir que el trabajador que contesta el cuestionario conteste a todos los ítems de la misma forma, sean de riesgo basal o de riesgo real. En este sentido, y para poder aplicar con garantías de medición ambas variables, cabe hacer un estudio de cuán bien los ítems de las escalas de riesgo basal y real discriminan con entre ambos conceptos. La validez convergente ya ha sido probada, ya que mediante los diversos análisis factoriales confirmatorios hemos visto cómo los factores propuestos estaban a la base, razonablemente, de la explicación de los ítems. En una palabra, que los ítems convergían adecuadamente con el factor hipotetizado. Ahora se trata de realizar un análisis factorial confirmatorio conjunto de las dos escalas. Este análisis factorial nos ofrecerá, básicamente, tres tipos de información extraordinariamente relevantes para el uso posterior de estas escalas. En primer lugar, la validez convergente de los ítems de riesgo basal con el factor de riesgo basal y de los ítems de riesgo real en el factor de riesgo real. En segundo lugar, la validez discriminante de los ítems de riesgo basal respecto del factor de riesgo real; y de los ítems de riesgo real respecto del factor de riesgo basal. Para que se cumpla la validez convergente los ítems de riesgo basal deben saturar significativamente en el factor de riesgo basal y los de riesgo real en el de riesgo real. Para que se cumpla la validez discriminante los factores de riesgo basal no deben saturar en el factor de riesgo real y los de riesgo real no deben saturar en el factor de riesgo basal. En tercer lugar, la correlación hipotetizada entre los dos factores del análisis factorial confirmatorio, riesgos basal y real, puede considerarse como un coeficiente de validez criterial para ambas escalas, que son criterios cada una de ellas de la otra. Así, este coeficiente de validez criterial tiene en cuenta los errores de medida que se han producido en los ítems para la medición de los factores, contrariamente a lo que sucede con la correlación entre las puntuaciones de las dos escalas mediante el coeficiente de correlación de Pearson.

Tabla 4.26. Índices de ajuste global para la estructura bifactorial hipotetizada para las escalas de riesgo basal y real, para comprobar la validez convergente-discriminante estimada por mínimos cuadrados generalizados. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Método de estimación: GLS=mínimos cuadrados generalizados.

		chi	g	p	F	LHR	chi/g	NFI	NNFI	CFI
2 FACTORES (RB y RR)	GLS	283.08	251	.0800	2.7	.253	1.127	.975	.997	.997

Como puede verse en la tabla 4.26, los resultados de ajuste global son muy adecuados, por lo que las estimaciones de las saturaciones factoriales y de la correlación, que es lo que nos interesa en este caso para poder estudiar la validez de las dos escalas pueden considerarse adecuadas. A continuación pueden consultarse en la tabla 4.27 los datos de las saturaciones factoriales. Las hipótesis de validez convergente y discriminante se cumplen en su totalidad.

Tabla 4.27. Saturaciones factoriales de los veinticuatro items de las escalas de riesgo basal y real en los dos factores hipotetizados para el estudio de la validez convergente-discriminante (\*\*= probabilidad <.01).

Nº Item y factor teórico	l con RB	l con RR
Item 1 (riesgo basal)	.387*	0
Item 2 (riesgo basal)	.581**	0
Item 3 (riesgo basal)	.692**	0
Item 4 (riesgo basal)	.631**	0
Item 5 (riesgo basal)	.728**	0
Item 6 (riesgo basal)	.583**	0
Item 7 (riesgo basal)	.421**	0
Item 8 (riesgo basal)	.361**	0
Item 9 (riesgo basal)	.630**	0
Item 10 (riesgo basal)	.824**	0
Item 11 (riesgo basal)	.822**	0
Item 12 (riesgo basal)	.894**	0
Item 13 (riesgo basal)	.800**	0
Item 14 (riesgo basal)	.796**	0
Item 15 (riesgo basal)	.681**	0
Item 16 (riesgo basal)	.715**	0
Item 17 (riesgo basal)	.666**	0
Item 18 (riesgo real)	0	.866**
Item 19 (riesgo real)	0	.744**
Item 20 (riesgo real)	0	.826**
Item 21 (riesgo real)	0	.861**
Item 22 (riesgo real)	0	.286**
Item 23 (riesgo real)	0	.729**
Item 24 (riesgo real)	0	.686**

Además, no se ha producido ninguna sugerencia ni en el LM test ni en el W test, por lo que se puede apoyar claramente la validez convergente de los items con sus factoes y la validez discriminante con respecto al otro factor. Se guarda una absoluta exclusividad y exhaustividad en todos los items de las dos escalas. Se confirma con ello el planteamiento teórico de distinción entre los dos tipos de riesgos, y parece razonable aplicarlo en el estudio de las causas de los accidentes.

Por su parte, y en lo que a validez criterial se refiere, la correlación entre los dos factores es alta y positiva (0.569\*\*), tal y como se esperaba por hipótesis.

En conclusión, y en base a todos los análisis, parece adecuado utilizar las dos escalas como medidas de constructos relacionados, pero identificables. Parece, además, que las escalas son medidas válidas y fiables de sus respectivos constructos teóricos.





## 5. Variables de Contexto Social.

### **1. Respuesta de los Superiores en seguridad**

#### **1.1. Introducción y medida**

La respuesta de los superiores en seguridad es usualmente considerada como una variable clave en el estudio de la seguridad. El trabajo día a día del trabajador se realiza en un contexto laboral y social determinado. Dentro de los actores sociales que lo conforman tenemos generalmente dos grupos, a los compañeros y al supervisor o supervisores inmediatos. Mientras que los compañeros son un grupo de referencia para los sujetos, que puede ejercer influencia informal, los supervisores son los encargados directa y formalmente de controlar, guiar, encauzar o modificar el comportamiento de los trabajadores, ejerciendo poder formal sobre ellos, y pudiendo ejercer influencia informal. De esta forma, los supervisores directos pueden considerarse como el grupo o la persona más importante de modificación o mantenimiento del comportamiento de los trabajadores. Dentro de la gama de comportamientos que exhibe el trabajador, el tipo que específicamente nos ocupa es el comportamiento de seguridad, y también sobre éste los supervisores pueden ejercer su influencia.

La respuesta de los supervisores puede verse desde otro punto de vista como el punto de unión entre las políticas organizacionales y el trabajador. La mayoría de las normativas, reglamentaciones e incluso acciones de promoción de la seguridad que realiza la dirección le llegan al trabajador de la mano de su supervisor. De esta forma, la política de seguridad e incluso en ocasiones la formación se pueden canalizar a través del supervisor, y el efecto que puedan tener éstas actuar con efecto indirecto a través de la respuesta de los superiores.

En cualquier caso, parece que la respuesta de seguridad de los supervisores debe convertirse en un importante predictor, si no el más importante, de la conducta segura o insegura que exhiba el trabajador en su desempeño.

En la formulación de los items se ha intentado medir tres aspectos de la respuesta de los supervisores o superiores: a) la conducta que el supervisor exhibe en materia de seguridad, es decir, las conductas seguras o inseguras que el propio superior exhibe en su desempeño laboral; b) las contingencias, positivas o negativas, que ofrece el supervisor ante la conducta segura o insegura del trabajador; y c) la actitud que presenta el supervisor ante la seguridad. Este último aspecto es probablemente el más problemático de medir. Si fuera el supervisor el que contesta el cuestionario sería posible preguntarle sobre su actitud, pero es el trabajador quien contesta sobre la actitud que presume en su supervisor. Esto evidentemente supone un problema, dado que el trabajador infiere la actitud del supervisor en base a su conducta para con él, con los otros compañeros, y de su propio desempeño, de forma que no se puede, en principio, desligar de los otros dos aspectos, conducta del supervisor y contingencias que ofrece. No obstante, los contenidos teóricos se corresponden con estos tres aspectos mencionados, aún con las dificultades que presenta su medición. A continuación se presentan los items de la escala, su contenido y escala de respuesta.

#### ESCALA:

Item 1. Si Ud. NO cumple alguna norma de seguridad, ¿su superior le llama la atención?

- Nunca u ocasionalmente .....1
- Casi siempre .....2
- Siempre lo hace.....3

Item 2 . ¿Qué respuesta recibe de sus superiores cuando Ud. realiza su trabajo en forma segura?

- Me piden que trabaje más aprisa..... 1
- No me dicen nada ..... 2
- Recibo comentarios positivos.....3
- Recibo incentivos económicos.....4

Item 3. ¿Apoya su superior el cumplimiento de las normas de seguridad?

- Nunca u Ocasionalmente.....1
- Casi siempre.....2
- Siempre lo hace .....3

Item 4. ¿Con qué frecuencia comentan Ud. y su superior sobre las medidas de seguridad que deben tomarse para prevenir un accidente?

- Nunca u Ocasionalmente..... 1
- Casi siempre..... 2
- Siempre lo hace..... 3

Item 5. ¿Cómo describiría la actitud de su superior hacia la seguridad en el trabajo?

- Negativa o Indiferente ..... 1
- Positiva ..... 2
- Muy Positiva .....3

Item 6. ¿Se esfuerza su superior por trabajar de forma segura?

- Nunca u Ocasionalmente..... 1
- Casi siempre..... 2
- Siempre lo hace.....3

Item 7. ¿Qué respuesta recibe de sus superiores cuando realiza su trabajo de forma peligrosa, con riesgo o insegura?

- Recibo presiones o sanciones ..... 1
- Recibo comentarios negativos..... 2
- No me dicen nada..... 3
- Recibo incentivos por trabajar más aprisa ..... 4

## 1. 2. Estructura factorial confirmatoria

Para contrastar la estructura factorial del cuestionario se han planteado tres configuraciones teóricas:

1) Una estructura unifactorial. Un solo factor subyace a la escala, y se basa en el supuesto teórico de que conductas del supervisor, actitud y contingencias no son discernibles, forman un único concepto global.

2) Una estructura de dos factores correlacionados. El primer factor estaría formado por los items 1, 2, 4 y 7, y se denominaría *contingencias por seguridad* que ofrece el superior. Incluye contenidos que hacen referencia a las contingencias positivas o negativas que el supervisor ofrece ante las conductas seguras o inseguras de los trabajadores. El segundo factor, *conducta de seguridad* del superior, incluye los otros tres items, números 3, 5 y 6. Estos items hacen referncia a la conducta de seguridad del supervisor en su desempeño. Además, se han incluido aquí los contenidos referidos a la actitud del supervisor en seguridad, bajo la idea de que

es precisamente de su conducta de donde el trabajador infiere principalmente su actitud, y por tanto serán más homogéneos con este factor.

3) Finalmente, se ha planteado teóricamente una estructura de tres factores. El primer factor quedaría conformado por los ítems 1, 2 y 7 y seguiría denominándose contingencias por seguridad. El segundo factor sería el de conducta del supervisor, con los ítems 4 y 6. Aquí se ha producido un cambio teórico, el ítem 3 se hipotetizaba como parte de las contingencias en la estructura de dos factores, mientras ahora se considera una conducta de seguridad del propio supervisor. El contenido del ítem hace referencia a la frecuencia de comentarios sobre aspectos de seguridad entre el supervisor y el trabajador. Si bien los comentarios pueden verse como una contingencia positiva que ofrece el supervisor al trabajador, dándole relevancia al tema de la seguridad, lo cierto es que es también una conducta de seguridad del propio supervisor. Este cambio de adscripción teórica del ítem tiene dos propósitos, por un lado evitar que un factor quede definido por un solo ítem y, por otro, contrastar qué hipótesis de relación con el factor es capaz de explicar mejor el ítem, si un factor de conducta o uno de contingencias.

Las tres estructuras confirmatorias han sido puestas a prueba mediante análisis factoriales confirmatorios, con métodos de estimación máximo verosímil y de mínimos cuadrados generalizados. Los resultados de ajuste global de las estructuras hipotetizadas pueden consultarse en la tabla 5.1.

Tabla 5.1. Índices de ajuste global de las tres estructuras hipotetizadas para la escala de respuesta de los superiores. Los índices son: *chi*=chi-cuadrado; *g*=grados de libertad del modelo; *p*=probabilidad; *F*=función de ajuste a minimizar; *LHR*=razón de verosimilitud escalada; *CAK*=Akaike information criterion; *CSK*=Schwarz information criterion; *NFI*=normed fit index; *NNFI*=non-normed fit index; *CFI*=comparative fit index. Métodos de estimación: *GLS*=mínimos cuadrados generalizados; *ML*=máxima verosimilitud.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
UN	GLS	12.8	14	.541	.124	.939	.914	.431	.838	.991	1.0	1
FACTOR	ML	14.7	14	.397	.142	.931	1.05	.449	.856	.943	.995	.997
DOS	GLS	8.76	13	.790	.085	.958	.673	.392	.799	.994	1.0	1
FACTORES	ML	9.79	13	.710	.095	.953	.753	.402	.809	.962	1.02	1
TRES	GLS	10.9	11	.446	.106	.948	.990	.413	.820	.992	1	1
FACTORES	ML	11.7	11	.383	.113	.945	1.06	.420	.827	.955	.994	.997

En general, los índices de ajuste global para los tres modelos hipotéticos son excelentes. Las pruebas de ji-cuadrado calculadas para comprobar el supuesto de que la matriz de covarianzas reproducida

por el modelo es igual a la matriz de covarianzas de las variables ha resultado en todos los casos estadísticamente no significativa con elevadas probabilidades. En otras palabras, las tres estructuras factoriales son capaces de representar adecuadamente los datos observados. En cuanto a los índices de comparación con un modelo nulo, menos restrictivos que la ji-cuadrado, muestran ganancias altísimas con respecto a un modelo basal de completa independencia. Por otro lado, y a nivel puramente estadístico es interesante comentar que los índices NFI, NNFI, y CFI muestran aquí alguno de los comportamientos típicos que ya comentamos en el capítulo 2. Así, a pesar de que los índices NNFI y CFI alcanzan sin problemas valores de uno, el índice NFI se presenta anormalmente bajo para el buen ajuste de los modelos, aunque siempre por encima de .95. Por su lado, el índice de ajuste no-normativo, que corrige este problema de acercamiento al uno de NFI supera en muchas ocasiones el límite superior, lo que también ha sido señalado como un defecto. Por su parte, el CFI muestra cómo resuelve los problemas de estos dos últimos índices, siendo alto, pero sin superar el valor de 1.

Estos resultados de ajuste global apuntan a que los ítems de la escala están fuertemente correlacionados entre sí, si bien hay que entrar en detalle al respecto de qué hipótesis sobre factores latentes refleja mejor los datos. En una primera aproximación pueden, como hemos visto, realizarse pruebas de diferencias de las ji-cuadrado para estudiar si se producen cambios estadísticamente significativos entre los modelos. Si volvemos de nuevo la tabla 5.1 apreciaremos que el modelo monofactorial presenta unas pruebas de ji-cuadrado muy similares a las del modelo de tres factores, y que es sólo el modelo de dos factores el que presenta valores más bajos de éstas, produciéndose un salto que es apreciable también en los niveles de significación estadística. Considerando, además, que el modelo de tres factores es el menos parsimonioso de los tres para representar a los datos, lo que procede es realizar una prueba de las diferencias de ji-cuadrado entre los modelos de uno y dos factores, que puede subsumir de algún modo el resto de pruebas.

$\Delta\chi^2 = 4.046$ , con  $\Delta gl = 1$  que resulta estadísticamente significativa ( $p < .05$ ),

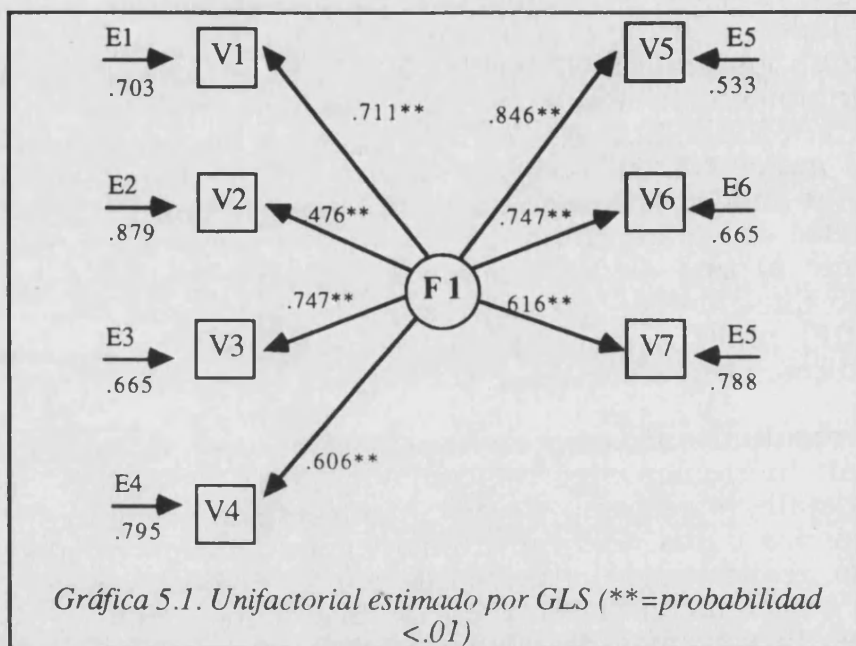
y para máxima verosimilitud:

$\Delta\chi^2 = 4.92$ , con  $\Delta gl = 1$  que también resulta significativa ( $p < .05$ ).

Puede afirmarse, por tanto, que en la explicación de los datos, el modelo de dos factores es estadísticamente mejor que el de un factor y que el de tres.

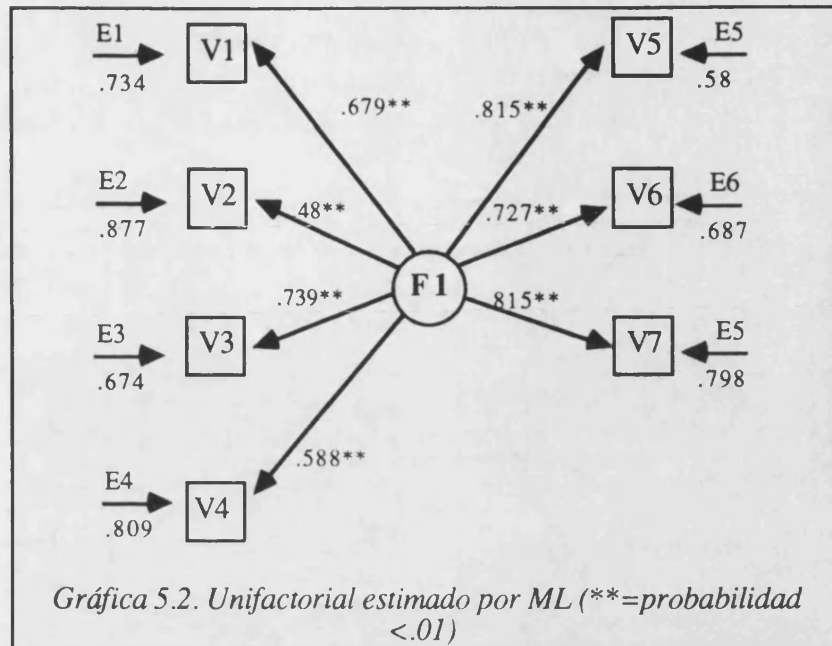
Además de estudiar los ajustes globales, es interesante analizar los parámetros de relación de los ítems con los factores. Así, en la gráfica 5.1 pueden consultarse los datos de ajuste analítico de modelo

unifactorial estimado mediante mínimos cuadrados generalizados. Todas las saturaciones factoriales resultan estadísticamente significativas, mostrando que un único factor hipotético de respuesta de los superiores es capaz de explicar razonablemente la varianza de los ítems. Sólo el ítem 2, aún siendo estadísticamente significativo, presenta una saturación algo más baja.

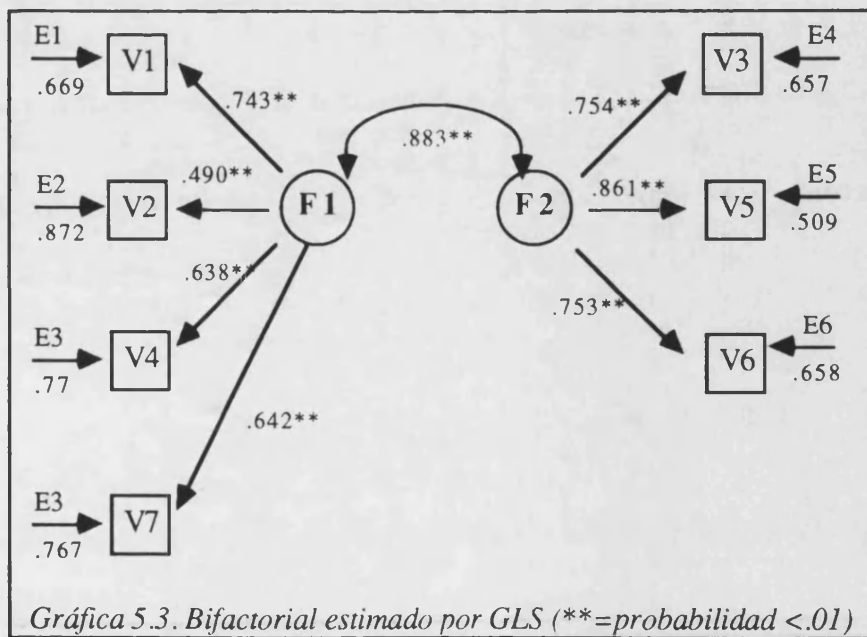


Los resultados del unifactorial confirmatorio para la estimación máximo verosímil se pueden observar en la gráfica 5.2. Los valores estimados para las diferentes saturaciones factoriales son muy similares a los hallados mediante GLS. Es también el ítem 2 el que presenta menor homogeneidad con el factor hipotetizado.



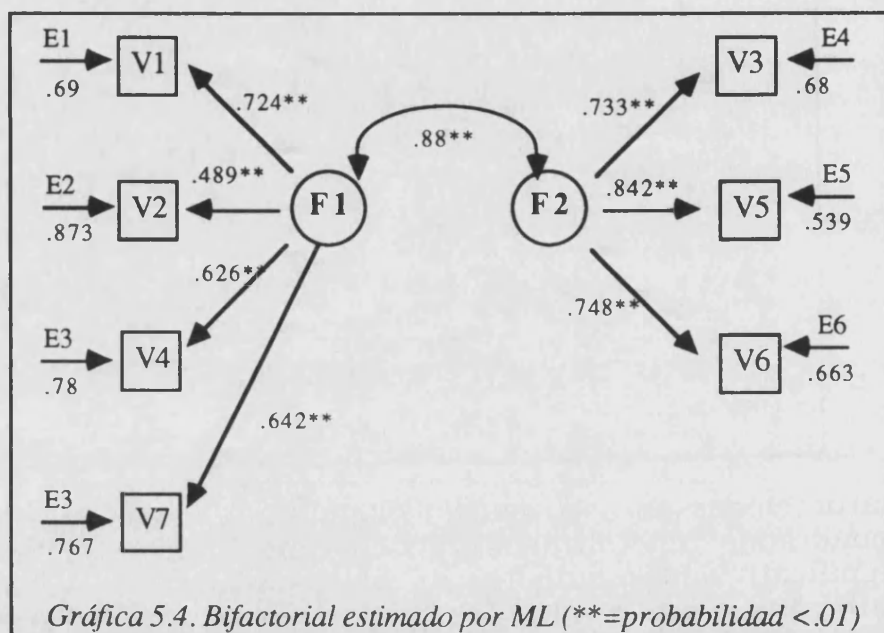


Las saturaciones para el modelo bifactorial estimado mediante mínimos cuadrados generalizados pueden consultarse en la gráfica 5.3. La significatividad estadística no ha cambiado, con todos los items significativamente explicados por los factores hipotetizados ( $p < .01$ ). Los cambios son ligeros en cuanto a los valores concretos, pero en general se producen ligeras mejoras en todos los items, pero en general son mínimas. Por su parte, los dos factores se encuentran altamente correlacionados, 0.883, lo que indica una gran relación entre la conducta de seguridad de los superiores y las contingencias que ofrecen a los trabajadores.



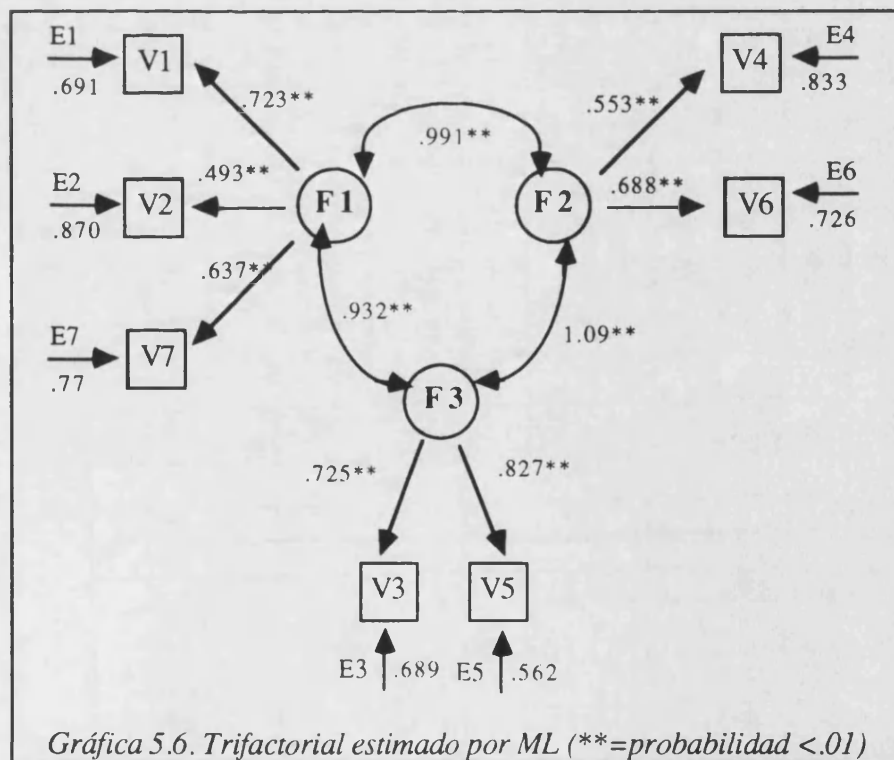
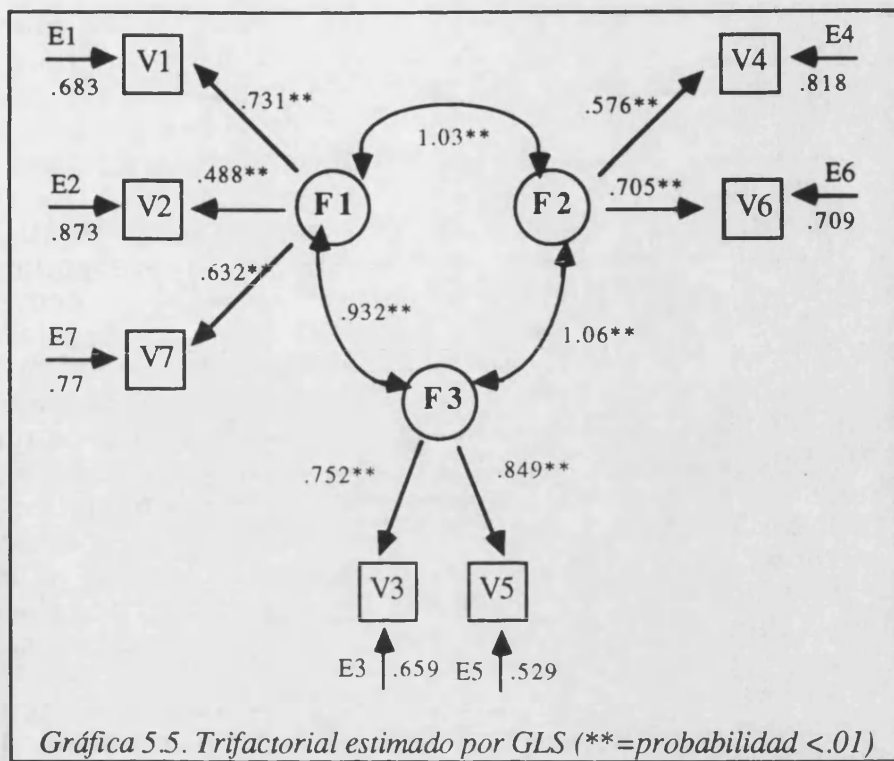


También resultan muy similares las estimaciones máximo verosímiles de las saturaciones factoriales, que pueden verse en la gráfica 5.4. En este caso, todas las saturaciones son algo superiores a las encontradas para el modelo unifactorial. La correlación entre los factores es muy alta, al igual que ocurría en la estimación por mínimos cuadrados generalizados.



Por último, considerar que el Lagrangian Multiplier Test no ha propuesto ningún cambio de adscripción de los ítems a los dos factores, ni en la estimación máximo verosímil ni en la mínimo cuadrática. La estructura factorial es simple, con un solo ítem adscrito a cada ítem, y es además la más parsimoniosa, dado que el modelo de un factor que es más parsimonioso explica peor los datos observados y no se retiene por ello.

En las gráficas 5.5 y 5.6 se presentan los resultados para el modelo de tres factores. Tan sólo resaltar que el modelo ajusta a los datos adecuadamente, como se vió en los resultados de ajuste global, y ésto se comprueba también en los datos de ajuste analítico. Todos los ítems están significativamente relacionados con los factores hipotetizados, y en general sus resultados son muy similares a los hallados tanto en el modelo unifactorial como en el bifactorial.



Una cuestión particular que nos habíamos planteado era el comportamiento del ítem 3 en la estructura de tres factores frente a

la de dos factores. Recordemos que mientras el ítem tres se planteaba como un ítem del factor contingencias en el bifactorial, ahora en el trifactorial se plantea como del factor de conductas. A nivel de capacidad de explicación de un factor u otro sobre el ítem 3, los resultados son muy similares. Las saturaciones factoriales en el bifactorial son, para GLS y ML respectivamente, 0.754 y 0.733; mientras que en el trifactorial son ligeramente inferiores, 0.752 y 0.725. En cualquier caso, hay que analizar en profundidad la cuestión. Por un lado, el factor en que se inscribe el ítem en el trifactorial tiene dos indicadores, mientras que en el bifactorial tienen tres indicadores, por lo que es normal que la saturación en el trifactorial sea un poco más elevada por el menor número de indicadores. Por otro lado, mientras en el bifactorial parece que los dos factores son discernibles; están muy correlacionados (0.8) pero no puede decirse que sean el mismo constructo. Sin embargo, la solución trifactorial arroja estimaciones altísimas para las relaciones entre los factores, superiores a 0.9 e incluso al 1, haciendo imposible distinguir entre factores y haciendo ver que estos tres factores hipotetizados están fuertemente confundidos en sus contenidos.

Por todos estos motivos, parece lo más razonable mantener la estructura bifactorial como mejor y más clara representación de los datos, desde los métodos de estimación basados en el supuesto de normalidad multivariable. Esto es, desde máxima verosimilitud y mínimos cuadrados generalizados. Otra cuestión que surge es la idoneidad de estos métodos para esta escala, que debe ser probada como hemos hecho hasta aquí. En primer lugar, cabe plantear la plausibilidad de la normalidad univariada, condición necesaria para la multivariada. En la tabla 5.2 pueden consultarse las pruebas de Kolmogorov-Smirnov realizadas para comprobar este punto. Como se puede apreciar, todas ellas han arrojado resultados estadísticamente significativos, indicando fuertes desviaciones de la normalidad para todos los ítems de la escala.

Tabla 5.2. Prueba de Kolmogorov-Smirnov de una muestra para comprobar el ajuste de los ítems de la escala de respuesta de los superiores a la distribución normal univariada.

ITEM	N muestral	Diferencia máxima	Probabilidad
1	119	.500	.000
2	122	.500	.000
3	118	.500	.000
4	122	.500	.000
5	120	.500	.000
6	120	.500	.000
7	117	.631	.000

Los datos de asimetría y curtosis univariada, tabla 5.3, también nos indican las desviaciones con respecto a la normalidad que se

producen en las variables, Todas ellas presentan fuerte asimetría o apuntamiento, o ambas tendencias.

Tabla 5.3. Datos de asimetría y curtosis para los items de la escala de respuesta de los superiores.

	Asimetría	Curtosis
Item 1	-.1281	-1.59
Item 2	.3733	.1454
Item 3	-.5446	-1.24
Item 4	1.077	.095
Item 5	.0649	-.377
Item 6	-.0833	-1.338
Item 7	.0645	-.384

Por último y aun cuando no es necesario abundar en ello por ser los análisis univariados suficientemente esclarecedores, el programa EQS calcula un estadístico descriptivo del alejamiento multivariado de la normalidad, el coeficiente de Mardia. En este caso su valor es también claro indicio de que no se cumple el supuesto de normalidad multivariable. Este es habitualmente el caso con items de escala de respuesta tipo Likert.

COEFICIENTE DE MARDIA MULTIVARIADO= 1.1598 Coeficiente esperado para la normal multivariada= 0.5268
--

En estas condiciones parecía adecuado plantear una análisis factorial a estimar mediante métodos arbitrarios. Hasta ahora habíamos aplicado los métodos arbitrarios de estimación sobre la solución que en los análisis por ML y GLS parecía más adecuada. Siguiendo esta lógica, deberíamos aplicar una solución bifactorial, que es la que parece representa mejor los datos. Hemos visto, no obstante, que en esta escala la solución factorial es compleja. Las tres estructuras hipotetizadas dan cuenta adecuadamente de los datos. Las saturaciones varían muy ligeramente a través de métodos y estructuras y los factores hipotetizados están altamente correlacionados. Todo ello hace difícil evaluar si es la estructura bifactorial, ligeramente superior en la explicación de porcentajes de varianza, o la unifactorial, más parsimoniosa, el ajuste adecuado. Hemos decidido, en consecuencia, probar ambos ajustes factoriales para ver si los resultados mediante métodos arbitrarios arrojan luz sobre la dimensionalidad de la escala. Los datos de ajuste global de los factoriales confirmatorios pueden consultarse en la tabla 5.4.

Tabla 5.4. Índices de ajuste global de las estructuras unifactorial y bifactorial para la escala de respuesta de los superiores. Los índices son: *chi*=chi-cuadrado; *g*=grados de libertad del modelo; *p*=probabilidad; *F*=función de ajuste a minimizar; *LHR*=razón de verosimilitud escalada; *CAK*=Akaike information criterion; *CSK*=Schwarz information criterion; *NFI*=normed fit index; *NNFI*=non-normed fit index; *CFI*=comparative fit index. Método de estimación *AGLS*=método arbitrario.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
1 FACTOR	AGLS	15.0	14	.374	.146	.929	1.07	.453	.860	.853	.981	.987
2 FACTORES	AGLS	12.0	13	.522	.117	.943	0.92	.424	.831	.882	1.019	1

GFI: 1 FACTOR=.978; 2 FACTORES=.982.

AGFI: 1 FACTOR=.956; 2 FACTORES=.962.

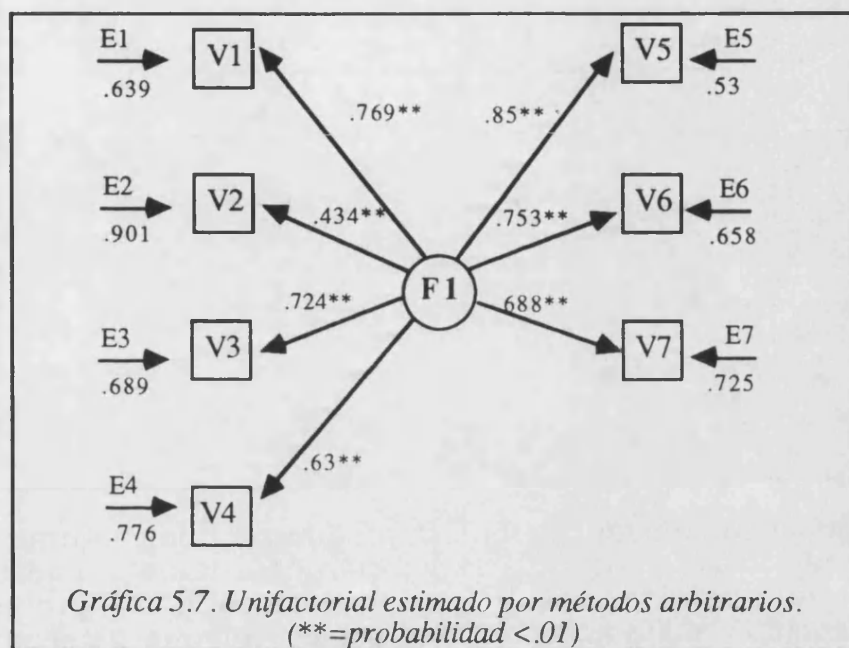
Como puede verse, los datos de ajuste no difieren sustancialmente de los hallados para los métodos de estimación máximo verosímil y mínimo cuadrático. Las dos estructuras factoriales representan adecuadamente los datos, a decir de la prueba de ji-cuadrado y el resto de índices. Además, ambas estructuras son altamente parsimoniosas, como muestra el índice ji-cuadrado partido los grados de libertad, en torno al uno. El índice de bondad de ajuste (GFI) y el ajustado (AGFI), que no ofrece el programa EQS para otras estimaciones, resultan también muy elevados, indicando un muy adecuado ajuste en ambas estructuras. Además, y como comentario puramente estadístico, se comprueba una vez más los comportamientos extraños del NFI, muy bajo para el resto de resultados de ajuste, y del NNFI, que excede de uno su límite superior teórico.

Queda abierta una vez más la cuestión de si las diferencias entre ambos ajustes son o no estadísticamente significativas. Si recordamos las pruebas de diferencias de ji-cuadrado llevadas a cabo para ML y GLS la estructura bifactorial representaba ligeramente mejor los datos que la estructura unifactorial ( $p < 0.05$ ). Para la estimación arbitraria, la situación es diferente. La proporción de varianza explicada, el ajuste, por ambas estructuras no difiere estadísticamente, como puede verse en la siguiente prueba de las diferencias.

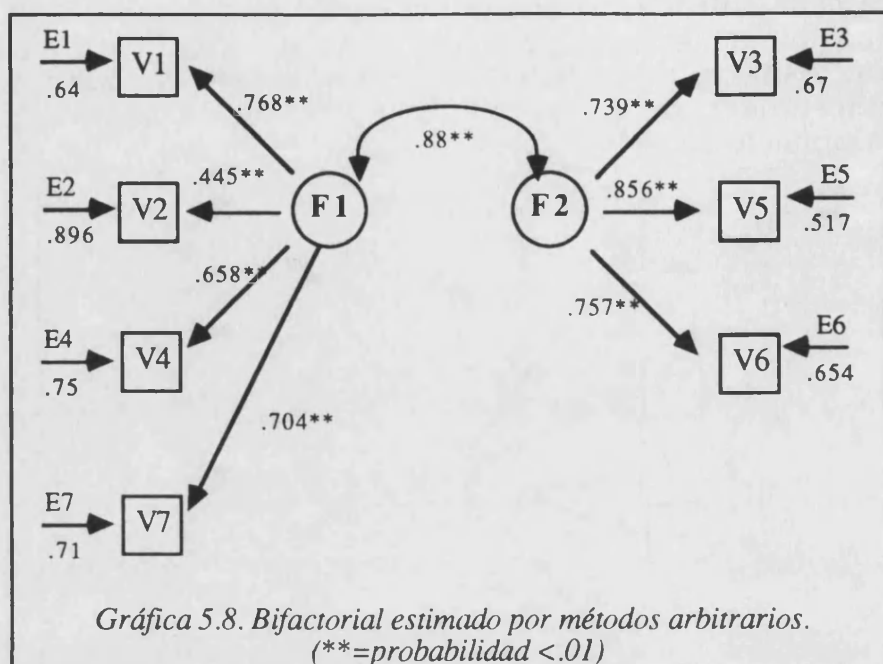
$\Delta\chi^2 = 2.999$ , con  $\Delta g = 1$  que no resulta estadísticamente significativa ( $p > .05$ ).

Ante un resultado así, cabe hablar de un mejor ajuste mediante la estructura unifactorial por su parsimonia, contrariamente a la conclusión que se extraía desde las estimaciones basadas en el supuesto de normalidad multivariable.

Ahora cabe entrar en profundidad en los ajustes analíticos, para observar posibles ganancias particulares en los ítems en una u otra estructura. Así, en la gráfica 5.7 se presentan las saturaciones factoriales estandarizadas, los errores y los ajustes analíticos de la estructura unifactorial.



Una primera cuestión, al respecto de la estructura unifactorial es comprobar los cambios que se han producido con respecto a los parámetros estimados mediante ML y GLS. La conclusión es muy clara, las diferencias son apenas apreciables, tanto frente a máxima verosimilitud como frente a mínimos cuadrados. Si hablamos en términos cuantitativo, las diferencias nunca alcanzan más allá del 0.08, y aún este valor es raro. La mayoría permanecen igual o ligerísimamente inferiores o superiores. Vemos qué ocurre para la estructura de dos factores (gráfica 5.8).



Si evaluamos las diferencias que se producen en las saturaciones entre la estimación arbitraria y la estimación basada en supuesto de normalidad multivariable -ML y GLS-, volvemos a comprobar la práctica ausencia de diferencias entre ellas. Las diferencias se sitúan en el orden de las centésimas y además sin un patrón uniforme, a veces a favor de un tipo de estimación, a veces a favor de otro.

Un último aspecto falta por evaluar, ¿produce saturaciones apreciablemente más elevadas el ajuste mediante dos factores que el unifactorial para la estimación arbitraria?. O dicho de otro modo, ¿se explica mejor los items desde dos factores que desde uno?. La respuesta es, una vez más clara, no. Las diferencias aún son menores que entre métodos, situándose generalmente a nivel del tercer decimal, de la milésimas. Tampoco se aprecia un patrón, no siempre son más elevadas las saturaciones de la estructura bifactorial que las de la unifactorial.

### 1. 3. Estructura factorial de segundo orden

Los resultados de los análisis factoriales apuntan hacia la existencia de una razonable homogeneidad en la escala, con los items muy correlacionados entre sí. Así, hemos visto que, aún cuando la estructura bifactorial era la que estadísticamente mejor representaba los datos, los dos factores estaban altamente relacionados. Además, la explicación de los items desde los factores latentes no varía sustancialmente entre la hipótesis de un factor y la de dos subyaciendo a la escala. Para complicar más la cuestión, si se aplica



estimación arbitraria no hay diferencias estadísticamente significativas entre los modelos de un factor y de dos, por lo que se retendría el unifactorial por ser más parsimonioso en la explicación de las covarianzas entre los items. Todos estos datos apuntan fuertemente a probar una estructura factorial de segundo orden. Así, se probará una estructura de dos factores de primer orden, los ya probados anteriormente, que serían indicadores de un factor de segundo orden denominado respuesta de los superiores en seguridad.

Esta estructura puede no estar identificada. Para que un modelo factorial de segundo orden esté identificado tienen que estar identificados las dos estructuras que lo componen (Rindskopf y Rose, 1989). Esto es, si entendemos el factorial de segundo orden como compuesto por dos partes, la que relaciona los factores de primer orden y las variables observables y la que relaciona los factores de segundo orden con los de primero, podemos decir que el modelo está identificado si se cumplen las reglas de identificación para ambas partes. Bentler (1989) plantea que para identificar un factor de segundo orden es necesario disponer de dos factores de primer orden, y "aún así con las restricciones adecuadas", consistentes en forzar las saturaciones de esos dos factores de primer orden con el de segundo a ser iguales. A pesar de la opinión de Bentler, que se aplica en un ejemplo en el manual de EQS, no parece ésta una opinión unánime. Así, Rindskopf y Rose (1989) sostienen una opinión radicalmente diferente. Para que un modelo de factorial de segundo orden sea identificado se necesitan, al menos, tres factores de primer orden como indicadores de un factor de segundo orden (cuando se hipotetiza un solo factor de segundo orden como es nuestro caso). Aún así, continúan los autores, cuando sólo se dispone de tres factores de primer orden el sistema se encuentra justamente identificado, por lo que los test de ajuste global no prueban la estructura de segundo orden. Esto indicará, a efectos prácticos, que los resultados de ajuste global serán los mismos que si no se hipotetizara estructura de segundo orden. Estos autores no contemplan la posibilidad de calcular estructuras de segundo orden con sólo dos indicadores.

Hemos calculado un factorial de segundo orden con sólo dos indicadores, factores de primer orden, del factor de segundo orden hipotetizado. Los parámetros de relación entre los factores de primer orden y el de segundo se constriñeron según las indicaciones de Bentler (1985) y se procedió a la estimación. No se produjo ningún mensaje de error y en la siguiente tabla pueden consultarse los índices de ajuste global calculados para la estimación por mínimos cuadrados generalizados.



Tabla 5.5. Índices de ajuste global de la estructura factorial de segundo orden hipotetizada para la escala de respuesta de los superiores. Los índices son: *chi*=chi-cuadrado; *g*=grados de libertad del modelo; *p*=probabilidad; *F*=función de ajuste a minimizar; *LHR*=razón de verosimilitud escalada; *CAK*=Akaike information criterion; *CSK*=Schwarz information criterion; *NFI*=normed fit index; *NNFI*=non-normed fit index; *CFI*=comparative fit index. Método de estimación: *GLS*=mínimos cuadrados generalizados.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
2º ORDEN	GLS	8.76	13	.790	.085	.958	.673	.392	.799	.994	1.0	1

Curiosamente los valores de ajuste global son los mismos que los hallados para el caso de estimación mínimo cuadrática en el factorial de primer orden. Los datos de estimación de las saturaciones factoriales y de los errores de la estructura que relaciona las variables observables con los factores de primer orden son asimismo iguales. Ofrecemos los valores de las saturaciones factoriales entre los factores de primer orden y el de segundo orden que son, para el primer factor de primer orden 0.923 ( $p < .01$ ), y para el segundo factor de primer orden 0.956 ( $p < 0.01$ ).

#### 1. 4. Fiabilidad de la escala y de los items

La escala muestra en su conjunto una buena consistencia interna, por todos los coeficientes. Así, un coeficiente alfa para el total de 0.836 indica una elevada covariación entre los items. Sólo el alpha para los items pares aparece un poco bajo, debido probablemente al menor número de items (tres frente a cuatro impares) y a que incluye dos de los items menos homogéneos, el 2 y el 4, ver tabla 5.7.

Tabla 5.6. Coeficientes de fiabilidad, consistencia interna, de la escala de respuesta de los superiores.

<i>Consistencia interna</i>	
Correlación dos mitades	.729
Coefficiente Spearman-Brown	.843
Coefficiente Guttman-Rulon	.804
Coefficiente alpha total	.836
Alpha items impares	.780
Alpha items pares	.591

Los datos del análisis de items se presentan en la tabla 5.7. Como ya venía ocurriendo en el análisis factorial, el item dos es el menos homogéneo con el total de la escala, alcanzando una correlación con el total de 0.556 (excluido el item de 0.444). Aunque no puede hablarse de una homogeneidad excesivamente baja, lo cierto es que es el de peor funcionamiento en este aspecto. Además y a nivel descriptivo, cabe destacar que los dos items que representan conducta de seguridad del supervisor ante la conducta segura del

trabajador (items 2 y 4) presentan la media más baja, mientras que otros items, como por ejemplo la conducta de seguridad del supervisor ante conductas inseguras del trabajador (item 7) presenta una media sustancialmente más alta, indicando mayor respuesta frente a la inseguridad que frente a la adopción de medidas seguras.

Tabla 5.7. Media, desviación típica, coeficiente de homogeneidad (correlación ítem-total), fiabilidad, homogeneidad excluyendo el ítem y coeficiente alpha que quedaría en la escala si se excluyera cada ítem, para cada ítem de la escala de respuesta de los superiores. Los items están escalados de 0 a 1, indicando 1 el máximo de respuesta segura.

ITEM	Media	D. T.	Homogeneidad	Fiabilidad	Homogen. sin el ítem	Alpha sin el ítem
1	0.534	0.423	.780	.330	.641	.808
2	0.391	0.238	.556	.132	.444	.834
3	0.639	0.401	.790	.317	.666	.801
4	0.221	0.309	.664	.205	.538	.821
5	0.442	0.305	.810	.247	.730	.793
6	0.524	0.388	.765	.297	.635	.807
7	0.596	0.194	.640	.124	.563	.826

La eliminación de ninguno de los items provocaría una mejora en el alfa de la escala, pero quitar el item dos no supondría una pérdida de consistencia interna, a pesar de la pérdida en el número de items que suele bajar el alfa.

Mirando la fiabilidad desde los resultados del análisis factorial confirmatorio podemos corroborar o matizar algo de lo visto en el análisis clásico de items. En las dos tablas siguientes pueden verse los resultados de fiabilidad desde las dos estructuras hipotetizadas que respondían bien a los datos, la unifactorial y la bifactorial respectivamente. Se ha calculado el cuadrado de la saturación factorial de cada ítem en cada estructura y por cada método de estimación como medida de la fiabilidad. Los resultados son paralelos a los hallados mediante el análisis de items. El ítem dos es explicado en torno al 20% por el factor hipotetizado, siendo el resto de explicaciones de los items satisfactorias.

Tabla 5.8. Saturaciones factoriales de la estructura unifactorial estimadas mediante mínimos cuadrados generalizados, máxima verosimilitud y estimación arbitraria elevados al cuadrado, como medida de la fiabilidad de los items. Los resultados indican la proporción de la varianza de cada ítem explicada por los factores latentes.

	GLS	ML	AGLS
Item 1	.505	.461	.591
Item 2	.226	.230	.188
Item 3	.558	.546	.524
Item 4	.367	.345	.396
Item 5	.715	.511	.722
Item 6	.558	.528	.567
Item 7	.379	.363	.473

Las diferencias entre los métodos de estimación son mínimas, como ya vimos al comentar las saturaciones factoriales, ligeramente aumentadas aquí por ser valores al cuadrado. Este resultado se mantiene para las dos estructuras factoriales.

*Tabla 5.9. Saturaciones factoriales estimadas mediante mínimos cuadrados generalizados, máxima verosimilitud y estimación arbitraria elevados al cuadrado, como medida de la fiabilidad de los ítems, para la estructura de dos factores. Los resultados indican la proporción de la varianza de cada ítem explicada por los factores latentes.*

	GLS	ML	AGLS
Item 1	.552	.524	.589
Item 2	.240	.239	.198
Item 3	.568	.537	.546
Item 4	.407	.391	.432
Item 5	.741	.708	.732
Item 6	.567	.559	.573
Item 7	.412	.412	.495

Por último, resaltar que entre la estructura hipotética de un factor y la de dos factores no se produce una ganancia significativa en la proporción de varianza de los ítems. Algunos ítems se explican mejor desde la unifactorial y otros desde la bifactorial, pero las diferencias son de centésimas.

### 1. 5. Validez criterial

En la tabla 5.10 pueden verse los resultados de correlacionar los siete ítems de la escala de respuesta con los principales criterios. Dado que se ha conceptualizado la respuesta de los supervisores como una variable clave en la unión entre los trabajadores y la empresa, se espera que los ítems del cuestionario consigan elevadas correlaciones con las medidas de estos aspectos. En concreto, las hipótesis de validez que se plantean son:

- Correlación positiva y elevada con la variable política organizacional. Los supervisores son los que directamente llevan a cabo o controlan muchas de las acciones, políticas, normativas, sanciones, recompensas, etc, que pone en marcha la empresa. Se espera, por tanto, que el trabajador que contesta el cuestionario vea cierta sintonía entre la actuación del supervisor y la política de la empresa en materia de seguridad.

- Correlación positiva y elevada con formación. En muchos casos son los supervisores los que controlan u ofrecen la formación de seguridad y, en cualquier caso, aunque no sean los responsables directos lo cierto es que la formación en seguridad es parte

Tabla 5.5. Índices de ajuste global de la estructura factorial de segundo orden hipotetizada para la escala de respuesta de los superiores. Los índices son: *chi*=chi-cuadrado; *g*=grados de libertad del modelo; *p*=probabilidad; *F*=función de ajuste a minimizar; *LHR*=razón de verosimilitud escalada; *CAK*=Akaike information criterion; *CSK*=Schwarz information criterion; *NFI*=normed fit index; *NNFI*=non-normed fit index; *CFI*=comparative fit index. Método de estimación: *GLS*=mínimos cuadrados generalizados.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
2º ORDEN	GLS	8.76	13	.790	.085	.958	.673	.392	.799	.994	1.0	1

Curiosamente los valores de ajuste global son los mismos que los hallados para el caso de estimación mínimo cuadrática en el factorial de primer orden. Los datos de estimación de las saturaciones factoriales y de los errores de la estructura que relaciona las variables observables con los factores de primer orden son asimismo iguales. Ofrecemos los valores de las saturaciones factoriales entre los factores de primer orden y el de segundo orden que son, para el primer factor de primer orden 0.923 ( $p < .01$ ), y para el segundo factor de primer orden 0.956 ( $p < 0.01$ ).

#### 1. 4. Fiabilidad de la escala y de los ítems

La escala muestra en su conjunto una buena consistencia interna, por todos los coeficientes. Así, un coeficiente alfa para el total de 0.836 indica una elevada covariación entre los ítems. Sólo el alpha para los ítems pares aparece un poco bajo, debido probablemente al menor número de ítems (tres frente a cuatro impares) y a que incluye dos de los ítems menos homogéneos, el 2 y el 4, ver tabla 5.7.

Tabla 5.6. Coeficientes de fiabilidad, consistencia interna, de la escala de respuesta de los superiores.

<i>Consistencia interna</i>	
Correlación dos mitades	.729
Coefficiente Spearman-Brown	.843
Coefficiente Guttman-Rulon	.804
Coefficiente alpha total	.836
Alpha ítems impares	.780
Alpha ítems pares	.591

Los datos del análisis de ítems se presentan en la tabla 5.7. Como ya venía ocurriendo en el análisis factorial, el ítem dos es el menos homogéneo con el total de la escala, alcanzando una correlación con el total de 0.556 (excluido el ítem de 0.444). Aunque no puede hablarse de una homogeneidad excesivamente baja, lo cierto es que es el de peor funcionamiento en este aspecto. Además y a nivel descriptivo, cabe destacar que los dos ítems que representan conducta de seguridad del supervisor ante la conducta segura del

trabajador (items 2 y 4) presentan la media más baja, mientras que otros items, como por ejemplo la conducta de seguridad del supervisor ante conductas inseguras del trabajador (ítem 7) presenta una media sustancialmente más alta, indicando mayor respuesta frente a la inseguridad que frente a la adopción de medidas seguras.

Tabla 5.7. Media, desviación típica, coeficiente de homogeneidad (correlación ítem-total), fiabilidad, homogeneidad excluyendo el ítem y coeficiente alpha que quedaría en la escala si se excluyera cada ítem, para cada ítem de la escala de respuesta de los superiores. Los ítems están escalados de 0 a 1, indicando 1 el máximo de respuesta segura.

ITEM	Media	D. T.	Homogeneidad	Fiabilidad	Homogen. sin el ítem	Alpha sin el ítem
1	0.534	0.423	.780	.330	.641	.808
2	0.391	0.238	.556	.132	.444	.834
3	0.639	0.401	.790	.317	.666	.801
4	0.221	0.309	.664	.205	.538	.821
5	0.442	0.305	.810	.247	.730	.793
6	0.524	0.388	.765	.297	.635	.807
7	0.596	0.194	.640	.124	.563	.826

La eliminación de ninguno de los items provocaría una mejora en el alfa de la escala, pero quitar el ítem dos no supondría una pérdida de consistencia interna, a pesar de la pérdida en el número de items que suele bajar el alfa.

Mirando la fiabilidad desde los resultados del análisis factorial confirmatorio podemos corroborar o matizar algo de lo visto en el análisis clásico de items. En las dos tablas siguientes pueden verse los resultados de fiabilidad desde las dos estructuras hipotetizadas que respondían bien a los datos, la unifactorial y la bifactorial respectivamente. Se ha calculado el cuadrado de la saturación factorial de cada ítem en cada estructura y por cada método de estimación como medida de la fiabilidad. Los resultados son paralelos a los hallados mediante el análisis de items. El ítem dos es explicado en torno al 20% por el factor hipotetizado, siendo el resto de explicaciones de los items satisfactorias.

Tabla 5.8. Saturaciones factoriales de la estructura unifactorial estimadas mediante mínimos cuadrados generalizados, máxima verosimilitud y estimación arbitraria elevados al cuadrado, como medida de la fiabilidad de los items. Los resultados indican la proporción de la varianza de cada ítem explicada por los factores latentes.

	GLS	ML	AGLS
Item 1	.505	.461	.591
Item 2	.226	.230	.188
Item 3	.558	.546	.524
Item 4	.367	.345	.396
Item 5	.715	.511	.722
Item 6	.558	.528	.567
Item 7	.379	.363	.473

Las diferencias entre los métodos de estimación son mínimas, como ya vimos al comentar las saturaciones factoriales, ligeramente aumentadas aquí por ser valores al cuadrado. Este resultado se mantiene para las dos estructuras factoriales.

*Tabla 5.9. Saturaciones factoriales estimadas mediante mínimos cuadrados generalizados, máxima verosimilitud y estimación arbitraria elevados al cuadrado, como medida de la fiabilidad de los ítems, para la estructura de dos factores. Los resultados indican la proporción de la varianza de cada ítem explicada por los factores latentes.*

	GLS	ML	AGLS
Item 1	.552	.524	.589
Item 2	.240	.239	.198
Item 3	.568	.537	.546
Item 4	.407	.391	.432
Item 5	.741	.708	.732
Item 6	.567	.559	.573
Item 7	.412	.412	.495

Por último, resaltar que entre la estructura hipotética de un factor y la de dos factores no se produce una ganancia significativa en la proporción de varianza de los ítems. Algunos ítems se explican mejor desde la unifactorial y otros desde la bifactorial, pero las diferencias son de centésimas.

### 1. 5. Validez criterial

En la tabla 5.10 pueden verse los resultados de correlacionar los siete ítems de la escala de respuesta con los principales criterios. Dado que se ha conceptualizado la respuesta de los supervisores como una variable clave en la unión entre los trabajadores y la empresa, se espera que los ítems del cuestionario consigan elevadas correlaciones con las medidas de estos aspectos. En concreto, las hipótesis de validez que se plantean son:

- Correlación positiva y elevada con la variable política organizacional. Los supervisores son los que directamente llevan a cabo o controlan muchas de las acciones, políticas, normativas, sanciones, recompensas, etc, que pone en marcha la empresa. Se espera, por tanto, que el trabajador que contesta el cuestionario vea cierta sintonía entre la actuación del supervisor y la política de la empresa en materia de seguridad.

- Correlación positiva y elevada con formación. En muchos casos son los supervisores los que controlan u ofrecen la formación de seguridad y, en cualquier caso, aunque no sean los responsables directos lo cierto es que la formación en seguridad es parte



importante del clima de seguridad y se espera la misma sintonía que ya se esperaba con política.

- Correlación moderada con estructura de seguridad. La estructura de seguridad no implica a los supervisores del trabajador, ya que es una estructura independiente. No obstante, por ser una parte del clima de seguridad se espera cierta correlación con la conducta del supervisor, más cuanto mayor sea la importancia real de la estructura en la seguridad de la empresa.

- Se espera que la respuesta de los supervisores esté fuertemente correlacionada con la conducta de los trabajadores, dado que son los directos responsables de su desempeño. Esto significa, con respecto a nuestras variables, que se espera una alta correlación con respuesta de los compañeros y con conducta hacia la seguridad de los trabajadores.

- La correlación con actitud hacia la seguridad es más compleja. Si recordamos, la medida de actitud es un cuestionario presentado por Leather (1988) es una medida de locus de control. Cuanta más puntuación en el cuestionario mayor control interno percibido por el sujeto. En nuestro caso de relación con los items de respuesta puede esperarse que cuanto mayor sea la respuesta de seguridad de los supervisores más implicados estén en fomentarla y el trabajador vea esto como un mayor control externo, como una mayor preocupación e implicación en seguridad de la empresa. Por ello, a mayor respuesta de los supervisores se espera mayor locus de control externo, esto es se espera una moderada correlación negativa.

- Con riesgo basal se espera correlación nula, dado que este tipo de riesgo es intrínseco al trabajo. Con riesgo real se espera relación nula o muy moderada, ya que creemos que afecta al riesgo real pero de forma indirecta a través de su efecto sobre la conducta del trabajador.

- Se espera relación negativa con tensión. Se espera que la tensión esté relacionada con el nivel de riesgos y con la percepción de facilidad de accidentarse. Si existe una elevada respuesta de seguridad de los supervisores esto puede provocar que el trabajador encuentre más ayuda al solucionar este aspecto disfuncional y esto rebaje su tensión. En cualquier caso esta es sólo una hipótesis inicial, y no es excesivamente importante en la validación criterial.

Tabla 5.10. Correlaciones de los items de la escala de respuesta de los superiores con los principales criterios. ES=estructura de seguridad; PO=política organizacional de seguridad; F=respuesta de los superiores en seguridad; RC=respuesta de los compañeros en seguridad; L=actitud hacia la seguridad; CHS=conducta hacia la seguridad; RB=riesgo basal; RR=riesgo real; T=tensión psicológica; y AC=sinestralidad. Símbolos de significatividad estadística de las correlaciones: n.s.= no significativo; #= marginalmente significativo; \*= probabilidad <0.05; y \*\*= probabilidad <0.01, extraídos de tablas (Welkowitz, Ewen y Cohen, 1981).

	ES	PO	F	RC	L	CHS	RB	RR	T	AC
1	.203*	.510**	.325**	.528**	-.12ns	.338**	.098ns	-.13ns	-.22*	-.04ns
2	.068ns	.366**	.317**	.437**	-.24**	.456**	-.09ns	-.17ns	-.23*	-.17ns
3	.128ns	.603**	.367**	.441**	-.37**	.355**	-.10ns	-.19*	-.25**	-.03ns
4	.228*	.618**	.409**	.434**	-.25**	.252**	-.07ns	-.14ns	-.25**	-.14ns
5	.207*	.645**	.453**	.463**	-.23*	.391**	-.19*	-.20*	-.33**	-.14ns
6	.223*	.555**	.437**	.468**	-.28**	.407**	-.08ns	-.29**	-.35**	-.20*
7	.151ns	.481**	.364**	.549**	-.31**	.351**	-.15ns	-.24**	-.12ns	-.07ns

Todas las hipótesis de relación se cumplen razonablemente. Las más importantes son las que relacionan los tres aspectos de clima positivamente con los items de respuesta de los supervisores y, por otro lado, la relación con la respuesta de los trabajadores (respuesta de los compañeros y conducta de seguridad). Salvo la relación con estructura, que es relativamente baja, el resto de correlaciones de éstos criterios con respuesta de los supervisores son muy elevadas, mostrando que la respuesta de los supervisores cumple un papel crucial en la relación trabajador-empresa, también en los aspectos de seguridad.

### 1. 6. Validez convergente-discriminante

Para el trabajador, ya hemos comentado que el portador directo de la mayoría de acciones de la empresa es el supervisor. Aún cuando nosotros hemos diferenciado entre la política de seguridad de la empresa y respuesta de los superiores como variables discernibles y de diferente nivel, puede que los trabajadores "engloben" ambas, considerándolas igualmente. De hecho, muchos autores conceptualizan ambos constructos como parte del clima de seguridad, junto con otros aspectos, y encuentran que forman parte del mismo factor (por ejemplo Dedobbeleer y Béland, 1991). Lo que nos preguntábamos es si los trabajadores que contestan al cuestionario diferencian, como nosotros, entre ambos conceptos y hemos sometido los items de la escala de política y los de la de respuesta de los superiores a un análisis factorial confirmatorio, para evaluar la convergencia-discriminancia de los items.



Tabla 5.11. Saturaciones factoriales de los quince items de las escalas de política y respuesta de los superiores en los dos factores hipotetizados (política y respuesta de los superiores) para el estudio de la validez convergente-discriminante (\*= probabilidad <.05; \*\*= probabilidad <.01).

Nº Item y factor teórico	$\lambda$ con PO	$\lambda$ con RS
Item 1 (política)	.559**	0
Item 2 (política)	.283*	0
Item 3 (política)	.302**	0
Item 4 (política)	.631**	0
Item 5 (política)	.458**	0
Item 6 (política)	.921**	0
Item 7 (política)	.658**	0
Item 8 (política)	.717**	0
Item 9 (respuesta)	0	.779**
Item 10 (respuesta)	0	.574**
Item 11 (respuesta)	0	.841**
Item 12 (respuesta)	0	.698**
Item 13 (respuesta)	0	.852**
Item 14 (respuesta)	0	.770**
Item 15 (respuesta)	0	.686**

Se hipotetizan dos factores subyaciendo a los items de ambas escalas, uno que representa la política organizacional y otro la respuesta de los superiores. La hipótesis de convergencias se concreta en que los ocho items de la escala de política deben saturar significativamente en el factor de política, mientras los siete de respuesta de los superiores deben saturar significativamente en el de respuesta de los superiores. La hipótesis de discriminancia de los items se concreta en que las saturaciones factoriales con los factores contrarios deben ser nulas. En la tabla 5.11 se pueden consultar las saturaciones factoriales resultado del análisis factorial confirmatorio. Como puede verse se satisfacen todas las hipótesis, y esto con un resultado de ajuste global satisfactorio:  $\chi^2=105.244$  con 89 grados de libertad ( $p=0.11512$ ), NFI=0.972, NNFI=0.995 y CFI=0.996. Por otra parte, aunque los items cargan claramente en los factores hipotetizados y no se ha producido confusión de contenidos los factores se encuentran altamente correlacionados (0.966), lo que se esperaba por hipótesis. A la vista de los resultados, puede mantenerse que no existe confusión de contenidos entre los constructos política organizacional y respuesta de los superiores, a pesar de ser variables muy correlacionadas.

### **1. 7. Conclusiones**

Los análisis realizados sobre los ítems de la escala de respuesta de los supervisores muestran una escala psicométricamente adecuada. Los resultados de estructura factorial muestran las altas covariaciones de los ítems. Se han probado estructuras unifactoriales, bifactoriales y trifactoriales, y todas han mostrado buenos ajustes, tanto globales como analíticos. Por motivos tanto teóricos como de resultados estadísticos parece que lo más razonable es mantener una estructura unifactorial. Un único factor es capaz de explicar tan bien la varianza de los ítems como estructuras más complejas y esto con una parsimonia mayor. Los análisis de ítems muestran una adecuada homogeneidad en todos los ítems. El ítem dos es el que resulta menos homogéneo con el total de la escala. No obstante, hay que señalar que resulta ser el de menor desviación pudiendo bajar ligeramente por este motivo la correlación. Si se tuviera que modificar la escala para conseguir una mayor homogeneidad podría eliminarse el ítem dos. Además es el que es peor explicado desde el factor latente hipotetizado en los factoriales confirmatorios. No obstante, hemos visto que el efecto que podría tener en el alfa no sería relevante y aunque la fiabilidad no es alta su relación con el factor es estadísticamente significativa. El contenido del ítem rastrea la conducta de seguridad del supervisor y es teóricamente relevante, por lo que no debería prescindirse de él. En cuanto a la validez criterial de los ítems hemos visto que el funcionamiento de todos los ítems, incluyendo el ítem dos, es conforme a hipótesis y han mostrado su relevancia en la cadena causal que puede concluir en un accidente. Por último, el análisis de validez convergente-discriminante muestra que los contenidos no están confundidos con un constructo con el que existían posibilidades de encontrarse confundido. En conjunto la escala tiene un muy buen comportamiento psicométrico y recortes en ella no estarían justificados, sobre todo a nivel teórico.

## 2. Respuesta de los Compañeros

### 2.1. Introducción y medida

Diversos son los autores que consideran a los compañeros como grupo social relevante en el estudio de los accidentes laborales. Hemos visto cómo la empresa, a través del clima, y los supervisores inmediatos del trabajador pueden afectar el comportamiento seguro-inseguro del trabajador, evidentemente entre los grupos de más influencia en el contexto laboral pueden contarse los compañeros. Petersen (1984) piensa que las presiones y normas no escritas del grupo de trabajo son el principal determinante de la conducta del trabajador. P. J. Leather (1987) piensa que las influencias que determinan la conducta son múltiples y diversos los agentes de esa influencia, pero entre ellos destaca también a los supervisores y el grupo de compañeros.

Lo que se pretende con la escala que ahora presentamos es medir la respuesta de los compañeros, específicamente en materia de seguridad, para posteriormente poder evaluar su impacto sobre la conducta del trabajador. Esta formada por un total de diez items que tratan de medir los mismos tres contenidos que se pretende, medir en la escala de respuesta de los supervisores pero refiriéndolos esta vez a los compañeros. Esto es: la actitud de los compañeros en seguridad, la conducta de seguridad que exhiben al desempeñar su trabajo, y la respuesta que ofrecen a los actos seguros o inseguros del trabajador.

#### ESCALA:

Item 1. ¿Qué respuesta recibe de sus compañeros cuando realiza su trabajo de forma peligrosa o arriesgada?

Me critican y me piden que trabaje de forma más segura.....	1
Recibo comentarios negativos.....	2
Les es igual y No me dice nada. ....	3
Me admiran por mi valor .....	4

Item 2. Cuando trabajo de forma arriesgada o peligrosa mis compañeros me alaban por mi valentía..... V F

Item 3. ¿Se esfuerzan sus compañeros trabajando con prácticas seguras?

Nunca u Ocasionalmente.....	1
Algunas veces.....	2
Muy Frecuentemente .....	3

## Capítulo 5. Variables de contexto social.

---

Item 4. Cuando trabajo peligrosamente mis compañeros se preocupan y me llaman la atención.....Sí No

Item 5. ¿Que actitud tienen sus compañeros hacia la seguridad en sus trabajos?

Negativa o Indiferente ..... 1  
Positiva..... 2  
Muy Positiva..... 3

Item 6. Cuando trabajo arriesgadamente mis compañeros no dicen nada o se muestran indiferentes..... V F

Item 7. ¿Con qué frecuencia hablan Ud. y sus compañeros sobre las medidas de seguridad que deben tomarse?

Nunca u Ocasionalmente..... 1  
Algunas veces..... 2  
Muy Frecuentemente ..... 3

Item 8. Cuando trabajo de forma insegura mis compañeros me dicen que trabaje con cuidado..... V F

Item 9. ¿Apoyan sus compañeros el cumplimiento de las normas de seguridad?

Nunca u Ocasionalmente.....1  
Casi siempre.....2  
Siempre lo hacen.....3

Item 10. ¿Qué respuesta recibe de sus compañeros cuando realiza su trabajo tomando medidas de seguridad?

Recibo críticas o comentarios negativos..... 1  
Dicen que no hace falta tanta medida de seguridad ..... 2  
Le es indiferente y No me dicen nada ..... 3  
Les parece bien y recibo comentarios positivos..... 4  
Me ayudan y me facilitan que trabaje de forma muy segura..... 5

### 2.2. Estructura factorial confirmatoria

Se han planteado tres posibles modelos factoriales para representar la estructura del cuestionario, dos de ellos teóricos y un tercero que se extrae de las sugerencias del LM test y del W test.

1) Un modelo unifactorial, que nos permite conocer hasta qué punto la escala puede representarse razonablemente con un único constructo, y que además permite su uso como modelo basal o competitivo, a partir del cuál evaluar estructuras más complejas.

2) Una estructura bifactorial, que básicamente recogería en el primer factor la respuesta de los compañeros ante los actos seguros-inseguros del trabajador que contesta la escala y un segundo factor que cuestiona sobre el desempeño seguro o inseguro de los compañeros y su actitud (que se muestra en su conducta). El primer factor incluiría los ítems 1, 2, 4, 6, 8, y 10. El segundo los cuatro restantes, 3, 5, 7, y 9. Se consideran los dos factores como correlacionados.

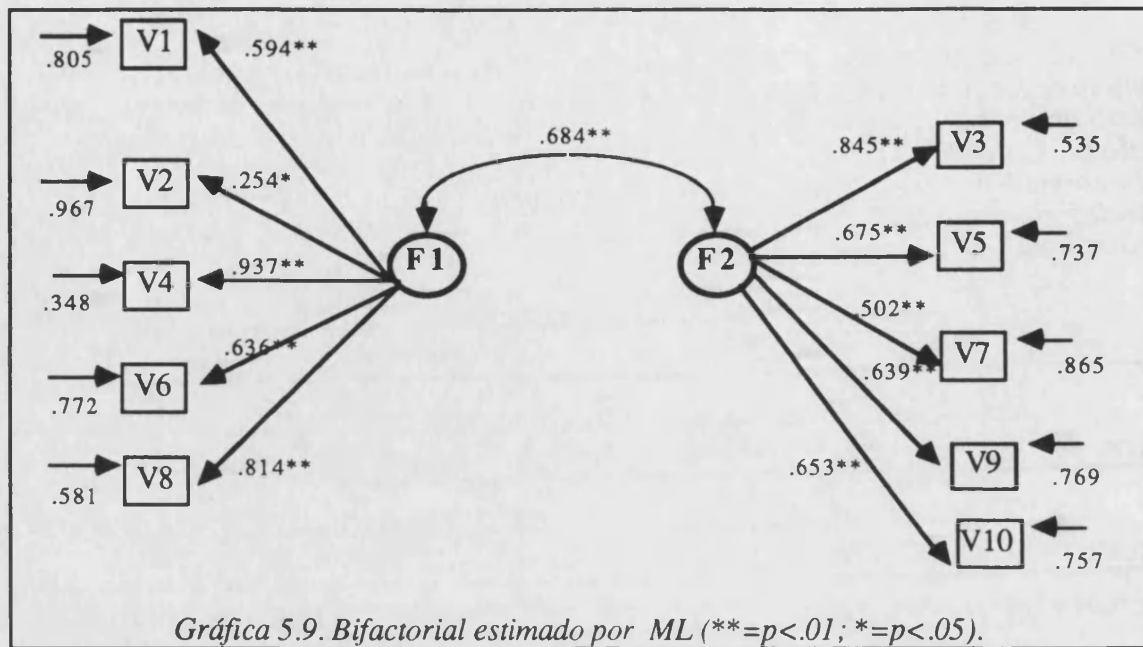
3) Una segunda estructura bifactorial producida en base a sugerencias estadísticas. Se conservarían los dos factores, pero un ítem cambia de adscripción. El ítem 10 pasa a ser explicado por el segundo factor. El proceso a través del cual se produjo el cambio fue en dos fases. Al calcular la estructura bifactorial teórica el LM test propuso la inclusión de un nuevo parámetro libre, a estimar, que se correspondía con la relación del factor 2 con el ítem 10. Así, se calculó esta estructura con el ítem 10 explicado por los dos factores. Este análisis factorial resultó en un ajuste relativamente adecuado y no lo presentamos en la tabla, junto con las tres estructuras señaladas, por que tan solo fue un paso intermedio. Todas las relaciones eran estadísticamente significativas salvo la del ítem 10 con el factor 1, que se había planteado teóricamente. Así, el W test mostraba que si se asumía cero no se perdería en el ajuste, y aumentarían los grados de libertad del modelo, por lo que el ajuste mejoraría previsiblemente, al menos en cuanto a la parsimonia. Así quedó establecida la estructura de dos factores ambos con cinco ítems a explicar, el primero con los ítems 1, 2, 4, 6 y 8; y el segundo con los ítems 3, 5, 7, 9, y 10.

Tabla 5.12. Índices de ajuste global de las tres estructuras hipotetizadas para la escala de respuesta de los compañeros. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; CAK=Akaike information criterion; CSK=Schwarz information criterion; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: GLS=mínimos cuadrados generalizados; ML=máxima verosimilitud.

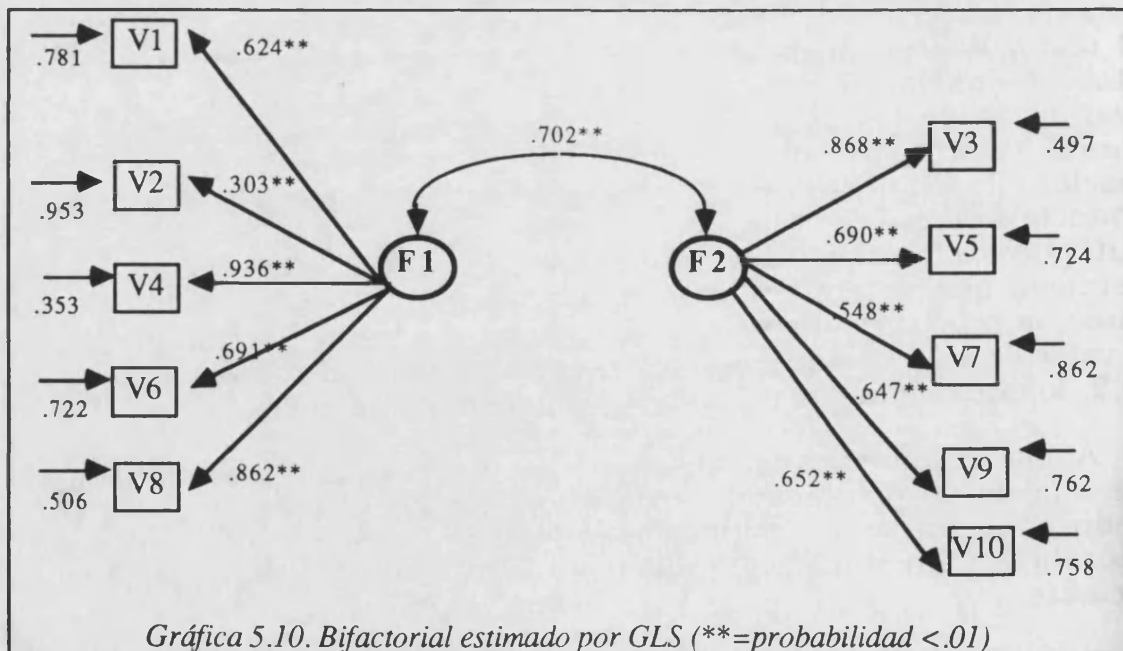
		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
UN	GLS	62.5	35	.002	.625	.731	1.78	1.02	1.53	.980	.988	.991
FACTOR	ML	100.	35	<.01	1.00	.603	2.88	1.39	1.91	.752	.766	.818
DOS	GLS	54.7	34	.013	.547	.760	1.61	.962	1.50	.982	.991	.993
FACTORES	ML	66.5	34	<.01	.665	.717	1.95	1.08	1.62	.837	.881	.910
2 FACTORES	GLS	44.1	34	.113	.441	.801	1.29	.856	1.40	.986	.996	.997
modificado	ML	44.1	34	.114	.441	.802	1.29	.856	1.40	.892	.963	.972

A la vista de los resultados de ajuste global del modelo (tabla 5.12) es la estructura bifactorial la que mejor representa los datos. El resto de modelos no ajustan por la prueba de ji-cuadrado, aunque sí por otros índices. Así, tanto el modelo unifactorial como el bifactorial teórico puede considerarse que tienen buen ajuste por la prueba de ji-cuadrado partido los grados de libertad y, en estimación mínimo cuadrática, por los índices de comparación con un modelo nulo. No obstante, al ver los datos de ajuste del modelo modificado en base al LM test y W test queda claro que es la mejor representación de los datos. Todos los índices muestran un elevada explicación de las covarianzas de los datos y un modelo parsimonioso. El índice LHR alcanza valores elevados, al igual que los de comparación con un modelo nulo (NFI, NNFI y CFI). Además tanto CAK como CSK indican ganancia frente a los otros dos modelos, si bien las diferencias son aquí poco importantes cuantitativamente a pesar de la elevada diferencia que se aprecia en la probabilidad de la ji-cuadrado. Por último, la prueba de ji-cuadrado partido los grados de libertad alcanza un valor de 1.29 bastante por debajo de los valores de ajuste mínimos 3 y 2, lo que nos indica un modelo altamente parsimonioso.

Ahora hay que prestar atención a los datos de ajuste analítico de este modelo de dos factores modificado. En las gráficas 5.9 y 5.10 pueden consultarse las estimaciones máximo verosímiles y mínimos cuadráticas para el modelo, junto con los datos de sus probabilidades asociadas.



Por un lado, y de forma general, puede apreciarse cómo en la estimación máximo verosímil todos los ítems se encuentran significativamente relacionados con los dos factores hipotetizados. Si observamos la cuantía de las saturaciones factoriales veremos que la práctica totalidad son elevadas, especialmente las correspondientes a los ítems 4, 8 (factor 1) y 3 (factor 2). El único ítem que no es bien explicado por su factor es el ítem 2.



Lo mismo que hemos visto que ocurría para la estimación máximo verosímil ocurre también para la estimación por mínimos

cuadrados generalizados. Todos los ítems se encuentran relacionados estadísticamente ( $p < 0.01$ ) con sus factores, pero ahora las saturaciones se han visto en líneas generales aumentadas, lo que resulta especialmente importante en el caso del ítem 2.

El siguiente paso en el análisis consiste en analizar si puede mantenerse el supuesto de normalidad multivariada y, caso de que no se mantenga, aplicar estimación por métodos arbitrarios para ver de confirmar las relaciones encontradas por ML y GLS. Como primer paso para comprobar el supuesto de normalidad multivariada se procederá a someter a las variables observables de la escala, los ítems, a sendas pruebas de Kolmogorov-Smirnov. Las pruebas de ajuste a una ley normal tienen un carácter univariado, son variable a variable, pero dado que la normalidad univariada es condición necesaria para la multivariada puede servirnos para la comprobación. En la tabla 5.13 pueden consultarse los cálculos de ajuste a la normalidad univariada.

Tabla 5.13. Prueba de Kolmogorov-Smirnov de una muestra para comprobar el ajuste de los ítems de la escala de respuesta de los compañeros a la distribución normal univariada.

ITEM	N muestral	Diferencia máxima	Probabilidad
1	117	.622	.000
2	117	.773	.000
3	120	.500	.000
4	120	.616	.000
5	121	.510	.000
6	117	.602	.000
7	122	.500	.000
8	116	.660	.000
9	120	.508	.000
10	118	.556	.000

Ninguno de los ítems de la escala se ajusta razonablemente a la distribución normal, lo cual es lógico teniendo en cuenta que la escala de respuesta es corta y los ítems son en todo caso de escala ordinal.



Tabla 5.14. Datos de asimetría y curtosis para los items de la escala de respuesta de los compañeros.

	Asimetría	Curtosis
Item 1	-.643	-1.21
Item 2	-3.39	9.50
Item 3	-0.27	-1.20
Item 4	-1.29	-0.313
Item 5	-0.00	-0.410
Item 6	-1.17	-0.631
Item 7	0.116	-1.03
Item 8	-1.596	0.547
Item 9	-0.304	-1.074
Item 10	-0.649	-0.09

Los datos de asimetría y curtosis y el coeficiente de Mardia nos muestran los fuertes alejamientos de la normalidad que se producen en las variables.

COEFICIENTE DE MARDIA MULTIVARIADO= 34.62  
 Coeficiente esperado para la normal multivariada= 11.229

Así las cosas, se pensó en métodos arbitrarios de estimación. Los índices de ajuste global mostraban, sin embargo, que la estructura que sí ajustaba de forma muy adecuada en ML y GLS no tienen un ajuste tan claro por métodos arbitrarios. Sólo los índices de bondad de ajuste (GFI) y de bondad de ajuste ajustado (AGFI) indican buen ajuste. Todos los índices calculados pueden consultarse en la siguiente tabla.

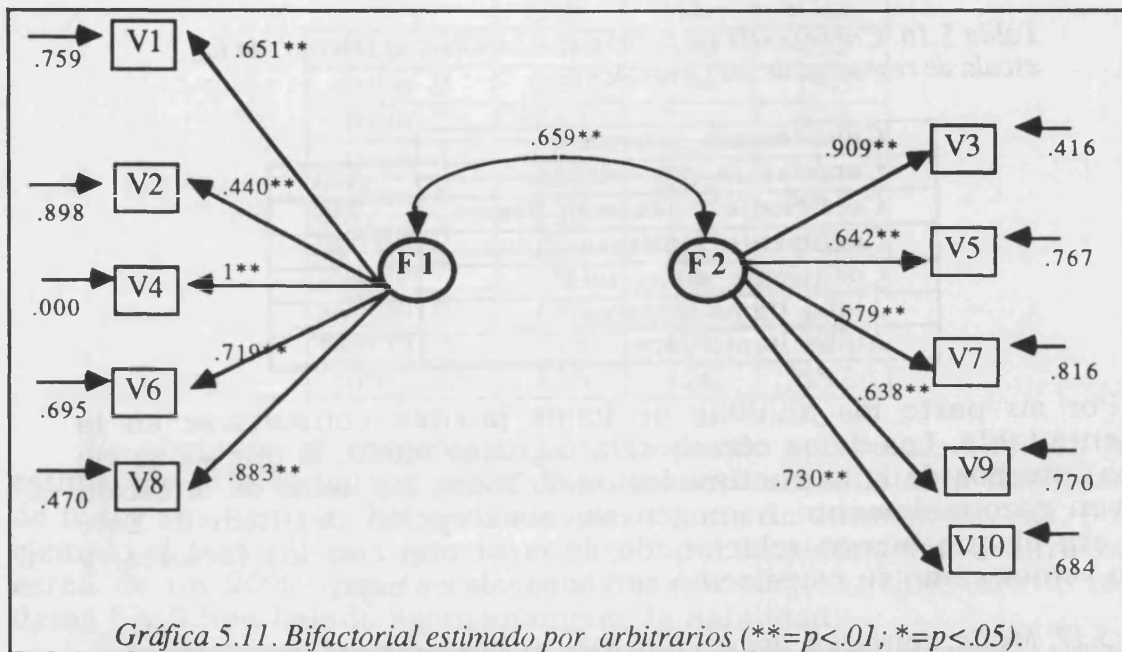
Tabla 5.15. Índices de ajuste global de la estructura bifactorial para la escala de respuesta de los compañeros. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Método de estimación AGLS=método arbitrario.

		chi	g	p	F	LHR	chi/g	NFI	NNFI	CFI
1 FACTOR	AGLS	115.	34	<.01	1.15	.561	3.4	.737	.726	.793

INDICE DE AJUSTE , GFI= .973  
 Índice de ajuste ajustado, AGFI= .957

A pesar de que el ajuste global no es satisfactorio, las estimaciones de las saturaciones factoriales y de las correlaciones

entre los factores son concordantes con lo hallado por ML y GLS, sólo que sistemáticamente más elevados. Todas las relaciones son estadísticamente significativas ( $p < 0.001$ ). Hasta el ítem 2 ha sufrido un aumento considerable en su relación con el factor, y algunos ítems son muy bien explicados por sus factores respectivos. Evidentemente los errores han disminuido considerablemente. En la gráfica 5.11 pueden consultarse todos estos datos.



### 2.3. Fiabilidad de la escala y de los ítems

En la tabla 5.16 pueden verse los datos de consistencia interna para la escala total. La homogeneidad de los ítems entre sí es elevada, como puede verse por la cuantía de los coeficientes. Por ejemplo, el coeficiente alfa para el total alcanza un 0.846, y también parece elevado el alfa para la mitad par y la impar.

Tabla 5.16. Coeficientes de fiabilidad, consistencia interna, de la escala de respuesta de los compañeros.

<i>Consistencia interna</i>	
Correlación dos mitades	.640
Coeficiente Spearman-Brown	.780
Coeficiente Guttman-Rulon	.780
Coeficiente alpha total	.846
Alpha ítems impares	.765
Alpha ítems pares	.757

Por su parte los análisis de ítems pueden consultarse en la siguiente tabla. Los datos corroboran, a grosso modo, lo que ya se vio en los estudios de la estructura factorial. Todos los ítems de la escala parecen razonablemente homogéneos, a excepción del ítem 2. Este ítem era el que menos relacionado se mostraba con los factores, y ahora vemos cómo su correlación con la escala es baja.

Tabla 5.17. Media, desviación típica, coeficiente de homogeneidad (correlación ítem-total), fiabilidad, homogeneidad excluyendo el ítem y coeficiente alpha que quedaría en la escala si se excluyera cada ítem, para cada ítem de la escala de respuesta de los compañeros.

ITEM	Media	D. T.	Homogeneidad	Fiabilidad	Homogen. sin el ítem	Alpha sin el ítem
1	0.772	0.273	.627	.171	.544	.832
2	0.931	0.254	.292	.074	.184	.856
3	0.579	0.377	.792	.299	.715	.814
4	0.772	0.419	.830	.348	.756	.808
5	0.505	0.311	.629	.195	.534	.832
6	0.752	0.432	.680	.294	.555	.832
7	0.460	0.357	.515	.184	.384	.846
8	0.812	0.391	.742	.290	.647	.821
9	0.594	0.364	.672	.244	.568	.829
10	0.644	0.264	.614	.162	.533	.834

Si calculamos la fiabilidad en lugar de por el análisis clásico de ítems por los resultados de los análisis factoriales, podremos comprobar una vez más el patrón de relación de los ítems. Es nuevamente el ítem 2 el de más baja fiabilidad para medir el factor.

*Tabla 5.18. Saturaciones factoriales de la estructura bifactorial modificada estimadas mediante mínimos cuadrados generalizados, máxima verosimilitud y estimación arbitraria elevados al cuadrado, como medida de la fiabilidad de los ítems. Los resultados indican la proporción de la varianza de cada ítem explicada por los factores latentes.*

	GLS	ML	AGLS
Item 1	.389	.352	.423
Item 2	.091	.064	.193
Item 3	.753	.714	.826
Item 4	.876	.877	1
Item 5	.476	.455	.412
Item 6	.477	.404	.516
Item 7	.300	.252	.335
Item 8	.743	.662	.779
Item 9	.418	.408	.407
Item 10	.425	.426	.532

No obstante, al contemplar la tabla puede apreciarse que con la estimación arbitraria, y a pesar de su inadecuado ajuste, la fiabilidad de todos sus ítems se ha visto considerablemente aumentada. Así, por ejemplo el factor 1 es capaz de explicar, por la estimación arbitraria, cerca de un 20%, que es una considerable mejora. Únicamente los ítems 5 y 9 han bajado ligerísimamente la fiabilidad.

#### **2.4. Validez criterial**

Las hipótesis de relación de la variable respuesta de los compañeros con los criterios son similares a las expuestas para respuesta de los superiores. Específicamente se espera una relación positiva y elevada con la respuesta de los superiores, dado que esta última se espera explique la conducta de los trabajadores, tanto del que contesta a los cuestionarios como de sus compañeros de trabajo. El segundo criterio, muy relevante también, es la conducta hacia la seguridad. Efectivamente lo que queremos es predecir los accidentes laborales y en un nivel justamente inmediato se encuentran los actos seguros-inseguros del trabajador. De esta forma, el interés de la variable respuesta de los compañeros es como predictora de la conducta de los trabajadores. Es por ello básico que los ítems se encuentren relacionados con la variable conducta hacia la seguridad. Menos relevantes son las relaciones con otras variables, pero sería interesante encontrar relaciones positivas, aunque moderadas, con política organizacional y negativas con accidentabilidad.

Tabla 5.19. Correlaciones de los ítems de la escala de respuesta de los compañeros con los principales criterios. ES=estructura de seguridad; PO=política organizacional de seguridad; F=respuesta de los superiores en seguridad; RS=respuesta de los superiores L=actitud hacia la seguridad; CHS=conducta hacia la seguridad; RB=riesgo basal; RR=riesgo real; T=tensión psicológica; y AC= sinestralidad. Símbolos de significatividad estadística de las correlaciones: n.s.= no significativo; #= marginalmente significativo; \*= probabilidad <0.05; y \*\*= probabilidad <0.01, extraídos de tablas (Welkowitz, Ewen y Cohen, 1981).

	ES	PO	F	RS	L	CHS	RB	RR	T	AC
1	.115ns	.287**	.139ns	.338**	-.16ns	.190*	-.10ns	-.02ns	-.178#	-.13ns
2	-.12ns	.080ns	.08ns	.114ns	-.13ns	.236*	-.07ns	-.09ns	-.11ns	-.198*
3	.087ns	.471**	.209*	.636**	-.36**	.499**	-.09ns	-.11ns	-.232*	-.15ns
4	.144ns	.303**	.210*	.467**	-.243*	.280**	-.07ns	-.09ns	-.212*	-.218*
5	.324**	.397**	.231*	.524**	-.233*	.395**	-.00ns	.00ns	-.200*	-.08ns
6	.073ns	.312**	.165#	.359**	-.185*	.326**	-.01ns	-.07ns	-.12ns	-.02ns
7	.214*	.321**	.196*	.246**	-.06ns	.176#	.01ns	-.06ns	-.09ns	-.13ns
8	.052ns	.288**	.134ns	.468**	-.220*	.331**	-.01ns	-.03ns	-.205*	-.00ns
9	.194*	.388**	.223*	.503**	-.220*	.369**	-.18*	-.05ns	-.25**	-.15ns
10	.154ns	.425**	.136ns	.498**	-.29**	.475**	-.04ns	-.184*	-.24**	-.176#

Todos los ítems se relacionan conforme a hipótesis con la variable respuesta de los superiores, a excepción del ítem 2. Este ítem, sin embargo, se muestra relacionado conforme a hipótesis con la variable conducta hacia la seguridad. Por otra parte, se muestra significativamente relacionada con la accidentabilidad, lo que resulta interesante, sobre todo teniendo en cuenta que sólo otro ítem de respuesta de los compañeros presenta este comportamiento. Además de estas relaciones básicas con respuesta de los superiores y con conducta hacia la seguridad la mayoría de los ítems muestran relaciones significativas con política organizacional, mostrando la capacidad predictiva que las acciones de la empresa pueden tener sobre la conducta de los trabajadores.

## 2.5. Conclusiones para la escala de respuesta de los compañeros

El ítem 2 era el de peor comportamiento psicométrico dentro de los análisis de estructura factorial y de consistencia interna. Parece que tampoco su comportamiento en validez es excepcional, más bien al contrario, por lo que podría quizá eliminarse. Sin embargo, el ítem se encuentra relacionado significativamente con conducta y con accidentabilidad y en la estimación por métodos arbitrarios sí se encontraba relacionado de forma bastante elevada con su factor, lo que elevaba su fiabilidad. El ítem, por otro lado, rastrea un contenido que consideramos relevante. Este contenido hace referencia a las alabanzas que reciben los trabajadores cuando realizan conductas peligrosas, que es una de las recompensas sociales más poderosas que pueden dar los compañeros ya que no suelen tener acceso a refuerzos materiales en el contexto laboral. Además sus bajas relaciones pueden deberse en parte a su corta escala de respuesta

(verdadero-falso). Por lo que respecta al resto de los items de la escala la conclusión general, que se extrae de los resultados y comentarios que se han ido desgranando hasta aquí, es clara: son todos ellos psicométricamente adecuados, fiables, válidos y con una adscripción clara a los factores que se pretenden medir.



## 6. Respuesta y Efectos sobre la seguridad del trabajador.

### **1. Introducción**

Se van a presentar en este punto diversas medidas de respuesta de los trabajadores y efectos de las variables anteriormente presentadas. Entre las respuestas de los trabajadores se considerarán dos: la conducta hacia la seguridad y la tensión. La tensión psicológica puede ser una importante consecuencia de los altos niveles de riesgo a que puede estar sometido el trabajador. Esta situación de tensión producida por los riesgos o por otras condiciones de trabajo, como el ritmo de trabajo, la mala actitud de los compañeros o los supervisores, etc puede a su vez tener importantes consecuencias. Efectivamente, un trabajador sometido a elevados niveles de tensión puede no responder de la misma forma ante las medidas de seguridad necesarias, y variables como riesgo real, conducta hacia la seguridad y, en definitiva, el número y gravedad de los accidentes de trabajo. Por su parte la variable conducta hacia la seguridad pretende medir el nivel de conducta seguras e inseguras que realiza el trabajador en su puesto de trabajo. Teniendo en cuenta que según diversos expertos el 90% de los accidentes se encuentran provocados por un tipo u otro de factor humano, se espera que la conducta hacia la seguridad sea capaz de explicar un parte



importante de los accidentes laborales. Por otra parte, se espera poder entender qué variables son capaces de predecir esas conductas seguras o inseguras del trabajador.

Por último se va a presentar un 'producto organizacional', la siniestralidad, producido por todas las condiciones físicas, sociales y personales del medio de trabajo y que es, junto con los otros productos un output del sistema organizacional. un producto indeseable, pero un producto. Buscaremos índices de siniestralidad que sean sensibles y adecuados para el contexto metodológico de modelos causales en que se enmarca este trabajo para poder comprobar adecuadamente su asociación con otras variables estudiadas.

## 2. Tensión psicológica

Para la medición de la tensión psicológica en el puesto de trabajo vamos a utilizar una versión reducida, de ocho items, de una escala presentada por Meliá et al. (1987) que originalmente tenía quince items. Se ha intentado que la reducción en siete items no supusiera una reducción paralela en el contenido que rastrea la escala y creemos que el constructo tensión psicológica queda bien representado por los items elegidos. Los ocho items que forman la versión reducida de la escala se presentan a continuación.

### 2.1. Medida

#### ESCALA:

Item 1. Me siento intranquilo e irritable a causa de mis jefes.....	V	F
Item 2. El nerviosismo que me produce mi trabajo afecta a mis relaciones familiares	V	F
Item 3. Mi trabajo me produce tensión y preocupación. ....	V	F
Item 4. El ritmo a que tengo que trabajar me fatiga y disgusta .....	V	F
Item 5. Me siento nervioso en mi trabajo .....	V	F
Item 6. Estoy ansioso, tenso o preocupado a causa de mi trabajo.....	VF	
Item 7. Estoy relajado y tranquilo mientras trabajo .....	V	F
Item 8. Me irrito con facilidad a causa de mi trabajo .....	V	F

### 2.2. Estructura factorial confirmatoria

A la luz de los contenidos teóricos que representan los ocho items de la escala se han planteado dos estructuras factoriales. La primera una estructura unifactorial, que como siempre nos sirve de modelo competitivo para representar los datos, y que además es en este caso una hipótesis razonable. La segunda estructura hipotetizada es una estructura de dos factores latentes. El primer factor incluye los items 1, 3, 4, 6 y 7, y se denominará causas de tensión, ya que los contenidos hacen referencia a posibles causantes de la tensión como el ritmo de trabajo, los jefes, etc. El segundo factor explicaría los items 2, 5 y 8, que se refieren a efectos de la tensión, como malas relaciones familiares, irritabilidad, nerviosismo, etc. Los dos factores se hipotetizan naturalmente correlacionados, ya que un elevado número de causantes de tensión conllevará probablemente un alto nivel de efectos de tensión.

Ambas estructuras fueron probadas en un primer momento mediante análisis factoriales confirmatorios estimados mediante máxima verosimilitud y mínimos cuadrados generalizados. Los datos de ajuste global de estas dos estructura por ambos tipos de estimación se pueden consultar en la tabla 6.1.

Tabla 6.1. Índices de ajuste global de las estructuras hipotetizadas para la escala de tensión psicológica. Los índices son: *chi*=chi-cuadrado; *g*=grados de libertad del modelo; *p*=probabilidad; *F*=función de ajuste a minimizar; *LHR*=razón de verosimilitud escalada; *CAK*=Akaike information criterion; *CSK*=Schwarz information criterion; *NFI*=normed fit index; *NNFI*=non-normed fit index; *CFI*=comparative fit index. Métodos de estimación: *GLS*=mínimos cuadrados generalizados; *ML*=máxima verosimilitud.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
UN	GLS	27.0	20	.132	.246	.884	1.354	.534	.924	.981	.993	.995
FACTOR	ML	1.67	20	.046	.287	.866	1.583	.575	.965	.884	.934	.953
DOS	GLS	25.4	19	.145	.231	.890	1.34	.537	.952	.982	.993	.995
FACTORES	ML	30.0	19	.044	.278	.870	1.582	.584	.992	.888	.930	.953

Ambas estructuras son adecuadas según los datos de ajuste global. Así, ambas pruebas de ji-cuadrado son no significativas en la estimación por mínimos cuadrados generalizados, mientras que para la estimación de máxima verosimilitud nos encontramos al borde de la significatividad. El resto de índices ofrecen resultados de ajuste muy adecuados para los dos tipos de estimación. Así, los cuatro índices de verosimilitud escalada están muy cercanos al uno que indica ajuste perfecto, y son muy bajos (indicando buen ajuste) los índices CAK y CSK. La prueba de ji-cuadrado partido los grados de libertad, con valores siempre por debajo de dos indica que los dos modelos presentan un alto grado de parsimonia. Siempre ligerísimamente superiores son los datos de ajuste global para la estructura bifactorial que para la de un factor. Queda por determinar si la ganancia que se produce de hipotetizar dos factores latentes a uno es estadísticamente significativa o es una complicación innecesaria de la estructura. Para determinar este último punto se han realizado dos pruebas de las diferencias de ji-cuadrado, una para cada tipo de estimación.

Para la estimación mediante mínimos cuadrados generalizados tenemos:

$$\Delta\chi^2 = 1.62, \text{ con } \Delta g = 1 \text{ que no resulta estadísticamente significativa (} p > 0.1 \text{),}$$

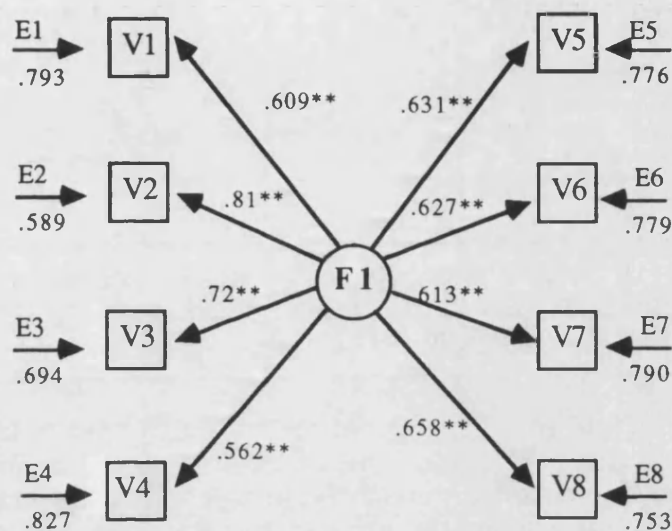
y para máxima verosimilitud:

$$\Delta\chi^2 = 1.61, \text{ con } \Delta g = 1 \text{ que tampoco resulta significativa (} p > 0.1 \text{).}$$

Ante los datos de las pruebas de diferencias de ji-cuadrado y los buenos resultados de ajuste global, el modelo unifactorial parece el que

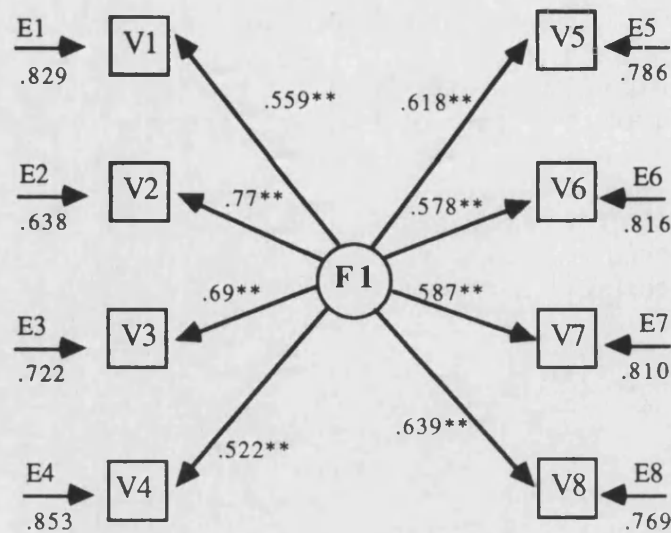
mejor representa los datos con la máxima parsimonia. Evidentemente esta conclusión surge a partir de los datos de ajuste global, y conviene estudiar también en profundidad los datos de ajuste analítico para ver que estructura ofrece mejores resultados ítem a ítem.

En las gráficas 6.1 y 6.2 aparecen los datos de las saturaciones factoriales, los errores y las pruebas z correspondientes para la estructura unifactorial.



Gráfica 6.1. Estructura factorial estimada mediante mínimos cuadrados generalizados para la escala de tensión psicológica. Ajuste analítico: \*\*= $p < .01$ ; \*= $p < .05$ ; n.s.= no significativo.

Todos los ítems son explicados significativamente por el factor tensión. Para la estimación por mínimos cuadrados generalizados la saturación más baja vale 0.609, lo que significa un apreciable 37% de varianza explicada. También para la estimación máximo verosímil los ítems se encuentran estadísticamente relacionados con el factor latente de tensión psicológica. Aquí, los valores de las saturaciones factoriales bajan ligeramente y la saturación más baja, para el ítem 1 igual que en mínimos cuadrados generalizados, se sitúa en 0.559, que en términos de porcentaje de varianza explicada representa un 31%. Las diferencias que se aprecian entre las saturaciones factoriales son ligeras, y no implican cambios de significatividad estadística, como ya vienen ocurriendo a través de los distintos análisis. Las diferencias, siempre a la baja en máxima verosimilitud, se sitúan en las centésimas.

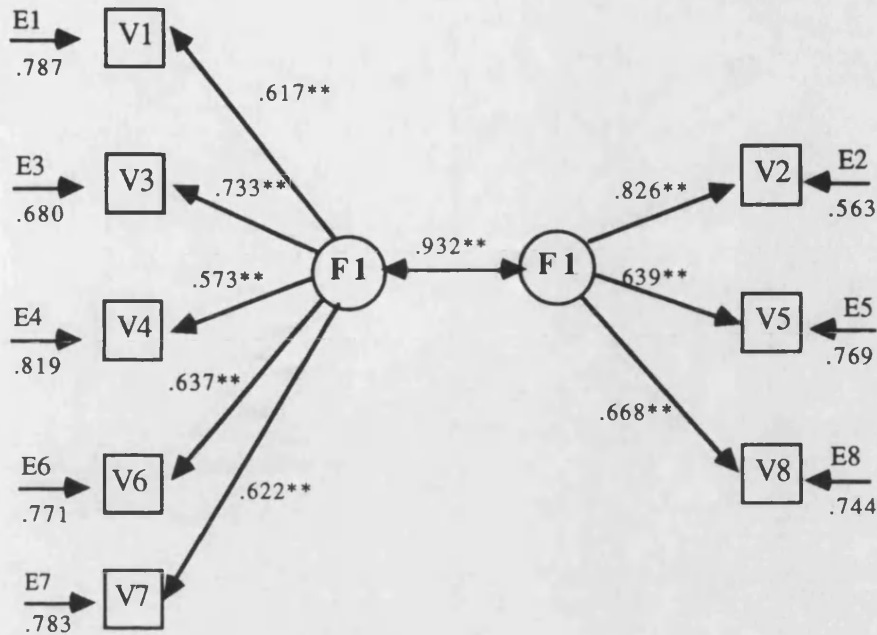


Gráfica 6.2. Estructura factorial estimada mediante máxima verosimilitud para la escala de tensión psicológica. Ajuste analítico: \*\*= $p < .01$ ; \*= $p < .05$ ; n.s.= no significativo.

Por otro lado, lo más interesante es comprobar las posibles diferencias que en la explicación de los items podían ocurrir entre las estructuras unifactoriales y las bifactoriales. En las gráficas 6.3 y 6.4 se encuentran los datos de ajuste analítico de las estructuras bifactoriales.

En la gráfica 6.3, a continuación, se pueden consultar las saturaciones factoriales para los ocho items de la escala estimadas mediante mínimos cuadrados generalizados. Al igual que ocurría en la estructura unifactorial todos los items se encuentran significativamente relacionados con los factores latentes, con una probabilidad menor de 0.01.

Al margen de los datos de ajuste analítico, repasando simplemente los valores de las saturaciones apreciaremos que no existe prácticamente diferencias entre la estructura unifactorial y la bifactorial estimadas por mínimos cuadrados generalizados. Nuevamente las diferencias se sitúan en las centésimas, con lo que no se gana nada apreciable al explicar la varianza de los items desde la más compleja estructura de dos factores. Ni siquiera en uno sólo de los items se produce una ganancia significativa. Por otro lado, la correlación entre los dos factores es muy elevada, pudiendo ésto significar que son parte de un mismo constructo.

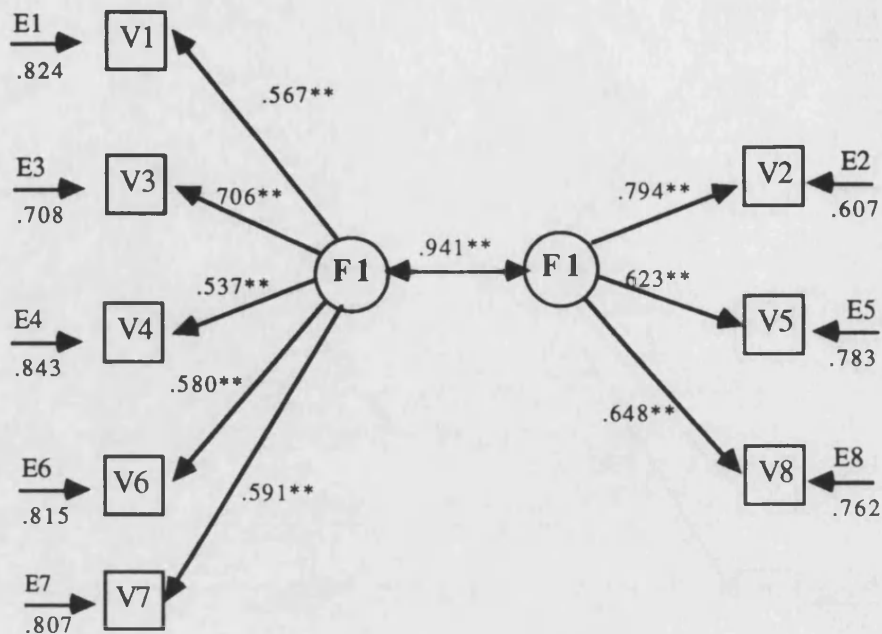


Gráfica 6.3. Estructura bifactorial estimada mediante mínimos cuadrados generalizados para la escala de tensión psicológica. Ajuste analítico: \*\*= $p < .01$ ; \*= $p < .05$ ; n.s.= no significativo.

Los mismos comentarios que ya hemos realizado en la comparación de las dos estructuras en mínimos cuadrados generalizados pueden aplicarse en la estimación máximo verosímil. Las diferencias entre las saturaciones factoriales en la estructura de un factor y las de la estructura bifactorial son apenas apreciables, representando poquísima ganancia en el porcentaje de varianza explicada. Nuevamente la correlación entre los dos factores latentes es elevadísima, 0.941.

Además, hay que comentar que ni el W test ni el LM test han ofrecido datos sobre modificaciones que pudieran mejorar el ajuste del modelo bifactorial. Esto es, ningún ítem es explicado por el factor no hipotetizado ni ninguna de las relaciones hipotetizadas es innecesaria.





Gráfica 6.4. Estructura bifactorial estimada mediante máxima verosimilitud para la escala de tensión psicológica. Ajuste analítico: \*\*= $p < .01$ ; \*= $p < .05$ ; n.s.= no significativo.

Tanto a la vista de los datos de ajuste global, como tras la revisión de las estimaciones concretas y los consiguientes ajustes analíticos parece claro que la estructura bifactorial, aunque sostenible es menos parsimoniosa que la unifactorial y no aporta ganancias relevantes. Por nuestro lado, disponer de un indicador unifactorial de la tensión es interesante, dado que en esta forma era como se planteaban las hipótesis de relación con otros constructos.

Como siempre, queda por determinar la idoneidad de las estimaciones realizadas. Todos los resultados de análisis factorial confirmatorio presentados hasta aquí han sido calculados con tipos de estimación basados en el supuesto de distribución normal multivariable. Los datos originales de a escala están codificados, como se vió, con una escala de respuesta de verdadero-falso (1-0) de nivel de medida ordinal. Otros cuestionarios, como el de riesgo basal, ya presentados tenían ese mismo tipo de escala de respuesta y ya vimos como no cumplían en absoluto el supuesto. Ahora comprobaremos una vez más el ajuste al supuesto. Para comprobar la normalidad multivariable, primero comprobaremos la normalidad univariada mediante una prueba de Kolmogorov-Smirnov para una muestra de los ocho items, dado que es condición necesaria para que se de la normalidad multivariada (tabla 6.2). Observaremos también los datos de curtosis y asimetría y calcularemos el coeficiente multivariado de Mardia (tabla 6.3).

Tabla 6.2. Prueba de Kolmogorov-Smirnov de una muestra para comprobar el ajuste de los items de la escala de tensión psicológica a la distribución normal univariada.

ITEM	N muestral	Diferencia máxima	Probabilidad
1	122	.500	.000
2	122	.500	.000
3	122	.500	.000
4	122	.500	.000
5	121	.500	.000
6	122	.500	.000
7	118	.500	.000
8	119	.500	.000

Como es lógico dada la escala de respuesta, ninguno de los items ajusta a una distribución normal, y los datos de asimetría y curtosis son elevados indicando claramente el alejamiento de la normalidad.

Tabla 6.3. Datos de asimetría y curtosis para los items de la escala de tensión psicológica.

	Asimetría	Curtosis
Item 1	.6796	-1.538
Item 2	.8063	-1.349
Item 3	-.052	-1.997
Item 4	.4454	-1.801
Item 5	1.433	.0554
Item 6	.8504	-1.276
Item 7	.2631	-1.930
Item 8	.5599	-1.686

COEFICIENTE DE MARDIA MULTIVARIADO= 4.5465

Coefficiente esperado para la normal multivariada= 1.9272

Ante los datos que se acumulan sobre el incumplimiento del supuesto de normalidad multivariada parece razonable calcular ambas estructuras mediante estimación arbitraria, para comprobar el efecto que el incumplimiento puede tener sobre las estimaciones y las probabilidades asociadas. Se han realizado, por tanto, dos nuevos análisis factoriales confirmatorios, uno hipotetizando una estructura unifactorial y otro bifactorial. Los datos de ajuste global pueden consultarse en la tabla 6.4.



Tabla 6.4. Índices de ajuste global para la escala de tensión psicológica. Los índices son: *chi*=chi-cuadrado; *g*=grados de libertad del modelo; *p*=probabilidad; *F*=función de ajuste a minimizar; *LHR*=razón de verosimilitud escalada; *CAK*=Akaike information criterion; *CSK*=Schwarz information criterion; *NFI*=normed fit index; *NNFI*=non-normed fit index; *CFI*=comparative fit index. Método de estimación *AGLS*=método arbitrario.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
UN FACTOR	AGLS	24.7	20	.212	.216	.897	1.235	.494	.871	.846	.950	.964
2 FACTORES	AGLS	24.6	19	.172	.216	.897	1.296	.509	.912	.846	.937	.957

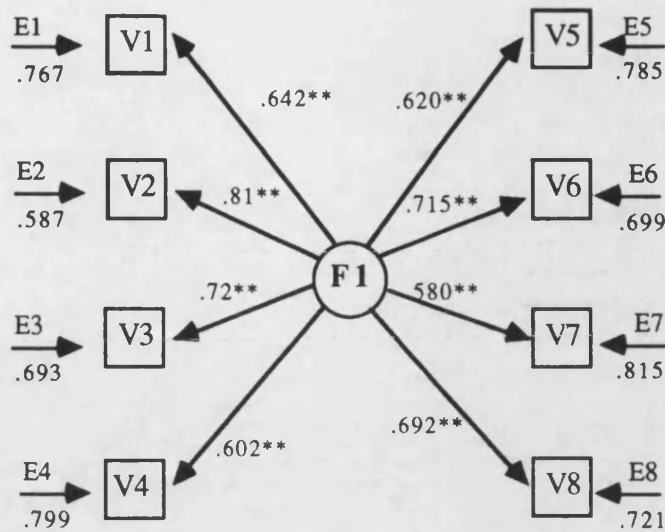
GFI: 1 FACTOR= 1; 2 FACTORES= 1.

AGFI: 1 Factor= 1; 2 Factores= 1.

Contrariamente a lo que ocurría con el uso de estimadores basados en la distribución normal multivariable, con los estimadores arbitrarios el ajuste de la estructura unifactorial es mejor que el de la estructura de dos factores. Si recordamos, los índices en ML y GLS eran consistentemente mejores para la hipótesis bifactorial, pero esas diferencias no eran estadísticamente significativas. Aquí lo que ha ocurrido es que el ajuste de la estructura unifactorial ha superado o igualado a la bifactorial por todos los índices.

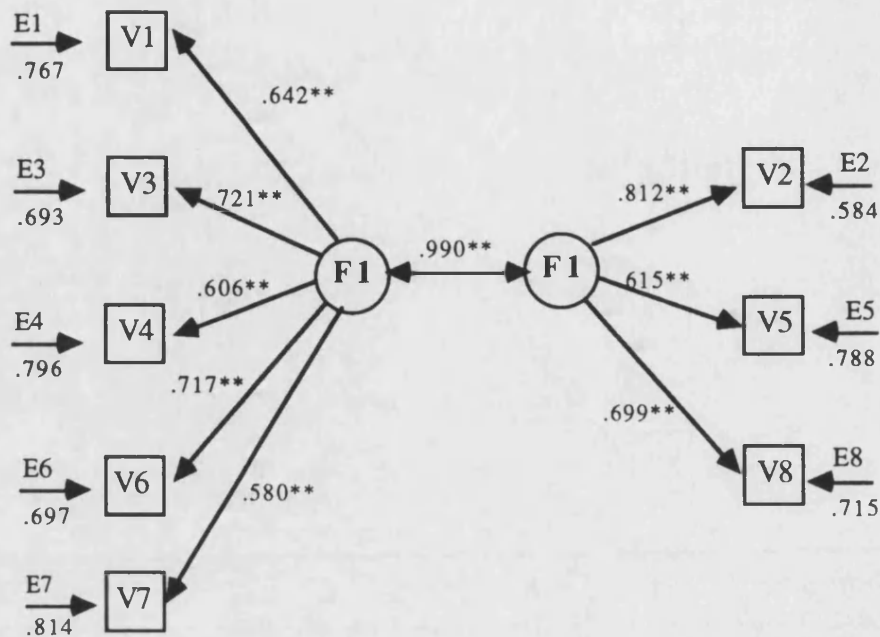
En otro orden de cosas, podemos apreciar cómo la evaluación de la probabilidad asociada a la prueba de ji-cuadrado ha variado sustancialmente. Mientras que se encontraba en torno al 0.14 en ML y GLS, ha pasado en estimación arbitraria al 0.21. Más relevante ha sido el cambio frente a la probabilidad por ML, que era estadísticamente significativa para las dos estructuras.

Además de comprobar los ajustes globales, vamos a analizar los datos de ajuste analítico y las estimaciones de las saturaciones factoriales para los modelos de uno y dos factores, gráficas 6.5 y 6.6 respectivamente.



Gráfica 6.5. Estructura factorial estimada mediante estimación arbitraria para la escala de tensión psicológica. Ajuste analítico: \*\*= $p < .01$ ; \*= $p < .05$ ; n.s.= no significativo.

Las estimaciones de las saturaciones factoriales en el modelo unifactorial no difieren prácticamente de las encontradas por los métodos de máxima verosimilitud y mínimos cuadrados generalizados. Aunque para algunos items aumentan ligeramente, para todos bajan. En cualquier caso estos cambios son poco importantes. Lo mismo ocurre para el modelo de dos factores, las saturaciones no se diferencian en exceso por el método de estimación empleado.



Gráfica 6.6. Estructura bifactorial estimada mediante método arbitrario para la escala de tensión psicológica. Ajuste analítico: \*\*= $p < .01$ ; \*= $p < .05$ ; n.s.= no significativo.

A través de análisis factoriales confirmatorios y de diferentes tipos de estimación parece claro que la estructura unifactorial representa adecuadamente a los datos. De las estructuras probadas es la que con más parsimonia representa a los datos y es capaz de explicar razonablemente bien todos los items.

### 2.3. Fiabilidad de la escala y de los items

Se ha estudiado la consistencia interna de la escala con los coeficientes de fiabilidad habituales. Existe, a la vista del coeficiente alfa para la escala total y para la parte par y la impar una alta covariación entre los items. Estos y otros coeficientes pueden verse en la tabla 6.5.

Tabla 6.5. Coeficientes de fiabilidad, consistencia interna.

<i>Consistencia interna</i>	
Correlación dos mitades	.698
Coefficiente Spearman-Brown	.822
Coefficiente Guttman-Rulon	.822
Coefficiente alpha total	.820
Alpha items impares	.704
Alpha items pares	.722

Se han calculado, también, los coeficientes de homogeneidad y fiabilidad de los items, que pueden consultarse en la tabla 6.6.

Tabla 6.6. Media, desviación típica, coeficiente de homogeneidad (correlación ítem-total), fiabilidad, homogeneidad excluyendo el ítem y coeficiente alpha que quedaría en la escala si se excluyera cada ítem.

ITEM	Media	D. T.	Homogeneidad	Fiabilidad	Homogen. sin el ítem	Alpha sin el ítem
1	0.330	0.470	.650	.306	.523	.814
2	0.313	0.464	.746	.346	.646	.798
3	0.513	0.500	.736	.368	.623	.801
4	0.391	0.488	.648	.316	.514	.816
5	0.209	0.406	.643	.261	.534	.813
6	0.304	0.460	.632	.291	.503	.817
7	0.426	0.495	.650	.322	.515	.816
8	0.365	0.481	.703	.339	.586	.806

La homogeneidad de los items resulta adecuada para todos los items, siempre por encima de 0.5 (excluyendo el ítem). Los índices de fiabilidad de los items también pueden considerarse elevados teniendo en cuenta que la escala es de respuesta dicotómica.

Tabla 6.7. Saturaciones factoriales estimadas mediante mínimos cuadrados generalizados, máxima verosimilitud y estimación arbitraria elevados al cuadrado, como medida de la fiabilidad de los items. Los resultados indican la proporción de la varianza de cada ítem explicada por el factor latente tensión psicológica.

	GLS	ML	AGLS
Item 1	.370	.312	.412
Item 2	.652	.592	.656
Item 3	.518	.478	.519
Item 4	.315	.272	.362
Item 5	.398	.381	.384
Item 6	.393	.334	.511
Item 7	.375	.344	.336
Item 8	.432	.408	.478

Por último, en la tabla 6.7 se presentan las fiabilidades de los items calculadas desde los resultados de los análisis factoriales

confirmatorios calculados desde la estructura unifactorial. El ítem peor explicado desde el factor es el 4, pero la capacidad explicativa de éste crece considerablemente con la estimación arbitraria, hasta el 36.2%. Las mejores estimaciones resultan ser las arbitrarias, que explican la mayor proporción de la varianza. En cualquier caso, casi todos los ítems son explicados en una parte importante por el único factor hipotetizado.

Los ítems de la escala son, en conjunto, homogéneos con la escala de tensión y tienen fiabilidades adecuadas para la medición de ésta. Desde un punto de vista psicométrico, los ítems miden adecuadamente el factor latente de tensión psicológica.

#### **2.4. Validez criterial**

La utilización de la variable tensión en este trabajo ha sido más exploratoria que confirmatoria. No están absolutamente perfiladas las hipótesis sobre su relación con otros constructos relevantes al estudio de la accidentabilidad, por lo que puede hablarse más de intuiciones que de hipótesis. No obstante, vamos a intentar esbozar algunas posibles relaciones:

a) Se hipotetiza una relación positiva con los niveles de riesgo real y basal. Esperamos que el nivel de riesgos y la probabilidad de accidentarse pueden aumentar el grado de tensión psicológica.

b) Se espera una relación negativa con la respuesta de los superiores y compañeros en seguridad, ya que un ambiente laboral en que los superiores y los compañeros colaboran y tienen en cuenta los aspectos de seguridad puede redundar en un mejor afrontamiento de los riesgos, y de la tensión que pueden generar, además de ser positivo en sí mismo para el bienestar psicológico.

c) Se espera relación negativa con la conducta de seguridad del trabajador por los mismos motivos que con respuesta de los compañeros y trabajadores.

d) Se espera que los ítems de tensión tengan cierto poder predictivo sobre el nivel de accidentes. Se hipotetiza que un sujeto tenso es más propenso a accidentarse.

Tabla 6.8. Correlaciones de los ítems de la escala de tensión psicológica con los principales criterios. ES=estructura de seguridad; PO=política organizacional de seguridad; F=respuesta de los superiores en seguridad; RC=respuesta de los compañeros en seguridad; CHS=conducta hacia la seguridad; RB=riesgo basal; RR=riesgo real; RS=respuesta de los superiores; y AC= sinestralidad. Símbolos de significatividad estadística de las correlaciones: n.s.= no significativo; #= marginalmente significativo; \*= probabilidad <0.05; y \*\*= probabilidad <0.01, extraídos de tablas (Welkowitz, Ewen y Cohen, 1981).

	ES	PO	F	RC	CHS	RB	RR	RS	AC
1	-.213*	-.39**	-.218*	-.29**	-.49**	.138ns	.24**	-.40**	.185*
2	.026ns	-.09ns	-.11ns	-.09ns	-.29**	.084ns	.30**	-.18*	.166ns
3	-.00ns	-.18*	-.08ns	-.27**	-.37**	.163ns	.29**	-.30**	.109ns
4	-.11ns	-.28**	-.19**	-.38**	-.44**	.24**	.24**	-.44**	.046ns
5	.101ns	-.20**	-.10ns	-.22**	-.28**	-.18*	.26**	-.26**	.063ns
6	.126ns	.018ns	.051ns	-.16ns	-.27**	-.00ns	.173ns	-.14ns	.039ns
7	-.02ns	-.08ns	-.09ns	-.23**	-.34**	.017ns	.185*	-.29**	.108ns
8	.049ns	.022ns	.031ns	.012ns	-.24**	.075ns	.189*	-.11ns	.075ns

Si repasamos hipótesis por hipótesis podremos comprobar cuáles parecen cumplirse para los ítems de la escala de tensión y cuáles no.

La primera de las hipótesis hablaba de una relación positiva con los riesgos real y basal. Así como sí hay una relación muy significativa con riesgo real, no ocurre así con riesgo basal. Si recordamos, el riesgo basal es el riesgo intrínseco al trabajo. Este puede ser un riesgo que no provoca tensión precisamente porque es siempre el mismo, está asumido por el trabajador. En cambio el riesgo real está provocado en muchos casos por malas políticas, inadecuada respuesta de los compañeros y lo superiores, falta de control ambiental, etc, con lo que de tensión pueden conllevar todas estas situaciones.

La segunda hipótesis planteaba relaciones negativas con la respuesta de superiores y compañeros, y sí parecen cumplirse. Casi todos los ítems se encuentran significativamente relacionados con estas dos variables. Parece que una inadecuada respuesta de seguridad de los compañeros y de los supervisores puede generar tensión en el trabajador. Quizá también los trabajadores con gran tensión puedan distorsionar este tipo de relaciones. También parece que se cumple la hipótesis de relación con conducta hacia la seguridad.

La hipótesis de relación de tensión con la variable accidentabilidad no ha sido refrendada por los datos de validez criterial de los ítems. No parece que un nivel alto de tensión conlleve mayor número de accidentes, al menos en esta muestra. No obstante hay que decir aquí, que muy probablemente la relación no sea lineal. Probablemente sólo muy altos niveles de tensión pueden llevar a problemas graves que puedan hacer que se desempeñe mal el trabajo y se tengan accidentes, y si se observan las medias de los ítems

(tabla 6.6) se puede apreciar como en general la muestra no presenta niveles elevados de tensión.

Por descontado esto sólo se refiere a la validez criterial de los items. La medida global de tensión puede ser más sensible y tener mayor variabilidad, con lo que las correlaciones con los criterios pueden variar sustancialmente, pero esto ya será estudiado mediante modelos causales en el próximo capítulo.

En resumen la tensión psicológica parece derivar más que de la siniestralidad de la empresa y de los riesgos intrínsecos del trabajo de las relaciones sociales de trabajo que mantiene el trabajador por las que parece es predecible.



### 3. Conducta hacia la seguridad de los trabajadores

La variable conducta hacia la seguridad es una de las más importantes de las planteadas en todo el trabajo. Efectivamente, muchos autores suponen que el factor humano, de una forma u otra es el causante primario de la mayoría de incidentes y accidentes laborales. Evidentemente este 'error humano' o estas conductas inseguras que pueden llevar a la ocurrencia de un accidente puede estar a su vez determinado por otras causas (ritmo de trabajo, pobre política de seguridad, presión, inexperiencia, etc) pero no cabe duda de que es capaz de explicar gran parte de los accidentes. En este contexto resulta muy interesante disponer de una medida sensible y adecuada de las conductas seguras e inseguras del trabajador. Por un lado, se podrá contar, si se consigue este propósito, un predictor de los accidentes y, por otro, se podrá evaluar el impacto que otras variables pueden tener sobre la conducta del trabajador. Así, se ha intentado construir una escala que con las suficientes garantías psicométricas nos permita la evaluación de todos estos aspectos. La escala ha quedado compuesta por diez items de escala de repuesta tipo Likert, que se presentarán a continuación, y que pretende ser una muestra de los diversos tipos de conductas seguras e inseguras que con frecuencia se han indicado en la literatura como antecedentes de un accidente o conductas preventivas de los mismos. Algunos de los items de la escala miden en el sentido de conducta de seguridad, mientras otras miden en el sentido inverso de conducta insegura. No obstante, estos últimos items han sido reescalados de forma que la escala mida la conducta de seguridad.

#### 3.1. Medida

##### ESCALA:

Item 1. Busco atajos para poder hacer más faena o para hacer la faena más aprisa.

0                      1                      2                      3                      4                      5  
Nunca    Muy Pocas veces    Algunas veces    Bastantes veces    Muchas veces    Siempre

Item 2. Mantengo mi lugar de trabajo (o mi equipo o maquinaria de trabajo) limpio y ordenado.

0                      1                      2                      3                      4                      5  
Nunca    Muy Pocas veces    Algunas veces    Bastantes veces    Muchas veces    Siempre



## Capítulo 6. Respuesta y efectos sobre la seguridad del trabajador.

---

Item 3. Tengo que trabajar rápido para conseguir resultados.

0                      1                      2                      3                      4                      5  
Nunca      Muy Pocas veces      Algunas veces      Bastantes veces      Muchas veces      Siempre

Item 4. Cumplo las normas de seguridad.

0                      1                      2                      3                      4                      5  
Nunca      Muy Pocas veces      Algunas veces      Bastantes veces      Muchas veces      Siempre

Item 5. Trabajo bajo fuerte cansancio o sueño.

0                      1                      2                      3                      4                      5  
Nunca      Muy Pocas veces      Algunas veces      Bastantes veces      Muchas veces      Siempre

Item 6. Si una máquina, vehículo o herramienta trabaja defectuosamente lo paramos enseguida.

0                      1                      2                      3                      4                      5  
Nunca      Muy Pocas veces      Algunas veces      Bastantes veces      Muchas veces      Siempre

Item 7. Tengo que incumplir alguna norma de seguridad para no entorpecer la tarea.

0                      1                      2                      3                      4                      5  
Nunca      Muy Pocas veces      Algunas veces      Bastantes veces      Muchas veces      Siempre

Item 8. Informo inmediatamente de cualquier problema que pueda causar un accidente de trabajo.

0                      1                      2                      3                      4                      5  
Nunca      Muy Pocas veces      Algunas veces      Bastantes veces      Muchas veces      Siempre

Item 9. Si tengo que levantar un objeto muy pesado pido ayuda.

0                      1                      2                      3                      4                      5  
Nunca      Muy Pocas veces      Algunas veces      Bastantes veces      Muchas veces      Siempre

Item 10. A veces utilizo herramientas defectuosas

0                      1                      2                      3                      4                      5  
Nunca      Muy Pocas veces      Algunas veces      Bastantes veces      Muchas veces  
Siempre

### **3.2. Estructura factorial confirmatoria**

Para estudiar la estructura factorial de la escala de conducta hacia la seguridad se plantearon cuatro estructuras hipotéticas. Cada una de ellas se probó en base a dos tipos de estimación, mínimos cuadrados generalizados y máxima verosimilitud, como hasta ahora. Dos estructuras de las planteadas eran de tipo racional, teóricas, mientras las dos restantes estaban basadas en análisis exploratorios. Pasamos a detallar los modelos calculados.

1) El primero de los modelos factoriales es unifactorial. Este modelo se utiliza como modelo basal a partir del cual evaluar modelos más sustantivos. Efectivamente, no creemos que un modelo de un solo factor pueda dar cuenta de la variedad de conductas seguras-inseguras que pueden darse en los diferentes contextos laborales que se recogen en la muestra.

2) Una estructura bifactorial, basada en una hipótesis teórica de agrupación del contenido de los ítems. El primer factor englobaría los ítems números 1, 3, 5, 6 y 7. Estos ítems hacen referencia a aspectos tales como el ritmo apresurado del trabajo, la presión por producir, la necesidad de incumplir normas para poder acabar el trabajo y similares. En resumen aspectos referidos a *condiciones del trabajo* que hacen al trabajador realizar conductas peligrosas. El segundo factor recoge los contenidos de los ítems 2, 4, 8, 9, y 10, que hacen referencia más a *conductas seguras e inseguras* que el propio trabajador realiza, tales como mantener en orden y limpio el lugar de trabajo, informar sobre problemas de seguridad, utilizar herramientas defectuosas, cumplir las normas de seguridad dictadas, etc. La diferencia entre ambos factores no viene dado por que uno de ellos mida conducta y el otro condiciones de inseguridad, si no más bien por el determinante de las conductas de seguridad e inseguridad, que en un caso son debidas a las condiciones del trabajo y en otro son más dependientes del propio trabajador.

3) La tercera estructura probada es una estructura tetrafactorial. Este modelo surge de un análisis factorial exploratorio realizado sobre la escala de conducta mediante el paquete estadístico Systat. El método elegido en este caso es el de componentes principales, único disponible en la versión 3 de Systat, con rotación varimax. La solución que se ofrecía en el factorial exploratorio no se esperaba que fuera una solución definitiva, dada la diferencia entre métodos, pero sí se esperaba que nos ofreciera un cierto patrón de covariación de los ítems, que si tenían sentido teórico pudieran guiar el establecimiento de la estructura de la escala. Así, el primer factor estaría formado por los ítems 1, 3, 5 y 7. Estos ítems rastrean contenidos referidos a *formas de trabajo seguro o inseguro*, tales como trabajar bajo cansancio, trabajar bajo presión de tiempo o teniendo que incumplir normas de seguridad para poder realizar el trabajo. El segundo factor

incluye dos items, 2 y 4, que hacen referencia a *cumplimiento de normas*. El tercer factor incluye los items 8 y 9 que agrupan contenidos de *conducta de comunicación* en torno a aspectos de seguridad. El cuarto y último factor explica los items 6 y 10 referidos a *conductas de seguridad frente a máquinas o herramientas defectuosas*.

4) La puesta a prueba de la estructura tetrafactorial produjo una indicación por parte del LM test sobre la adscripción de un ítem a dos factores. Se mantenía la estructura tetrafactorial ya planteada, pero el tanto en la estimación máximo verosímil como en la de mínimos cuadrados generalizados, aparecía el ítem siete como posiblemente relacionado con el factor cuatro, además de con el uno. El LM test ofrecía la información de que este cambio supondría una ganancia estadísticamente significativa en la capacidad explicativa del modelo tetrafactorial.

Para una primera evaluación del modelo, cabe estudiar los análisis de ajuste global de las cuatro estructuras factoriales planteadas. Estos datos pueden consultarse en la tabla 6.9, y ofrecen apoyo básicamente a las dos estructuras tetrafactoriales.

Tabla 6.9. Índices de ajuste global de las estructuras hipotetizadas para la escala de conducta hacia la seguridad. Los índices son: *chi*=chi-cuadrado; *g*=grados de libertad del modelo; *p*=probabilidad; *F*=función de ajuste a minimizar; *LHR*=razón de verosimilitud escalada; *CAK*=Akaike information criterion; *CSK*=Schwarz information criterion; *NFI*=normed fit index; *NNFI*=non-normed fit index; *CFI*=comparative fit index. Métodos de estimación: *GLS*=mínimos cuadrados generalizados; *ML*=máxima verosimilitud.

		chi	g	p	F	LHR	chi/g	CAK	CSK	NFI	NNFI	CFI
UN	GLS	68.1	35	<.01	.608	.737	1.94	.961	1.44	.956	.972	.978
FACTOR	ML	100.	35	<.01	.897	.638	2.87	1.25	1.73	.589	.577	.671
DOS	GLS	68.9	34	<.01	.599	.741	2.02	.97	1.47	.957	.970	.977
FACTORES	ML	98.5	34	<.01	.856	.651	2.89	1.22	1.73	.607	.585	.686
CUATRO	GLS	38.4	29	.113	.342	.842	1.32	.802	1.42	.975	.990	.994
FACTORES	ML	40.2	29	.080	.359	.835	1.38	.819	1.44	.835	.913	.944
CUATRO	GLS	32.0	28	.272	.286	.866	1.14	.763	1.41	.979	.996	.997
FACT. (2)	ML	31.4	28	.296	.281	.868	1.12	.758	1.41	.871	.972	.983

Las pruebas de ji-cuadrado ofrecen valores estadísticamente significativos para los modelos de un factor y de dos factores, tanto para la estimación máximo verosímil como para la mínimo cuadrática. En cambio, sus valores no alcanzan la significatividad estadística, indicando buen ajuste del modelo, en las dos estructuras tetrafactoriales propuestas, también tanto para máxima verosimilitud como en mínimos cuadrados generalizados. El ajuste es considerablemente mejor, en lo que a evaluación de la probabilidad del modelo se refiere, cuando se introduce el cambio sugerido por el LM test, que sube las probabilidades a 0.27 y 0.29, respectivamente, para cada tipo de estimación. El cambio es especialmente relevante para el caso de la estimación máximo verosímil, que sufre una

ganancia más elevada que la sufrida por mínimos cuadrados generalizados.

La situación, que aparece meridianamente clara para la prueba de ji-cuadrado es menos clara si se miran otros índices. Aunque todos ofrecen ganancias para las estructuras tetrafactoriales, las diferencias son menos drásticas. La razón de verosimilitud escalada sí aprecia la diferencia entre los modelos, siempre con los tetrafactoriales por encima de .83 y el resto algo más bajos. Pero el CAK y el CSK no son capaces de discriminar bien entre las estructuras, en algunos casos hasta ofrecen valores iguales para estructuras que, a decir de otros índices, son muy desiguales en el porcentaje de varianza de los datos que explican. Lo mismo ocurre con los índices NFI, NNFI y CFI en el caso de la estimación mínimo cuadrática que rápidamente alcanzan valores muy elevados indicativos de adecuado ajuste (más de .95), aún en la estructura de uno y dos factores, claramente deficientes por otros índices. No ocurre lo mismo para estos últimos índices cuando se trata de la estimación máximo verosímil. Efectivamente aquí incluso puede verse una vez más que el índice NFI adopta valores inusualmente bajos cuando NNFI y CFI están cercanos al ajuste perfecto.

En cualquier caso, la evaluación del ajuste global a la luz de los diversos índices parece clara. Las estructuras tetrafactoriales ajustan mejor a los datos observados que la bifactorial y la unifactorial. Aún cuando el LM test ya nos ha ofrecido una prueba de que la estructura tetrafactorial modificada en base a sus sugerencias es estadísticamente más explicativa que la tetrafactorial planteada en un principio, puede generarse una prueba de las diferencias de ji-cuadrado para examinar este punto.

Para la estimación mediante mínimos cuadrados generalizados tenemos:

$$\Delta\chi^2 = 6.34, \text{ con } \Delta gl = 1 \text{ que no resulta estadísticamente significativa (} p < .05 \text{),}$$

y para máxima verosimilitud:

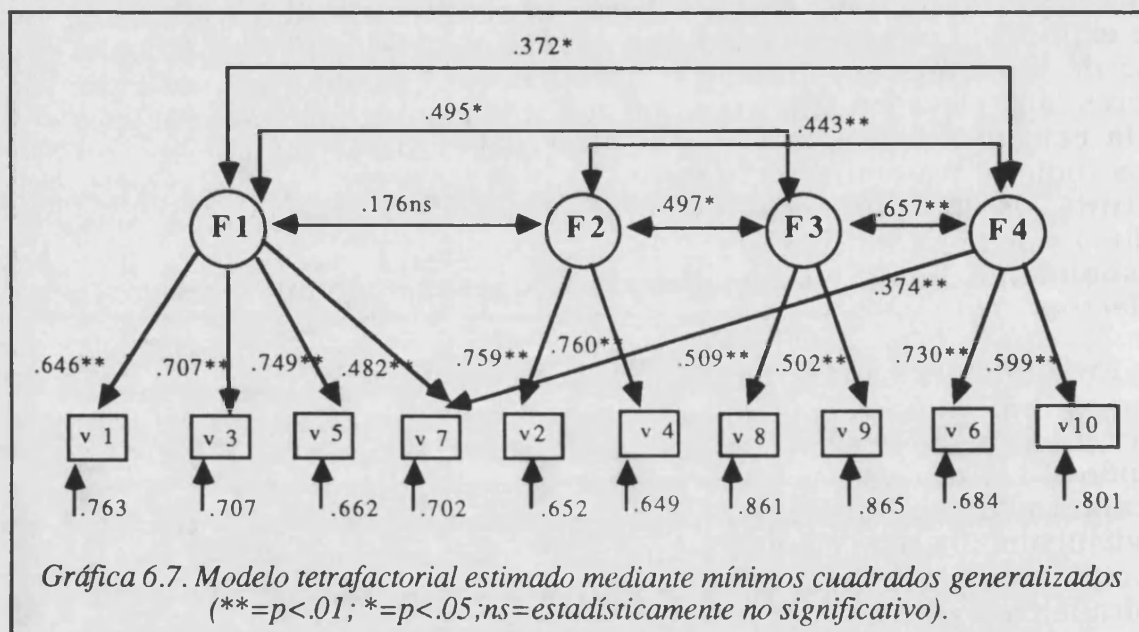
$$\Delta\chi^2 = 8.75, \text{ con } \Delta gl = 1 \text{ que tampoco resulta significativa (} p < .05 \text{).}$$

Y nos permite comprobar una apreciación que ya habíamos realizado en cuanto a que se había producido una mejora mayor en la estimación máximo verosímil con la inclusión de la nueva relación, que la que se ha producido en la estimación mínimo cuadrática con la misma modificación.

Queda por observar el ajuste analítico de la estructura de mejor ajuste, en este caso la tetrafactorial modificada, tanto por mínimos cuadrados generalizados como por máxima verosimilitud. Las

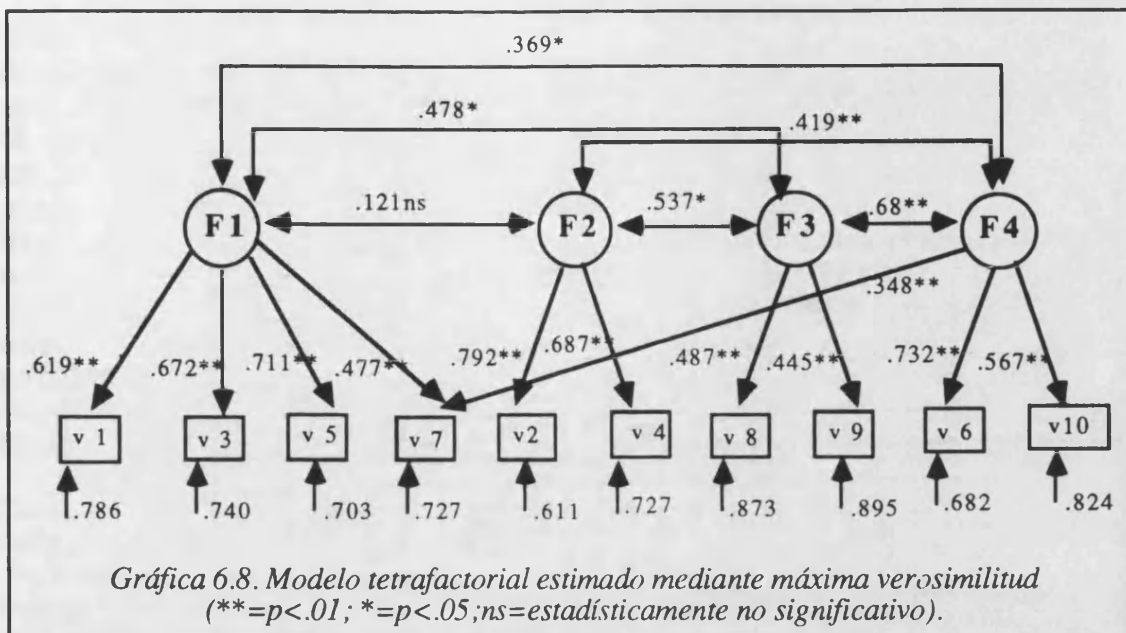
saturaciones factoriales estimadas para el resto de estructuras planteadas se presentarán también, pero en el apartado de fiabilidad y elevadas al cuadrado como medida de la fiabilidad de los items. Allí comentaremos algunos cambios interesantes en la fiabilidad de los items que se producen a través de las diferentes estructuras hipotetizadas.

Los datos de ajuste analítico para la estructura tetrafactorial modificada por el LM test se ofrecen en las gráficas 6.7 y 6.8, junto con las estimaciones de las saturaciones, los errores asociados, y las correlaciones entre los factores.



Puede verse en la gráfica 6.7 cómo todos los items se encuentran significativamente relacionados, al menos con un factor. Quizá los items peor explicados, si observamos los errores podremos apreciarlo con mayor claridad, sean los explicados por el factor tres, y esto a pesar de que la proporción de varianza explicada supera en ambos casos el 25%, que puede ser considerada, sino como adecuada, sí sustancial.

Por su parte, los datos de estimación máximo verosímil pueden verse en la gráfica 6.8. El patrón es muy similar al encontrado para GLS, y siguen siendo los items 8 y 9 los peor explicados por su factor, el tres.



Un dato importante son las correlaciones entre los factores. En general, las saturaciones son estadísticamente significativas, si bien no excesivamente elevadas, salvo en el caso de los factores uno y dos. Este resultado, concordante a través de estimaciones, indica que no hay una asociación significativa entre formas de trabajo seguras (factor 1, con alta presencia del contenido presión de tiempo) y cumplimiento de normas y orden y limpieza (factor 2).

Como hemos venido haciendo hasta aquí, vamos a realizar pruebas para comprobar el ajuste de las variables a la normalidad. Se intenta optimizar la estimación por medio de métodos arbitrarios, caso de que no pueda mantenerse los supuestos de las estimaciones máximo verosímiles y mínimo cuadráticas. Para ello comenzaremos probando la normalidad univariada de los items de la escala mediante pruebas de Kolmogorov-Smirnov, requisito indispensable para que se cumpla la normalidad multivariada. Los datos de esta prueba pueden consultarse en la siguiente tabla.

Tabla 6.10. Prueba de Kolmogorov-Smirnov de una muestra para comprobar el ajuste de los items de la escala de conducta hacia la seguridad a la distribución normal univariada.

ITEM	N muestral	Diferencia máxima	Probabilidad
1	123	.839	.000
2	122	.953	.000
3	124	.687	.000
4	122	.895	.000
5	122	.895	.000
6	122	.838	.000
7	121	.928	.000
8	122	.912	.000
9	122	.797	.000
10	122	.936	.000

Ninguno de los items de la escala cumple razonablemente con la normalidad univariada. Además, los datos de asimetría y curtosis muestran estas desviaciones de la normalidad y apoyan los datos de la prueba de ajuste -tabla 6.11-.

Tabla 6.11. Datos de asimetría y curtosis para los items de la escala de conducta hacia la seguridad.

	Asimetría	Curtosis
Item 1	-.617	-.569
Item 2	-1.03	-.117
Item 3	-.118	-1.27
Item 4	-.423	-1.07
Item 5	-1.01	.578
Item 6	-1.03	-.43
Item 7	-1.27	1.11
Item 8	-1.59	1.29
Item 9	-.406	-1.27
Item 10	-1.13	.750

El coeficiente de Mardia multivariado, que ofrece el programa EQS, también da muestra de la fuerte desviación de la normalidad multivariada que presentan los items de la escala.

COEFICIENTE DE MARDIA MULTIVARIADO=23.4608  
 Coeficiente esperado para la normal multivariada=8.1552

A la luz de los datos descriptivos y de las pruebas de ajuste realizadas está justificado emplear métodos arbitrarios de estimación por el alejamiento de los supuestos. Así pues, se estimó el modelo tetrafactorial hipotético que encontraba buen ajuste mediante



mínimos cuadrados y máxima verosimilitud. En lugar de probar directamente la mejor solución factorial, que recordemos venía de una modificación de la solución tetrafactorial sugerida por el LM test, se optó por poner a prueba la solución tetrafactorial originaria basada en el análisis factorial exploratorio. Al actuar de esta forma se pretendía comprobar si la sugerencia del LM test de relacionar el ítem siete con el factor cuatro, coincidentes en los anteriores dos métodos de estimación, también era sugerida ahora al cambiar de método. Efectivamente fue así, y era necesaria para alcanzar un adecuado ajuste. Así pues, se realizó un nuevo análisis factorial confirmatorio con las mismas relaciones que el tertafactorial modificado ya expuesto y que llevó a los datos de ajuste global de la tabla 6.12.

Tabla 6.12. Índices de ajuste global para la escala de conducta hacia la seguridad. Los índices son: *chi*=chi-cuadrado; *g*=grados de libertad del modelo; *p*=probabilidad; *F*=función de ajuste a minimizar; *LHR*=razón de verosimilitud escalada; *NFI*=normed fit index; *NNFI*=non-normed fit index; *CFI*=comparative fit index. Método de estimación *AGLS*=método arbitrario.

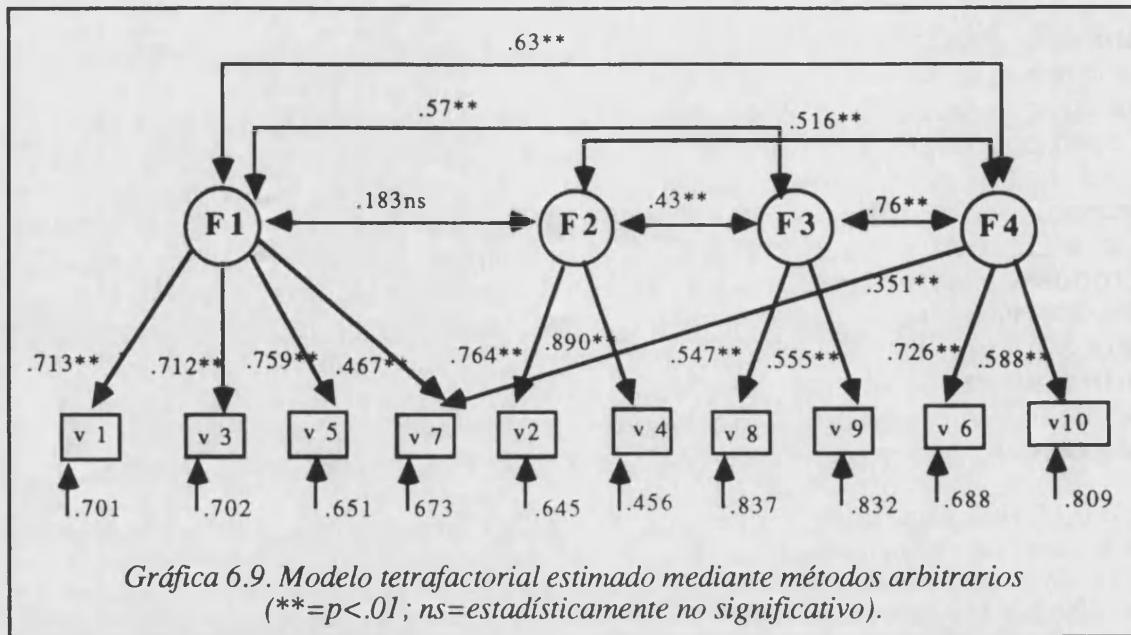
		chi	g	p	F	LHR	chi/g	NFI	NNFI	CFI
4 FACTORES (2)	AGLS	42.1	28	.042	.366	.832	1.50	.844	.899	.937

INDICE DE BONDAD DE AJUSTE, GFI= .961  
 Índice de Bondad de Ajuste Ajustado, AGFI= .923

Aunque la prueba de ji-cuadrado es estadísticamente significativa, se encuentra próximo al 0.05. Por su parte los índices LHR, ji-cuadrado partido los grados de libertad, CFI, GFI y AGFI indican ajuste adecuado y muestran una estructura parsimoniosa. Por su parte, NFI, y NNFI no muestran adecuado ajuste a los datos, no pasa de 0.9.

En la gráfica 6.9. pueden consultarse las saturaciones factoriales, los errores, la correlación entre los factores y lo datos de ajuste analítico (pruebas z para las relaciones).





Aunque las estimaciones son muy similares a las encontradas mediante los métodos basados en el supuesto de normalidad, ciertos cambios sí son apreciables. A pesar de que los datos de ajuste global son peores que los hallados para ML y GLS, las saturaciones y las correlaciones han tendido en general a ascender, con lo que los porcentajes de varianza no explicados de todas las variables del modelo han disminuido. Especialmente son significativos el aumento en las correlaciones que hace cambiar la significatividad estadística de varias de ellas hasta alcanzar el nivel de probabilidad de 0.01, y algunos cambios en las saturaciones relevantes, como la del ítem 4 que pasa de saturaciones entorno a 0.7 a cerca de 0.9. No obstante los cambios no son excesivamente importantes y mantienen la coherencia con los resultados de máxima verosimilitud y mínimos cuadrados generalizados.

### 3.3. Fiabilidad de la escala y de los ítems

Los datos de consistencia interna de la escala, calculados por diversos métodos, pueden verse en la tabla 6.13. El coeficiente alpha alcanza un valor de .73, que puede considerarse adecuado, pero el resto de coeficientes bajan relativamente. Así, Guttman-Rulon y Spearman-Brown pasan ligeramente del 0.5.

Tabla 6.13. Coeficientes de fiabilidad, consistencia interna.

<i>Consistencia interna</i>	
Correlación dos mitades	.346
Coeficiente Spearman-Brown	.514
Coeficiente Guttman-Rulon	.507
Coeficiente alpha total	.730
Alpha ítems impares	.702
Alpha ítems pares	.638

Los datos relativamente bajos de consistencia interna de la escala puede corroborarse en la tabla 6.14 que presenta los análisis de ítems. Puede verse cómo la tónica de homogeneidad de los ítems con el total de la escala es moderada. No obstante, todos los ítems se comportan de forma similar, sin grandes oscilaciones en su comportamiento, salvo los ítems 5 y 7 que presentan una correlación más elevada con el resto de la escala. Recordemos que la explicación del ítem siete se hallaba compartida por dos factores, el uno y el cuatro.

Tabla 6.14. Media, desviación típica, coeficiente de homogeneidad (correlación ítem-total), fiabilidad, homogeneidad excluyendo el ítem y coeficiente alpha que quedaría en la escala si se excluyera cada ítem.

ITEM	Media	D. T.	Homogeneidad	Fiabilidad	Homogen. sin el ítem	Alpha sin el ítem
1	3.224	1.543	.581	.897	.429	.702
2	4.129	1.149	.430	.494	.301	.721
3	2.647	1.802	.570	1.028	.385	.712
4	3.552	1.404	.486	.682	.333	.717
5	3.569	1.372	.609	.836	.482	.695
6	3.828	1.652	.612	1.010	.454	.698
7	3.922	1.327	.671	.891	.562	.684
8	4.207	1.374	.475	.653	.324	.718
9	3.293	1.732	.510	.883	.321	.723
10	4.052	1.159	.474	.549	.349	.715

A continuación se presentarán las tablas con las saturaciones factoriales de todos los análisis factoriales confirmatorios realizados para la escala de conducta hacia la seguridad elevadas al cuadrado. Esto nos permitirá apreciar los cambios en la explicación de la varianza del ítem que se producen al cambiar de modelo hipotético y de método de estimación, además de ofrecernos un cálculo de la fiabilidad de los ítems para medir los diversos factores o aspectos de la conducta hacia la seguridad. En la tabla 6.15, pueden consultarse las saturaciones al cuadrado de las cuatro estructuras propuestas cuando se estiman mediante máxima verosimilitud.

Tabla 6.15. Saturaciones factoriales elevadas al cuadrado para la escala de conducta estimadas mediante máxima verosimilitud, para la cuatro estructuras hipotetizadas como medida de fiabilidad de los ítems.

	1 factor	2 factor.	4 factor.	4 correg.
Item 1	.284	.355	.351	.383
Item 2	.058	.073	.692	.627
Item 3	.279	.400	.396	.451
Item 4	.109	.141	.427	.471
Item 5	.410	.481	.511	.505
Item 6	.238	.160	.616	.535
Item 7	.512	.559	.440	.471
Item 8	.101	.110	.234	.237
Item 9	.131	.128	.200	.198
Item 10	.164	.179	.278	.321

Tras una inspección rápida, lo primero que aparece claro es que las fiabilidades de los ítems son muy superiores cuando se pasa a los modelos de cuatro factores. Esto no es demasiado elevado en los ítems que forman parte en la estructura tetrafactorial del primer factor, ya que son cuatro ítems y por su número y altas intercorrelaciones pueden forzar el contenido del factor tanto en la estructura unifactorial como en la bifactorial, su peso específico en estas estructuras es mayor. No obstante para el resto de ítems como el 2, 4, 8, 9 10 la mejora es sustancial, dado que ahora quedan incluidos en factores más acordes a sus contenidos específicos, factores que además se encuentran en algunos casos no excesivamente correlacionados entre sí. Estudio aparte merece el ítem 7, que se mantienen con similares fiabilidades a través de los cuatro modelos, incluso algo superiores en el unifactorial y el bifactorial. Ya hemos visto que éste ítem satura en dos factores y puede, aunque no alcance la significatividad estadística, estar parcialmente solapado su contenido con otros, lo que explicaría este comportamiento. Por otro lado, las diferencias en las fiabilidades entre la estructura tetrafactorial y la tetrafactorial modificada por el LM test no son excesivas, aumentan un poco en algunos casos y bajan en otros.

Tabla 6.16. Saturaciones factoriales elevadas al cuadrado para la escala de conducta estimadas mediante mínimos cuadrados generalizados, para la cuatro estructuras hipotetizadas como medida de fiabilidad de los items.

	1 factor	2 factores	4 factores	4 corregido
Item 1	.356	.401	.394	.417
Item 2	.220	.179	.927	.576
Item 3	.376	.442	.478	.499
Item 4	.309	.305	.321	.577
Item 5	.446	.483	.522	.561
Item 6	.309	.279	.567	.532
Item 7	.549	.597	.492	.507
Item 8	.134	.121	.567	.259
Item 9	.150	.133	.293	.252
Item 10	.216	.195	.250	.358

Los comentarios ya realizados para el caso de la estimación máximo verosímil valen para la estimación mínimo cuadrática -tabla 6.16-. Sólo apuntar que el item dos queda excepcionalmente explicado por su factor en el caso de la tetrafactorial. Este resultado es llamativo pero correcto, dado que ha sido revisado en repetidas ocasiones, y después, en el tetrafactorial vuelve a situarse en valores cercanos a los hallados para la estimación máximo verosímil. En cuanto a las diferencias de fiabilidad entre los métodos de estimación no puede hablarse de un patrón claro. En algunos casos es más fiable el resultado por ML, en otros por GLS, con diferencias no muy importantes sin embargo. Y esto a pesar de que al ser valores al cuadrado se agrandan las diferencias, que como hemos visto al nivel de las saturaciones eran muy bajas.

Tabla 6.17. Saturaciones factoriales elevadas al cuadrado para la escala de conducta estimadas mediante métodos arbitrarios, para la estructura tetrafactorial corregida como medida de fiabilidad de los items.

	4 factores	4 correg.
Item 1	.394	.508
Item 2	.925	.584
Item 3	.479	.507
Item 4	.323	.792
Item 5	.523	.576
Item 6	.565	.527
Item 7	.494	.547
Item 8	.230	.299
Item 9	.294	.308
Item 10	.251	.346

Para terminar con este apartado de fiabilidad y análisis de items, presentar los datos de fiabilidad de los items en la estimación arbitraria, tabla 6.17. Los resultados vuelven a ser similares a los de



los dos métodos anteriores, pero hay dos importantes comentarios a hacer. Primero, que la elevada fiabilidad que hemos visto para el caso del ítem 2 en la estructura tetrafactorial estimada por mínimos cuadrados generalizados se ve corroborada aquí en la estimación arbitraria. Por otro lado, que sí podemos hablar de un patrón relativamente constante en la comparación de los resultados de la estimación arbitraria con los otros dos métodos. Los resultados de fiabilidad son algo mejores. Así, al comparar la estimación arbitraria con la estimación mínimo cuadrática se producen ligeras mejoras en siete de los ítems, mejoras importantes en uno de ellos (ítem 4) , otro queda igual y sólo uno baja ligerísimamente. La comparación del método arbitrario con máxima verosimilitud ofrece resultados similares: ocho ítems mejoran y dos empeoran ligeramente en fiabilidad.

### **3.4. Validez criterial**

Las correlaciones de los ítems de la escala de conducta con los principales criterios se presentan en la tabla 6.18. Dos son los criterios principales para la escala de conducta:

a) Se espera una correlación elevada y positiva con respuesta de los superiores. La respuesta de los superiores se supone un antecedente importante de la conducta de seguridad.

En este mismo sentido se espera la correlación con la variable política organizacional, si bien más moderada dado que el principal canalizador de la política de seguridad de la empresa son los supervisores inmediatos, y por tanto el efecto de ésta puede ser indirecto a través de la respuesta de los superiores. También se espera que la respuesta de los compañeros pueda tener cierto efecto sobre la conducta del trabajador, aunque menor ya que los supervisores determinan en el contexto laboral mucho más la conducta que los compañeros.

b) Se espera una correlación negativa y fuerte con el nivel de riesgo real. Se hipotetiza que las conductas seguras tienen efecto rebajando el nivel de probabilidad de accidentarse, lo que define una relación negativa entre las dos variables.

Otras relaciones son también posibles, pero no se consideran criterios tan relevantes como los anteriores. Así, se espera relación negativa con los accidentes y con la tensión. No se espera relación con el nivel de riesgo basal. Se espera relación positiva entre la formación de seguridad y conducta, ya que una mayor formación debería redundar en una mayor conducta segura, precisamente para ello se realiza la formación. También se espera una cierta relación positiva con la escala de actitud hacia la seguridad de P. J. Leather,

aunque no se tiene una hipótesis demasiado clara porque la escala mide locus de control de los accidentes más que actitud.

Tabla 6.18. Correlaciones de los ítems de la escala de conducta hacia la seguridad con los principales criterios. ES=estructura de seguridad; PO=política organizacional de seguridad; F=respuesta de los superiores en seguridad; RC=respuesta de los compañeros en seguridad; L=actitud hacia la seguridad; RS=respuesta de los superiores; RB=riesgo basal; RR=riesgo real; T=tensión psicológica; y AC= sinestralidad. Símbolos de significatividad estadística de las correlaciones: n.s.= no significativo; #= marginalmente significativo; \*= probabilidad <0.05; y \*\*= probabilidad <0.01, extraídos de tablas (Welkowitz, Ewen y Cohen, 1981).

	ES	PO	F	L	RS	RC	RB	RR	T	AC
1	.155#	.200*	.097ns	-.191*	.176#	.201*	-.05ns	-.31**	-.46**	-.24**
2	.067ns	.180*	.094ns	-.27**	.237*	.302**	.022ns	-.13ns	-.03ns	.057ns
3	.012ns	.186*	.088ns	-.29**	.37**	.253**	-.12ns	-.34**	-.52**	-.126ns
4	.174#	.344**	.209*	-.26**	.37**	.379**	-.161#	-.29**	-.22*	-.03ns
5	.055ns	.248**	.207*	-.10ns	.37**	.160#	-.14ns	-.19*	-.40**	-.001ns
6	.113ns	.277**	.023ns	-.27**	.28**	.206*	-.01ns	-.198*	-.211*	-.16#
7	.117ns	.184*	.050ns	-.191*	.38**	.326**	-.13ns	-.26**	-.33**	-.09ns
8	.057ns	.176#	.034ns	-.188*	.31**	.316**	.002ns	-.181*	-.25**	-.118ns
9	.115ns	.222*	.072ns	-.24**	.35**	.369**	.071ns	-.03ns	-.168#	-.045ns
10	-.027ns	.107ns	-.09ns	-.15ns	.239*	.248**	-.198*	-.14ns	-.204*	-.112ns

Las relaciones hipotetizadas, tanto con los principales criterios como con los secundarios, se cumplen razonablemente, lo que muestra ítems relativamente válidos. Los criterios principales son respuesta de los superiores y riesgo real. Todos los ítems se muestran relacionados positiva y significativamente con respuesta de los superiores, salvo el ítem 1. No obstante, éste ítem se muestra especialmente relacionado con riesgo real, por lo que no puede decirse que esté falto de validez criterial. Los ítems 2, 9 y 10 no están relacionados con riesgo real, pero por contra están adecuadamente relacionados con respuesta de los superiores.

Menos importantes, pero muy interesantes, son las correlaciones con tensión y respuesta de los compañeros. La relación con tensión es como se hipotetizó, elevada y negativa. Por su parte parece que los ítems de conducta muestran una consistente relación con la respuesta de los compañeros en seguridad. La política organizacional también se encuentra relacionada en alguna medida con los ítems de conducta, especialmente los número 4, 5, y 6, probablemente debido a los contenidos de los ítems, que hacen referencia a las normas de seguridad, si se desconectan máquinas cuando funcionan mal, etc, y si se trabaja con cansancio o sueño, conductas de seguridad todas ellas que tienen mucho que ver con la política de seguridad que tenga la empresa más que con la voluntad del trabajador.

### 3.5. Conclusiones

La escala de conducta hacia la seguridad ha mostrado a través de los análisis que es fiable y razonablemente válida. Quizá se esperaba una mayor relación de los items con el nivel de riesgo real, pero los datos no serán definitivos hasta que no se estudie la relación de las medidas globales, en lugar de los items, y en un marco multivariado, en lugar de bivariado. No obstante, la estructura factorial de la escala, aún cuando se ha obtenido una buena solución a nivel global y los datos de ajuste analítico mostraban relaciones altamente significativos parece algo compleja. Cuatro factores, algunos correlacionados y otros no, items relacionados con más de un factor, etc parecen razones para detenernos en este punto. En la tabla 6.19 pueden consultarse las correlaciones entre los items, observándolas veremos cosas interesantes que explican, en parte, los resultados.

Tabla 6.19. de correlaciones entre los items de la escala de conducta hacia la seguridad de los trabajadores.

	1	2	3	4	5	6	7	8	9	10
1	1.000									
2	0.043	1.000								
3	<u>0.499</u>	-0.045	1.000							
4	0.125	<u>0.510</u>	0.069	1.000						
5	<u>0.363</u>	0.068	<u>0.444</u>	0.107	1.000					
6	0.138	0.225	0.125	0.195	0.238	1.000				
7	<u>0.377</u>	0.072	<u>0.333</u>	0.239	<u>0.488</u>	<u>0.333</u>	1.000			
8	0.086	0.239	0.127	0.241	0.179	0.249	0.198	1.000		
9	0.262	0.121	0.151	0.033	0.155	0.268	0.279	0.240	1.000	
10	0.094	0.174	0.148	0.202	0.162	<u>0.378</u>	<u>0.354</u>	0.089	0.124	1.000

En la tabla de correlaciones se han subrayado aquellas que excedían de 0.30. De las cuarenta y cinco relaciones entre los items de la escala sólo 10 alcanzan esa cifra. Lo que estas bajas bajas correlaciones puede indicar es que los trabajos que componen la muestra son muy diferentes y en consecuencia los trabajos a desempeñar, las conductas a realizar y las conductas de seguridad o inseguridad a desarrollar distintos. De hecho puede darse casos de que cuando se den un tipo de conductas no se den otro, etc, de que

ciertas conductas sean comunes en ciertos trabajos e inexistentes otras. Todo este comentario nos hace entender la dificultad de encontrar una estructura factorial que no plasmara esta complejidad y la ausencia de relación o baja relación entre factores. Al igual que se comentó para el caso de los riesgos, otra alternativa hubiera sido operacionalizar la variable como check-list, como un inventario de conductas. Esta es efectivamente una posibilidad, y habrá que probar qué forma de medida puede llevar a una mejor explicación de la accidentabilidad. No obstante este comentario más que para llegar a la conclusión de que hay que cambiar de medida sirve para valorar más los resultados obtenidos para la escala, aún teniendo en cuenta sus dificultades intrínsecas.



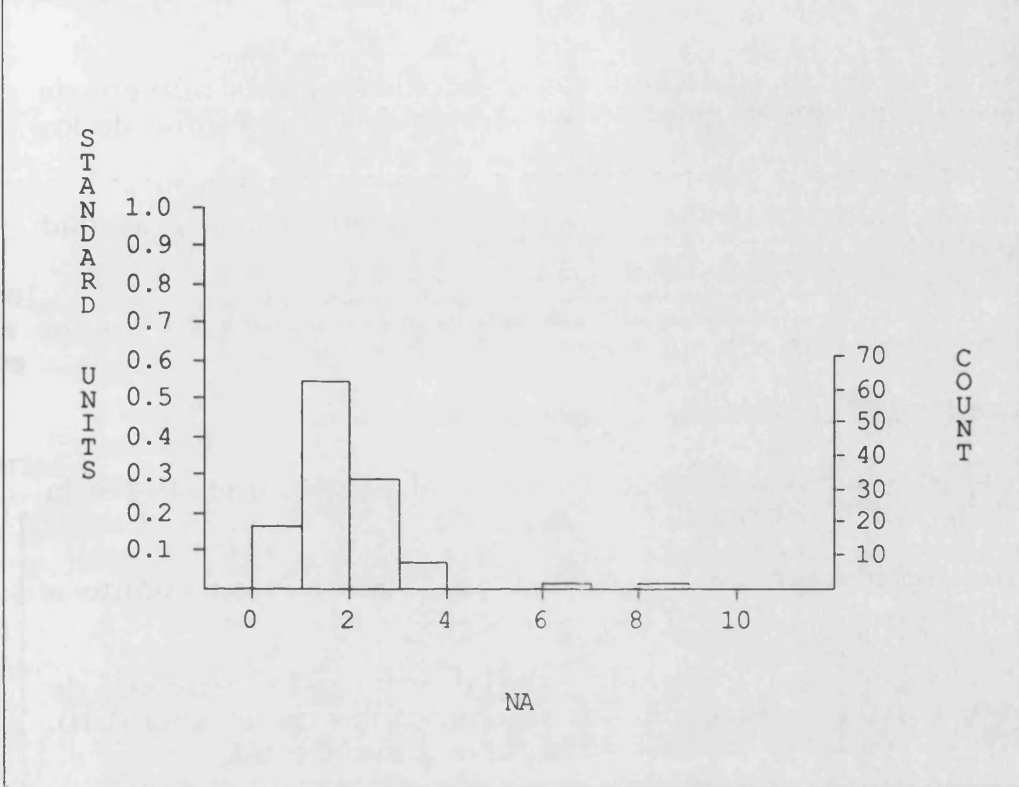
#### **4. Índices para la medición de la siniestralidad**

La medida más usual para operacionalizar la siniestralidad en los estudios psicológicos es el número de accidentes ocurridos para los trabajadores de la muestra durante un cierto periodo de tiempo. Esta variable así definida presenta dos problemas estadísticos graves (Hansen, 1989). El primero de los problemas es la poca variabilidad y el segundo la extremada asimetría. Aún cuando se escojan periodos de tiempo muy elevados, la ocurrencia de un accidente es un hecho muy poco probable. De esta forma, lo común en una muestra de trabajadores extraída al azar es encontrar pocos accidentados, de los pocos accidentados la mayoría tienen un único accidente y el resto se van distribuyendo, cada vez con menos probabilidad en dos, tres, cuatro etc. Esta situación puede paliarse en parte escogiendo períodos de tiempo muy amplios, pero esto a su vez presenta graves inconvenientes, como los cambios que se han producido en los trabajos, en las empresas y en los mismos sujetos en ese tiempo.

Por otro lado, recoger únicamente el número de accidentes puede esconder otros aspectos importantes que pueden conceptualizarse también como accidentabilidad, como por ejemplo la gravedad de los accidentes, el número de incidentes, etc, restando sensibilidad a la medición de la siniestralidad.

Si observamos la distribución de la variable número de accidentes en nuestra muestra de 124 trabajadores observaremos estos problemas mencionados (gráfica 6.10). Sólo una apreciación queda por hacer. Hemos dicho que el comportamiento usual de esta variable es encontrar muchos ceros, porque el número de accidentados es relativamente bajo. Esto no parece ocurrir en nuestra muestra, pero esto es debido al procedimiento de recogida de información. Al ser a través de los trabajadores que iban a los servicios de un mútua de accidentes, por definición, éstos eran accidentados. Por lo demás puede comprobarse fácilmente el comportamiento altamente asimétrico positivo. Las categorías de uno y dos accidentes acaparan la grandísima mayoría de casos.

Gráfica 6.10. Histograma para la variable número de accidentes en la muestra de 124 trabajadores.



Los problemas estadísticos y psicométricos asociados a la variable accidentabilidad puede afectar los resultados hallados en estudios correlacionales. La falta de variabilidad y la grave asimetría puede afectar la cuantía y la evaluación de la probabilidad de las correlaciones y los resultados de los estudios con modelos causales, según el tipo de estimación de que se trate. Una de las formas de solucionar estos problemas ha sido la creación de índices compuestos de accidentabilidad que incluyeran diversos aspectos como la gravedad de los accidentes, etc. Se pretendía aumentar la sensibilidad, aumentar la variabilidad y acercar a la normalidad o por lo menos eliminar parte de la asimetría. Así, por ejemplo Hansen (1989) en un estudio de modelos causales sobre accidentes plantea la medida que el denomina consistencia de los accidentes, que consiste en el número de accidentes de trabajo más el número de años en que el trabajador tuvo al menos un accidente. Lo que supone Hansen es que esta medida operacionaliza la consistencia de la tendencia a accidentarse del trabajador al unir una medida de frecuencia temporal al número de accidentes.

En nuestro trabajo hemos probado diversas medidas de accidentabilidad, de entre ellas hemos escogido tres por ser las de mejor funcionamiento para un estudio exhaustivo. Las medidas son:

1) Número de accidentes. Operacionalizada como número de accidentes en los últimos cinco años.

2) AC4: Se define como número de accidentes más número de años diferentes en que se produjo un accidente más gravedad de los accidentes.

3) AC16: Se define como número de accidentes más gravedad de los accidentes.

Para evaluar la idoneidad de las medidas tendremos en cuenta los siguientes criterios.

1) Que tengan el máximo de variabilidad.

2) Que se ajusten a una ley normal o, al menos, que tengan la mínima asimetría y curtosis.

3) Que mantengan un comportamiento adecuado en cuanto a validez criterial.

Para evaluar la normalidad se han realizado pruebas de Kolmogorov-Smirnov. Los resultados se pueden ver en la tabla 6.10. Ninguna de las variables puede considerarse como normal.

*Tabla 6.20. Pruebas de Kolmogorov-Smirnov realizadas sobre los datos de las tres medidas de accidentabilidad para probar el ajuste a la normalidad.*

Variable	N muestral	Diferencia máxima	Probabilidad
Número ACC	124	.805	.000
AC 4	124	.688	.000
AC 16	124	.784	.000

Tras estos resultados se calcularon los estadísticos descriptivos de las tres variables, que se presentan en la siguiente tabla. De especial interés son la asimetría, la curtosis y la variabilidad.

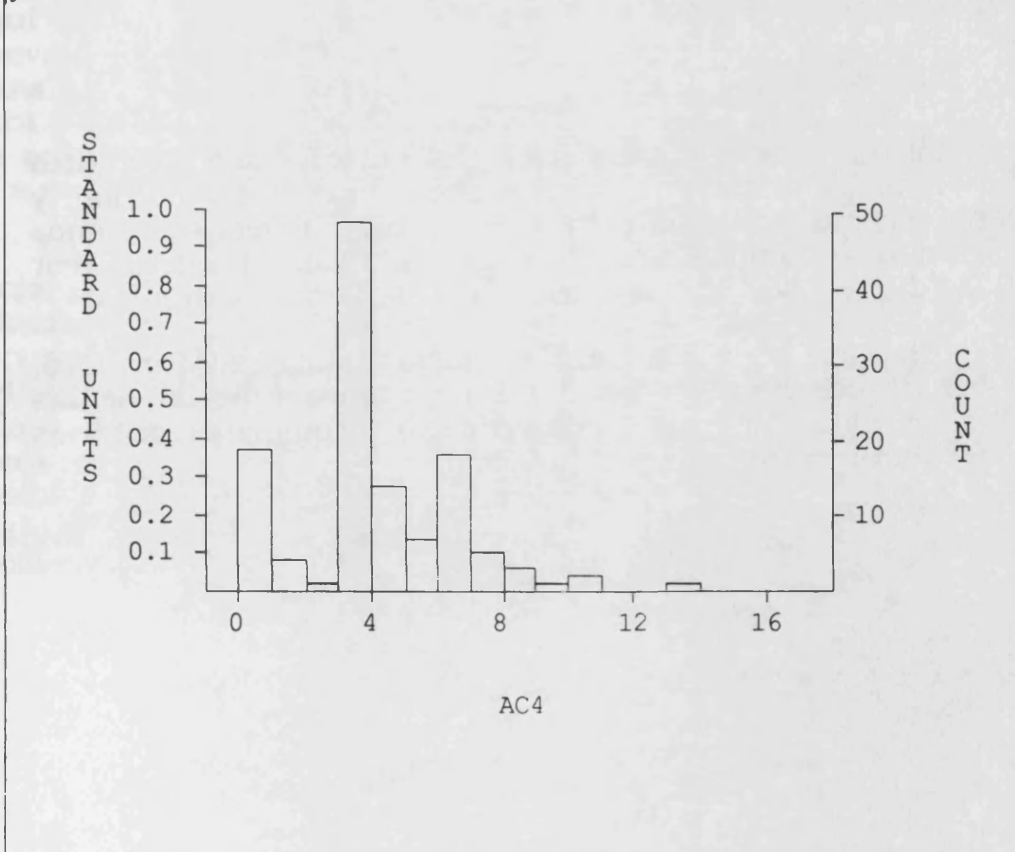
Tabla 6.21. Estadísticos descriptivos de los tres índices de accidentabilidad.

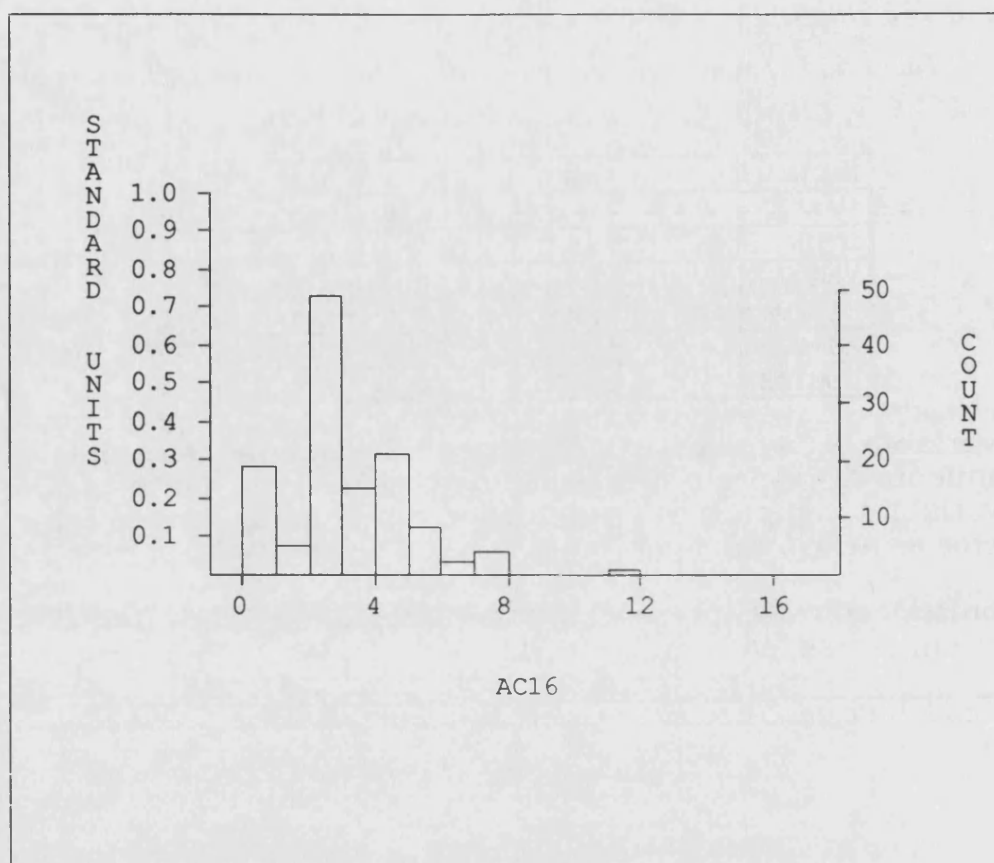
	AC4	AC16	NA
MINIMO	0.000	0.000	0.000
MAXIMO	13.000	11.000	8.000
MEDIA	3.653	2.605	1.339
VARIANZA	5.919	3.412	1.169
DESVIACION TIPICA	2.433	1.847	1.081
ASIMETRIA	0.649	1.049	2.552
CURTOSIS	1.108	2.622	12.417

La variable AC4 y la AC16 tienen claramente un mejor comportamiento al respecto de variabilidad, asimetría y curtosis. De las dos variables compuestas la de mejor comportamiento en estos dos aspectos es AC4.

A continuación se presentan los histogramas para las dos variables compuestas, en la gráfica 6. 11.

Gráfica 6.11. Histograma para las variables compuestas de accidentabilidad AC4 y AC16.





No obstante, aunque la asimetría y la curtosis son diferentes parece que ambas variables incumplen claramente la ley normal, y esas diferencias en apuntamiento y asimetría pueden ser menos importantes que su comportamiento en cuanto a validez criterial. Por su parte, las diferencias en desviación típica no parecen extremas.

Para la elección de la medida de accidentabilidad, AC4 o AC16, se debe basar más bien en los datos de validez criterial de las variables. Los datos de correlación con los principales criterios pueden consultarse en la tabla a continuación.

Tabla 6.22. Correlación de los tres índices de accidentabilidad con los criterios más relevantes. (\*\*= .01; \*= .05; #=marginamente significativo; ns= no significativo).

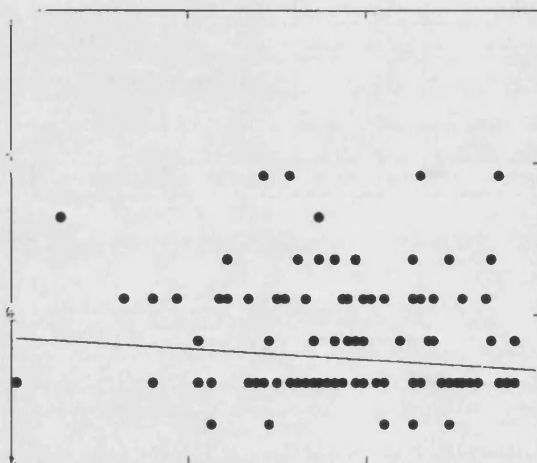
	NA	AC4	AC16
Política	-.094ns	-.101ns	-.116ns
Formación	-.119ns	-.131ns	-.147ns
Estructura	.019ns	-.003ns	.000ns
R. Superiores	-.122ns	-.162ns	-.152ns
R. Compañeros	-.161ns	-.205*	-.181#
Tensión	.167#	.158ns	.164#
Conducta	-.074ns	-.084ns	-.084ns
Riesgo Real	.236*	.240*	.254**
Riesgo Basal	.243**	.267**	.280**

Aunque en general las correlaciones son muy bajas, puede verse cómo AC16 presenta una mayor relación con los criterios. Especialmente importante como criterios son el nivel de riesgo real, antecedente inmediato de los accidentes y el riesgo basal, condiciones de partida del trabajo para que se produzcan accidentes. Aquí la cuantía de las relaciones no difiere excesivamente, pero es nuevamente AC16 la de comportamiento ligeramente superior. Ante estas bajas correlaciones puede resultar crítica la elección de medida para poder encontrar relaciones estadísticamente significativas con los accidentes al evaluar modelos causales. Por ello, la medida elegida va a ser AC16, en adelante accidentabilidad.

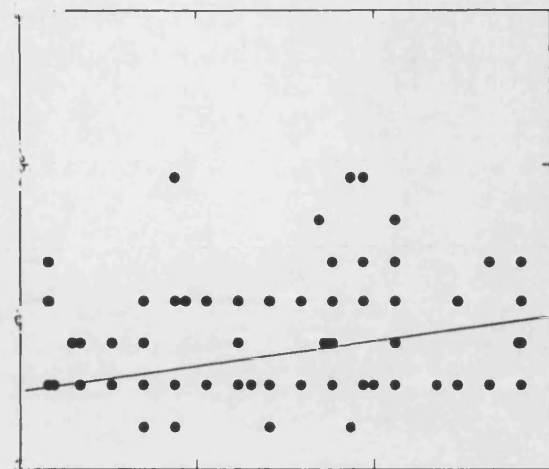
Un punto importante cara a la utilización de modelos causales y otras técnicas como regresión lineal múltiple es evaluar el tipo de relación que mantienen las variables. Hemos visto cómo las correlaciones de accidentabilidad con los criterios son bajas. Evidentemente esto no implica que no exista relación, sino solamente que no existe relación lineal. Podría ocurrir que la relación fuera de tipo no lineal. Así en algunas ocasiones se han planteado posibles relaciones curvilíneas entre los riesgos y los accidentes. Para evaluar este punto se ha graficado la relación bivariada entre accidentabilidad y las variables riesgo real, riesgo basal y conducta.

Gráfica 6.12. Relación entre la medida de accidentabilidad y sus principales criterios, conducta, riesgo real y riesgo basal.

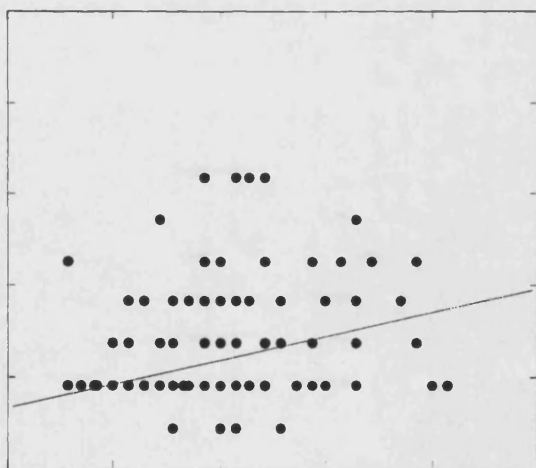
a) relación entre conducta y accidentabilidad



CHSN



RB



RR

A  
C  
1  
6



Como puede apreciarse en las gráficas que relacionan accidentabilidad con los criterios las relaciones son débiles pero lineales. No aparece ningún patrón curvilíneo ni de otro tipo de forma clara, por lo que es razonable probar el ajuste de modelos causales con accidentabilidad como variable dependiente, a explicar. Este será el objeto del próximo capítulo.

## 7. Modelos causales explicativos de los accidentes y conclusiones.

### 1. Relaciones bivariadas

Se van a presentar a continuación las correlaciones entre las diez variables implicadas en el modelo. Estas correlaciones se han ordenado en la tabla 7.1, donde pueden consultarse.

Las correlaciones cumplen en este contexto dos funciones. Primero, sirven de validación criterial de los cuestionarios. Segundo, los resultados de relación bivariada entre las variables del modelo, junto con las hipótesis teóricas de relación entre las variables nos permitirán poner a prueba modelos causales de explicación causal de los accidentes. Estos modelos causales cumplen, además, un importante papel en la validación nomológica de los cuestionarios.

Variable a variable iremos comprobando ambos aspectos, la validez criterial y las relaciones causales que puede mantener cada variable con el resto en el modelo. Evidentemente, las relaciones en un marco bivariado no tienen porqué mimetizar los patrones de relación en un marco multivariado, donde para cada relación se

ejerce control estadístico sobre el resto de variables. No obstante, comprobar algunos patrones de relación tomados dos a dos ayuda en el posterior planteamiento de hipótesis causales multivariadas. Digamos que es un paso exploratorio necesario, a menos que se tenga teóricamente claras todas las relaciones hipotéticas a priori, lo que no es nuestro caso.

Tabla 7.1. Tabla de correlaciones entre las diez variables consideradas en los modelos causales explicativos de los accidentes laborales. ES= estructura de seguridad; F= Formación-información sobre seguridad; PO= política de seguridad; RS= respuesta de los superiores; RC= respuesta de los compañeros; CHS= conducta hacia la seguridad del trabajador; RR= riesgo real; RB= riesgo basal; T= tensión; AC= medida de accidentabilidad. (\*\*=  $p < 0.01$ ; \*=  $p < 0.05$ ; #= marginalmente significativo; ns= no significativo).

	ES	F	PO	RS	RC	CHS	RR	RB	T	AC
ES	1.000									
F	.184#	1.000								
PO	.416**	.588**	1.000							
RS	.230*	.516**	.749**	1.000						
RC	.178#	.269**	.496**	.651**	1.000					
CHS	.056ns	.146ns	.382**	.581**	.516**	1.000				
RR	.030ns	-.097ns	-0.195*	-0.279*	-.110ns	-.33**	1.000			
RB	.000ns	-.077ns	-.201*	.099ns	-.080ns	-.033ns	.478**	1.000		
T	-.016ns	-.126ns	-.216*	-0.39**	-.31**	-.46**	.338**	.115ns	1.000	
AC	.00ns	-.147ns	-.116ns	-.152ns	-.181#	-.084ns	.254**	.280**	.153ns	1.000

La variable estructura de seguridad se pensaba que se encontraría relacionado fuertemente con los otros aspectos del clima, política de seguridad y formación, y que estaría afectando principalmente a la respuesta que los supervisores dan en materia de seguridad, por el papel de control que según la ley pueden ejercer. A la vista de las correlaciones estas hipótesis sólo se cumplen parcialmente. Existe una relación relativamente elevada con la variable política y una baja relación con la formación, que parece apuntar hacia la no relación de ambas variables. Por su parte la correlación con la respuesta de los superiores existe pero no es elevada ( $p < 0.05$ ), el impacto habrá que evaluarlo mejor en un marco multivariado. No presenta relaciones relevantes con el resto de variables del modelo.

El siguiente aspecto del clima, la variable formación, se hipotetiza como relacionada con los otros aspectos de clima y con la respuesta de los superiores y trabajadores, incluyendo la conducta hacia la seguridad del trabajador, ya que son los que están sometidos a entrenamiento e información sobre la seguridad. Las relaciones con política son fuertes y con la respuesta de los superiores también, ambas estadísticamente significativas ( $p < 0.01$ ). Parece que los esfuerzos de formación tienen que ver con una adecuada política de seguridad y que afectan el comportamiento de los superiores. También se aprecia un cierto impacto sobre la respuesta de los compañeros, pero es mucho menor éste sobre la conducta del trabajador. El resto de relaciones no son estadísticamente significativas.

El comportamiento de la variable de política es totalmente conforme a hipótesis. La política organizacional se espera que esté muy relacionada con la respuesta de los superiores, y efectivamente es así (0.749,  $p < 0.01$ ). También se espera que las acciones de la empresa en seguridad tengan influencia en los trabajadores, ya sea en la respuesta de los compañeros o en la conducta del trabajador. Esas dos relaciones son también elevadas. Además de estas hipótesis teóricamente determinantes para evaluar la validez de la medida, se puede apreciar en la tabla 7.1 cómo también se encuentra relacionada, aunque menos fuertemente, con las variables de riesgo y tensión (negativamente como se espera por hipótesis). La relación con accidentabilidad no alcanza la significatividad estadística.

Las variables de contexto social, respuesta de los superiores y respuesta de los compañeros, tienen un importante papel teórico al ser el nexo de unión entre la empresa y el trabajador, y son determinantes inmediatos de la conducta del trabajador. Estas variables cumplen con esta definición teórica. Efectivamente ambas están relacionadas positiva y fuertemente con la conducta de seguridad del trabajador, 0.651 y 0.516 ( $p < 0.01$ ), respectivamente. Y además, ambas se encuentran también relacionadas, 0.496 ( $p < 0.01$ ). Ambas variables se encuentran relacionadas con el nivel tensión de forma negativa. Este resultado apunta a que el nivel de tensión elevado del trabajador le hace percibir un contexto social con una menor respuesta de seguridad. también considerar que la respuesta de los compañeros tienen una relación marginalmente significativa con la accidentabilidad.

La variable conducta hacia la seguridad tienen tres criterios principales. La respuesta de los superiores y de los compañeros, como antecedentes inmediatos, cuya relación ya hemos visto era conforme a hipótesis. Como consecuente principal de la conducta de seguridad tenemos a la tercera variable criterio, el nivel de riesgo real. Por la definición del riesgo real, se espera que la conducta segura-insegura del trabajador sea un importante de este tipo de riesgo. La relación entre ambas variables es, efectivamente,

estadísticamente significativa y negativa, 0.331. Cuanto mayor es la conducta de seguridad más descende la probabilidad de accidentarse. Muy interesante es la relación entre conducta y tensión. Aunque no es teóricamente claro el sentido de la relación explicativa, sí es claro que existe una relación significativa. Queda por decidir si la tensión puede predecir una parte de las conductas inseguras que se realizan o es el verse obligado a tomar riesgos y realizar conductas inseguras lo que genera tensión. Cabe también la posibilidad que una tercera variable explique la covariación entre ambas variables, como un elevado nivel de riesgo basal.

Las variables de riesgo tienen un comportamiento criterial adecuado. Hemos visto que riesgo basal debe estar relacionado con el nivel de riesgo real, y esta relación es efectivamente alta, 0.478 ( $p < 0.01$ ). Por ser el nivel intrínseco de riesgos puede afectar en cierta medida al resto de variables pero no parece que sea el caso en esta muestra, ya que el resto de relaciones de riesgo no son elevadas, a excepción de la relación con accidentabilidad. El riesgo real, por su parte cumple bien con sus relaciones criterioales, al estar relacionada con la conducta de seguridad, el nivel de riesgo basal y la tensión, posibles antecedentes y ser la variable más relacionada con accidentabilidad (0.280,  $p < 0.01$ ).

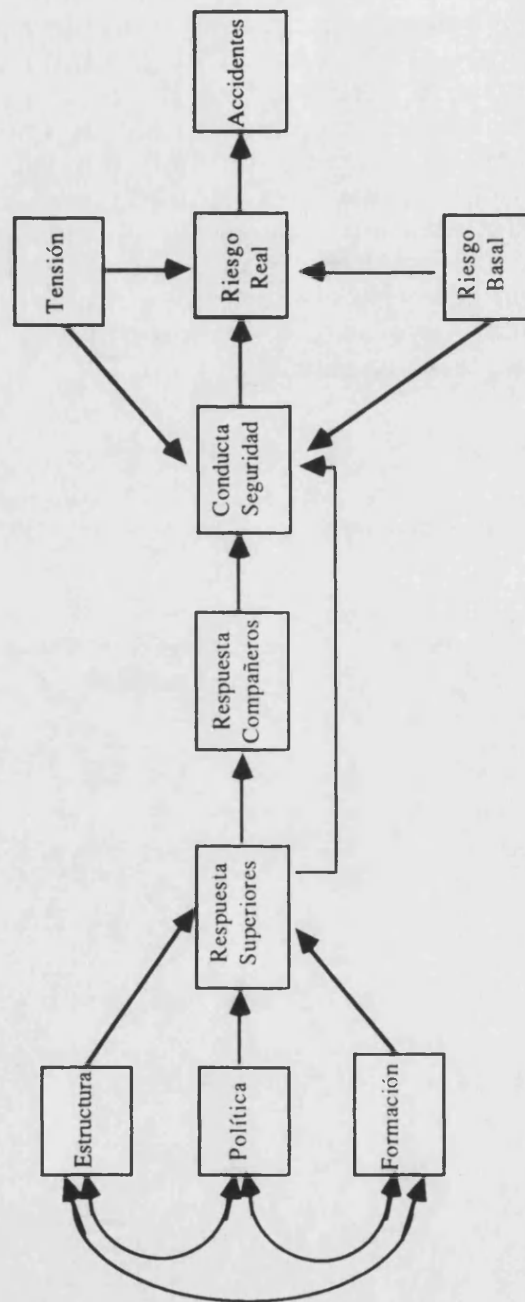
## **2. Modelo hipotético**

A la vista de las relaciones bivariadas y las consideraciones teóricas puede plantearse un primer acercamiento multivariado a las relaciones causales entre las variables, y plantear un primer modelo causal hipotético. La base del mismo es una cadena que va desde los niveles empresariales, que afectan a las variables del contexto social del trabajador, éstas afectando a la conducta, y la conducta junto con los riesgos explicando los accidentes. Si depuramos un poco esta cadena causal e incluimos en una primera aproximación el papel de la tensión en el conjunto obtenemos un modelo causal como el que se presenta en la gráfica 7.1.

Las tres variables de clima de seguridad se consideran variables exógenas, y se espera que guarden una cierta covariación positiva entre ellas. Estos aspectos de clima determinan directamente la conducta de los supervisores, que son el nexo de unión entre empresa y trabajador. Son los supervisores los encargados concretar la política de la empresa y afectar con estas medidas la respuesta de los trabajadores. ello se hipotetiza una influencia causal de los supervisores sobre los compañeros del trabajador y sobre la propia conducta del trabajador. Los compañeros tienen un papel importante al determinar la conducta segura-insegura del trabajador, por ser un importante grupo de referencia de éste. Se hipotetiza también una cierta relación del riesgo basal con la conducta. La conducta de seguridad, la tensión y el riesgo basal determinan, por su parte, el

riesgo real de tener un accidente, y éste es el antecedente directo de los accidentes. La tensión se asume, de momento, como afectando a la conducta, pero la relación podría darse al contrario. De esta forma, las variables tensión y riesgo basal se asumen como exógenas, su explicación es externa al modelo. Una de las razones por las que no se plantea, en principio, la relación causal de conducta a tensión es precisamente porque la variable tensión, bastante estudiada en la literatura se encuentra muy relacionada con una diversidad de variables no consideradas aquí, de estrés, satisfacción, etc, por lo que difícilmente puede considerarse como determinada por la conducta de seguridad únicamente. No obstante es un camino de causación que puede probarse más adelante para cuantificar el efecto, aunque previsiblemente no sea muy elevado.

Gráfica 7.1. Modelo causal hipotético de la ocurrencia de accidentes.



El diagrama muestra un modelo de ecuaciones estructurales recursivo. Esto indica que no hay ningún camino de doble causación, ni bucles de causación, ni correlaciones entre los errores (Bollen, 1989). Las variables exógenas, cuya explicación es externa al modelo, son la tensión, el nivel de riesgo basal, y las tres variables de clima, a saber, política, formación y estructura de seguridad. Las variables endógenas, por su parte, son respuesta de los superiores y compañeros, conducta hacia la seguridad, riesgo real, y como variable dependiente última, la accidentabilidad. El modelo, por ser

recursivo, tiene garantizada la identificación de los parámetros (Bollen, 1989; Long, 1983; Visauta, 1986).

### 3. Modelos causales

El modelo original presentado en la gráfica 7.1 ha sido sometido a contraste estadístico por dos métodos de estimación, máxima verosimilitud y mínimos cuadrados generalizados. Los resultados de ajuste global pueden verse en la tabla a continuación.

Tabla 7.2. Índices de ajuste global para el modelo original. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: GLS=mínimos cuadrados generalizados; ML=máxima verosimilitud.

		chi	g	p	F	LHR	chi/g	NFI	NNFI	CFI
MODELO	GLS	37.8	30	.154	.343	.842	1.26	.989	.997	.998
ORIGINAL	ML	47.0	30	.024	.427	.807	1.56	.879	.926	.951

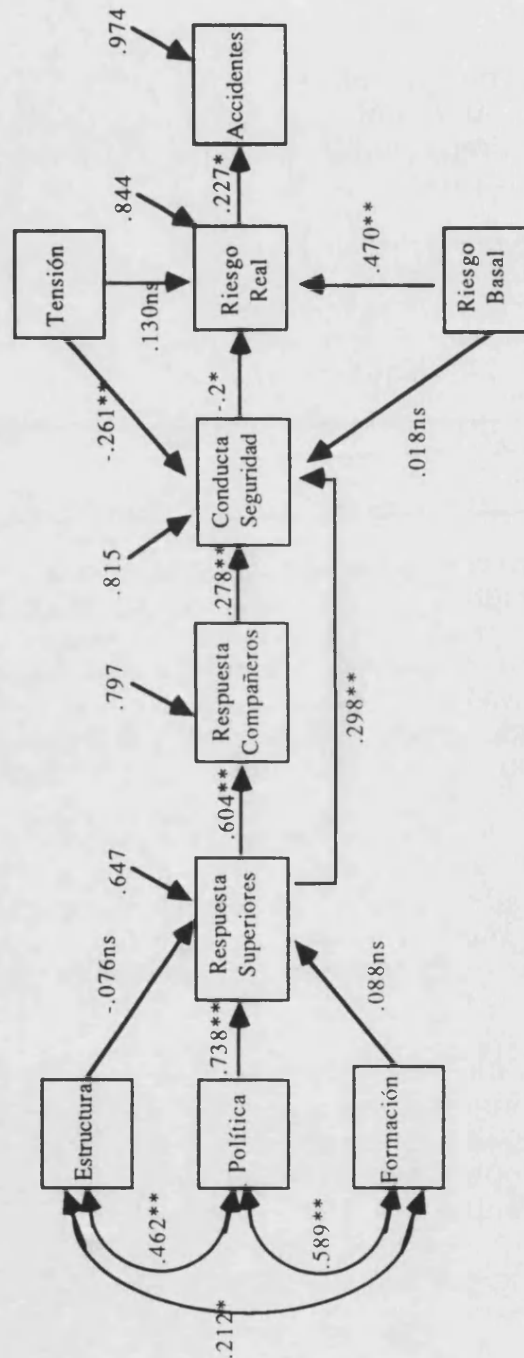
En la tabla parece claro que la estructura representa bien los datos, de forma global, en el caso de la estimación mínimo cuadrática. La prueba de ji-cuadrado no es estadísticamente significativa el resto de índices muestran también una claro ajuste, LHR elevado, NFI, NNFI y CFI por encima de 0.9. Por su parte, la ji-cuadrado partido los grados de libertad muestra la parsimonia del modelo, la reducción que supone en la complejidad de las relaciones.

La evaluación del ajuste no es tan clara desde la estimación máximo verosímil. La prueba de ji-cuadrado es significativa, indicando ajuste inadecuado. y el índice NFI, no alcanza el 0.9 límite inferior teórico del buen ajuste. El resto de índices dan apoyo a la estructura teórica, con valores que indican un alto grado de parsimonia.

Queda por evaluar los datos de ajuste analítico; las pruebas z sobre la significatividad de cada parámetro y los datos que aporte el LM test y el W test. Las estimaciones del modelo para la estimación mínimo cuadrática pueden verse en la gráfica 7.2, junto con los resultados de probabilidad de las pruebas z asociadas a cada parámetro.



Gráfica 7.2. Modelo causal hipotético de la ocurrencia de accidentes estimado mediante mínimos cuadrados generalizados. (\*\*=prob<0.01; \*=prob<0.05; #=marginalemente significativo; ns= estadísticamente no significativo).



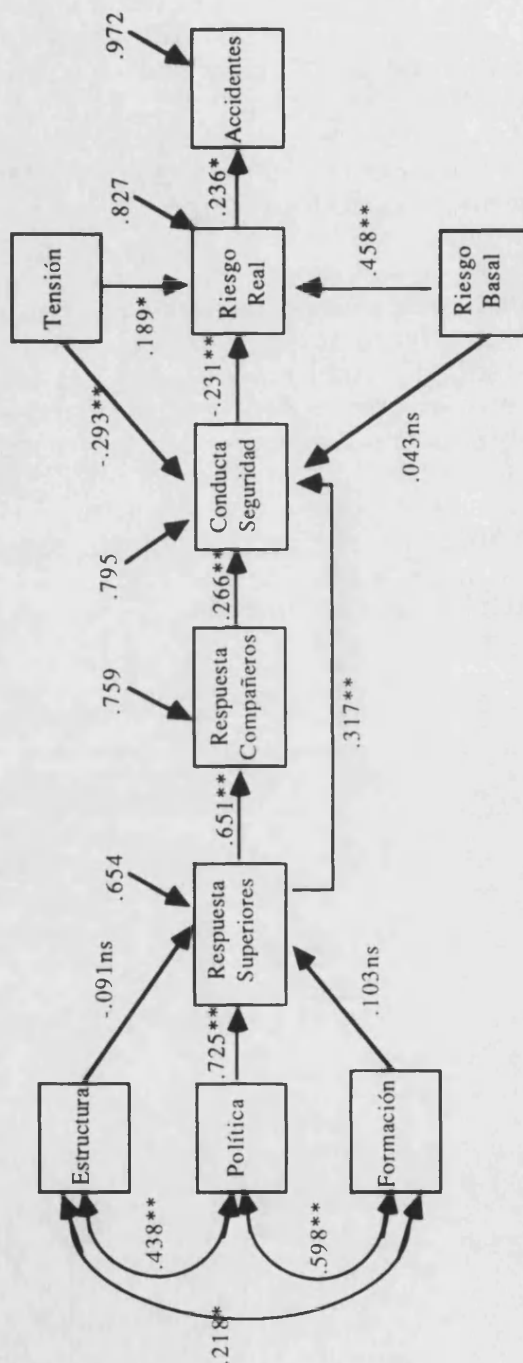
Las hipótesis de relación de las variables se cumplen en su mayoría. No obstante, algunas de las relaciones hipotetizadas no alcanzan la significatividad estadística. Puede apreciarse cómo, en general, aunque las relaciones son en su mayoría significativas y

conforme a hipótesis los valores alcanzados no son excesivamente elevados. Las tres variables de clima se encuentran correlacionadas entre sí, sobre todo en lo que respecta a la correlación de política con estructura y con formación. La correlación entre estas dos últimas es menor. De todas las variables de clima sólo la política organizacional es capaz de predecir la respuesta de los superiores. El papel de la estructura de seguridad y de la formación dentro de la cadena de accidentes queda en entredicho, dado que salvo la relación que presentan con clima no son capaces de predecir ninguna otra variable de la cadena de accidentes, por lo que su efecto podría en todo caso ser indirecto.

Otras variables que no cumplen la hipótesis de relación planteadas tienen un efecto menos importante sobre la concepción de la cadena causal de accidentes. Así, la tensión no predice significativamente el nivel de riesgo real, mientras el riesgo basal no tienen un efecto sobre la conducta hacia la seguridad.

Además de los aspectos comentados, y como resultado especialmente relevante, puede verse cómo la relación de riesgo real con accidentes es más baja de lo esperado teóricamente. La capacidad predictiva de los accidentes se encuentra en un pobre 5.15%.

Gráfica 7.3. Modelo causal hipotético de la ocurrencia de accidentes estimado mediante máxima verosimilitud. (\*\*=prob<0.01; \*=prob<0.05; #=marginalemente significativo; ns= estadísticamente no significativo).



Los resultados de la estimación máximo verosímil son muy similares a los encontrados mediante mínimos cuadrados generalizados. En general puede verse que los valores de la estimaciones han ascendido ligeramente. Así, la tensión es ahora un predictor estadísticamente significativo de los accidentes, mientras

que la relación de riesgo real con accidentes ha aumentado algo, de 0.227 a 0.236. En general, no obstante, se mantienen los mismos sentidos y cuantía de relaciones, no hay ningún cambio digno de mención.

Ante los resultados que mostraban cómo algunos caminos de causación no eran correctos tal y como se hipotetizaron, y ante la falta de un ajuste plenamente satisfactorio por máxima verosimilitud, se decidió acudir a los datos del LM test y del W test. Las sugerencias desde ambos tipos de estimación eran coincidentes. Por parte del W test se proponía eliminar las relaciones de estructura de seguridad y formación con respuesta de los superiores y la de riesgo basal con conducta de seguridad. Estos cambios se supone no tendrán un efecto significativo sobre el ajuste global del modelo. Estas relaciones ya habían sido evaluadas como no significativas por la prueba z aplicada a cada relación. Mientras, el LM test proponía la existencia de una relación estadísticamente significativa e importante entre tensión y la respuesta de los superiores. Esta relación es problemática. Por un lado, puede ocurrir que se deba a la percepción del sujeto, que al estar sometido a tensión evalúa como menos positiva la respuesta de seguridad de los superiores. Por otro, puede darse la relación contraria ya que una respuesta inadecuada por parte del superior puede generar tensión en el sujeto. Ambas hipótesis serán probadas.

Dado el poco efecto de las variables estructura de seguridad y formación en la cadena causal de accidentes y tomando en cuenta las sugerencias para eliminar sus relaciones hipotetizadas se decidió eliminarlas del análisis como tales variables. Además se eliminó la relación de riesgo basal con conducta, con lo que el papel del riesgo basal es únicamente predictor del riesgo real. El modelo así modificado fue sometido de nuevo a estimación por máxima verosimilitud y mínimos cuadrados generalizados. Los datos de ajuste global se pueden consultar en la siguiente tabla.

Tabla 7.3. Índices de ajuste global para el modelo modificado en base a sugerencias del LM test y del W test. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: GLS=mínimos cuadrados generalizados; ML=máxima verosimilitud.

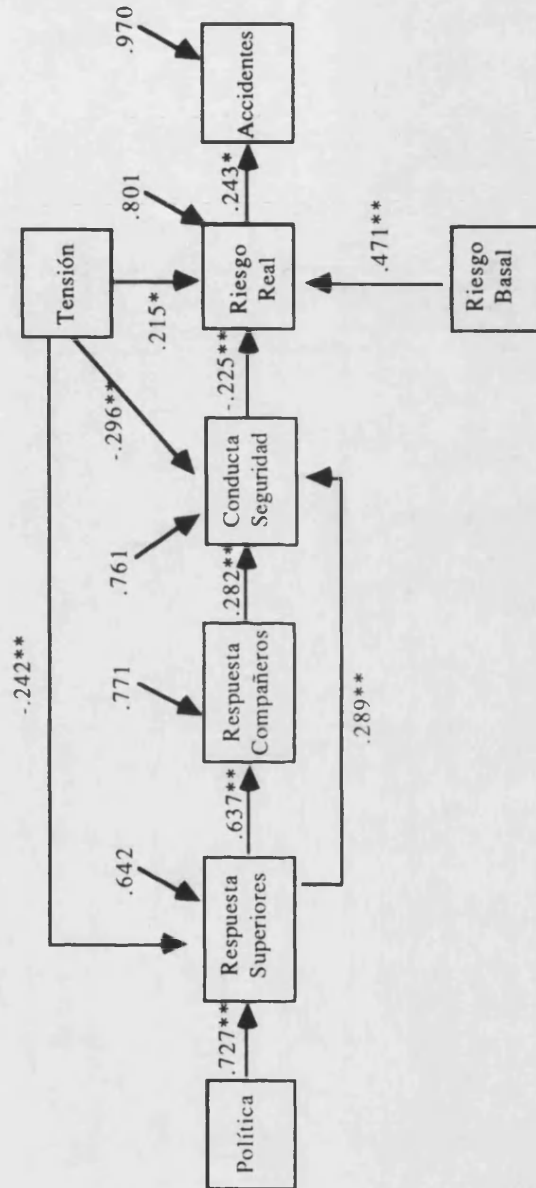
		chi	g	p	F	LHR	chi/g	NFI	NNFI	CFI
MODELO	GLS	20.74	18	.292	.188	.910	1.152	.992	.998	.999
MODIFICADO	ML	26.91	18	.080	.244	.884	1.495	.912	.950	.968

La nueva estructura planteada ajusta, para ambos métodos de estimación, según todos los índices. Para la estimación por mínimos cuadrados generalizados ni el LM test ni el W test hacen ninguna sugerencia. Sin embargo, el LM test hace una sugerencia teóricamente plausible, y que plantearemos para encontrar un mejor

ajuste. Propone una relación entre riesgo basal y política, que plantearemos como una covarianza entre variables exógenas. El resultado de ajuste global para la estructura así hipotetizada y calculada por máxima verosimilitud es bastante más explicativa. Así, la prueba de ji-cuadrado pasa a valer 21.565 que para 17 grados de libertad significa una probabilidad de 0.2020. El resto de índices también indican ajuste adecuado, como los de comparación con un modelo nulo, siempre por encima del límite inferior de ajuste: NFI=0.93; NNFI=0.973; y CFI=0.981. Por su parte la razón de verosimilitud escalada es muy elevado, 0.906, mostrando buen ajuste. Por último, también la razón de ji-cuadrado por los grados de libertad indica un modelo altamente parsimonioso, 1.268.

Dado que este cambio es menor, pero interesante, se van a presentar en las gráficas 7.4 y 7.5 los resultados de mínimos cuadrados generalizados sin la covarianza entre política y riesgo basal y los de máxima verosimilitud con ella, respectivamente.

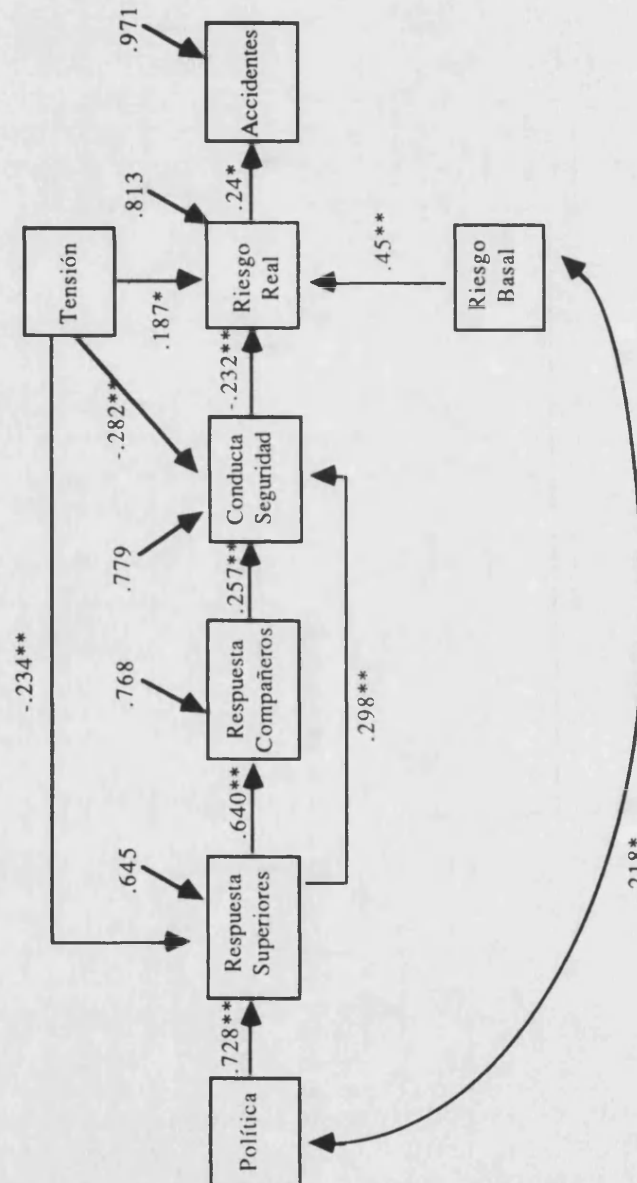
Gráfica 7.4. Modelo causal hipotético de la ocurrencia de accidentes estimado mediante mínimos cuadrados generalizados. (\*\*=prob<0.01; \*=prob<0.05; #=marginalemente significativo; ns= estadísticamente no significativo).



Esencialmente, el patrón de relaciones no se ha visto modificado y las estimaciones son muy similares a las encontradas en los modelos de diez variables que incluían los otros dos aspectos de clima, estructura y formación. El panorama, no obstante, se ha aclarado bastante. Lo mismo ocurre con máxima verosimilitud, solo que aquí aparece una relación interesante, aunque poco elevada que muestra una relación negativa entre el nivel de riesgos y la política de seguridad. Este dato es interesante dado que implica que cuanto mayor es el nivel de riesgos, menos acciones emprende la empresa. Claro que este resultado puede deberse a la percepción del

trabajador, que al estar sometido a mayor nivel de riesgos percibe una menor preocupación por parte de la empresa, lo mismo que ocurría entre tensión y respuesta de los superiores.

Gráfica 7.5. Modelo causal hipotético de la ocurrencia de accidentes estimado mediante máxima verosimilitud. (\*\*=prob<0.01; \*=prob<0.05; #=marginamente significativo; ns= estadísticamente no significativo).



Como era posible que la relación entre tensión y respuesta de los superiores fuera bidireccional, tal y como se ha comentado más arriba, se probó esta posibilidad. Se planteó un modelo causal no recursivo, exactamente igual que el presentado en la gráfica 7.4, salvo en que permitía una doble dirección causal entre la variable tensión y la respuesta de los superiores. Los resultados no apoyaron la hipótesis

de doble dirección causal, ya que aún cuando de forma global ajustaba ( $\chi^2=19.361$ ,  $g=17$ ,  $p=0.30$ ), los datos de ajuste analítico evaluaban la relación causal de respuesta de los superiores a tensión (-.158) como no significativa.

Como hemos venido haciendo hasta aquí se va a plantear la idoneidad de las estimaciones basadas en el supuesto de normalidad multivariable. Si este supuesto no es razonable se pasará a emplear métodos arbitrarios de estimación y a comprobar las diferencias inter-métodos. Las variables que quedan en el análisis tras la eliminación de formación y estructura de seguridad han sido sometidas a las pruebas habituales para comprobar la plausibilidad de la multinormalidad. Realizadas las consiguientes pruebas de Kolmogorov-Smirnov para probar los supuestos de normalidad variable a variable, los resultados indican que ninguna variable es razonablemente normal.

Los datos descriptivos y las pruebas de normalidad aplicadas a las variables muestran la inadecuación de los datos al supuesto de normalidad multivariada. Sin embargo las desviaciones no son tan elevadas como en casos anteriores.

Tabla 7.4. Datos de asimetría y curtosis para las siete variables de los modelos causales.

	Asimetría	Curtosis
Política	0.1606	-0.7411
R. Super.	-.01022	0.9259
R. Compa.	-0.7030	-0.6640
Conducta	-0.7427	0.6429
R. Real	0.4915	-0.1978
R. Basal	0.2158	-1.0688
Tensión	0.4672	-1.0039
Acciden.	1.1034	3.0786

COEFICIENTE DE MARDIA MULTIVARIADO= 2.78  
 Coeficiente esperado para la normal multivariada= 1.158

Ante los datos de incumplimiento del supuesto de normalidad multivariable se decidió aplicar métodos arbitrarios de estimación. En la siguiente tabla pueden consultarse los datos de ajuste global para este método de estimación.

Tabla 7.5. Índices de ajuste global para el modelo estimado por métodos arbitrarios. Los índices son:  $\chi^2$ =chi-cuadrado;  $g$ =grados de libertad del modelo;  $p$ =probabilidad;  $F$ =función de ajuste a minimizar; LHR=razón de verosimilitud escalada; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index.

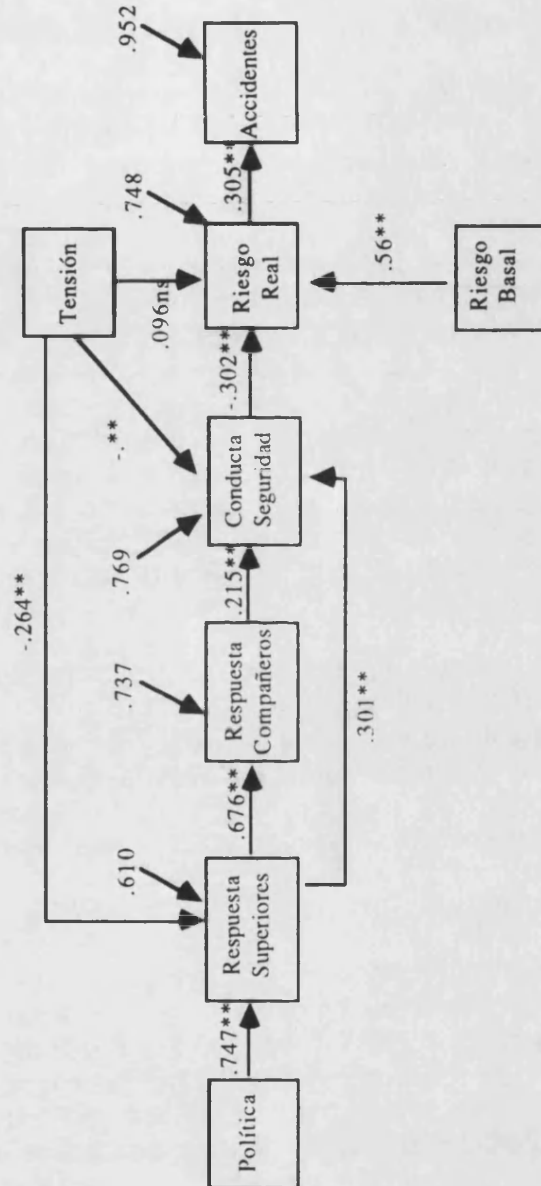


	chi	g	p	F	LHR	chi/g	NFI	NNFI	CFI
AGLS	34.41	18	.011	.312	.855	1.911	.837	.855	.907

INDICE DE BONDAD DE AJUSTE, GFI= .959  
Indice de bondad de ajuste ajustado, AGFI= .918

Los datos no muestran un claro ajuste del modelo a los datos, aunque algunos índices sí ofrecen valores adecuados. Queda por ver el efecto que en las estimaciones concretas se ha producido, y en el ajuste analítico. La gráfica con el diagrama path resultante de la estimación arbitraria puede consultarse a continuación.

Gráfica 7.6. Modelo causal hipotético de la ocurrencia de accidentes estimado mediante métodos arbitrarios. (\*\*=prob<0.01; \*=prob<0.05; #=marginamente significativo; ns=estadísticamente no significativo).



Aunque los datos de ajuste global no mostraban un adecuado ajuste por todos los índices, lo cierto es que al observar los datos de ajuste analítico puede verse cómo los patrones de relación y la significatividad estadística de los coeficientes estructurales se mantiene ampliamente. De hecho se produce una mejora generalizada de los coeficientes estimados, salvo en dos casos, tensión al predecir riesgo y respuesta de los compañeros con conducta. El resto de relaciones obtienen mejoras, que en algunos casos son considerables. Así, la capacidad predictiva de los accidentes es ahora más fuerte

(0.305,  $p < 0.01$ ), lo que supone una capacidad predictiva del 9%, frente al 5% obtenido para mínimos cuadrados generalizados.

#### 4. Modelo considerando los errores de medida de las variables

Hemos visto que los modelos de path analysis se basan en un cierto número de supuestos como la linealidad, aditividad, etc. Y entre estos supuestos que *las variables están medidas sin error*. Esta asunción es muy a menudo incumplida, e incumplida severamente. No se realizan a menudo consideraciones sobre los efectos que el error de medida tiene en las estimaciones, y se siguen aplicando las técnicas acriticamente esperando que los errores cometidos no sean elevados. No obstante, se ha demostrado que no considerar los errores de medida puede llevar a errores de diferente signo y cuantía (Bollen, 1989). Así, algunas de las conclusiones más relevantes de este autor sobre el efecto del error de medida sobre los modelos de path analysis son que el sentido del sesgo que se produce y su cuantía es difícil de predecir. Algunos coeficientes no se ven prácticamente afectados, mientras que otros aumentan o pueden verse atenuados. El impacto de estos cambios en el ajuste global resulta también impredecible.

Bollen presenta un procedimiento para incluir explícitamente los errores de medida en el modelo, de forma que los resultados sean más acordes a las verdaderas relaciones y no estén sesgados por la existencia de error. Para aplicar la fórmula es necesario conocer la fiabilidad de las variables observables y su varianza y de esta forma poder estimar la varianza de error de cada variable.

El procedimiento que propone Bollen (1989) consiste en tratar cada variable como un *indicador con error de medida de un factor latente que se encuentra libre de este error*. En resumen, se trata de estimar la varianza de error de cada variable observable y fijar a la unidad la relación del factor latente con su único indicador, con lo que se da al factor la misma escala de medida que la de la variable observable. En lugar de hipotetizar relaciones estructurales entre las variables observables se plantean éstas entre los factores latentes. Queda por estimar, para cada variable, su varianza de error. Si se dispone de una estimación de la fiabilidad de la medida, Bollen propone la siguiente fórmula para el cálculo:

$$\text{varianza de error} = (1 - \rho_{yy'}) \text{var}(y)$$

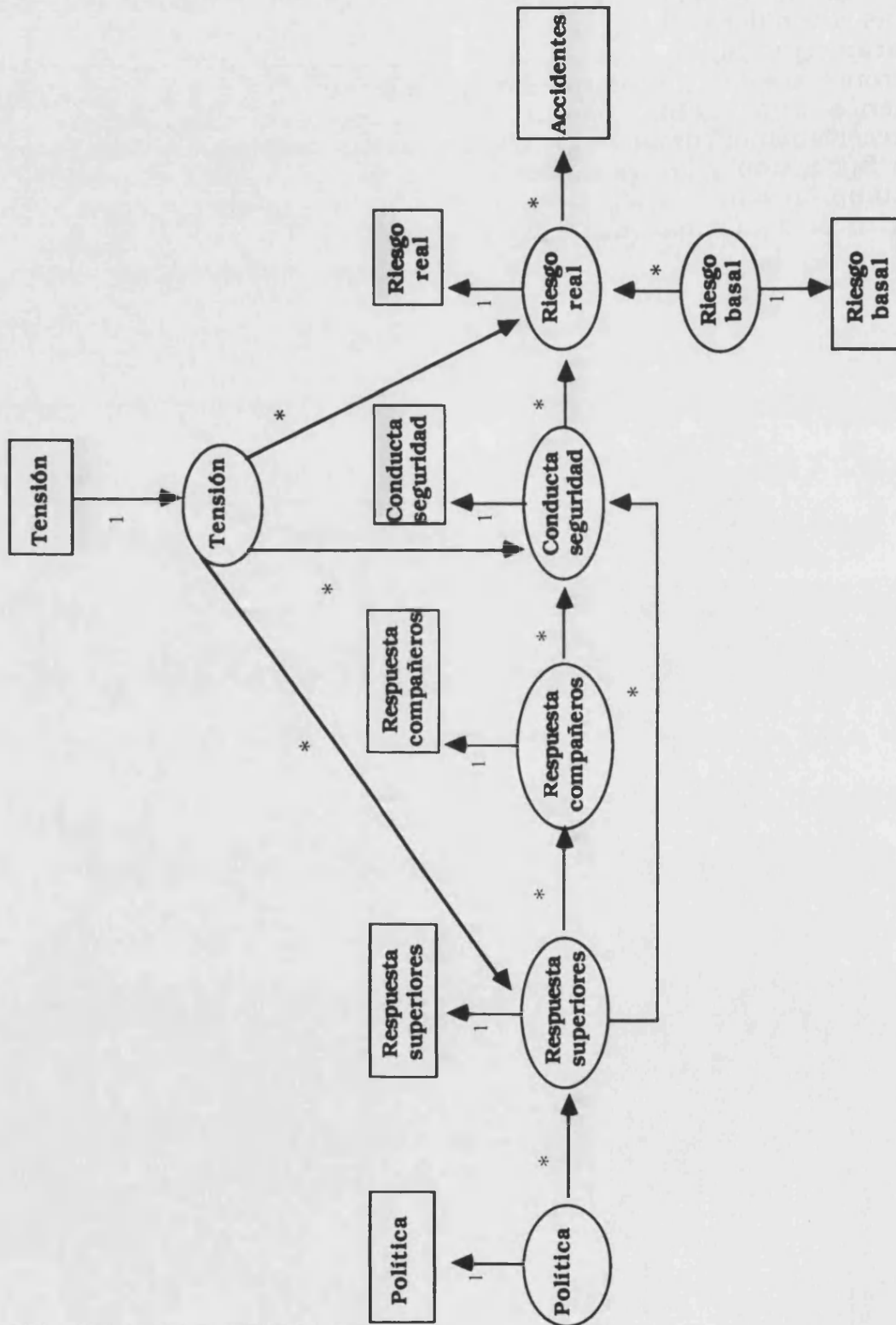
Lo que indica la fórmula es la proporción que de la varianza total de la variable corresponde a varianza de error.

Utilizaremos el coeficiente alfa como medida de la fiabilidad de cada variable, dado que es un límite inferior de la fiabilidad. La

variable accidentabilidad no tienen disponible su fiabilidad, por lo que será considerada como medida sin error.

En la siguiente gráfica se puede consultar el modelo resultante tras considerar los errores de medida. Los asteriscos indican parámetros libres a estimar en el modelo. Se han eliminado los errores, uno para cada variable dependiente (a la que le llegue por lo menos una flecha), porque de otra forma el diagrama se hace excesivamente complejo. La relación de las variables con sus factores se fija a uno y las varianzas de error no las estima el modelo, las estima en base a la fórmula presentada por lo que se consideran a partir de aquí como valores fijados.

Gráfica 7.7. Modelo causal de accidentes considerando los errores de medida de las variables. Los asteriscos indican parámetro a estimar. Se han eliminado los errores de variables y factores por sencillez.



Aplicando la fórmula para el cálculo de las varianzas de error se obtienen los resultados de la tabla 7.7.

Tabla 7.6. Estimación de la varianza de error de las variables observables del modelo causal de accidentes medidas por cuestionario.

Variable	alfa	1-alfa	varianza variable	varianza error
Política	0.641	0.359	2.320	0.832
R. Sup.	0.836	0.164	0.061	0.010
R. Com.	0.846	0.154	0.053	0.008
Conducta	0.730	0.270	216.78	58.53
R. Real	0.845	0.155	0.711	0.110
R. Basal	0.890	0.110	0.082	0.009
Tensión	0.820	0.180	6.718	1.209

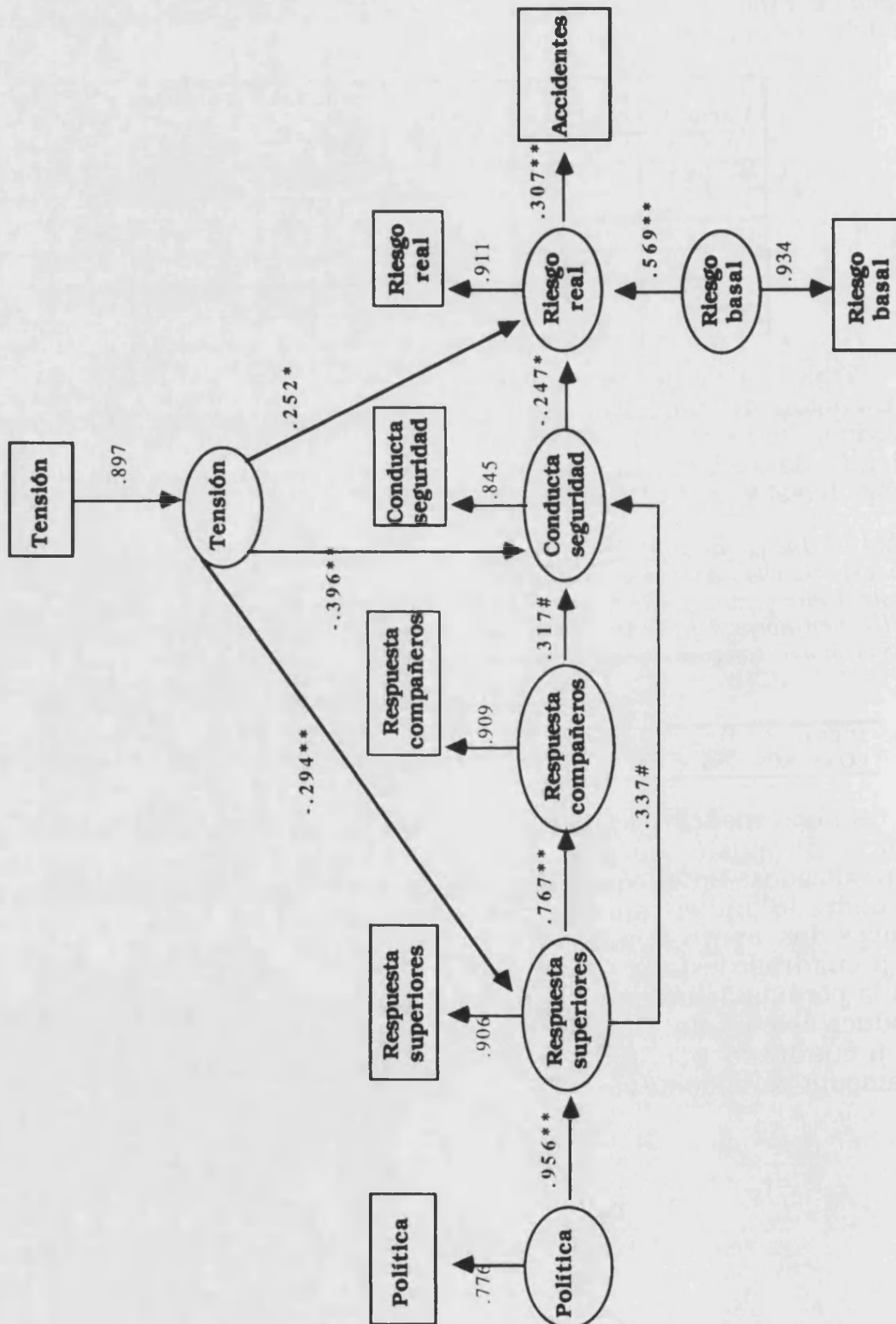
Tras el cálculo de las varianzas de error se ha procedido a la estimación del modelo por mínimos cuadrados generalizados y por máxima verosimilitud. En máxima verosimilitud se ha planteado además la covarianza que fue encontrada en otros modelos entre riesgo basal y política, ambas variables exógenas.

Tabla 7.7. Índices de ajuste global para el modelo considerando los errores. Los índices son: chi=chi-cuadrado; g=grados de libertad del modelo; p=probabilidad; F=función de ajuste a minimizar; LHR=razón de verosimilitud escalada; NFI=normed fit index; NNFI=non-normed fit index; CFI=comparative fit index. Métodos de estimación: GLS=mínimos cuadrados generalizados; ML=máxima verosimilitud.

		chi	g	p	F	LHR	chi/g	NFI	NNFI	CFI
MODELO	GLS	28.73	18	.051	.261	.877	1.59	.989	.994	.996
CON ERROR	ML	28.60	17	.038	.260	.878	1.68	.097	.931	.958

Ambos modelos ajustan a los datos de modo general. Todos los índices de ajuste son adecuados en el caso de mínimos cuadrados generalizados. En el caso de máxima verosimilitud, sólo la prueba de ji-cuadrado indica ajuste inadecuado, mientras que el resto de índices dan apoyo a la estructura hipotetizada. No obstante, el valor de ji-cuadrado está cercano a no ser significativo, y probablemente sea la pérdida de un grado de libertad que supone la covarianza la que produce este efecto. En ambos casos hay que considerar que la razón de ji-cuadrado por los grados de libertad indica una estructura altamente parsimoniosa.

Gráfica 7.8. Modelo causal de accidentes considerando los errores de medida de las variables. Se han eliminado los errores de variables y factores por sencillez. (\*\*=p<0.01; \*=p<0.05; #=marginamente significativo; ns= no significativo). En negrita los coeficientes estructurales de relación entre los factores.

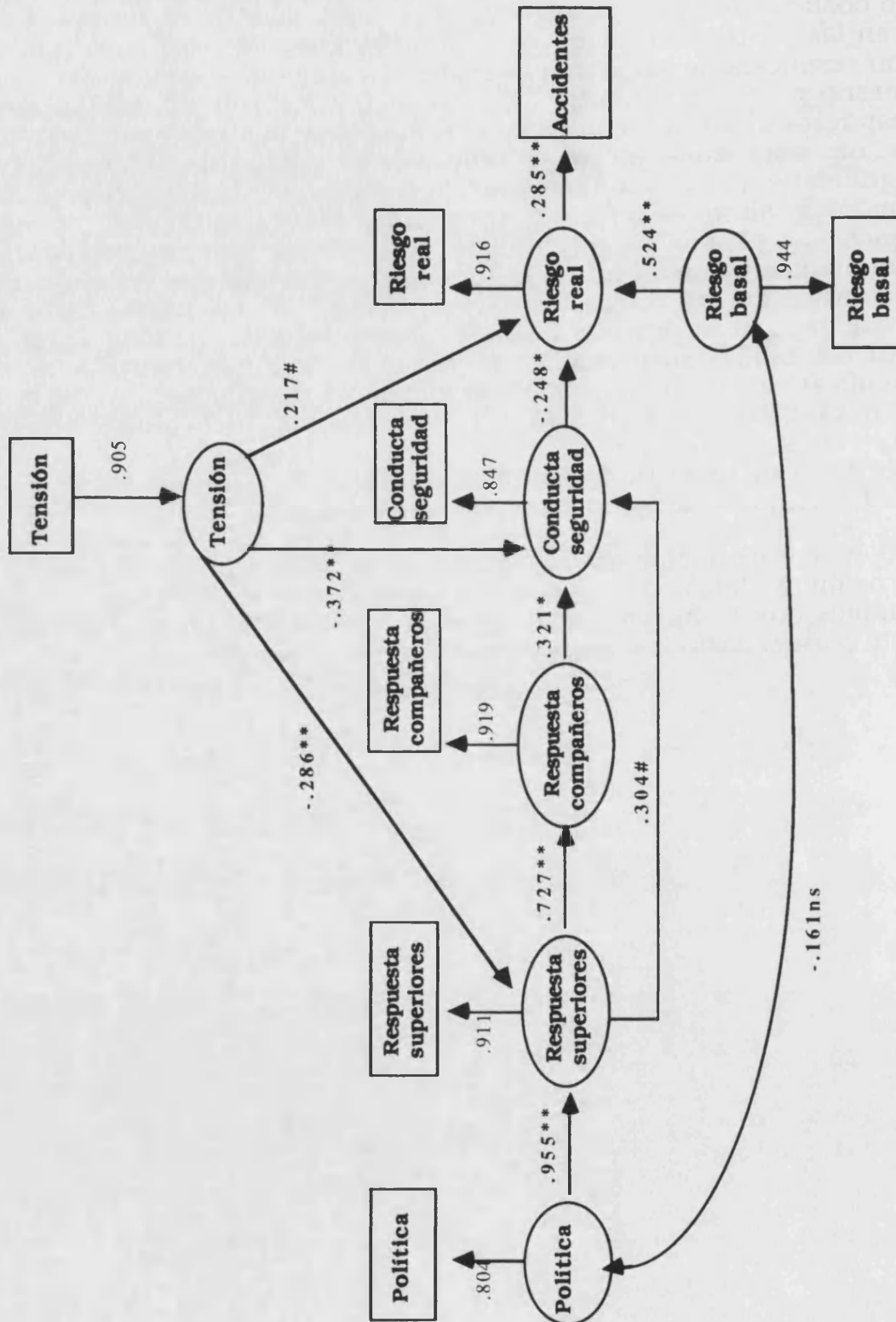


En la estimación por mínimos cuadrados generalizados, cuyos valores pueden consultarse en la gráfica 7.8, se mantiene la lógica de las relaciones ya encontrada para los modelos sin considerar el error. No obstante se han producido cambios en la cuantía de las relaciones y en las pruebas de significatividad asociadas. La relación de política con respuesta de los superiores: ahora es más alta, aunque al ser ya anteriormente altamente significativa no cambia. La relación de respuesta de los superiores con respuesta de los compañeros: ahora es un poco más elevada, tampoco cambia apreciablemente la significatividad. La relación de respuesta de los superiores con conducta: ahora es más alta, pero al cambiar la estimación del error típico, la prueba de significatividad señala que la relación es marginalmente significativa. Esto mismo ocurre para la relación entre respuesta de los compañeros y conducta. Por su parte, todos los predictores de riesgo real han aumentado también de valor, pero no han cambiado las probabilidades evaluadas desde las pruebas de significatividad. Las relaciones de tensión han cambiado de diversas maneras: ha quedado igual la relación con respuesta de los superiores, es más alta la relación con riesgo real y conducta, pero sólo la evaluación de la significatividad de ésta última relación ha cambiado, mostrándose ahora más significativa.

A continuación se presentarán las estimaciones máximo verosímiles del modelo que considera los errores, gráfica 7.9. Los cambios que se ha producido son muy similares a los que ya hemos visto para el caso de mínimos cuadrados generalizados.



Gráfica 7.9. Modelo causal de accidentes considerando los errores de medida estimado mediante máxima verosimilitud. Errores de variables y factores eliminados por sencillez. (\*\*= $p < 0.01$ ; \*= $p < 0.05$ ; #=*marginalmente significativo*; ns= *no significativo*). En **negrita** los coeficientes entre los factores.



Evaluando los cambios que se producen entre el modelo que consideran el error y el que no lo considera puede verse que los cambios producidos son relevantes. La relación de política con respuesta de los superiores, así como el de respuesta de los superiores con respuesta de los compañeros se han visto incrementados. Las relaciones de los diferentes predictores de conducta con ésta se han visto también incrementados, pero no así las probabilidades asociadas, que han bajado hasta la marginalidad en el caso de respuesta de los superiores y al 0.05 en el caso de respuesta de los compañeros. Más altos son también los coeficientes estructurales de riesgo real con sus diferentes predictores, pero ha bajado la significatividad de tensión como predictor. Más elevada son las relaciones de tensión con respuesta de los superiores y conducta. Por último, también ha aumentado en cuantía y en significatividad estadística la relación de riesgo real con la variable de accidentabilidad.

## **5. Conclusiones**

Los objetivos presentados en el capítulo 2 deben ahora evaluarse a la luz de los resultados obtenidos. Allí dividimos los objetivos en generales y específicos, manteniéndose esta división aquí.

### *Objetivos generales*

a) Presentación de cuestionarios para la medición de variables implicadas en la ocurrencia de accidentes. Este objetivo se ha cumplido. Se han presentado escalas para evaluar algunas de las variables psicosociales más empleadas en la literatura para la predicción de los accidentes. Así, han sido validados cuestionarios para la medición de las variables:

- Clima de seguridad, dividiendo el constructo en sus tres factores teóricos principales -estructura, formación-información y política- para estudiar diferencialmente sus efectos en la cadena causal de accidentes.

- Riesgos laborales. Se ha diferenciado entre dos aspectos de los riesgos, el riesgo basal y el riesgo real. Esta distinción, que no ha aparecido anteriormente en la literatura, parece prometedora a la luz de los resultados.

- Variables de contexto social. Las respuestas de seguridad dos importantes puntos de referencia para la conducta de seguridad del trabajador, como son los supervisores y los propios compañeros, han sido sometidos a medición y sus efectos diferenciales podrán también ser estudiados separadamente mediante modelos causales.

- Se han presentado dos medidas de respuesta de propio trabajador, su conducta segura o insegura y la tensión que puede derivarse de estar sometido a riesgos de accidente.

- Se han estudiado en cierta profundidad diversos indicadores de accidentabilidad, tanto simples como complejos, siguiendo sugerencias que aparecen en la literatura de medición de los accidentes (Hansen, 1989).

Todos estos constructos se han diferenciado teóricamente, han sido sometidos a análisis separados, y no han sido tratados de forma global, como a menudo ocurre en la literatura. Dentro de los contenidos rastreados por los items se ha intentado mantener un compromiso entre contenidos específicos y cuestiones generales, de forma que las escalas puedan ser útiles tanto en diagnóstico como en

investigación. Por otro lado, se han aportado definiciones novedosas de constructos, como los de clima, mucho más restringidos ahora, o la diferenciación entre los dos tipos de riesgo, basal y real.

b) Construir modelos causales de la ocurrencia de accidentes laborales. Este objetivo también ha sido cumplido. A pesar de las dificultades metodológicas, se ha conseguido establecer un modelo que concatena causas y efectos de diversos niveles organizacionales hasta la ocurrencia de accidentes.

La línea principal de causación de accidentes queda relativamente perfilada a través de los diversos modelos causales probados:

En primer lugar, los factores organizacionales, principalmente la política organizacional se encuentran muy relacionados y son capaces de predecir una parte importante de la respuesta de los supervisores. Esta relación estructural entre política y respuesta de los superiores obedece a la idea de que los supervisores son principalmente transmisores de las reglas y órdenes y formas de trabajar que plantea la dirección, de forma que son los que directamente las aplican al contexto directo del trabajador. Por tanto, encontrar una relación positiva era previsible desde la teoría.

No se esperaba, sin embargo, que la relación de los otros constructos de clima, la estructura de seguridad y especialmente la formación-información. Estos dos constructos se hipotetizaba que podrían tener efectos diferentes a diversos niveles. Por ejemplo, la estructura de seguridad puede tener una misión de cierto control de la respuesta de los supervisores, pero no parece que esta hipótesis se refrende en los datos. Por su parte, la variable de formación siempre se ha considerado en la literatura como especialmente relevante. Se esperaba que pudiera afectar la conducta de los trabajadores (el trabajador que contesta los cuestionarios y los compañeros), e incluso que pudiera tener cierto efecto sobre los supervisores. No obstante, todas estas hipótesis no se han mantenido. La variable formación se encuentra covariada con la política de la empresa, como era de esperar, pero no tiene ningún efecto significativo sobre otras variables.

La variable política afecta, por su parte de forma indirecta el resto de la cadena causal, a través de la respuesta de los supervisores.

La respuesta de los superiores se hipotetizaba como afectando a la respuesta de los trabajadores y los compañeros. Efectivamente, es el supervisor el primer responsable del desempeño de los trabajadores en seguridad y en cualquier otro aspecto. Hay un efecto directo claro, a través de los distintos métodos de estimación, sobre la respuesta de los compañeros. Por su parte, también hay un efecto

directo sobre la conducta de seguridad del trabajador, y, además, un efecto indirecto, que se produce a través de la respuesta de los compañeros. Estos resultados se mantienen también a través de métodos de estimación y de procedimientos

Por su parte, la conducta de seguridad y el riesgo basal son los principales predictores del riesgo real. Este riesgo real, probabilidad de tener accidentes, se vislumbra no solo como el predictor directo de accidentes sino como una variable importante a predecir en sí misma. Efectivamente, la probabilidad de accidentarse puede ser un índice de accidentabilidad, más sencillo de predecir que el número de accidentes y severidad de los mismos que es un suceso altamente azaroso.

Por último, puede hablarse de una capacidad de predicción significativa de los accidentes. Aunque la predictibilidad es más bien modesta en la muestra que hemos estudiado (un máximo del 9%), sí hay efectos directos e indirectos que son apreciables. Además, encontrar elevados porcentajes de explicación en fenómenos tan complejos como éste y con una muestra tan variada en su composición es difícil.

Además, los modelos se han probado con una variedad de métodos de estimación e incluso de procedimientos causales, y los resultados son coincidentes salvo ligeros cambios de cuantía en algunos parámetros.

c) Los modelos causales de predicción de accidentes permiten también evaluar la capacidad predictiva sobre la conducta segura-insegura del trabajador. Principalmente tres variables se han mostrado como predictores significativos de la conducta: respuesta de los superiores y de los compañeros, tal y como se hipotetizaba, y tensión, que en principio no se había considerado, pero parece tener cierto impacto. No obstante, existe la posibilidad de una doble vía causal entre conducta y esta última variable. No obstante, esta última hipótesis no se ha visto apoyada por los datos de esta muestra. Además, otras variables tienen cierto efecto indirecto, principalmente la política de seguridad de la empresa.

### *Objetivos específicos*

a) Se han obtenido datos psicométricos sobre los cuestionarios e indicadores de las variables implicadas. Se han conseguido instrumentos psicométricamente fiables y válidos, teniendo en algunos casos que depurarse para conseguirlo. Además, se han

realizado diversos estudios de validez convergente-discriminante que han mostrado la adecuación empírica de las distinciones teóricas realizadas.

b) Por último, se han cumplido los objetivos de estudio del efecto de la variable tensión. Las ausencias de relación con accidentes muestran que no puede mantener esta variable como consecuente de los accidentes, al menos en este tipo de estudio transversal. Parece más bien, que la tensión afecta básicamente a las conductas seguras-inseguras, y puede afectar a algunas percepciones del ambiente laboral que tienen el trabajador.



## REFERENCIAS METODOLOGIA.

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716-723.
- Anderson, T. W. (1973). Asymptotically efficient estimation of covariance matrices with linear structure. *Annals of Statistics*, 1, 135-141.
- Anderson, T. W. y Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, 49, 155-173.
- Arbous, A. G. y Kerrich, J. E. (1951). Accident statistics and the concept of accident proneness. *Biometrics*, 7, 340-432.
- Bagozzi, R.P. (1981). An examination of the validity of two models of attitude. *Multivariate Behavioral Research*, Vol. 16, pp: 323-359.
- Bearden W.O., Sharma S. y Teel J.R. (1982). Sample size effects on chi-square and other statistics used in evaluating causal models. *Journal of Marketing Research*, 19, 425-530.
- Béland, F. (1993). Comunicación personal.
- † Bentler, P. M. (1976). Multistructural statistical model applied to factor analysis. *Multivariate Behavioral Research*, 11, 3-25.
- \* Bentler, P. M. (1980). Multivariate analysis with latent variables: causal modeling. *Annual Review of Psychology*, 31, 419-456.
- Bentler, P. M. (1983). Some contributions to efficient statistics for structural models: Specification and estimation of moment structures. *Psychometrika*, 48, 493-517.
- † Bentler, P. M. (1988). *EQS. Estructural Equations Program*. BMDP, Statistical Software.
- Bentler, P. M. (1988b). Comparative fit indexes in structural equation models. *Psychological Bulletin*, in press.
- Bentler, P. M. y Bonnet, D. G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin*, 80, 588-606.
- Bentler, P. M. y Dijkstra, T. (1985). Efficient estimation via linearization in structural models. In P. R. Krishnaiah (Ed.), *Multivariate analysis VI* (pp 9-42). Amsterdam, North Holland.
- Bentler, P. M. y Weeks, D. G. (1980). Linear structural equations with latent variables. *Psychometrika*, 45, 289-308.
- Bisquerra, R. (1989). *Introducción conceptual al análisis multivariable*. Barcelona, 2 Vol, PPU.
- Bollen, K. A. (1989). *Structural equations with latent variables*. Wiley series in probability and mathematical statistics. New York: John Wiley & sons.
- Boosma, A. (1983). *On the robustness of LISREL (maximum likelihood estimation) against small sample size and nonnormality*. Ph. D. Thesis, University of Groningen.



- Botella, J.; León, O. G. y San Martín, R. (1992). *Análisis de datos en psicología*. Madrid, Pirámide.
- Boyle, R. P. (1970). Path analysis with ordinal data. *American Journal of Sociology*, 75, 461-480.
- Brooke, P. P., Russell, D. W. y Price, J. L. (1988). Discriminant validation of measures of job satisfaction, job involvement, and organizational commitment. *Journal of Applied Psychology*, 73, 139-145.
- Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, Vol 3, 296-322.
- Brown, R. L. y Holmes, H. (1986). The use of a factor-analytic procedure for assessing the validity of an employee safety climate model. *Accident Analysis and Prevention*, 18, 455-470.
- Browne, M. W. (1974). Generalized least-squares estimators in the analysis of covariance structures. *South African Statistical Journal*, 8, 1-24.
- Browne, M.W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37, 62-83.
- Carmines, E. G. y McIver, J. P. (1981). Analyzing models with unobserved variables: analysis of covariance structures. En G. W. Bornstedt y E. F. Borgatta (Eds), *Social Measurement: current issues*. Beverly Hills: Sage.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Cuadras, C. M. (1981). *Métodos de análisis multivariable*. Barcelona, Eunibar.
- Cuadras, C. M. (1991). *Métodos de análisis multivariable* (Ed. rev.). Barcelona, PPU.
- Cudeck, R. y Browne, M. W. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research*, 9, 443-448.
- Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of Econometrics*, 18, 5-46.
- Dedobbeleer, N. y Béland, F. (1991). A safety climate measure for construction sites. *Journal of Safety Research*, 22, 97-103.
- Dillon, W.R. y Goldstein, M. (1984). *Multivariate analysis. Methods and application*. Nueva York: John Wiley.
- Duncan, O. D. (1966). Path analysis: Sociological examples. *American Journal of Sociology*, 72, 1-16.
- Fernández Ballesteros, R. (1987). *Psicodiagnóstico* (Tomo I). Madrid: UNED.
- Gnanadesikan, R. (1977). *Methods for Statistical Data Analysis of Multivariate Observations*. New York: John Wiley & Sons.
- González-Roma, V. (1990). *Path Analysis en modelos recursivos*. Mimeo, Universitat de València.
- González-Romá, V., Luna, R. y Lloret, S. (1992). Estructura factorial del cuestionario de comportamiento extra e intrarol. *Psicológica*, 13, nº1.

- Guttman, L. (1945) A basis for analyzing test-retest reliability. *Psychometrika*, 10, 255-282.
- Kadane, J. B.; y Anderson, T. W. (1977). A comment of the test of overidentifying restrictions. *Econometrica*, 45, 1027-31.
- Hansen, C. P. (1989). A causal model of the relationship among accidents, biodata, personality, and cognitive factors. *Journal of Applied Psychology*, 74, nº1, 81-90.
- Harlow, L. L. (1985). *Behavior of some elliptical theory estimators with nonnormal data in a covariances structures framework: A Monte Carlo study*. Ph.D. Thesis, University of California, Los Angeles.
- Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46, 1251-71.
- Hocevar, D.; Khattab, A.; y Michael, W. B. (1987). Significance testing and efficiency in LISREL Measurement Models. *Educational and Psychological Measurement*, Vol.47, 45-49.
- Jöreskog, K. G. (1967). Some contributions to maximum likelihood factor analysis. *Psychometrika*, 34, 183-202.
- Jöreskog, K. G. (1969). A general approach to confirmatory factor analysis. *Psychometrika*, 34, 183-202.
- Jöreskog, K. G. (1973). A general method for estimating a linear structural equation system. En A. S. Goldberg y O. D. Duncan (Eds). *Structural Equations Models in the Social Sciences*. New York, Seminar.
- Jöreskog, K. G. y Goldberger, A. S. (1972). Factor analysis by generalized least squares. *Psychometrika*, 37, 243-260.
- Jöreskog, K. G. y Sörbom, D. (1979). *Advances in factor analysis and structural equations models*. Cambridge, Mass.
- Jöreskog, K. G. y Sörbom, D. (1983). *LISREL V: Analysis of linear structural relationships by the method of maximum likelihood*. Chicago. International Educational Services.
- 6 Jöreskog, K. G. y Sörbom, D. (1989). *LISREL 7, A guide to the program and applications*. Chicago, SPSS.
- Kadane, J. B.; y Anderson, T. W. (1977). A comment of the test of overidentifying restrictions. *Econometrica*, 45, 1027-31.
- Krishnaiah, P. R. (1980). *Handbook of statistics*. (6 volúmenes). Amsterdam: North Holland Publishing.
- Kuder, G.F. y Richardson, M.W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2, 151-160.
- Long, J. S. (1983). *Confirmatory Factor Analysis: A preface to LISREL*. Beverly Hills, CA: Sage.
- Long, J. S. (1983b). *Covariance Structure Models. An introduction to LISREL*. Sage University Papers series on Quantitative Applications in the Social Sciences, series no. 07-304. Beverly Hills and London. Sage Publications.



- Lord, F.M. y Novick, M.R. (1968) *Statistical Theories of Mental Tests Scores*, Reading-Mass., Addison-Wesley.
- McDonald, R.P. (1986). Describing the elephant: structure and function in multivariate data. *Psychometrika*, 51 (4), 513-534.
- ¶ Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57, 519-530.
- ¶ Mardia, K. V. (1980). Test of univariate and multivariate normality. En *Handbook of Statistics*, Vol. 1, P. R. Krishnaiah, Ed. North Holland Pub. Co., Amsterdam.
- Marsh, H.W., Balla, J. y McDonald, R.P. (1988). Goodness-of-fit in confirmatory factor analysis: the effect of sample size. *Psychological Bulletin*, 103, 391-410.
- Marsh, H. W. y Hocevar, D. (1985). The application of confirmatory factor analysis to the study of self-concept: First and higher order factor structures and their invariance across age groups. *Psychological Bulletin*, 97, 562-582.
- Meliá, J.L.; Peiró, J.M.; Zornoza, A; Sanz, M.J; Morte, M.P; Zurriaga, R.(1987) Una escala para la evaluación de la tensión de rol: estructura factorial, fiabilidad y validez. El cuestionario de tensión T3/15. *II Congreso de Evaluación Psicológica*, Madrid.
- Meliá, J. L. (1990). *La construcción de la psicometría como ciencia teórica y aplicada*. València, Gráficas Alhorí.
- Meliá, J. L. e Islas, M. E. (1991). *Batería de Seguridad Laboral V3*. València.
- Meliá, J.L.; Sesé, A.; Tomás, J.M.; Oliver, A. (1992). Propiedades psicométricas y análisis factorial confirmatorio del cuestionario de respuesta hacia la seguridad de supervisores y mandos intermedios (RS3/7). *Psicológica*, Vol. 13, nº3.
- Meliá, J.L.; Tomás, J.M.; Oliver, A. e Islas, M.E. (1992). Organisational and psychological variables as antecedents of work safety: A causal model. *Conference on Safety & Well-being at Work: A Human Factors Approach*, Loughborough, United Kingdom.
- Meliá, J.L.; Tomás, J.M.; Oliver, A. (1992). Concepciones del Clima organizacional hacia la Seguridad Laboral: Replicación del modelo confirmatorio de Dedobeleer y Béland. Sugerencias sobre el papel del Análisis Factorial Confirmatorio en el análisis de cuestionarios. *Revista de Psicología del Trabajo y de las Organizaciones*, vol 9, nº 22.
- Meliá, J.L.; Oliver, A. y Tomás, J.M. (1993). An organisational and psychological causal model of occupational accidents. *Sixth European Congress on the Psychology of Work and Organization*, Alicante.
- Mulaik, S.A.; James, L.R.; Van Alstine, J.; Bennett, N.; Lind, S. and Stilwell, C.O. (1989). Evaluation of Goodness-of-fit Indexes for Structural Equations Models. *Psychological Bulletin*, 105. 3, 430-445.
- Muñiz, J. (1992). *Teoría Clásica de los Tests*. Madrid, Pirámide.
- Oliver, A.; Tomás, J.M. y Meliá, J.L. (1993). Una nueva medida del clima organizacional de seguridad: la escala C3/18. En revisión en *Investigaciones Psicológicas*.
- Petersen, D. (1984). Analyzing the behavioral system. En *Analyzing Safety Performance*. Garland Publishing, Inc. New York, 81-97.

- Rindskopf, D. y Rose T. (1988) Some Theory and Applications of Confirmatory Second-Order Factor Analysis. *Multivariate Behavioral Research*, 23, 51-67.
- Rulon, P. J. (1939) A simplified procedure for determining the reliability of a test by split-halves. *Harvard Educational Review*, 9, 99-103.
- Santisteban, C. (1990). *Psicometría. Teoría y práctica en la construcción de tests*. Madrid: Norma.
- Satorra, A. (1989). Alternative test criteria in covariance structure analysis: A unified approach. *Psychometrika*, 54, 131-151.
- Satorra, A. y Bentler, P. M. (1986). Some robustness properties of goodness of fit statistics in covariance structure analysis. *American Statistical Association: Proceedings of the Business & Economic Statistics Section*, 594-554.
- Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.
- Siegel, S. (1983). *Estadística no paramétrica*. México, Trillas.
- Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, Vol 3, 271-295.
- Tanaka, J. S. (1984). *Some results on the estimation of covariance structure models*. Ph. D. Thesis, University of California, Los Angeles.
- Tanaka, J. S. y Bentler, P. M. (1985). Quasi-likelihood estimation in asymptotically efficient covariance structure models. *American Statistical Association: 1984 Proceedings of the Social Statistics Section* (658-662).
- Tatsuoka, M. M. (1971). *Multivariate analysis: Techniques for educational and psychological research*. New York: John Wiley.
- Tucker, L. R. y Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1-10.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading-Mass: Addison-Wesley.
- Visauta, B. (1986). *Técnicas de investigación social. Modelos Causales*. Ed. Hispano-Europea, Barcelona.
- Welkowitz, J.; Ewen, R. B.; y Cohen, J. (1981). *Estadística aplicada a las Ciencias de la Educación*. Madrid, Santillana Aula XXI.
- Werts, C. E.; Linn, R. L. y Jöreskog, K. G. (1974). Interclass reliability estimates: testing structural assumptions. *Educational and Psychological Measurement*, Vol. 34, pp: 25-33.
- Werts, C. E.; Rock, D. A. y Grandy, J. (1979). Confirmatory factor analysis applications: Missing data problems and comparisons of path models between populations. *Multivariate Behavioral Research*, 14, 199-213.
- Wheaton, B.; Muthén, B.; Alwin, D. F. y Summers, G. F. (1977). Assessing reliability and stability in panel models. En D. R. Heise (Ed), *Sociological Methodology*. San Francisco, Jossey Bass.
- Wiley, D. E. (1973). The identification problem for structural equations models with unmeasured variables. En A. S. Goldberg y O. D. Duncan (Eds). *Structural Equations Models in the Social Sciences*. New York, Seminar.

Wilkinson, L. (1986). *SYSTAT, the system for statistics*. Evanston, Inc.

Winer, B.J. (1971) *Statistical Principles in Experimental Design*. Nueva York, McGraw-Hill.

Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, 5, 161-215.

## REFERENCIAS SEGURIDAD LABORAL.

- Adams, E. (1976). Accident causation and the management system. *Professional safety*, october, pp. 26-29.
- Andersson, R. , Johansson, B. y Linden, K. (1978). Development of a model for research on occupational accidents. *Journal of Occupational Accidents*, 1, pp. 341-352.
- Argyris, C. (1957). *Personality and organizations*. Harper & Row, New York.
- Arvey, R. y Jones, A. (1985). The use of discipline in organizational settings: A framework for future research. *Research in organizational Behavior*, (7), 367-408.
- Benner, L. (1975). Accidents investigations, Multilinear events sequencing methods. *Journal of Safety Research*, 7, pp. 67-73.
- Bürgy, A. (1976). *Die Analyse von Berufen und Berufsansforderungen*. Stuttgart: Kohlhammer.
- Caillard, J.F. (1976). *Cahiers des Comites de prevention des batiments et des travaux publics*, 6, 280-282.
- Cohen, A. (1977). Factors in succesfull occupational safety programs. *Journal of Safety Research*, 9, pp. 168-178.
- Cohen, H., Templer, J. y Archea, J. (1985). An analysis of occupational stair accident patterns. *Journal of Safety Research*, (16),171-181.
- Cox, S. y Cox, T. (1991). The structure of employee attitudes to safety: a european example. *Work and Stress*, 5, 93-107.
- Chhokar, J. S. (1990). Behavioural Safety Management. *Vikalpa, the Journal for Decision Making*, Vol. 15, n°1.
- Dejoy, D. (1985). Attributional processes and hazard control management in industry. *Journal of Safety Research*, 16, pp. 61-71.
- Dejoy, D. (1990). Toward a comprehensive human factors model of workplace accident causation. *Professional Safety*.
- Denton, K. (1986). *Seguridad Industrial*. México, McGraw-Hill.
- Duhon, D.; Knouse, S.; Robert, R. & Walling, J. (1989). Improvent oilfield safety performance by behavior modification techniques. *Academy of Management 1989 Annual Meeting*, Washington, D.C.
- Dunne, J. G. (1971). *Safety Psychology: A review of the literature*. A. P. Report 35, Applied Psychology Department, University of Aston.
- Dwyer, T. y Raftery, A. E. (1991). Industrial accidents are produced by social relations of work: A sociological theory of industrial accidents. *Applied Ergonomics*, 22, 167-179.
- Evan,W. M. (1976). *Organizations Theory Structures, Systemsand Environments*. New York, Wiley.
- Ferguson, J. L.; McNally, M. S. and Both, R. F. (1984). Individual characteristics as predictors of accidental injuries in naval personnel. *Accidents Analysts & Prevention*, 16, 47-54.

- Fernández, I. Nogareda, C. y Oncins, M. (1988). Encuesta Nacional de Condiciones de Trabajo 1987. *Salud y Trabajo*, 70, 26-36.
- Forehand, G.A. y Gilmer, B.H. (1964). Environmental variation in studies of organizational behavior. *Psychological Bulletin*, 62, 361-382.
- Fox, C. J. y Sulzer-Azaroff, B. (1987). Increasing completion of accident reports. *Journal of Safety Research*, (18), 65-71.
- Frieling, E. y Hoyos, C. G. (1978). *Fragebogen zur Arbeitsanalyse. (FAA)* (Bern: Huber).
- Frieling, E.; Kannheiser, W.; Facaoaru, C.; Wöcherl, H. y Dürholt, E. (1984). Entwicklung eines theorie-geleiteten, standardisierten, verhaltenswissenschaft-lichen Verfahrens zur Tätigkeitsanalyse. *Forschungsbericht 01 HA 029-2A-TAP 0015*. (Universität München)
- Geller, E.S., Rudd, J. R., Kalsher, M. J., Streff, F. M. y Lehman, G. R. (1987). Employer-based programs to motivate safety belt use: A review of short-term and long-term effects. *Journal of Safety Research*, (18), 1-17.
- Glendon, I. (1991). Influencing behaviour: a framework for action. *Journal of Health & Safety*, 6, 23-38.
- Goldberg, A.I.; Dar-El, EM; & Rubin, A. E. (1991). Threat perception and the readiness to participate in safety programs. *Journal of Organizational Behavior*, Vol. 12, 109-122.
- Goodman, P. S. y Garber, S. (1988). Absenteeism and Accidents in a Dangerous Environment: Empirical Analysis of Underground Coal Mines. *Journal of Applied Psychology*, vol. 73, nº 1, pp. 1-86.
- Haas, J. (1977). Learning real feelings a study of high steel ironworkers' reactions to fear and danger. *Sociology of Work and Occupation*, 4, (2).
- Hale, A. R. y Glendon, A. J. (1987). *Individual Behavior in the control of danger*. (Amsterdam: Elsevier).
- Hale, A.R. y Hale, M. (1972). *A review of the industrial accident literature*. London: Her Majesty's Stationery Office.
- Halpin, A. W. y Croft, D. B. (1963). *The organizational climate of schools*. Chicago. The University of Chicago. Mydwerl Administrative Centers.
- Hammer, W. (1972). *Handbook of system and product safety*. Prentice Hall, Englewood Cliffs, New Jersey.
- Hansen, C. P. (1989). A causal model of the relationship among accidents, biodata, personality, and cognitive factors. *Journal of Applied Psychology*, 74, 81-90.
- Harris, G. (1992). The cost of accidents. *Conference on Safety & Well-being at work: A human factors approach*. Loughborough, U.K.
- Heinrich, H.W. (1931). *Industrial accident prevention*. McGraw Hill, New York.
- Heinrich, H.W.; Petersen, D. y and Roos, N. (1980). Principles of accident prevention. In: *Industril accident prevention*. New York, McGraw-Hill.
- House, R.J. y Rizzo, J.R. (1972). Toward the measure of organizational practices: scale development and validation. *Journal of Applied Psychology*, 56, 388-396.
- Hoyos, K. y Zimolong, B. (1988). *Occupational Safety and Accident Prevention*. Elsevier.

- Islas, M. E. y Meliá, J. L. (1991). Accidentes de trabajo: Intervención y propuestas teóricas. *Revista Latinoamericana de Psicología*. Vol. 23, nº3, pág. 323-348.
- Islas, M.E.; Vera, A. y Ocaño, C. (1986). Un estudio exploratorio de algunos aspectos disposicionales que afectan la seguridad industrial. *La psicología social en México*. México: Trillas.
- James, L. R. y Jones, A. P. (1974). Organizational Climate: A review of theory and research. *Psychological Bulletin*, 1974, 81, 1096-1112.
- Keefer, W. D. (1953). *Cost Accident*. Serie: Industrial Safety en R. P. Blake Ed.: New York.
- Kjellén, U. y Larsson, J. J. (1981). Investigating accidents and reducing risks: A dynamic approach. *Journal of Occupational Accidents*, 3, 129-140.
- Krause, T. R. (1984). Behavioral science applied to accident prevention. *Professional Safety Journal*, 29, 21-27.
- Landau, K.; Luczak, H. y Rohmert, W. (1975). *Arbeits-wissenschaftlichen Erhebunsbogen zur Tätigkeitsanalyse*. In Rohmert and Rutenfranz (Eds.) Bonn: ohne Verlag.
- Landy, F.J. y Trumbo, D.A. (1980). *Psychology of Work Behavior*. Hornewood: Dorsey Press.
- Lawrence, A. C. (1974). Human error as a cause of accidents in gold minning. *Journal of safety Research*, 6, 78-88.
- Leather, P. J. (1987). Safety and accidents in the construction industry: a work design perspective. *Work & Stress*, 1, 167-174.
- Leather, P. J. (1988). Attitudes towards safety performance on construction work: an investigation of public and private sector differences. *Work & Stress*, 2, 155-167.
- Leather, P. J. (1992). Rethinking the nature of attitudes to workplace safety. *Conference on Safety & Well-being at work: A human factors approach*. Loughborough, U.K.
- Lee, T.R. (1981). The public's perception of risk and the question of the irrationality. *Proceedings of the Royal Society, London*, 376, pp. 5-16.
- Lee, T.R. (1983). The perception of risks. In a Study Group Report on Risk Assesment (*The Royal Society, London*.)
- Leigh, J. P. (1986). Individual and job characteristics as predictors of industrial accidents. *Accident Analysis & Prevention*, 18, 209-216.
- Leveson, H.; Hirschfeld, M. L.; and Hirschfeld, A. H. (1980). Industrial accidents and recent life events. *Journal of Occupational Medicine*, 22, 53-57.
- Litwin, G.H y Stringer, R.A. Jr (1968). *Motivation and organizational climate*. Boston: Division of Research, Harvard Business School.
- McAfee, R. B. y Winn, A.R. (1989). The use of incentives feedback to enhance work place safety: A critique of the literature. *Journal of Safety Research*, (20), 7-19.
- McDonald, G.L. (1972). *The involvement of tractor design in accidents*. Research Report 3/72. Dep. of Mechanical Engineering, University of Queensland.



- Melamed, S.; Luz, J.; Najenson, T; Jucha, E.; Green, M. (1989). Ergonomic stress levels, personal characteristics, accident occurrence and sickness absence among factory workers. *Ergonomics*, Vol. 32, no. 9, 1101-1110.
- Meliá, J.L., Ricarte, J.J. y Arnedo, M.T. (1993). La seguridad en el trabajo desde la Psicología: Una revisión de las principales perspectivas y modelos teóricos. En revisión en *Revista de Psicología General y Aplicada*.
- Monteau, M. (1977). *Praktische Methode zur Untersuchung von Unfallfaktoren, Grundsätze und Anwendungen im Versuch*. Luxemburg: Kommission der Europäischen Gemeinschaften.
- Nohl, J. y Thiemecke, H. (1988). *Systematik zur Durchführung von Gefährdungsanalysen*. Fb 536 (Bremerhaven: Wirtschaftsverlag NW).
- Pasmore, W. y Friedlander, F. (1982). An action-research program for increasing employee involvement in problem solving. *Administrative Science Quarterly*, 27, (3), 343-362.
- Peiró, J. M. (1983). Dimensiones psicosociales del comportamiento del individuo en la organización. *Revista de Psicología General y Aplicada*, 38, 625-666.
- Peiró, J. M. (1990). *Psicología de la Organización*. Madrid: UNED.
- Pekkarinen, A. y Anttonem, H. (1989). The comparison of accidents in a foreign construction project with construction in Finland. *Journal of Safety Research*, (20), 187-195.
- Peña, M. (1990). *Dirección de Personal: Organización y Técnicas*. Barcelona: ESADE.
- Perrow, C. (1984). *Normal accidents*. Basic Books: New York.
- Peters, R. H. (1991). Strategies for encouraging self-protective employee behavior. *Journal of Safety Research*, (22), 53-70.
- Petersen, D. (1984). *Human-error reduction and safety management*. Aloray: New York.
- Porter, L.W. y Lawler, E.E. (1965). Properties of organization structure in relation to job attitudes and job behavior. *Psychological Bulletin*, 64, 23-51.
- Porter, L.W.; Lawler, E.E. y Hackman, J.R. (1975). *Behavior in organizations*. New York: McGraw Hill.
- Porter, C. S. y Corlett, E.N. (1989). Performance differences of individuals classified by questionnaire as accident prone or non-accident prone. *Ergonomics*, 32, 317-333.
- Ramsey, J.A. (1987). *Ergonomic Support of consumer product safety*. American Industrial Hygiene Association Conference.
- Raven, B.H. (1988). *French & Raven 30 years later: Power, Interaction and Interpersonal Influence*. Meetings of the XXIV International Congress of Psychology, Sidney, Australia.
- Raymond, V. (1952). Cause des accidents du travail: le geste nefaste. *Archives des maladies professionnelles*, 13, (5), 449-453.
- Rigby, L.V. (1970). *The nature of human error*. Annual Technical Conference Transactions of the A.S.Q.C. pp.457-466. Milwaukee.

- Rosentock, I.M. (1984). The health believe model and preventive health behavior. *Health Education Monographs*, 2, 354-366.
- Ruppert, F. y Hoyos, C. (1989). Safety diagnosis in industrial plants: concepts and preliminary results. En *Recent Developments in job analysis*. Landau y Rohmert (Eds).
- Saarela, K. L. (1989). A poster campaign for improving safety on shipyard scaffolds. *Journal of safety Research*, (20) , 177-185.
- Saarela, K. L. (1990). An intervention program utilizing small groups: A comparative study. *Journal of Safety Research*, (21), 149-156.
- Salminen, S. (1993). The specific accident factor of older employees. *Accident Analysis and Prevention*. (25), nº 1, pp. 99-102
- Salminen, S. y Saari, J. (1992). Measures to improve productivity and safety as evaluated by those involved in serious occupational accidents. *Conference on Safety & Well-being at work: A human factors approach*. Loughborough, U.K.
- Schneider, B. (1975). Organizational climates: An essay. *Personnel Psychology*, 28, 447-479.
- Schneider, B. y Bartlett, C.J. (1970). Individual differences and Organizational climate. Measurement of Organizational climate by the multi-trait multi-rater matrix. *Personnel Psychology*, 23, 493-512.
- Schneider, B. (1973). The perception of organization: The customers' view. *Journal of Applied Psychology*, 57, 248-256.
- Schafer, S. R. & Kotrlik, J. W. (1986). Factors affecting farm safety practices. *Journal of Safety Research*, 17, 123-127.
- Sells, S.B. (1968). An approach to the nature of organizational climate. En R. Tagiuri y G. Littwin, *Organizational Climate*. Boston: Division of research, Harvard Bussiness School Press.
- Skiba, R. (1974). *Stand und Tendenzen des Arbeitsschutzes in der industrie*. Verkehrs und Wirtschaftsverlag Dr. Borgmann, Dortmund, pp. 5-18.
- Smillie, R.J. y Ayoub, M.A. (1976). Accident causation theories: A simulation approach. *Journal of Occupational Accidents*, 1, 47-68.
- Surry, J. (1969). A human engineering appraisal. *Industrial Accident Research*. Universidad de Toronto.
- Tagiuri, R. (1968). The concepts of organizations. En R. Tagiuri y G. Littwin, *Organizational Climate*, Boston: Division of research, Harvard Bussiness School Press.
- Taylor, C.W. (1972). *Climate for creativity*. New York: Pergamon Press.
- Tuttle, T.C.; Wood, G.D.; Grether, C.B. y Reed, D.E. (1974). *Psychological Behavioral strategies for accident control: a system for diagnosis and intervention*. Final Report for the National Institute for Occupational Safety and Health, Columbia: Behavioral Safety Center.
- Vojtecky, M.A. y Schmitz, M.F. (1986). Program evaluation and health and safety training. *Journal of Safety Research*, 17, 57-63.

Walters, V. y Haines, T. (1988). Workers use and knowledge of the internal responsibility system: Limits to participation in occupational- health and safety. *Canadian Public Policy-Analyse de Politiques*, 14, 411-423.

Weaver, D. (1971). Symptoms of operational error. *Professional Safety*, october.

Webb, G. R., Redman, S., Wilkinson, C. y Sanson-Fisher, R. W. (1989). Filtering effects in reporting work injuries. *Accident Analysis and Prevention*, (21), 115-123.

Winsemius, W. (1960). Paper to the international Conference on the Human Factor in Industrial Accidents. *Soziale Welt*, 9, (1).

Zohar, D. (1980). Safety climate in industrial organizations: theoretical and applied implications. *Journal of Applied Psychology*, Vol. 65, nº. 1, 96-102.

