

ACTITUDES SOBRE DERECHOS FUNDAMENTALES PROCESAL-PENALES: UNA DIMENSIÓN INEXPLORADA DE LA PUNITIVIDAD

Un análisis de clases latentes

Alfonso Serrano Maíllo

Profesor Titular de Derecho Penal y Criminología. UNED

SERRANO MAÍLLO, Alfonso. Actitudes sobre derechos fundamentales procesal-penales: una dimensión inexplorada de la punitividad. Un análisis de clases latentes. *Revista Electrónica de Ciencia Penal y Criminología* (en línea). 2011, núm. 13-05, p. 05:1-05:38. Disponible en internet: <http://criminet.ugr.es/recpc/13/recpc13-05.pdf> ISSN 1695-0194 [RECPC 13-05 (2011), 17 ago]

RESUMEN: La punitividad se ha convertido en uno de los principales objetos de estudio de la Criminología. Sin embargo, el concepto, naturaleza y medición de la «punitividad» son cuestiones muy complejas. Una prueba de ello son las dudas que encierra la dimensionalidad de la punitividad. En particular, puede argumentarse que elementos relativos al proceso penal forman parte de la misma, por ejemplo de un país. Utilizando datos de una encuesta nacional del Centro de Investigaciones Sociológicas que incluye algunas preguntas sobre la materia, en el presente trabajo se trata de testar la hipótesis de que, a nivel individual, *existe una variable latente unidimensional de punitividad con indicadores categóricos de actitudes ciudadanas sobre derechos fundamentales-procesales*. Se discute si esta variable es categórica o continua. Por

limitaciones de los datos a nuestra disposición, sin embargo, no es posible analizar la relación de este constructo con otros relacionados con el castigo. A tal fin, se recurre aquí a una herramienta estadística poderosa y cada vez más importante en ciencias humanas y sociales, sobre todo en el caso de indicadores categóricos u ordinales, como son los análisis de clases latentes. Esta estrategia ofrece la oportunidad de testar la hipótesis de unidimensionalidad, a la vez que también permite una aproximación al error de medición. Entre sus dificultades se encuentran sus relativamente fuertes asunciones, a cuya comprobación debe asegurarse. En el presente trabajo también se testa la existencia de una clase latente de «ideólogos» y un modelo de rasgo latente según el cual nuestra variable sería en realidad continua y se distribuiría de modo normal. Finalmente, se valora la influencia de los datos perdidos en nuestros modelos.

PALABRAS CLAVE: Punitividad, actitudes sobre derechos fundamentales procesal-penales, análisis de clases latentes, error de medición.

Fecha de publicación: 17 agosto 2011

SUMARIO: 1. INTRODUCCIÓN. 2. EL PRESENTE ESTUDIO. 2.1. Hipótesis. 2.2. Datos. 2.3. Estrategia analítica. 3. RESULTADOS. 3.1. Análisis ordinarios. 3.2. Una aproximación al error de encuesta mediante análisis de clases latentes (M₁). 3.3. ¿Existen «ideólogos» en nuestra muestra? 3.4. Comprobación de las asunciones del modelo de clases latentes (M₁). 3.5. Análisis de rasgo latente. 3.6. Análisis con datos perdidos. 4. DISCUSIÓN Y CONCLUSIONES. BIBLIOGRAFÍA.

Para Paul Biemer, con mi admiración y agradecimiento.

1. INTRODUCCIÓN

Los importantes cambios que se vienen observando desde hace algunas pocas décadas en las legislaciones jurídico-penales y sancionadoras en general, en las tasas de personas privadas de libertad, en algunos derechos fundamentales y garantías individuales, en los Sistemas de Administración de Justicia y su actuación, etc., y que se han agrupado bajo el término *punitividad*, se han convertido en uno de los objetos de estudio más importantes de la Criminología contemporánea¹. La punitividad, sin embargo, es un concepto muy complejo². Así, en particular su definición, naturaleza y medición siguen despertando serias dudas en la literatura³ y, más aún, muchos de los enfoques utilizados pueden calificarse de superficiales o pobres por lo que a estos elementos se refiere⁴ –también para niveles superiores al individual, que son los que con más aplicación han sido estudiados⁵. Ante este escenario, algunos ilustres criminólogos como nada más y nada menos que Helmut Kury han llegado incluso a mostrar un sano escepticismo sobre las posibilidades de encontrar en la punitividad un objeto susceptible de estudio científico⁶.

Algunas de las dudas tienen que ver con el contenido de la punitividad, y más concretamente con su grado de *homogeneidad* o su *dimensionalidad*. Para empezar, es menester reconocer que su análisis abarca varios niveles –verbigracia estatal, agregado, individual, etc.– y que en particular no está claro que se encuentren significativamente relacionados entre sí⁷. Ello sugiere ya una cierta heterogeneidad. Dudas semejantes se presentan desde un punto de vista interno. Así, una práctica casi habitual como las preguntas únicas y a menudo limitadas al castigo o a alguno

¹ Serrano Maillo, 2006: 245-252; Serrano Maillo y Serrano Gómez, 2009: 293-315; Zugaldía Espinar, 2006: 1347-1382.

² Kury y Obergfell-Fuchs, 2008: 277-302. Incluso su nombre puede ser discutible.

³ Kury et al., 2004: 52-64; Stalans, 2002: 15-29; Viki y Bohner, 2009: 96-115.

⁴ Mayhew y van Kesteren, 2002: 66-67; Tonry y Farrington, 2005: 31.

⁵ Frost, 2006: 117-118.

⁶ Kury y Putkaradze, 2009.

⁷ Kuhn, 1993: 271; Kury et al., 2004: 51-52.

de sus aspectos evocan la cuestión de si de este modo se mide, se capta algo que intuitivamente parece más complejo cualitativamente⁸. Verbigracia, no es raro mencionar única y exclusivamente las tasas de encarcelamiento –para niveles de análisis estatales o sistémicos– como medida de la punitividad, cuando otros elementos –a estos mismos niveles– como la aplicación de la pena de muerte, las tasas de arrestos, imputaciones, procesamientos, condenas, etc., las condiciones de cumplimiento de las penas privativas de libertad, la gravedad de las sanciones administrativas, las tasas de policías de diversos cuerpos, el recurso gubernamental a estados excepcionales como el de alarma... en fin, la lista podría ser mucho más larga, podrían también ser relevantes. Algo semejante puede decirse para el nivel de los individuos –el que aquí nos interesa–, verbigracia si se les pregunta por su opinión para cada uno de los temas recién mencionados. Dicho con otras palabras, aquí estamos en parte preocupados con si la punitividad es un constructo uni- o pluridimensional, lo cual ha sido hasta ahora insuficientemente explorado⁹.

Centrándonos ya en los derechos fundamentales procesal-penales –y las actitudes individuales hacia los mismos–, éstos plantean, en la línea expuesta, la cuestión de su naturaleza y su relación con aproximaciones más habituales a la punitividad, por regla relacionadas con el castigo. En efecto, un sujeto puede tener una opinión más o menos favorable a distintos derechos fundamentales de este tipo, habida cuenta del conflicto que suele subyacer a los mismos entre libertad y seguridad y habida cuenta de que los derechos encuentran siempre límites. Por ejemplo, un individuo –también una legislación, un gobierno, un grupo...– puede ser más o menos favorable al establecimiento de límites en la actuación policial; a fijar unos requisitos exigentes para la admisión de algo como prueba; o bien a instaurar un sistema de recursos más o menos amplio o estrecho. A mi modo de ver, es plausible hipotetizar a nivel (al menos) individual una *conexión íntima* entre estas opiniones o actitudes y la punitividad tal y como es estudiada habitualmente en la actualidad –esto es, en referencia al castigo o a partes del mismo¹⁰. Sin ir más lejos, igual que los derechos fundamentales vigentes en un país deben influir en sus tasas de detenciones, de condenas, de personas privadas de libertad, etc.; quienes opinan que los delincuentes deben ser castigados con firmeza quizá igualmente estén a favor de una rebaja en determinados derechos fundamentales, lo cual puede favorecer el castigo. Por ejemplo, Harrendorf escribe que «Para analizar y evaluar el punitivismo global de un sistema legal, las comparaciones [...] deberían incluir el sistema legal y sancionador

⁸ Cullen et al., 2009: 77; Keil y Vito, 1991: 455; Kury et al., 2002: 4 y 170; Kury et al., 2004: 97; Mayhew y van Kesteren, 2002: 67.

⁹ Vid. al respecto Keil y Vito, 1991: 457; Kury y Oberfell-Fuchs, 2008: 288 y 299; Kury et al., 2002: 3; Kutateladze, 2011: 157-160.

¹⁰ El concepto de actitud es discutido y existen concepciones más o menos amplias y estrechas. Entendidas por ejemplo como «representaciones cognitivas de las evaluaciones positivas o negativas de una persona de diferentes “objetos”», se ajustan perfectamente a nuestro objeto de estudio. Vid., de entre una bibliografía inabarcable, Roberts, 2010: 196-197, cita procede de 196.

completo así como el modo en que los casos son procesados»¹¹. Kurki mantiene que las salvaguardias procesales de los acusados son un indicador de la punitividad¹². Kutateladze utiliza un concepto de punitivismo estatal muy amplio que sin duda incluye aspectos procedimentales, aunque más bien luego pasan desapercibidos entre sus medidas¹³. Trabajando también en el nivel estatal de análisis, Frost interpreta la obra de algunos importantes teóricos como que, al menos implícitamente, se refieren a *algo más* que la (mera) existencia de más o menos castigo¹⁴; y desde luego no es irrazonable pensar que, verbigracia, una *cultura del control* es consistente con una disminución de importantes derechos y libertades de naturaleza procesal-penal¹⁵. Por otro lado, la misma Frost insiste en las distintas dimensiones de la punitividad y, en particular, en una de ellas como es la *propensión al castigo*: «aquellos Estados que han incrementado [...] *el riesgo de encarcelamiento* [...] son los más punitivos por definición»¹⁶. No es la única, desde luego, que evoca conceptos u operacionalizaciones de la punitividad que deberían como mínimo encontrarse influenciados de modo directo por algunos derechos fundamentales¹⁷. Parece claro, entonces, que una rebaja en muchos de ellos tenderá a repercutir en *un aumento del riesgo* de recibir una condena y una sanción, incluso privativa de libertad. Una argumentación semejante puede hacerse al nivel de los individuos. Así las cosas, en efecto, no es descabellado hipotetizar que esta actitud sobre los derechos fundamentales procesal-penales constituye bien una dimensión de la punitividad (o firmeza/benevolencia frente al delito) o bien incluso forma parte de la misma.

Puesto que no conozco pruebas definitivas en la actualidad para preferir una de estas dos alternativas –esto es, si la punitividad es un constructo uni- o pludimensional–, y tampoco se dispone de datos en el presente estudio como para evaluar esta cuestión, por simple parsimonia asumiré la segunda de ellas y me referiré a nuestras variables sobre derechos fundamentales procesal-penales directamente como *punitividad*. Con ello, sin embargo, no se toma partido por la naturaleza de la misma.

Preocupaciones como éstas evocan, desde un punto de vista metodológico, naturalmente, el enfoque de las variables latentes¹⁸, esto es la idea de que existe un factor o variable subyacente, no manifiesto del que se observan ciertos indicadores¹⁹. Esta será, en efecto, la estrategia analítica que seguiremos en la presente investigación.

Así las cosas, en el presente trabajo, como acaba de decirse, nos centraremos en el

¹¹ Harrendorf, 2011: 130 (énfasis añadido).

¹² Kurki, 2001: 333-368.

¹³ Kutateladze, 2011: 154-155.

¹⁴ Frost, 2006: 107, 111, 118-120 y 130.

¹⁵ Vid. en el sentido del texto, por ejemplo, Garland, 2001: 175, aunque también 170; el mismo, 2002: 9-10.

¹⁶ Frost, 2006: 119 (énfasis añadido).

¹⁷ Vid., así, Blumstein et al., 2005: 349 y 353; Tonry y Farrington, 2005: 9 y 31.

¹⁸ Biemer, 2011: 116-119; Bollen, 1989: 11-20; Powers y Xie, 2000: 9-11.

¹⁹ Agresti, 2002: 441 y 539.

nivel individual de análisis para estudiar una potencial dimensión o elemento de la punitividad comparativamente poco estudiada, como es la relativa a las opiniones o actitudes sobre los derechos fundamentales procesal-penales. Nuestro estudio se ocupa, en particular, de la *identificación* del constructo subyacente que denominamos, por parsimonia, *punitividad* a partir de una serie de variables categóricas –en realidad ordinales– que se hipotetiza están causadas por aquélla, esto es que son indicadores suyos. Dicho con otras palabras, nuestra (primera) hipótesis sustantiva nuclear es la existencia de una variable latente categórica unidimensional de punitividad –cuyos indicadores, como acaba de señalarse, son también categóricos (H₁). Como también se verá, este test abre las puertas, dentro de unos márgenes, a una valoración del *error de medición* –algo sobre lo que nos extenderemos más abajo con un cierto detalle. Una vez valorada la primera, testaremos la hipótesis de que nuestra (potencial) variable latente tiene en realidad una naturaleza continua y se distribuye normalmente (H₃)²⁰. A mayor abundamiento, la literatura ha sugerido la existencia de una clase latente específica de individuos con opiniones muy consistentes en determinadas materias, a los cuales denomina «ideólogos»²¹ –una hipótesis (H₂), pues, que también testaremos.

Como se advirtió, sin embargo, no es posible aquí estudiar algo tan nuclear como la relación entre nuestra potencial variable latente de opiniones o actitudes hacia los derechos fundamentales y otras medidas más comunes de estimación de la punitividad.

2. EL PRESENTE ESTUDIO

2.1. Hipótesis

Siguiendo la lógica descrita en el apartado precedente, nuestras hipótesis sustantivas son las siguientes:

H₁. Existe una variable latente categórica unidimensional de punitividad con indicadores también categóricos de actitudes ciudadanas sobre derechos fundamentales procesal-penales.

H₂. Existe en la variable latente categórica punitividad una clase latente de «ideólogos».

H₃. La variable latente punitividad tiene en realidad una naturaleza continua y se distribuye normalmente.

2.2. Datos

En el estudio del Centro de Investigaciones Sociológicas (CIS) *Los ciudadanos y el Estado (II)* (número 2671, de 2007) aparecen varias preguntas que se relacionan con derechos fundamentales procesal-penales. Se trata de una encuesta de ámbito nacional

²⁰ Nótese que esta H₃ tiene un contenido empírico mayor que el de H₁, puesto que toda variable continua puede segmentarse en subgrupos susceptibles de ser ordenados.

²¹ McCutcheon y Mills, 1998: 87.

y un universo compuesto por los residentes en España de 18 años de edad o más. El muestreo fue polietápico, estratificado por conglomerados, con selección de las unidades primarias y secundarias de modo aleatorio-proporcional y de las unidades últimas (individuos) mediante selección aleatoria nominal realizada sobre el marco de viviendas generado por el Instituto Nacional de Estadística a partir del padrón continuo de septiembre de 2006 (*sic*)²². Como es habitual en el CIS, el modo seleccionado fue la entrevista cara a cara en la vivienda del encuestado. Informa asimismo el CIS, con una fórmula bien conocida, del error muestral: para un nivel de confianza del 95,5% y $P=Q$, el error real es de $\pm 1,99\%$ para el conjunto de la muestra y para el supuesto de muestreo aleatorio simple (*sic*). Aquí, en todo caso, no asumimos representatividad ni tampoco nuestra preocupación se centra en el error *muestral*. El trabajo de campo se llevó a cabo entre el 15 de enero y el 15 de marzo de 2007.

De 4000 entrevistas diseñadas, el tamaño muestral final asciende a 2517. De dicho total de entrevistas, 39 no incluían una contestación sustantiva a ninguno de los ítems aquí utilizados. A nuestros intereses, estos casos son equiparables a *no respuestas al nivel de la unidad* y, por lo tanto, es como si no hubieran participado²³. Del total (corregido) de 2478, esto es excluyendo las 39 encuestas mencionadas, sólo han podido utilizarse 1881 en la presente investigación porque las contestaciones sustantivas a las preguntas no habían sido exhaustivas. Ello se traduce en que los sujetos excluidos por el problema de los datos perdidos alcanzan un elevado 24,09%. Esta situación conlleva potenciales problemas, de modo que más adelante se estudiará esta cuestión con algo más de detalle²⁴. Utilizamos de inicio, pues, los casos con información completa.

Las preguntas del CIS 2671 aquí de interés son las número 4, 9.01, 9.02 y 9.03. La Tabla 1 ofrece el texto de las mismas y sus categorías originarias de respuesta, así como la distribución de frecuencias y porcentajes de respuesta.

²² Al seguir un diseño muestral *proporcional* no se ha considerado aquí la inclusión de ponderaciones, vid. Vermunt y Magidson, 2007: 87-107.

²³ Nótese que, como se verá más abajo, nuestro enfoque para los datos perdidos no incluye su imputación en la matriz de modo preliminar.

²⁴ Biemer, 2011: 311-315; Collins y Lanza, 2010: 25, 39, 80-81 y 85; Vermunt et al., 2008: 369-391; Winship et al., 2002: 408-430.

Tabla 1 – Distribución de frecuencias y porcentajes para cuatro preguntas del Estudio CIS 2671: PP 4, 9.01, 9.02 y 9.03

P4. «En todos los sistemas judiciales se cometen errores. En su opinión, ¿qué es peor: condenar a una persona inocente o dejar libre a una que es culpable?»

<i>Categorías de respuesta</i>	<i>Fre</i>	<i>%</i>	<i>% válido</i>
«Condenar a una persona inocente»	1549	61,5	73,1
«Dejar libre a una persona culpable»	569	22,6	26,9
No puede elegir	382	15,2	-
No contesta	17	0,7	-
<i>Total</i>	<i>2517</i>	<i>100</i>	<i>100</i> <i>(N=2118)</i>

P9.01. «Imagine que el Gobierno sospecha que está a punto de cometerse un atentado terrorista. ¿Cree usted que las autoridades deberían tener derecho a mantener a personas detenidas tanto tiempo como quieran sin llevarlas a juicio?»

<i>Categorías de respuesta</i>	<i>Fre</i>	<i>%</i>	<i>% válido</i>
«Sin ninguna duda»	530	21,1	23,3
«Probablemente deberían tener derecho»	598	23,8	26,3
«Probablemente no deberían tener derecho»	532	21,1	23,4
«No, de ninguna manera deberían tener derecho»	617	24,5	27,1
No puede elegir	180	7,2	-
No contesta	60	2,4	-
<i>Total</i>	<i>2517</i>	<i>100</i>	<i>100</i> <i>(N=2277)</i>

P9.02. «Imagine que el Gobierno sospecha que está a punto de cometerse un atentado terrorista. ¿Cree usted que las autoridades deberían tener derecho a grabar las conversaciones telefónicas de la gente?»

<i>Categorías de respuesta</i>	<i>Fre</i>	<i>%</i>	<i>% válido</i>
«Sin ninguna duda»	535	21,3	23,1
«Probablemente deberían tener derecho»	601	23,9	26,0
«Probablemente no deberían tener derecho»	498	19,8	21,5
«No, de ninguna manera deberían tener derecho»	680	27	29,4
No puede elegir	154	6,1	-
No contesta	49	2	-
<i>Total</i>	<i>2517</i>	<i>100</i>	<i>100</i> <i>(N=2314)</i>

P9.03. «Imagine que el Gobierno sospecha que está a punto de cometerse un atentado terrorista. ¿Cree usted que las autoridades deberían tener derecho a detener y registrar a personas por la calle al azar?»

<i>Categorías de respuesta</i>	<i>Fre</i>	<i>%</i>	<i>% válido</i>
«Sin ninguna duda»	296	11,8	12,7
«Probablemente deberían tener derecho»	424	16,8	18,2
«Probablemente no deberían tener derecho»	546	21,7	23,4
«No, de ninguna manera deberían tener derecho»	1063	42,2	45,6
No puede elegir	142	5,6	-
No contesta	46	1,8	-
<i>Total</i>	<i>2517</i>	<i>100</i>	<i>100</i> <i>(N=2329)</i>

La primera de nuestras cuestiones que aparece en el cuestionario del CIS, la número 4, la hemos denominado CONDENAR (D). La misma es dicotómica y figura de modo independiente. Las otras tres, DETENER (A), GRABAR (B) y REGISTRAR (C), conforman una serie y comparten la primera parte de los enunciados. También tienen en común las cuatro categorías de respuesta²⁵. Para nuestro estudio se han eliminado las categorías de respuesta que carecían de un contenido sustantivo de interés para nuestra misión, las cuales coincidían en los cuatro casos: «No puede elegir» y «No contesta». Las mismas, pues, constituyen los datos perdidos para cada una de las cuatro cuestiones utilizadas en los presentes análisis. Como puede observarse, las cuatro hacen referencia a actitudes acerca de derechos fundamentales recogidos en la Constitución española de 1978, actualmente vigente, tales como la libertad (artículo 17.1, 2 y 3), el secreto de las comunicaciones (artículo 18.3) o la tutela judicial efectiva (artículo 24.1 y 2). Como puede observarse, su naturaleza es al menos parcialmente jurídico-procesal.

Las preguntas 9.01, 9.02 y 9.03 se han dicotomizado²⁶ para evitar el problema de *escasez* (sparseness) que se produce cuando, en determinados análisis estadísticos, algunas celdas no contienen ninguna observación, esto es $n=0$ –o bien cuando contienen números muy bajos²⁷. Este es un problema bien conocido ya en el análisis de tablas de contingencia de dos variables. Por diversos motivos que incluyen desde tamaños muestrales pequeños e incluso moderados a *ceros estructurales* que se refieren a celdas que representan situaciones imposibles (como *hombres embarazados*) o al menos no observables (individuos que forman parte de una población teórica pero *no aparecen en un censo ni tienen teléfono*), algunas celdas pueden no incluir ningún caso²⁸. Por ejemplo, García España y sus asociadas incluyen en su libro *Encuestas a víctimas en España* una tabla de contingencia con el fin de testar la relación entre *frecuencia con que se reciben noticias sobre delincuencia y percepción del crecimiento de la delincuencia en España en los últimos años*²⁹. Cada una de las preguntas tiene cuatro categorías de respuesta sustantiva posibles. Puesto que la inmensa mayoría de los entrevistados (92,8%) se ubican en la categoría de respuesta «Casi a diario» para la primera pregunta, dos celdas (12,5%) contienen cero observaciones –puesto que otras cuatro celdas

²⁵ Estas variables observadas, en realidad, tienen una naturaleza claramente ordinal: quienes creen que las autoridades deberían tener derecho bajo ciertas circunstancias a mantener personas detenidas sin llevarlas a juicio, a grabar conversaciones telefónicas o a registrar a personas al azar o bien creen que dejar libre a un culpable es peor que condenar a un inocente son, desde mi punto de vista, más punitivos respecto a los derechos fundamentales que quienes son de una opinión opuesta. Ello tiene todas las características de las variables ordinales. Algo más habrá que decir sobre esta cuestión y, en particular, sobre los análisis de clases latentes ordinales más abajo.

²⁶ Se han refundido las categorías de respuesta «No, de ninguna manera deberían tener derecho» y «Probablemente no deberían tener derecho» por un lado; y «Sin ninguna duda» y «Probablemente deberían tener derecho» por otro. Los valores asignados han sido, respectivamente, 1 y 2.

²⁷ Agresti, 2002: 542; Biemer, 2011: 188-190, 326-327 y 330; Collins y Lanza, 2010: 85-86; Collins et al., 1996: 135-137 y 144-145; Kaplan, 2009: 188.

²⁸ Acerca de la distinción, Agresti, 2002: 392.

²⁹ García España et al., 2009: 157 Gráfico 87.

parecen problemáticas por alojar una baja frecuencia, el total de celdas con menos de cinco observaciones para esta tabla de contingencia pareciera ser seis (37,5%)³⁰. Bajo estas circunstancias y sin ulteriores acciones correctoras es altamente problemático el recurso a estadísticos como χ^2 , el cual no puede interpretarse sin un conocimiento preciso de su distribución³¹.

Esta dificultad no aparece, como es sabido, en análisis también habituales como los factoriales o los mínimos cuadrados ordinarios. Ello es debido a que en estos supuestos se asume que las correlaciones de segundo orden y superior son cero. En nuestro caso de análisis de clases latentes también existen correlaciones que no son bivariadas sino de órdenes superiores y por eso requieren mayores tamaños muestrales³².

En efecto, un problema paralelo al descrito para las tablas de contingencia se presenta en el análisis de clases latentes. En realidad, en este tipo de análisis estas situaciones son habituales debido a que, al incluirse por regla más de dos variables, el número de celdas se dispara³³. En este sentido, el método de las clases latentes es *intensivo al nivel de los casos* –o sea que suele exigir tamaños muestrales elevados³⁴. Bajo estas circunstancias y en particular para las celdas con cero observaciones, los tests de χ^2 se vuelven problemáticos, como acaba de señalarse, porque la verdadera distribución del estadístico no es aproximada por la distribución teórica del χ^2 de modo correcto³⁵. Lo mismo cabe decir para L^2 . La literatura propone la razón N/k (siendo k el número de celdas) para valorar cuándo una tabla puede considerarse *escasa*³⁶. Para nuestro caso y utilizando todas las categorías de respuesta posibles de las preguntas del CIS 2671, esta razón no es especialmente baja. En efecto, puesto que $N=1881$ y $k=4*4*4*2=128$, $N/k=14,695$. Sin embargo, una inspección más detallada de la tabla de referencia arroja siete celdas con ceros muestrales y más de otras cuarenta con menos de cinco observaciones. Ello apunta claramente a problemas de escasez y a la conveniencia de reducir las categorías de respuesta para las tres preguntas utilizadas con cuatro de ellas.

³⁰ Digo *pareciera* porque no se informa de la frecuencia de cada celda, por lo que la información ha sido reconstruida artesanalmente a partir de otras tablas. Sobre el problema de celdas con ceros inferiores a 5 que superan el 20%, Agresti, 2002: 396.

³¹ Sobre dicho problema en tablas de contingencia, su alcance y posibles respuestas, vid. Agresti, 2002: 391-398, en particular 395-396 sobre cómo queda afectada la distribución de χ^2 y 397-398 sobre las soluciones.

Ello no quiere decir necesariamente, sin embargo, que no exista una relación, sino únicamente que el expediente probablemente no está justificado en la forma concreta en que se emplea. La reconstrucción aludida en la nota precedente, que sin duda debe contener errores, de la tabla de contingencia –unificando categorías de respuesta hasta evitar celdas con $n<5$ – sugiere que, en efecto, existe evidencia de tal relación. La precariedad de la reconstrucción de los datos, sin embargo, excluye el recurso por ejemplo a los *residuos ajustados tipificados* para comprobar dónde reside la relación o cualquier valoración de la fuerza de la misma. A la vez, esta relación, bivariada como es y, sobre todo, ambigua respecto a la direccionalidad, impide extraer conclusiones causales como las que sugieren las autoras, García España et al., 2009: 156 y, sobre todo, 168.

³² Vid. sobre ello Basilevsky, 1994: 611-612.

³³ Agresti, 2002: 392.

³⁴ Sobre el tamaño muestral en estos supuestos, vid. Collins et al., 1996: 135-136.

³⁵ Langeheime et al., 1996: 492-514, proponen métodos de remuestreo como el *bootstrap* para evaluar los parámetros en casos como estos de tablas con *escasez*.

³⁶ Biemer, 2011: 188 y 329.

Aparte de lo anterior, también es importante que, de acuerdo con Alwin, el hecho de que las categorías de respuesta de las variables manifiestas coincida con el de clases latentes es, como mínimo, ventajoso desde el punto de vista de la fiabilidad³⁷.

2.3. Estrategia analítica

Como se ha adelantado, un instrumento que cuenta con una cierta tradición en ciencias sociales³⁸ y que es cada vez más utilizado es el *análisis de clases latentes*³⁹. A menudo la literatura apunta que en una de sus versiones, probablemente la más habitual a día de hoy, representa un equivalente de los análisis factoriales en el caso de variables manifiestas y latentes categóricas⁴⁰. Dicho con otros términos más técnicos, uno de los usos más habituales es la identificación de variables latentes que pueden ser medidas a partir de una serie de variables manifiestas u observadas. No importa repetir que en estos análisis tanto las variables latentes como las manifiestas están medidas a nivel categórico –u ordinal. Esto representa un *modelo de medición*⁴¹. Ésta será la orientación que inicialmente seguiremos aquí, en consonancia con nuestras hipótesis.

En nuestra disciplina se pueden encontrar algunos interesantes ejemplos de análisis de clases latentes como los aquí propuestos. Así, Britt utilizó una estrategia múltiple que incluía análisis de clases latentes para defender la versatilidad de los delincuentes y, por lo tanto, las dificultades para establecer tipologías. Encontró que el mejor modelo era el de dos clases: delincuentes y no delincuentes⁴². D'Unger y sus colegas aplicaron el mismo esquema a las carreras criminales, aunque con notables particularidades analíticas, y encontraron entre cuatro y cinco clases latentes de criminales en las distintas muestras que tuvieron a su disposición⁴³. Kruttschnitt y Macmillan utilizaron análisis de clases latentes para clasificar a una muestra de mujeres que habían sido víctimas de violencia en relaciones personales de distinta naturaleza y a continuación utilizaron las clases como la variable dependiente en modelos de regresión logit multinomiales⁴⁴. Aunque, por supuesto, no son comportamientos delictivos, Sacco y sus compañeros identificaron tres clases de bebedores en una elevada muestra de sujetos de 60 o más años de edad del *National*

³⁷ Alwin, 2007: 266 y 272. Vid., en otro sentido, Biemer, 2011: 163.

³⁸ Acerca de su evolución histórica, vid. Collins y Lanza, 2010: 7-8; Goodman, 2002: 5 y 27-31; Skrondal y Rabe-Hesketh, 2007: 714-715.

³⁹ Collins y Lanza, 2010: 3-9 y 23-47; McCutcheon, 1987: 5-27. Es importante, por lo tanto, insistir en la existencia de una familia de análisis de variables latentes, a la cual pertenecen, sin duda, los análisis de clases latentes, Kline, 2005: 16-17.

⁴⁰ Así, Basilevsky, 1994: 611-612; Biemer, 2011: xiii y 117; Collins y Lanza, 2010: 6; Dayton, 1998: 1; McCutcheon, 1987: 7; el mismo, 2002: 56; McCutcheon y Hagenars, 1997: 266; Vermunt, 2003: 24. También puede verse como un análogo a los análisis de conglomerados, Uebersax, 1993: 1.

⁴¹ Bollen, 1989: 16-20; Goodman, 2002: 21; Skrondal y Rabe-Hesketh, 2007: 714. Acerca de los problemas y falta de flexibilidad de las tablas de contingencia de 3 y más variables, vid. Powers y Xie, 2000: 87 y 129-135.

⁴² Britt, 1994: 180 y 186-189.

⁴³ D'Unger et al., 1998: 1593-1625.

⁴⁴ Kruttschnitt y Macmillan, 2006: 147-161.

*Epidemiologic Survey on Alcohol and Related Conditions*⁴⁵; mientras que Timberlake encontró cinco clases de fumadores utilizando una también amplia muestra procedente del conocido estudio *Add Health*⁴⁶. En ambos casos se utilizaron análisis de clases latentes. Vaughn y sus asociados observaron mediante esta misma herramienta un modelo de cinco clases de niños según su nivel de autocontrol, con una quinta clase que nítidamente incluía a los más bajos en esta variable⁴⁷. Confío en que estos ejemplos sirvan como muestra. Nótese, en todo caso, que algunas de las aplicaciones precedentes utilizan datos longitudinales, con lo cual las diferencias metodológicas y analíticas con el presente estudio son significativas.

Una importante advertencia que habrá hecho el crítico lector es que nuestro análisis –que en lo fundamental trata de estimar una variable no observada– utiliza sólo cuatro ítems o preguntas –cuando en el ámbito de la medición, por ejemplo bajo la influencia de la Psicología clásica, suelen utilizarse muchos más. Esto es importante porque llama la atención sobre cómo en el fondo de nuestro enfoque general de variables latentes figuran muchas importantes diferencias con las formas más tradicionales de medición⁴⁸. Una revisión de esta cuestión, ni que decir tiene, excede la misión de este artículo; pero una de estas consecuencias es que escalas de medición cortas pueden arrojar mediciones plausibles e incluso superiores a otras más largas⁴⁹. A mayor abundamiento, Uebersax alerta contra el abuso de análisis de clases latentes con muchas variables manifiestas, y escribe que «Con más de dos o tres [...] los resultados [...] pueden ser difíciles de interpretar»⁵⁰.

Aunque, de modo sorprendente, la investigación y la literatura suelen prestarles poca atención, los modelos de clases latentes llevan a cabo una serie de importantes asunciones –algunas de las cuales pueden considerarse, además, relativamente fuertes. Biemer, a quien seguiremos en este punto con fidelidad, ofrece una taxonomía de las mismas: el modelo exige que las observaciones hayan sido recogidas mediante *muestreo aleatorio simple*; los indicadores deben ser *localmente independientes*, esto es independientes una vez controlada la variable latente de que se trate; las probabilidades de respuesta deben ser *homogéneas*, lo cual quiere decir que las probabilidades sean iguales para cualesquiera dos unidades seleccionadas; y los indicadores *unívocos*, o bien que lo sean todos ellos de la misma variable latente⁵¹. Aunque ahora mismo me limito a su enumeración, habrá que volver con más detalle sobre la comprobación de las mismas más abajo.

Quizá no está de más recalcar que, de entre estas asunciones, una que puede

⁴⁵ Sacco et al., 2009: 829-837.

⁴⁶ Timberlake, 2008: 709-715.

⁴⁷ Vaughn et al., 2009: 18-21, 26 y 28 sobre todo.

⁴⁸ Embretson y Reise, 2000: 3 y 13-39; Reeve y Mâsse, 2004: 248, 256-258 y 271.

⁴⁹ Embretson y Reise, 2000: 18-21; Reeve y Mâsse, 2004: 256-257.

⁵⁰ Uebersax, 1993: 2.

⁵¹ Biemer, 2011: 76, 83, 116, 125-128, 196-203 y 325-331.

considerarse nuclear es la de independencia local o condicional⁵². Como señala con gran claridad McCutcheon, esta situación se da cuando las relaciones entre un conjunto de variables es cero dentro de las categorías de alguna otra variable X; en estos casos puede decirse que aquellas variables son localmente independientes respecto de la variable latente subyacente X que puede considerarse explicativa de las primeras. Y esto es decisivo porque este criterio «proporciona un método para determinar si las relaciones entre un conjunto de medidas observadas se deben a alguna variable explicatoria no medida»; cuando se cumple esta asunción, entonces, «decimos que la variable adicional “explica” las relaciones observadas»⁵³. Por ello, así las cosas, más abajo dedicaremos un esfuerzo especial al test de esta hipótesis.

Finalmente, debe mencionarse que los análisis de clases latentes han sido también acreedores de algunas importantes consideraciones críticas. El propio Biemer recoge y discute las que siguen: no hay pruebas de que la variable latente estimada realmente sea la que subyace a los indicadores y es de interés; realiza, como he dicho, asunciones bastante fuertes; en casos de escasez sus resultados son pobres; sus hallazgos son difíciles de replicar; o, finalmente, sus resultados son fácilmente mal interpretados⁵⁴.

Los cálculos han sido realizados utilizando los programas LEM⁵⁵ y, de modo subsidiario, GLLAMM⁵⁶, el cual trabaja en conjunción con Stata. También se ha utilizado CONDEP⁵⁷ para testar la hipótesis de independencia local en nuestro modelo de dos clases latentes. Es muy importante subrayar que los tres son gratuitos y accesibles fácilmente por internet.

3. RESULTADOS

3.1. Análisis ordinarios

Nuestra primera hipótesis sustantiva, como se recordará, es que *existe una variable latente categórica unidimensional de punitividad con indicadores también categóricos de actitudes ciudadanas sobre derechos fundamentales procesales penales* (H₁). Procedemos a su test mediante un análisis de clases latentes y comenzamos seleccionando el modelo, en particular su número de clases. La orientación es aquí básicamente exploratoria. La siguiente Tabla número 2 ofrece nuestros estadísticos de bondad de ajuste para tres modelos de clases latentes⁵⁸, los cuales permitirán valorar cada uno de los mismos y compararlos entre sí.

⁵² Basilevsky, 1994: 609; Collins y Lanza, 2010: 44-47; Heinen, 1996: 6-9; McCutcheon, 1987: 14-16; McCutcheon y Mills, 1998: 84-85; Uebersax, 1993: 1-2.

⁵³ McCutcheon, 1987: 16.

⁵⁴ Biemer, 2011: 324-328.

⁵⁵ Vermunt, 1997a: *passim*; el mismo, 1997b: 4-6 y *passim*.

⁵⁶ Rabe-Hesketh et al., 2004a: 17-32; los mismos, 2004b: 169-170.

⁵⁷ Uebersax, 2009: *passim*.

⁵⁸ Como deber ser evidente, nuestros modelos contienen *una única variable latente*, Anderson y Vermunt, 2000: 85-89.

Tabla 2 – Estadísticos de bondad de ajuste para 3 modelos de clases latentes

<i>Modelo</i>	χ^2 [<i>p</i>]	L^2 [<i>p</i>]	<i>BIC</i> (L^2)	<i>AIC</i> (L^2)	<i>GL</i>
<i>Modelo de independencia</i> (M_0)	932,4547 [<,00001]	795,2258 [<,00001]	712,2907	773,2258	11
<i>2 clases latentes</i> (M_1)	8,1924 [,2243]	7,9882 [,239]	-37,2492	-4,0118	6
<i>3 clases latentes</i> (M_2)	,3568 [,5503]	,3597 [,5487]	-7,1798	-1,6408	1

N=1881.

Como puede observarse, el modelo de mutua independencia o de homogeneidad completa M_0 –en el que nuestras variables observadas se asumen mutuamente independientes entre sí y, de este modo, se especifica la existencia de una única clase latente– no se ajusta bien a los datos y debe rechazarse categóricamente (no sólo p para $\chi^2 < 0,00001$; sino que también existen numerosos residuos típicos superiores a $|2|$, e incluso a 10)⁵⁹. Se testan a continuación dos modelos de clases latentes con 2 (M_1) y 3 (M_2) clases respectivamente. Ambos, pues, cuentan con una única variable latente con varias categorías o clases latentes. Ninguno de los dos incluye restricción alguna y, como acabo de decir, pueden considerarse exploratorios⁶⁰. Ambos modelos se ajustan de manera más que aceptable a los datos según estadísticos bien conocidos y ampliamente utilizados en análisis de esta naturaleza. De entre los dos, M_1 es claramente superior según BIC y AIC, en cuanto que sus valores (negativos) son más bajos. Estos estadísticos, como es sabido, permiten comparaciones incluso entre modelos no anidados –siempre que los datos utilizados sean idénticos⁶¹. También es un modelo más parsimonioso tanto desde el punto de vista *empírico* de que estima menos parámetros y por ello cuenta con más grados de libertad como del *lógico* puesto que clasifica las observaciones recurriendo a un número menor de clases⁶².

Modelos con más clases latentes no pueden estimarse sin restricciones debido a que no se encuentran identificados, y en el momento actual de la discusión sobre nuestro constructo *punitividad sobre los derechos fundamentales* no existen

⁵⁹ Sobre los modelos de independencia, vid. en particular Agresti, 2002: 542 y 544.

⁶⁰ McCutcheon, 2002: 58; Vermunt, 1997b: 32-33. Van der Heijden et al., 2002: 127, sugieren con razón que una buena estrategia es, precisamente, decidir primero qué número de clases latentes se precisa para describir correctamente los datos. A falta de hipótesis con base teórica, esta estrategia será la seguida en este trabajo.

⁶¹ Aunque estos modelos pueden considerarse anidados, en estos casos no es posible utilizar la partición de L^2 . McCutcheon y Mills, 1998: 89, añaden que BIC y AIC negativos son indicativos de un modelo superior al saturado.

⁶² Acerca de la valoración de los modelos y, en particular, de su ajuste, vid. Agresti, 2002: 216-217 y 257; Biemer, 2011: 158-163, 169 y 190-191; Collins y Lanza, 2010: 82-88 y 109; Dayton, 1998: 14-24; Heinen, 1996: 39-43; Long, 1997: 109-113; McCutcheon, 2002: 66-69 y 85 nota 7; Powers y Xie, 2000: 65-71, 92, 99-101, 104-107 y 145-146; Vermunt y Magidson, 2002: 98. Una mirada más cautelosa sobre algunos de los estadísticos de bondad de ajuste aquí utilizados en Rudas, 2002: 345-348.

criterios teóricos claros para imponer dichas restricciones⁶³. En efecto, fuera de estos casos no es posible técnicamente testar modelos con más clases puesto que el modelo sin restricciones debe satisfacer:

$$I*J*K*L-1 > T(I+J+K+L-3)-1$$

siendo T el número de clases e I, J, K y L el número de categorías de cada uno de los items utilizados. En todo caso, tampoco existe justificación para ello –en puridad tampoco para M_2 – al haber hallado un modelo más parsimonioso con un ajuste aceptable⁶⁴.

Como acaba de decirse y puede comprobarse, el modelo con 2 clases latentes M_1 se ajusta francamente bien a los datos (p para $\chi^2=0,2243$; p para $L^2=0,239$; índice de disimilitud $d=0,0142$; residuos típicos $\leq |1,878|$). El modelo, además, es parsimonioso⁶⁵ y tiene una muy sencilla interpretación⁶⁶ y sentido desde un punto de vista teórico. Nada sugiere que no haya habido convergencia de EM⁶⁷. Igual que sus hermanos, no existen indicios de que M_1 no se encuentre identificado (por ejemplo, autovalores $\geq 101,045$)⁶⁸. No parece haber problemas tampoco con *soluciones* u *óptimos locales*⁶⁹. De hecho, se han utilizado distintos valores de inicio ($J=10$), siempre con los mismos resultados, lo cual sugiere ausencia de problemas evidentes para diversas cuestiones⁷⁰. No se han encontrado indicios de *sobreajuste* (overfitting)⁷¹. Sí se ha producido una situación de *intercambio de clases latentes* (flippage), pero para modelos relativamente sencillos como los nuestros es muy fácil de detectar y corregir⁷².

Ya hemos mencionado que nuestros indicadores pueden considerarse en puridad medidos a nivel ordinal y, así las cosas, puede pensarse en estrategias analíticas alternativas⁷³. Sin embargo, el carácter dicotómico de aquéllos minimiza la utilidad de éstas. Además, lo que es más importante es que incluso bajo estas circunstancias existen razones que justifican e incluso aconsejan mantenerse en el modelo tradicional de clases latentes⁷⁴.

⁶³ Anderson y Vermunt, 2000: 94-95; Böckenholt, 2002: 165.

⁶⁴ Collins y Lanza, 2010: 99.

⁶⁵ Collins y Lanza, 2010: 82.

⁶⁶ La interpretación es, en efecto, un importante elemento evaluador, McCutcheon, 1987: 13-14; McCutcheon y Mills, 1998: 88 y 90. Vid., por ejemplo, Kruttschnitt y Macmillan, 2006: 152-153 para unas clases a mi modesto entender difíciles de interpretar.

⁶⁷ Biemer, 2011: 89.

⁶⁸ Biemer, 2011: 183-185 y 330; Collins y Lanza, 2010: 89-92; Hagenaars, 1990: 112; McCutcheon, 1987: 25-26; el mismo, 2002: 66; Rabe-Hesketh y Skrondal, 2001: 1256-1259.

⁶⁹ Biemer, 2011: 89, 91-92, 192-194 y 330; McCutcheon, 2002: 65; Vermunt y Magidson, 2002: 97. Con pocas clases, como es nuestro caso, los óptimos locales son menores que cuando aquéllas aumentan, Böckenholt, 2002: 165.

⁷⁰ Agresti, 2002: 541; Collins y Lanza, 2010: 92.

⁷¹ Biemer, 2011: 160. En la Tabla 6.2 en Kruttschnitt y Macmillan, 2006: 149, existen valores que sugieren sobreajuste.

⁷² Biemer, 2011: 194-195.

⁷³ Agresti, 2010: 282-288; Biemer, 2011: 231-235; Powers y Xie, 2000: 119.

⁷⁴ Así, Alwin, 2007: 263 y 317. Como se dijo algo de pasada, teniendo en cuenta que nuestras variables manifiestas se encuentran medidas en realidad a nivel ordinal, se han llevado a cabo algunos análisis de clases latentes con ciertas adaptaciones para estos casos, vid. Rabe-Hesketh et al., 2004a: 87-108; Vermunt, 1997b: 36-38. Los resultados (no

Nuestras dos clases latentes pueden denominarse, respectivamente, «No punitivos» ($X=1$) y «Punitivos» ($X=2$). Esto no debe interpretarse necesariamente como que existen en la población dos grupos discretos, cualitativamente distintos, sino que es perfectamente posible que sean la consecuencia de un corte en una variable que se distribuye de modo continuo –más habrá que decir sobre esta cuestión. Los sujetos que son clasificados entre los «No punitivos» tienen una probabilidad relativamente elevada de ofrecer respuestas favorables a los derechos fundamentales (que coinciden con las categorías de respuesta numeradas como «1»); mientras que los «Punitivos» tienen una probabilidad relativamente elevada de ofrecer la respuesta opuesta («2»), aunque más tímidamente para C y con la excepción de D –también éstos es más probable que consideren más grave condenar a un inocente que dejar libre a un culpable. La clase de los «Punitivos» es menor en tamaño en nuestro estudio ya que existe una probabilidad más baja de pertenecer a la misma ($=0,4204$) en comparación con los «No punitivos» ($=1-0,4204=0,5796$). Estos datos pueden transformarse en porcentajes con facilidad, y para nuestro estudio la frecuencia esperada de cada clase es, respectivamente, 1090,723 y 790,277 –no se olvide la naturaleza probabilística del enfoque. Informa de estas cuestiones la Tabla 3, así como la 4, que figura algo más abajo. Las mismas igualmente ofrecen información valiosa para la valoración del error.

mostrados) coinciden con los reportados en el texto, sin nada digno de reseñar, incluyendo el caso de un enfoque no-paramétrico que LEM ofrece, vid. Vermunt, 1999: 187-210 y 214-218, sobre todo 214-217.

Tabla 3- Diagnósticos de *Punitivo*

Patrón de respuesta	Frecuencias			Probabilidad de pertenencia a clase		Clase pronosticada	Probabilidad de clase	Frecuencias esperadas	
	Frecuencia observada	Frecuencia estimada	Residuos típicos	$P(X=1)$	$P(X=2)$			$X=1$	$X=2$
ABCD									
1111	495	493,818	,053	,994	,006	1	,994	492,03	2,97
1112	147	148,906	-,156	,99	,01	1	,99	145,53	1,47
1121	43	42,943	,009	,88	,12	1	,88	37,84	5,16
1122	15	13,82	,317	,822	,178	1	,822	12,33	2,67
1211	148	142,714	,442	,767	,233	1	,767	113,516	34,484
1212	42	48,824	-,977	,673	,327	1	,673	28,266	13,734
1221	65	61,846	,401	,136	,864	2	,864	8,84	56,16
1222	26	28,128	-,401	,09	,91	2	,91	2,34	23,66
2111	157	160,753	-,296	,908	,092	1	,908	142,556	14,444
2112	56	50,94	,709	,86	,14	1	,86	48,16	7,84
2121	26	35,039	-1,527	,321	,679	2	,679	8,346	17,654
2122	22	14,78	1,878	,229	,771	2	,771	5,038	16,962
2211	190	186,497	,257	,175	,825	2	,825	33,25	156,75
2212	81	83,547	-,279	,117	,883	2	,883	9,477	71,523
2221	249	249,389	-,025	,01	,99	2	,99	2,49	246,51
2222	119	119,054	-,005	,006	,994	2	,994	0,714	118,286
Total	1881	1880,998		,5796 (ET=,02)	,4204 (ET=,02)			1090,723	790,277
		N de casos	1881						Correctos=1727,733
		N de casos invariantes	614						
		Proporción de invariantes	,326						
		N de observaciones	7524						
		N de observaciones congruentes	6957						
		Prop. observaciones congruentes	,925						
		Fiabilidad del conjunto	,919						

Texto de las preguntas y categorías de respuestas (agrupadas para A, B y C) válidas:

[A=] DETENER: «Imagine que el Gobierno sospecha que está a punto de cometerse un atentado terrorista. ¿Cree usted que las autoridades deberían tener derecho a mantener a personas detenidas tanto tiempo como quieran sin llevarlas a juicio?» (P9.01):

1. «No, de ninguna manera deberían tener derecho»/«Probablemente no deberían tener derecho».
2. «Sin ninguna duda»/«Probablemente deberían tener derecho».

[B=] GRABAR: «Imagine que el Gobierno sospecha que está a punto de cometerse un atentado terrorista. ¿Cree usted que las autoridades deberían tener derecho a grabar las conversaciones telefónicas de la gente?» (P9.02):

1. «No, de ninguna manera deberían tener derecho»/«Probablemente no deberían tener derecho».
2. «Sin ninguna duda»/«Probablemente deberían tener derecho».

[C=] REGISTRAR: «Imagine que el Gobierno sospecha que está a punto de cometerse un atentado terrorista. ¿Cree usted que las autoridades deberían tener derecho a detener y registrar a personas por la calle al azar?» (P9.03):

1. «No, de ninguna manera deberían tener derecho»/«Probablemente no deberían tener derecho».
2. «Sin ninguna duda»/«Probablemente deberían tener derecho».

[D=] CONDENAR: «En todos los sistemas judiciales se cometen errores. En su opinión, ¿qué es peor: condenar a una persona inocente o dejar libre a una que es culpable?» (P4):

1. «Condenar a una persona inocente».
2. «Dejar libre a una persona culpable».

La misma Tabla 3, que contiene una rica información, muestra también la clase con la que se relaciona de modo probabilístico cada patrón de respuesta —existen 2^k (=16) patrones de respuesta⁷⁵. Aunque debido a esta naturaleza probabilística existe un riesgo apreciable si se desean llevar a cabo análisis posteriores basados en la clasificación de observaciones particulares, la verdad es que las probabilidades de clase $P(X)$ son relativamente altas. Dicho al revés, las probabilidades de clasificación errónea son bajas en general. Solamente cuatro patrones de respuesta tienen una dicha probabilidad de clasificación errónea superior al 0,2 y, de éstos, sólo para dos la probabilidad es superior al 0,3. En particular, se trata de $y_j=(1, 2, 1, 1)$ y $(2, 1, 2, 2)$ para el primer caso, e $y_j=(1, 2, 1, 2)$ y $(2, 1, 2, 1)$ para el segundo. No hay nada sorprendente respecto a estos patrones concretos. Es importante así mismo señalar que, afortunadamente, esta situación sólo se espera que ocurra, respectivamente, en un 12,9% y un 4,5% de los casos respectivamente. Esto es, que menos de uno de cada veinte casos tiene una probabilidad de ser mal clasificado superior al 0,30, aunque también inferior al 0,33⁷⁶. No se pierda de vista, en todo caso, nuestra vocación ahora mismo evaluadora.

También podemos valorar las tasas de error para los items particulares. En efecto, cada uno de ellos puede ser visto como *indicadores medidos con error* de la variable latente no observada⁷⁷. La Tabla 4 nos ofrece esta información, siempre para M_1 .

Tabla 4 – Estimaciones para el diagnóstico de *Punitivo* para M_1

	<i>Clase 1 (No punitivo)</i>		<i>Clase 2 (Punitivo)</i>		<i>Var.</i>
	Estimación (E.T.)	Probabilidad <i>1-Especificidad</i>	Estimación (E.T.)	Probabilidad <i>Sensibilidad</i>	
<i>A [DETENER]</i>	-1,212 (,017)	,22926	1,531 (,0207)	,82215	1,8339
<i>B [GRABAR]</i>	-1,5004 (,0194)	,18236	2,3397 (,0193)	,91211	3,5929
<i>C [REGISTRAR]</i>	-2,5633 (,0116)	,07154	,47236 (,0238)	,61594	2,2453
<i>D [CONDENAR]</i>	-1,2027 (,014)	,231	-,73567 (,0184)	,32395	,05314
		<i>1-Prev.</i> ,5796		<i>Prev</i> ,4204	
	<i>E</i> = ,0823				
	<i>λ</i> = ,8043				

⁷⁵ McCutcheon, 2002: 60.

⁷⁶ Vid., sobre las probabilidades de clasificación incorrecta y su interpretación, Skrondal y Rabe-Hesketh, 2004: 291.

⁷⁷ Goodman, 2002: 19.

Así, las tasas de error para los ítems A, B, C y D son 0,2293, 0,1824, 0,0715 y 0,231 (falsos negativos), respectivamente, para quienes pertenecen a la clase latente «No punitivos»; mientras que la tasa de error es 0,1779, 0,0879, 0,3841 y 0,6761, respectivamente, para los «Punitivos» (falsos positivos). Partiendo de la base de que lo natural es pronosticar a los «Punitivos», las primeras probabilidades, en efecto, representan *falsos negativos* y las segundas *falsos positivos*. Algo más tendremos que decir sobre estos resultados, pero la impresión general para $X=1$ es buena y más moderada para $X=2$, en particular muy pobre para CONDENAR (D), que, como vamos a ir viendo, se va a revelar como el ítem más problemático. De todos modos, no es sorprendente en absoluto que, debido a la radicalidad de la cuestión, incluso quienes son ubicados por nuestro análisis en la clase «Punitivos» tengan una probabilidad relativamente alta de pensar que es más grave condenar a un inocente que dejar libre a un culpable. Con un par de salvedades, ambas para el caso de $X=2$, las tasas de error a este nivel son asumibles en lo tocante a la *evaluación del modelo*.

La Tabla 5 ofrece ulterior información valiosa para la evaluación de la relación entre cada uno de nuestros ítems y nuestra variable latente. Se trata de una serie de pseudocoefficientes de determinación⁷⁸ –que tienen la interpretación habitual–, siempre para nuestro modelo de 2 clases latentes M_1 .

Tabla 5 – Pseudocoefficientes de determinación para M_1

	$P(A X)$	$P(B X)$	$P(C X)$	$P(D X)$
	Pseudo- R^2			
<i>Entropía</i>	,2649	,4221	,2976	,0091
<i>Varianza cualitativa</i>	,3432	,5193	,3436	,0107
<i>Error de clasificación</i>	,566	,7084	,3244	,0001
<i>-2/N*Log-verosimilitud</i>	,2649/,2683	,4221/,3691	,2976/,2668	,0091/,0105
<i>Verosimilitud^{-2/N}</i>	,307/,4096	,4429/,5906	,305/,4323	,0105/,0153

Es fácil apreciar cómo D queda claramente infraexplicado en nuestro modelo, desde luego en comparación con sus compañeros. Dicho con otras palabras, D es, de todos nuestros ítems particulares, el más problemático con diferencia. Aunque estos pseudocoefficientes deben utilizarse con cautela⁷⁹, la impresión que ofrecen coincide con el conjunto de nuestras estimaciones. Por nuestra parte no se plantea la posibilidad de excluir CONDENAR (D) de nuestros modelos debido no sólo a que no estamos interesados aquí en la construcción de éstos, sino a que nuestra orientación *a priori*, inicial ha sido teórica⁸⁰. Desde este punto de vista, básico por otro lado para la validez, este ítem se relaciona con lo que hemos denominado *punitividad frente a los derechos fundamentales*. Aunque D puede aportar su granito de información, en una estrategia de construcción de modelos podría

⁷⁸ Sobre los mismos, en general, vid. Long, 1997: 102-109.

⁷⁹ Long, 1997: 102; Powers y Xie, 2000: 71.

⁸⁰ En realidad, no está claro que D pertenezca a la variable latente que hemos identificado.

considerarse su remoción. La Tabla 5 sugiere que los estadísticos para cada uno de los tres restantes ítems es aceptable.

En resumidas cuentas, esta impresión aceptablemente positiva –para la evaluación– de la clasificación con los patrones e ítems individuales en mente es confirmada por los estadísticos globales de *errores de clasificación* ($E=0,082$) y de *reducción en la proporción de errores* ($\lambda=0,804$)⁸¹, de los que informa la Tabla 4.

3.2. Una aproximación al error de encuesta mediante análisis de clases latentes (M_1)

Como se anunció y probablemente resulta evidente después de los párrafos precedentes, el enfoque de clases latentes abre enormes posibilidades, potencialmente revolucionarias, para la evaluación del *error de medición*⁸². Hasta ahora, sin embargo, este camino ha sido menos explorado por los analistas que el de la reducción de la información, la construcción de tipologías o el test de ciertas hipótesis. Paul Biemer, a quien está dedicado este artículo⁸³, se extiende en los detalles de esta utilización particular de los modelos de clases latentes⁸⁴. En concreto, escribe que se trata de un «método estadístico para la predicción de la verdadera clasificación de individuos de acuerdo con sus clasificaciones observadas. Adicionalmente, el LCA [análisis de clases latentes] proporcionará estimaciones de las probabilidades de los encuestados de ser mal clasificados por la pregunta», «Estas estimaciones pueden ser utilizadas para identificar ítems de cuestionario que sean problemáticos, y en última instancia para la mejora de la encuesta. Un objetivo secundario es la estimación de la prevalencia de características de una población corregidas por el error de clasificación»⁸⁵. Sin embargo, el *enfoque de análisis de error* sólo puede ser seguido aquí de modo *aproximado*. El motivo es que nuestro análisis no asume que la variable latente con dos clases que hemos identificado en M_1 sea la característica o variable latente *verdadera* que se buscaba. Sólo en este caso, por supuesto, es posible establecer con seguridad las probabilidades de error de cada ítem particular, y ello exigiría conocer con precisión y *a priori* la variable latente X y sus clases, y para que esta situación se dé, X debe poder ser medida directamente y sin error⁸⁶. Por supuesto, esto no excluye que sea posible, para nuestra investigación en particular, una

⁸¹ Sobre su cálculo e interpretación, vid. McCutcheon, 1987: 36-37.

⁸² Biemer, 2004: 226-228 y 242-246 sobre todo; el mismo, 2011: 1-24 y 115-125; McCutcheon y Mills, 1998: 82-83.

⁸³ Existe ya mucha investigación en esta línea, pero la reciente aparición de *Latent class analysis of survey error* de Biemer constituye un paso significativo, tanto por la información que recopila como por la investigación que añade. Mi trabajo ha recibido una fuerte influencia e inspiración de este autor, y la preocupación que destilan mis modestas investigaciones por la metodología de encuesta, la calidad y el error no muestral –algo que por lo general recibe una atención relativamente escasa en Criminología– recibió un fuerte impulso durante mi asistencia como oyente a las clases que impartían Biemer y Groves, durante el semestre de verano (2001) en *The Joint Program in Survey Methodology*. Como reconocimiento de su generosidad y excelencia científica, así como con mi agradecimiento, este artículo está dedicado a este autor, en el año de aparición de su monografía.

⁸⁴ Biemer, 2011: 1-24 sobre todo.

⁸⁵ Biemer, 2011: 22-23.

⁸⁶ Biemer, 2011: 118, 126 y 324-325.

aproximación a las probabilidades de error en que incurrimos. Con esta importante advertencia en mente, pues, pasamos a explorar el error en nuestros datos.

Nuestro enfoque permite, por lo pronto, una valoración de la fiabilidad en el caso de variables latentes dicotómicas –un terreno relativamente atrasado en comparación con las continuas. En efecto, los enfoques de variables latentes son muy convenientes para la estimación de la fiabilidad, e incluso superiores a otros habituales con medidas repetidas como el porcentaje de acuerdo en las respuestas a lo largo del tiempo –lo cual depende, naturalmente, del número de categorías⁸⁷, por no mencionar los efectos de la memoria. Nuestro modelo de clases latentes, para empezar, es uno de «múltiples indicadores»⁸⁸ –frente a los de múltiples medidas. Varios estadísticos son aquí de utilidad, y en términos generales nos permiten ser relativamente optimistas. Alwin describe el procedimiento de calcular el número de casos invariantes –esto es de observaciones en las que todas las contestaciones a los ítems coinciden⁸⁹– y la proporción que representan⁹⁰. Para nuestro modelo, la Tabla 3 informa de que la *proporción de invariantes*=0,33. Naturalmente, este procedimiento también depende del número de categorías y de ítems, pero haber acumulado estos dos patrones de respuesta invariantes hasta un tercio del total de las observaciones no puede valorarse de modo desfavorable. Una estrategia semejante es la que utiliza la proporción de observaciones congruentes, aquí un elevado 0,925 que invita al optimismo. Una información superior, «*más sensitiva*»⁹¹ es la que ofrece el estadístico de *fiabilidad del conjunto*⁹², que para nuestro caso rebasa el 0,9 –del que también informa la Tabla 3.

Algo semejante puede decirse para nuestros ítems particulares. Las especificidades para cada uno de los parámetros, como muestra la Tabla 4, más arriba, son relativamente altas. Las menores, por ejemplo, que corresponden a DETENER y CONDENAR (D), son de 0,771 y 0,769 respectivamente. Las sensibilidades estimadas son también iguales o superiores a 0,77, excepto para el caso de CONDENAR, que se queda en un modesto 0,324. Esto último quiere decir que el 77% o más de los casos punitivos contestan 2 en los tres primeros ítems, pero sólo el 32,3% en CONDENAR. Como ya sabemos, pues, el ítem D es problemático –en particular si se utilizara en solitario.

Así las cosas, pese a la evaluación que realizamos más arriba de nuestro modelo exploratorio de dos clases latentes, no puede olvidarse que, pese a que es posible utilizar este enfoque para la clasificación individual de sujetos, esto es su adscripción a una u otra clase, con la idea en mente, verbigracia, de utilizar esta información para ulteriores análisis, la clasificación que se hace es probabilística⁹³,

⁸⁷ Alwin, 1992: 105.

⁸⁸ Alwin, 2007: 61-67.

⁸⁹ Dicho de otro modo, $y_j=(1, 1, 1, 1)$ o $(2, 2, 2, 2)$.

⁹⁰ Alwin, 2007: 270.

⁹¹ Alwin, 2007: 270.

⁹² Vid. Clogg y Manning, 1996: 173-174.

⁹³ Goodman, 2002: 22; Vermunt y Magidson, 2002: 91; los mismos, 2003: 531-537.

de modo que por regla no es aconsejable seguir este procedimiento. La clave se encuentra, por supuesto, en las tasas de error en que se incurre.

3.3. ¿Existen «ideólogos» en nuestra muestra?

Como se mencionó, algunos autores han sugerido que, en determinadas situaciones, algunos encuestados muestran un nivel de consistencia en sus respuestas muy elevado⁹⁴. En particular, estos individuos tienen una *probabilidad perfecta* ($p=1$) de ofrecer una determinada respuesta a un ítem dado. Puede pensarse en estar siempre «Muy de acuerdo» con una serie de afirmaciones en una escala o, para nuestro caso, con seleccionar siempre las respuestas más punitivas. Dicho de otro modo más intuitivo, me refiero a sujetos que siguen un patrón de respuesta $y_j=(2, 2, 2, 2)$ con $p=1$. Al parecer, esta idea procede nada más y nada menos que de Otis Duncan, quien denominó «ideólogos (ideologues)» a estas personas⁹⁵. Se recordará que ésta era precisamente nuestra segunda hipótesis sustantiva de que *existe en la variable latente categórica punitividad una clase latente de «ideólogos»* (H_3).

La investigación criminológica ha realizado alguna referencia, si no al concepto, sí a figuras cercanas. Por ejemplo, Kuhn sugiere la existencia de un pequeño grupo de encuestados «particularmente punitivo» que estarían afectando los estadísticos descriptivos de medidas de punitividad⁹⁶. Quizá más próximo a nuestra hipótesis, Varona Gómez divide su muestra entre quienes habían contestado a todos los escenarios de su cuestionario de modo punitivo y el resto. Aunque este autor no sigue un modelo probabilístico ni de variables latentes y asume que su clase (observada) está compuesta por los más punitivos, aquí se despierta el interrogante de si este grupo podría tener alguna particularidad de naturaleza cualitativa⁹⁷.

Para testar esta hipótesis es preciso construir un modelo que incluyan una tal clase latente de «ideólogos» –en realidad «ideólogos-punitivos». La opción más natural, a la luz de nuestros hallazgos previos, es hipotetizar dos clases semejantes a las de M_1 más una tercera para los ideólogos⁹⁸. A tal fin, es menester introducir restricciones en el modelo fijando una respuesta punitiva para cada uno de nuestros cuatro ítems para el caso de los «ideólogos», esto es construir una clase latente con probabilidades condicionales de 1,0 para los cuatro ítems⁹⁹. Abandonamos aquí, entonces, nuestra orientación exploratoria por otra confirmatoria. A continuación,

⁹⁴ McCutcheon y Mills, 1998: 87.

⁹⁵ Vid. McCutcheon y Mills, 1998: 87.

⁹⁶ Kuhn, 2002: 123.

⁹⁷ Varona Gómez, 2008: 24.

⁹⁸ La opción de un modelo de sólo 2 clases latentes (M_3 bis), una para los «ideólogos» y otra para todos los demás parece intuitivamente poco plausible. De todos modos hemos testado esta opción con resultados muy poco halagüeños para la misma ($\text{Chi}^2=772,1257$ [$p<0,00005$]; $L^2=637,7074$ [$p<0,00005$]; $gl=10$ –devueltos mediante matriz de diseño–; índice de disimilitud $d=0,2363$; $\text{BIC}[L^2]=9584,7707$; $\text{AIC}[L^2]=617,7074$).

⁹⁹ Sobre la imposición de restricciones en general, vid. Collins y Lanza, 2010: 79-80 y 102-106; McCutcheon, 1987: 37-44.

la Tabla 6 ofrece los estadísticos de bondad de ajuste y los grados de libertad para el modelo de «ideólogos» y nuestro modelo de 2 clases latentes sin restricciones, que hemos aceptado de momento como el mejor.

Tabla 6 – Estadísticos de bondad de ajuste para dos modelos de clases latentes (M₁ y M₃)

Modelo	Chi ² [p]	L ² [p]	BIC (L ²)	AIC (L ²)	GL
2 clases latentes (M ₁)	8,1924 [,2243]	7,9882 [,239]	-37,2492	-4,0118	6
3. ^a clase de «ideólogos» (M ₃)	8,1976 [,1457]	7,989 [,1568]	-29,7088	-2,011	5 [†]

N=1881.

[†]Grados de libertad para M₃ correctos, devueltos mediante matriz de diseño en LEM¹⁰⁰.

El modelo de 3 clases latentes que incluye una tercera de «ideólogos» o «ideólogos-punitivos» (M₃) se ajusta razonablemente bien a los datos (p para Chi²=0,1457; p para L²=0,1568). Verbigracia, el *índice de disimilitud* (d=0,0141), que se interpreta como la menor proporción de observaciones que habría que cambiar de unas celdas a otras para un ajuste perfecto, y que Biemer¹⁰¹ considera indicativo de un buen ajuste cuando es inferior a 0,05 –o bien 0,01 en otras situaciones– invita al optimismo. Sin embargo, el modelo es claramente inferior a M₁ atendiendo a BIC y AIC. Además es obviamente menos parsimonioso. Finalmente pero igual de importante es que la probabilidad estimada de pertenencia a dicha tercera clase es, nuestro modelo M₃, de sólo 0,0002, esto es que su tamaño requeriría una población elevada como para contar con algún miembro. Por todo lo anterior debe rechazarse esta opción.

Ni que decir tiene que, en contraposición, también puede pensarse en una clase latente de sujetos *no punitivos* que responden de modo perfectamente consistente. Testado este modelo de «ideólogos-no punitivos», el veredicto es el mismo que para los «ideólogos-punitivos»¹⁰².

Así las cosas, al menos para nuestro estudio, no existen pruebas de la existencia de una clase latente significativa de «ideólogos»¹⁰³. Con ello no se encuentra apoyo para nuestra hipótesis sustantiva H₂.

¹⁰⁰ Vid. Biemer, 2011: 168.

¹⁰¹ Biemer, 2011: 160.

¹⁰² Chi²=8,1489 (p=0,1482); L²=7,9445 (p=0,1593); gl=5; d=0,0134; BIC(L²)=-29,7533; AIC(L²)=-2,0555; P(X=1)=0,0104.

¹⁰³ Quizá se pueda acusar a este modelo de demasiado determinista. Es también posible, mediante la imposición de restricciones, construir un modelo en el que la tercera clase de «ideólogos-punitivos» simplemente tenga una probabilidad relativamente alta e idéntica de contestar de modo punitivo a cada uno de nuestros cuatro ítems. Este modelo de 3 clases latentes M₃ ter, como su hermano, también se ajusta relativamente bien a los datos (p para Chi²=0,5058; p para L²=0,5005; d=0,0121; BIC[L²]=-26,8046; AIC[L²]=-4,6464). La probabilidad de respuesta punitiva para cada uno de los cuatro ítems y para esta clase es 0,7682. Como en el caso discutido en el texto, de nuevo este modelo es inferior a M₁ y la probabilidad de pertenencia a la clase latente «ideólogos», aunque asciende a 0,0591, sigue siendo muy baja. El veredicto, por lo tanto, es el mismo.

3.4. Comprobación de las asunciones del modelo de clases latentes (M_1)

Biemer, a quien como se dijo seguiremos en este punto, ofrece una excelente discusión de las asunciones del modelo de clases latentes –algo que suele pasar más bien desapercibido en el caso de un elevado número de analistas. Para empezar, el modelo debe realizarse a partir de una muestra extraída de una amplia población mediante muestreo aleatorio simple sin reemplazo. Esta condición no se cumple en nuestro estudio, y aunque existen formas relativamente sofisticadas de tener en cuenta los aspectos complejos del muestreo¹⁰⁴, en la práctica de nuestra disciplina, al menos hasta ahora, rara vez se recurre a estos expedientes aunque son asunciones de un conjunto de técnicas habituales. Más importante es que, como reconoce Biemer, todavía se sabe relativamente poco de la aplicación del análisis de clases latentes a datos de encuesta complejos¹⁰⁵. Eso sí, la base de nuestra muestra es el azar.

La asunción nuclear de nuestro enfoque es, como se dijo, la independencia condicional dada la pertenencia a una clase latente particular: los indicadores deben ser *localmente independientes*, o sea independientes una vez que se controla la variable latente X de que se trate¹⁰⁶. Cuando no se cumple la asunción de independencia local, tanto los estadísticos de bondad de ajuste como los errores típicos pueden resultar demasiado elevados, con lo que sus efectos a la hora de evaluar el modelo pueden ser críticos, aunque difíciles de predecir. Del mismo modo, las estimaciones de los parámetros pueden ser incorrectas. Lo más importante, sin embargo, es que el modelo mismo es puesto en duda. Se han propuesto, eso sí, formas de modelado de la dependencia condicional¹⁰⁷.

Existen varias formas, siguiendo siempre a la literatura, para detectar la dependencia condicional o local –o sea para testar nuestra asunción. Un diagnóstico muy interesante es el propuesto por Uebersax, que utiliza una versión modificada del *test de la razón de las log-ventajas* (LORC) originariamente desarrollado por Garret y Zeger. Aunque para los detalles he de remitirme a los trabajos originales, en realidad LORC trata de estimar bien conocidos valores z para cada par de items o variables, para lo cual utiliza la diferencia entre las razones de las log-ventajas observadas y esperadas, divididas por el error típico de las segundas. Naturalmente, la comprobación de si los valores z exceden los valores críticos habituales –en nuestro caso $|1,96|$ (o, si se prefiere, $|1,645|$)– nos ofrece evidencia de que tales items son condicionalmente dependientes. Esto es, uno de los procedimientos más habituales en análisis estadísticos, por lo tanto bien conocido. El programa CONDEP permite realizar esta operación con una relativa sencillez y rapidez. La

¹⁰⁴ Biemer, 2011: 209-230 y 330-331.

¹⁰⁵ Biemer, 2011: 230 y 333.

¹⁰⁶ Basilevsky, 1994: 609; Collins y Lanza, 2010: 44-47; Heinen, 1996: 6-9; Kaplan, 2009: 185-186; McCutcheon, 1987: 14-16; el mismo, 2002: 58; McCutcheon y Mills, 1998: 84-85; Skrondal y Rabe-Hesketh, 2004: 74; Uebersax, 1993: 1-2; Vermunt, 2003: 23.

¹⁰⁷ Vid., por ejemplo, Uebersax, 1999: 283-295.

Tabla 7 nos ofrece los resultados de los tests de dependencia condicional entre nuestros pares de items, por supuesto para M_1 .

Tabla 7 – Diagnósticos para dependencia condicional (M_1)

Tests		G^2	Razón de las log-ventajas esperada	E.T.	Razón de las log-ventajas observada	Valor z
<i>i</i>	<i>j</i>					
A	B	,01	1,8	,102	1,81	,05
A	C	,07	1,6	,112	1,57	-,25
A	D	1,02	,27	,104	,38	1,01
B	C	,01	2,07	,122	2,08	,07
B	D	1,45	,34	,104	,21	-1,2
C	D	,48	,29	,111	,37	,69

Tests definidos en Tabla 3.

Como puede observarse, $z \leq |1,2|$, de modo que no existen en estos análisis pruebas de una violación de la asunción de referencia (por ejemplo para $p < 0,05$, tests a dos colas), cuyo carácter es, no importa repetirlo, nuclear en análisis de clases latentes.

Biemer ofrece una manera de testar la hipótesis de *homogeneidad* –una asunción importante para la identificabilidad del modelo y la estimación de los parámetros– mediante la introducción de *variables de grupo* o de *agrupación*. Una opción habitual, a la par que intuitivamente natural y que seguiremos aquí es elegir el sexo (S) como variable de grupo. Nuestro autor sugiere aquí especificar el modelo $\{XS\} \{XA\} \{XB\} \{XC\} \{XD\}$, dejando fuera XSA XSB ... Esta exclusión implica homogeneidad de grupo para cada una de nuestras probabilidades de error. Este modelo M_{hom} se ajusta aceptablemente bien a los datos ($\text{Chi}^2=37,6619$ [$p=0,0097$]; $L^2=36,0279$ [$p=0,0153$]; $gl=20$; $d=0,0411$; $\text{BIC}[L^2]=-114,7632$; $\text{AIC}[L^2]=-3,9721$)¹⁰⁸. De este modo, nuestros hallazgos apuntan a que XA ... no dependen de S , o lo que es lo mismo que existe homogeneidad, como exige nuestro modelo de clases latentes¹⁰⁹. A mayor abundamiento, una mirada a los resultados de probabilidades de clase latente de la Tabla 8 sugiere consistencia asimismo con este diagnóstico.

¹⁰⁸ No está claro, sin embargo, si es superior a su alternativo que incluye los términos interactivos referidos ($\text{Chi}^2=16,2334$ [$p=0,1808$]; $L^2=15,4257$ [$p=0,219$]; $gl=12$; $d=0,0199$; $\text{BIC}[L^2]=-75,049$; $\text{AIC}[L^2]=-8,5743$). Decisivo a mi modo de ver es el criterio de la parsimonia. Acerca de supuestos en los que los estadísticos de ajuste no coinciden, como aquí, McCutcheon y Mills, 1998: 90. Adviértase, por otro lado, que los datos no coinciden con el caso de M_1 , ya que han debido adaptarse para añadir S, con lo que no es posible llevar a cabo comparaciones.

¹⁰⁹ Sobre todo ello, vid. Biemer, 2011: 148 sobre todo, así como 72-77, 126 y 144-155.

Tabla 8 – Probabilidades de clase latente para M_{hom}

	$X=1$	$X=2$
$S=1$ (mujer)	,5798	,4202
$S=2$ (hombre)	,4819	,4902
	,5181	,5098

La *univocalidad* es otra importante asunción de los modelos de clases latentes. La misma implica que todos los indicadores lo son de la misma variable latente. De nuevo Biemer ofrece una estrategia para testar esta hipótesis añadiendo una segunda variable latente (Y). A y B se modelan como indicadores de X y, por otro lado, C y D de Y¹¹⁰. El modelo M_{mult} así estimado se ajusta bastante bien a los datos ($\chi^2=7,0563$ [$p=0,1329$]; $L^2=6,8765$ [$p=0,1426$]; $gl=4$; $d=0,0141$; $BIC[L^2]=-23,2819$; $AIC[L^2]=-1,1237$). Sin embargo, es claramente inferior a M_1 desde diversas ópticas. Es importante añadir que la correlación entre X e Y debe ser elevada –lo cual contribuye a que el modelo esté identificado–, lo cual también es consistente con la superioridad de nuestro modelo con una única variable latente categórica y dos clases latentes (M_1). Así las cosas, tampoco hemos encontrado pruebas de una potencial infracción de la asunción de univocalidad.

3.5. Análisis de rasgo latente

Hasta aquí, pues, nuestros análisis de clases latentes. Nuestras dos clases latentes en M_1 identifican, como ya sabemos bien, dos grupos de modo claro: los «punitivos» ($X=2$) y los «no punitivos» ($X=1$) –o, si se prefiere, firmes y benévolo frente al delito o desfavorables a los derechos fundamentales procesal-penales y sus opuestos, respectivamente en ambos casos. Adviértase que se trata de clases que no son, ni mucho menos, estrictamente nominales, sino que *guardan un orden entre ellas*. Esto es, que las dos clases latentes representan «diferentes niveles de una escala ordenada»¹¹¹. Croon señala incluso cómo en ocasiones las clases pueden ordenarse *a lo largo de un continuo* en el sentido de que éstas sólo representan la división de una población. El mismo autor aclara que los análisis ordinarios de clases latentes no son suficientes para confirmar este orden¹¹². Así las cosas, aquí podemos, incluso, dar un paso más e hipotetizar que nuestra variable latente X de interés es no ya ordinal, sino continua¹¹³. Ello no sólo ofrecería una mejor descripción de nuestra variable de interés, sino que en hipotéticos análisis posteriores se podría utilizar más información. De este modo procedemos a testar

¹¹⁰ Biemer, 2011: 126, 203 sobre todo y 329.

¹¹¹ Dayton, 1998: 2; también Collins y Lanza, 2010: 34.

¹¹² Croon, 2002: 135-136 y 160-161, así como 141-152 para un modelo de clases ordinales.

¹¹³ También puede pensarse en *diferencias cualitativas o (meramente) cuantitativas*, vid. Yang et al., 2005: 194-196.

nuestra tercera hipótesis: *la variable latente punitividad tiene en realidad una naturaleza continua y se distribuye normalmente* (H_3)¹¹⁴.

En efecto, es bien sabido que las variables latentes pueden perfectamente ser continuas, naturalmente también en el caso de ítems categóricos u ordinales. Estos modelos que combinan variables manifiestas categóricas u ordinales con variables latentes que se encuentran en su origen de tipo continuo son denominados de *rasgo latente* –o *teoría de respuesta al ítem*¹¹⁵. Los mismos también pueden entenderse como una forma de análisis factorial cuyos indicadores son de nuevo categóricos u ordinales y el o los factores continuos. Se trata de una herramienta muy flexible y puede utilizarse, como nos interesa aquí, para valorar hipótesis sobre la naturaleza y distribución de un constructo, o bien simplemente para la reducción de datos o para otras formas exploratorias¹¹⁶.

Procedemos, pues, a testar nuestro modelo de rasgo latente (M_4), utilizando los datos a nuestra disposición. La siguiente Tabla 9 ofrece los estadísticos de bondad de ajuste más los grados de libertad para el mismo, así como para nuestro modelo de dos clases latentes (M_1), sobre cuyo buen ajuste nos hemos extendido más arriba.

Tabla 9 – Estadísticos de bondad de ajuste para un modelo de clases latentes y un modelo de rasgo latente

<i>Modelo</i>	χ^2 [<i>p</i>]	L^2 [<i>p</i>]	$BIC (L^2)$	$AIC (L^2)$	<i>GL</i>
<i>2 clases latentes</i> (M_1)	8,1924 [,2243]	7,9882 [,239]	-37,2492	-4,0118	6
<i>Rasgo latente</i> (M_4)	10,047 [,1859]	9,958 [,191]	-42,8189	-4,042	7

N=1881.

Como puede observarse, el modelo de rasgo latente (M_4) se ajusta francamente bien a los datos¹¹⁷. No sólo cuenta con el BIC más bajo de todos los modelos ensayados en este trabajo, sino que destaca también por su parsimonia, con hasta 7 grados de libertad¹¹⁸. Igualmente $d=0,0142$ ofrece una lectura positiva. Tampoco

¹¹⁴ Habrá apreciado el amable y crítico lector que nuestra hipótesis H_3 en realidad incluye dos elementos o subhipótesis a la vez, que serán testadas simultáneamente. Por ello y por parsimonia he preferido definir de este modo algo heterodoxo nuestra última hipótesis sustantiva.

¹¹⁵ Biemer, 2011: 119; Embretson y Reise, 2000: 40-61; Heinen, 1996: 26-28, 91-120, 93-97 sobre todo y 150-190; Rabe-Hesketh et al., 2004a: 13-14; Reeve y Måsse, 2004: 249-155; Skrandal y Rabe-Hesketh, 2007: 715-717 y 727-728; Uebersax, 1997: 188-193; Vermunt, 1997b: 38-40.

¹¹⁶ Heinen, 1996: 26-28, 55-60, 69-70, 91-120 y 151-152.

¹¹⁷ Sobre la evaluación de estos modelos, con un enfoque bastante sofisticado, vid. Embretson y Reise, 2000: 233-238.

¹¹⁸ Como era de esperar dada su naturaleza, los estadísticos globales para la clasificación *errores de clasificación* ($E=0,512$) y *reducción en la proporción de errores* ($\lambda=0,2693$) son ahora relativamente pobres, desde luego en comparación con M_1 . Por supuesto que se incurre en muchos más errores cuando, digamos, existen muchas más clases latentes – esto es cuando se asume la distribución a lo largo de un continuo.

presentan problemas dignos de mención los residuos típicos ($\leq |1,802|$), salvo para el caso de $y_i=(2, 1, 2, 2)$ en el que se superan ligeramente los límites de la significación ($=2,01$) –ello sugiere que el modelo no ofrece para dicho patrón una buena estimación, pero, por lo demás, no debería ser muy problemático. No se aprecian problemas de identificación (autovalores $\geq 10,437$), de óptimos locales o de otra naturaleza. Como se dijo, hemos testado M_4 bajo la asunción de distribución normal¹¹⁹, de modo que nuestros hallazgos, pues, favorecen la misma.

Los modelos de rasgo latente también realizan una serie de asunciones¹²⁰. En particular, exigen *unidimensionalidad* y *un buen ajuste* a los datos. Ambas cosas parecen haber sido alcanzadas en nuestro estudio. De nuevo, la asunción de independencia local es nuclear. Reeve y Mâsse, sin embargo, mantienen que ningún programa estadístico es capaz de testar esta hipótesis, a la vez que recomiendan una estrategia que se aproxima a la construcción de modelos¹²¹ –algo que, ya se ha dicho, no forma parte de nuestros intereses.

En particular nos interesa la comparación entre M_1 y M_4 , y aquí el modelo de rasgo latente se muestra claramente superior tanto empírica como lógicamente según diversos criterios estadísticos y lógicos¹²². Con ello, pues, encontramos evidencia favorable a nuestra tercera hipótesis de que *la variable latente punitividad tiene en realidad una naturaleza continua y se distribuye normalmente* (H_3).

3.6. Análisis con datos perdidos

Como se dijo, los presentes análisis han sido realizados sin tomar en cuenta la existencia de datos perdidos. La Tabla 1 mostraba que el número de datos perdidos para cada uno de los ítems es el siguiente: A=240; B=203; C=188; y D=399. Cuando se combinan a la vez, resulta que, como se dijo, de los 2478 sujetos que habían contestado de modo sustantivo a alguno de los cuatro ítems de interés, sólo han podido utilizarse 1881. Una pérdida de datos del 24,09% es preocupante y puede conllevar potenciales problemas, de modo que se hace preciso explorar hasta qué punto esta situación afecta a nuestros modelos principales, esto es M_1 y M_4 . La literatura ha dedicado una importante atención a esta circunstancia ya en el caso de datos categóricos¹²³. Una primera opción que viene a la mente de modo intuitivo es la de imputar los datos perdidos en la matriz originaria mediante algún procedimiento válido para variables dicotómicas y, a partir de ahí, calcular las

¹¹⁹ En particular, hemos especificado que el método para aproximar la distribución normal sea el de *densidad reescalada* y que el rango de la distribución normal sea -3 a +3. Esto se puede realizar mediante la opción «nor(·)» de LEM.

¹²⁰ Sobre las asunciones de estos modelos, vid. Embretson y Reise, 2000: 231-233; Reeve y Mâsse, 2004: 255-256.

¹²¹ Reeve y Mâsse, 2004: 256.

¹²² Aunque se muestra cauteloso, Uebersax afirma que modelos de clases latentes y de rasgo latente pueden compararse entre sí utilizando índices como BIC, 1997: 191, si bien recomienda utilizar otros criterios.

¹²³ Biemer, 2011: 311-315; Collins y Lanza, 2010: 25, 39, 80-81 y 85; Vermunt et al., 2008: 369-391; Winship et al., 2002: 408-430.

celdas para el análisis de clases latentes¹²⁴. Vermunt, por el contrario, sugiere una interesante estrategia alternativa consistente en especificar tipos diferentes de tablas de frecuencias observadas correspondientes a subgrupos diferentes de individuos, subgrupos compuestos por aquellos para los que se dispone del mismo tipo de información¹²⁵. La lógica es muy simple¹²⁶: es posible que algunos sujetos no entren a formar parte de la muestra si su información está perdida para algún ítem, pero no para los demás; de este modo, es posible recuperar la información disponible mediante tablas que incluyan los ítems a los que sí hayan contestado. De este modo, en efecto, se pueden especificar subgrupos de individuos según sus patrones de respuesta. En nuestros análisis para datos perdidos nos hemos limitado a añadir la Tabla (y el subgrupo) ABC debido a que al agregar otras se producía un problema de escasez que se hacía cada vez más serio, a la par que la información que se iba añadiendo era cada vez menor. Con nuestra estrategia, la menor frecuencia de una celda es seis. Ello es debido en buena medida a que D es, como vimos, el ítem con más datos perdidos. Siguiendo este procedimiento propuesto por Vermunt se utilizan finalmente 2197 casos de un total de 2478 individuos que había contestado a alguno de los cuatro ítems de interés. Ello representa un 88'66% del total, o lo que es lo mismo, los datos perdidos se han rebajado a un 11,34% del total, esto es a menos de la mitad de los primeros análisis¹²⁷. La Tabla 10 ofrece los estadísticos de bondad de ajuste y los grados de libertad para nuestros modelos de referencia M_1 y M_4 .

Tabla 10 – Estadísticos de bondad de ajuste para un modelo de clases latentes y un modelo de rasgo latente, análisis para datos perdidos (Tablas ABCD y ABC)

<i>Modelo</i>	χ^2 [p]	L^2 [p]	$BIC (L^2)$	$AIC (L^2)$	GL
<i>2 clases latentes (M₁)</i>	26,8875 [,0132]	50,2139 [,0116]	-72,8191	1,2139	13
<i>Rasgo latente (M₄)</i>	28,871 [,0109]	52,0915 [,01]	-78,5833	1,1445	14

N=2197.

Es claro que nuestro interés ahora se centra en comprobar si la pérdida de datos está afectando a nuestros análisis principales limitados a datos con información completa. Esta es una vocación fundamentalmente confirmatoria. Ambos modelos

¹²⁴ Vid. una estrategia general en esta línea en Allison, 2002: 19-27; Collins y Lanza, 2010: 80-81; Enders, 2006: 315-339.

¹²⁵ Vermunt, 1997b: 49-50.

¹²⁶ *Muy simple* si se considera la forma en que se introducen habitualmente los datos en LEM mediante tablas de contingencia. Existen, eso sí, otras maneras de indicar al programa qué datos debe utilizar.

¹²⁷ Puede ser discutible nuestra consideración de la categoría de respuesta «No puede elegir» como dato perdido – igual, entonces, que «No contesta». No sólo es imaginable que ambas categorías de respuesta hayan sido en alguna medida utilizadas de modo intercambiable por los encuestados; sino que, sobre todo, el volumen de «No contesta» es comparativamente tan bajo que probablemente estas contestaciones no representan un problema serio por ellas mismas.

de 2 clases latentes sin restricciones (M_1) y de rasgo latente (M_4) –y ya estamos familiarizados con los dos, especialmente con el primero de ellos– se ajustan bien a los datos. Pese a que para ambos χ^2 $0,05 > p > 0,01$ y L^2 $0,05 > p \geq 0,01$, no puede olvidarse que estos estadísticos tienden a ser, a nuestros intereses, *conservadores* aquí y que, por ese motivo, incluso buenos modelos podrían ser rechazados conforme a ellos. En todo caso, debe insistirse en que χ^2 (sólo) es significativo al nivel $\alpha=0,05$, pero no $0,01$; y que nuestra muestra, superior a 2000, no puede considerarse pequeña. A mayor abundamiento, el índice de disimilitud es para ambos modelos inferior a $0,05$ (d para $M_1=0,0348$; d para $M_4=0,0339$), lo cual es indicativo de un buen ajuste. De este modo, la presente evidencia refuerza nuestra confianza en nuestros análisis principales así como en nuestras pruebas a favor de las hipótesis H_1 y H_3 . De nuevo, este esfuerzo sugiere, recurriendo a expedientes ya utilizados más arriba, la superioridad de M_4 sobre M_1 , o sea que nuestra variable latente punitividad parece tener un carácter continuo y distribuirse de modo normal¹²⁸.

4. DISCUSIÓN Y CONCLUSIONES

A pesar de que la punitividad se ha convertido en uno de los objetos de estudio más importantes de la Criminología contemporánea, todavía queda mucho por hacer en ámbitos básicos tan relevantes como su definición, operacionalización o medición. En el presente estudio hemos estado fundamentalmente preocupados con la identificación de una variable latente a la que, por pura simplicidad, hemos denominado *punitividad* a partir de actitudes sobre derechos fundamentales procesal-penales, a nivel individual. Así, hemos recurrido a modelos de clases latentes para testar, con evidencia favorable a la misma, la hipótesis H_1 de que existe una variable latente categórica (u ordinal), unidimensional con indicadores categóricos (u ordinales) de actitudes ciudadanas sobre derechos fundamentales. En particular, el modelo de 2 clases latentes (M_1) se ajusta bien a los datos y no se han encontrado problemas de ningún tipo ni tampoco violaciones de las asunciones más importantes. Análisis con datos perdidos han confirmado los hallazgos ordinarios. Ulteriores análisis de rasgo latente M_4 sugieren, por su parte, que nuestra variable latente puede tener en realidad una naturaleza continua y que se distribuye normalmente (H_3). Recuérdese, sin embargo, que hemos denominado a nuestro constructo latente «punitividad» por parsimonia, pero sin tomar partido sobre su verdadera relación con otras medidas más habituales. En efecto, si la punitividad es empíricamente un constructo unidimensional o si, por el contrario, tiene una naturaleza mixta, heterogénea –y la opinión sobre los derechos fundamentales

¹²⁸ Puesto que los datos utilizados no coinciden, no es posible utilizar BIC y AIC para comparar estos análisis con los realizados con datos con información completa. Ello carece de mayor interés en el conjunto de nuestra estrategia.

constituye una de sus dimensiones— está todavía por ver¹²⁹ y aquí no ha podido testarse por falta de información. La investigación es también en este punto escasa y limitada. En un estudio que casi puede considerarse una excepción, Kutateladze dedica mucha atención a este tema utilizando 44 ítems sobre punitividad estatal. Sin embargo, primero informa de unidimensionalidad usando un análisis de fiabilidad mediante el alfa de Cronbach y más adelante de multidimensionalidad a partir de una matriz de correlaciones con los ítems agrupados en 5 conjuntos¹³⁰. Este enfoque, así las cosas, es insuficientemente sofisticado para una cuestión ciertamente compleja y no puede, por lo tanto, asumirse¹³¹. Utilizando análisis de componentes principales sobre punitividad para 24 actos delictivos muy heterogéneos entre sí, Kury y Obergfell-Fuchs informan de 5 factores distintos¹³², si bien el carácter exploratorio de esta herramienta estadística no debe pasarse por alto. Keil y Vito, por su parte, informan en su estudio de unidimensionalidad, pero sus indicadores se limitaban a la pena de muerte¹³³. En buena medida, se dirá con razón, esta cuestión depende de la definición que otorguemos a nuestro concepto, pero la verdad es que reina aquí un gran desacuerdo, aunque sin duda muchas de las propuestas existentes incluyen sin gran esfuerzo potenciales aspectos jurídico-procesales y constitucionales de la punitividad y el castigo¹³⁴.

Nuestro estudio, entonces, se ha centrado en el modelo de medición, y a tal fin ha recurrido con especial énfasis a los *análisis de clases latentes* —aunque, como debería resultar ya obvio, hemos seguido la recomendación general de utilizar varias técnicas de modelaje¹³⁵. Debe aclararse desde ahora que nuestro modelo de rasgo latente no excluye en absoluto la relevancia y utilidad del de clases latentes¹³⁶, en el cual hace hincapié esta investigación. Estos últimos son especialmente prometedores en nuestra disciplina y en las ciencias humanas y sociales en general debido a que asumen que sus indicadores se encuentran medidos únicamente a nivel categórico u ordinal. No puede exagerarse la importancia de esta circunstancia habida cuenta de que los criminólogos no siempre disponen de variables medidas a nivel de razón o de intervalo¹³⁷, a pesar de lo cual recurren habitualmente a herramientas de análisis estadístico que asumen, entre otras cosas, este nivel de medición e incluso distribución normal multivariante¹³⁸.

¹²⁹ Kury et al., 2002: 3.

¹³⁰ Kutateladze, 2011: 156 nota 7 y 158 sobre todo.

¹³¹ Bollen, 1989: 226-232, por ejemplo y entre otros pasajes. La estrategia de la autora quizá no pueda calificarse ni siquiera de exploratoria.

¹³² Kury y Obergfell-Fuchs, 2008: 288 en especial.

¹³³ Keil y Vito, 1991: 457 sobre todo.

¹³⁴ Vid., por ejemplo y en esta línea, Brown, 2006: 306.

¹³⁵ Uebersax, 1997: 188.

¹³⁶ Uebersax, 1997: 192.

¹³⁷ Clarke y McCutcheon, 2009: 713.

¹³⁸ Debe añadirse que otras herramientas, como es el caso de los modelos de ecuaciones estructurales, están avanzando mucho en la incorporación de variables que no son continuas ni se distribuyen de modo normal, Finney y DiStefano, 2006: 269-301.

Del mismo modo, en ocasiones es imaginable que se hipotetice una variable latente categórica, por no mencionar el estudio de la heterogeneidad no observada. Esta técnica es flexible y poderosa, en el sentido de que abre la puerta al test de muchos modelos distintos, con y sin base teórica; y, a la vez, permite relajar asunciones de métodos más tradicionales¹³⁹. Entre sus principales desventajas, por el contrario, se encuentra que lleva a cabo asunciones relativamente fuertes¹⁴⁰. Su flexibilidad, así, permite su utilización en diversos frentes, como los que siguen: reducción de datos, estudio del error de medición, análisis de heterogeneidad no observada, refinamiento de escalas de medición, investigación de datos perdidos, test de determinadas hipótesis, exploración o confirmación de la existencia de variables latentes unidimensionales, etc.¹⁴¹. Así las cosas no puede sorprender que haya llegado a proponerse su utilización de modo rutinario¹⁴².

Dos de las funciones de esta herramienta han sido especialmente provechosas en el presente estudio y, a mi modo de ver, pueden ser particularmente relevantes para la Criminología en general y para el estudio contemporáneo de la punitividad en particular. En primer lugar, como acaba de decirse, aquí hemos identificado una variable latente de naturaleza unidimensional de actitudes sobre derechos fundamentales. En efecto, el estudio de la dimensionalidad es uno de los usos más prometedores de estos métodos en el caso de indicadores categóricos¹⁴³. Nuestro hallazgo plantea la relación del mismo con concepciones más habituales de la punitividad a nivel individual centradas en el castigo. Como se dijo, los derechos fundamentales entran de lleno en el complejo libertad-seguridad que subyace al problema y, por otro lado, los mismos sin duda pueden influenciar por ejemplo la probabilidad del castigo de modo más o menos directo. Así las cosas, nuestro estudio sugiere que la punitividad podría ser un constructo mucho más complejo incluso de lo que a menudo ya se considera, y que deberían tomarse más en cuenta potenciales ulteriores dimensiones.

Los modelos de clases latentes –tampoco los de rasgo latente– no deben interpretarse demasiado literalmente¹⁴⁴, pero aquí parece existir un orden o incluso un continuo a lo largo de los cuales se puede bien *ordenar* bien *colocar a distancias conocidas* a los grupos de individuos. ¿Es nuestra variable latente esencialmente ordinal o continua? Existe una cierta polémica acerca de si las variables categóricas lo son de modo inherente o si, por el contrario, una mayoría de variables son conceptualmente continuas, si bien a menudo sólo son observables

¹³⁹ Biemer, 2004: 227, 243 y 245; McCutcheon, 1987: 7-8 y 79-80; McCutcheon y Hagenaars, 1997: 266 y 276.

¹⁴⁰ Biemer, 2011: 326; vid. también el mismo, 2004: 227.

¹⁴¹ McCutcheon y Mills, 1998: 94 nota 2.

¹⁴² Uebersax y Grove, 1990: 559.

¹⁴³ Embretson y Reise, 2000: 229-230; McCutcheon, 1987: 8; Uebersax, 1997: 190-191. El mismo Uebersax propone una forma alternativa de valorar la dimensionalidad de una variable: llevar a cabo análisis de clases latentes con un relativamente elevado número de clases, 1993: 11 –ya vimos que modelos de este tipo no se encuentran identificados en nuestro estudio sin ulteriores acciones restrictoras.

¹⁴⁴ Agresti, 2002: 544; Uebersax, 1997: 191.

o mensurables de modo categórico¹⁴⁵. Collins y Lanza tienen que decir lo siguiente al respecto: «Nuestro punto de vista es que muchos fenómenos pueden tener características tanto continuas como categóricas», «Más que un debate acerca de si el consumo de alcohol es verdaderamente continuo o categórico, más bien consideraríamos si una operacionalización continua o categórica es más relevante para las preguntas de la investigación de que se trate». Desde este punto de vista, que parece razonable, tanto el modelo de 2 clases latentes como el de rasgo latente pueden ser plenamente útiles. Concluyen los autores con que «simplemente hacemos la modesta asunción de que, dado un conjunto particular de preguntas de investigación y un conjunto de datos disponible, un modelo categórico puede proporcionar información útil»¹⁴⁶. Así ha debido de quedar clara la utilidad de los modelos de clases latentes.

Por otro lado, sin embargo, las reflexiones precedentes no deben hacer olvidar las dificultades que presenta la utilización de clasificaciones como las mencionadas –que son probabilísticas– en análisis posteriores, por ejemplo para la construcción de variables dependientes en análisis de regresión. El motivo reside, naturalmente, en la limitada certeza que se puede tener en la clasificación¹⁴⁷.

Una segunda función importante de los enfoques de variables latentes que puede resultar especialmente útil para el estudio de la punitividad –en realidad para infinitísimos objetos de estudio– tiene que ver con el *error de medición*¹⁴⁸. Al combinar varios indicadores –que a menudo han sido medidos a nivel categórico u ordinal–, permiten una evaluación del error en que se incurre con cada uno de los indicadores individuales. El razonamiento es fácil de seguir. Cuando se utiliza un único indicador manifiesto, esto es de modo aislado, se asume que se trata de un *indicador perfecto*, no afectado por error de medición¹⁴⁹. Es bien sabido tanto que esta asunción tiende a ser insostenible porque toda medición de este tipo conlleva un cierto grado de error¹⁵⁰, como que suele ser la estrategia habitual en la investigación criminológica en la materia. Por el contrario, el enfoque de variables latentes, al permitir varios indicadores de modo simultáneo y extraer lo que de común tienen, permite una valoración del error de cada uno de ellos en particular¹⁵¹. Desde esta perspectiva, los indicadores individuales ya no se asumen mediciones perfectas, sino «mediciones falibles de una variable latente»¹⁵².

Nuestro análisis sobre el error –con la limitación, como se dijo, de que no se asume que la variable latente de interés teórico haya sido verdaderamente medida

¹⁴⁵ Powers y Xie, 2000: 7-11, 8 en particular.

¹⁴⁶ Collins y Lanza, 2010: 10.

¹⁴⁷ Collins y Lanza, 2010: 28 y 67-68.

¹⁴⁸ Biemer, 2004: 226-228 y 242-246 sobre todo; Collins y Lanza, 2010: 74-75.

¹⁴⁹ McCutcheon y Mills, 1998: 82.

¹⁵⁰ Vid. Collins y Lanza, 2010: 27, 45 y 47.

¹⁵¹ Ni que decir tiene que la familia de los modelos de variables latentes ofrece más opciones, incluyendo la identificación de variables latentes con control del error en los indicadores.

¹⁵² Biemer, 2004: 227.

sin error– ofrece varias conclusiones notables. Para comenzar deja claras las limitaciones de las mediciones de la punitividad basadas en una única cuestión o pregunta¹⁵³. No importa repetir una vez más que esto es, por otro lado, habitual en Criminología. Por supuesto, existen importantes excepciones. En una de ellas, Kury y sus asociados escriben lo que sigue: «Los resultados dejan clara la complejidad del constructo punitividad y que, por lo tanto, también es problemática su operacionalización con el ítem sobre la pena de muerte o con cualquier otro procedimiento de medición reducido»¹⁵⁴. Así que la familia de los análisis de variables latentes ofrece una estrategia claramente preferible. Pero adviértase que ello no obliga a renunciar a indicadores individuales. Nuestro análisis, sin ir más lejos, muestra que algunas variables manifiestas parecen bastante pobres, pero abre la puerta a la selección de otras que contengan menos error y, sobre todo, al perfeccionamiento y refinamiento de ítems y preguntas para la medición de la punitividad. Además, si somos capaces de replicar determinados hallazgos con operacionalizaciones y mediciones distintas de un mismo constructo –incluidas las basadas en un único ítem–, entonces nuestra confianza en dichos hallazgos se verá reforzada. Es lo que Braithwaite denomina *la concurrencia de las debilidades*¹⁵⁵. Verbigracia, a menudo los investigadores desean utilizar datos secundarios con muestras elevadas y una información relativamente rica, pero con mediciones pobres para constructos nucleares, como es el caso de la punitividad. Así, a modo de ejemplo, algunas encuestas del CIS incluyen una pregunta sobre la pena de muerte¹⁵⁶. Ya sabemos que ésta es una estimación pobre de la punitividad, pero, por un lado, análisis de clases latentes pueden ofrecer una aproximación al error de medición en que incurren; y, por otro, un conjunto de estudios empíricos independientes con medidas modestas pero distintas que arrojan unos mismos resultados, pueden ofrecer resultados verosímiles cuando coinciden.

Finalmente y a mayor abundamiento, los modelos de clases latentes tienen consecuencias mucho más profundas de lo que aquí hemos sugerido. Así, evocan la vieja idea de que, dada la gran complejidad de la definición y medición de actitudes como la punitividad –o la firmeza/benevolencia frente al delito–, una complejidad que el presente trabajo sólo ha reflejado de modo parcial, en la tradición de Lazarsfeld, las variables observables que actúan como indicadores y los conceptos que representan podrían encontrarse unidos por relaciones probabilísticas mucho más que deterministas¹⁵⁷.

¹⁵³ Cullen et al., 2009: 77; Keil y Vito, 1991: 455; Kury et al., 2002: 4 y 170; Kury et al., 2004: 97; Mayhew y van Kesteren, 2002: 67.

¹⁵⁴ Kury et al., 2002: 170.

¹⁵⁵ Braithwaite, 1979: 22.

¹⁵⁶ Vid. Serrano Maíllo y Kury, 2008: 328-329.

¹⁵⁷ Dayton, 1998: 1.

BIBLIOGRAFÍA

- Agresti, A. 2002. *Categorical data analysis*, 2.^a ed. Hoboken, NJ: John Wiley and sons.
- 2010. *Analysis of ordinal categorical data*, 2.^a ed. Hoboken, NJ: John Wiley and sons.
- Allison, P.D. 2002. *Missing data*. Thousand Oaks, Ca. [etc.]: Sage.
- Alwin, D.F. 1992. «Information transmission in the survey interview: number of response categories and the reliability of attitude measurement». *Sociological Methodology*, 22, 83-118.
- 2007. *Margins of error. A study of reliability in survey measurement*. Hoboken, NJ: John Wiley and sons.
- Anderson, C. y J.K. Vermunt. 2000. «Log-multiplicative association models as latent variable models for nominal and/or ordinal data». *Sociological Methodology*, 30, 81-122.
- Basilevsky, A. 1994. *Statistical factor analysis and related methods. Theory and applications*. New York [etc.]: John Wiley and sons.
- Biemer, P.P. 2004. «Modeling measurement error to identify flawed questions». Pp. 225-246 en *Methods for testing and evaluating survey questionnaires* (S. Presser et al. eds.). Hoboken, NJ: John Wiley and sons.
- 2011. *Latent class analysis of survey error*. Hoboken, NJ: John Wiley and sons.
- Blumstein, A., M. Tonry y A. van Ness. 2005. «Cross-national measures of punitiveness». *Crime and Justice*, 33 – *Crime and punishment in Western countries, 1980-1999*, 347-376.
- Böckenholt, U. 2002. «Comparison and choice. Analyzing discrete preference data by latent class scaling models». Pp. 163-182 en *Applied latent class analysis* (J.A. Hagenars y A.L. McCutcheon eds.). Cambridge [etc.]: Cambridge University Press.
- Bollen, K.A. 1989. *Structural equations with latent variables*. New York [etc.]: John Wiley and sons.
- Braithwaite, J. 1979. *Inequality, crime, and public policy*. London [etc.]: Routledge and Kegan Paul.
- Britt, C.L. 1994. «Versatility». Pp. 173-192 en *The generality of deviance* (T. Hirschi y M.R. Gottfredson eds.). New Brunswick, NJ y London: Transaction Publishers.
- Brown, E.K. 2006. «The dog that did not bark: punitive social views and the “professional middle classes”». *Punishment and Society*, 2006, 287-312.
- Clarke, H.D. y A.L. McCutcheon. 2009. «The dynamics of party identification reconsidered». *Public Opinion Quarterly*, 73, 704-728.
- Clogg, C.C. y W.D. Manning. 1996. «Assessing reliability of categorical measurements using latent class models». Pp. 169-182 en *Categorical variables in developmental research. Methods and analysis* (A. von Eye y C.C. Clogg eds.). San Diego, Ca. [etc.]: Academic Press.
- Collins, L.M. y S.T. Lanza. 2010. *Latent class and latent transition analysis with applications in the social, behavioral, and health sciences*. Hoboken, NJ: John Wiley and sons.
- Collins, L.M., P.L. Fidler y S.E. Wugalter. 1996. «Some practical issues related to the estimation of latent class and latent transition parameters». Pp. 133-146 en *Categorical variables in developmental research. Methods and analysis* (A. von Eye y C.C. Clogg eds.). San Diego, Ca. [etc.]: Academic Press.
- Croon, M. 2002. «Ordering the classes». Pp. 137-162 en *Applied latent class analysis* (J.A. Hagenars y A.L. McCutcheon eds.). Cambridge [etc.]: Cambridge University Press.

- Cullen, F.T., J.D. Unnever, K.R. Blevins, J.A. Pealer, S.A. Santana, B.S. Fisher y B.K. Applegate. 2009. «The myth of public support for capital punishment». Pp. 73-95 en *Public opinion and criminal justice* (J. Wood y T. Gannon eds.). Cullompton: Willan.
- Dayton, C.M. 1998. *Latent class scaling analysis*. Thousand Oaks, Ca. [etc.]: Sage.
- D'Unger, A.V., K.C. Land, P. McCall y D.S. Nagin. 1998. «How many latent classes of delinquent/criminal careers? Results from mixed poisson regression analysis of the London, Philadelphia, and Racine cohorts studies». *American Journal of Sociology*, 103, 1593-1630.
- Embretson, S.E. y S.P. Reise. 2000. *Item response theory for psychologists*. Mahwah, NJ y London: Lawrence Erlbaum Associates.
- Enders, C.K. 2006. «Analyzing structural equation models with missing data». Pp. 315-344 en *Structural equation modeling. A second course* (G.R. Hancock y R.O. Mueller eds.). Greenwich, CT: IAP.
- Finney, S.J. y C. DiStefano. 2006. «Nonnormal and categorical data in structural equation modeling». Pp. 269-314 en *Structural equation modeling. A second course* (G.R. Hancock y R.O. Mueller eds.). Greenwich, CT: IAP.
- Frost, N.A. 2006. *The punitive State. Crime, punishment, and imprisonment across the United States*. New York: LFB Scholarly Publishing.
- García España, E., F. Pérez Jiménez, M.J. Benítez Jiménez y A. Cerezo Domínguez. 2009. *Encuesta a víctimas en España* (J.L. Díez Ripollés y E. García España directores). Málaga: IAIC y Fundación Cajazol.
- Garland, D. 2001. *The culture of control. Crime and social order in contemporary society*. Chicago y Oxford: The University of Chicago Press.
- 2002. «Ideology and crime: a further chapter». Pp. 3-19 en *Ideology, crime and criminal justice. A symposium in honour of Sir Leon Radzinowicz* (A. Bottoms y M. Tonry eds.). Cullompton: Willan.
- Goodman, L.A. 2002. «Latent class analysis: the empirical study of latent types, latent variables, and latent structures». Pp. 3-55 en *Applied latent class analysis* (J.A. Hagenaars y A.L. McCutcheon eds.). Cambridge [etc.]: Cambridge University Press.
- Hagenaars, J.A. 1990. *Categorical longitudinal data: log-linear panel, trend, and cohort*. Newbury Park, Ca.: Sage.
- Harrendorf, S. 2011. «How to measure punitiveness in global perspective: what can be learned from international survey data». Pp. 125-148 en *Punitivity. International developments, 1 - Punitiveness – a global phenomenon?* (H. Kury y E. Shea eds.). Bochum: Universitätsverlag Dr. Brockmeyer.
- Van der Heijden, P.G.M., L.A. van der Ark y A. Mooijaart. 2002. «Some examples of latent budget analysis and its extensions». Pp. 107-136 en *Applied latent class analysis* (J.A. Hagenaars y A.L. McCutcheon eds.). Cambridge [etc.]: Cambridge University Press.
- Heinen, T. 1996. *Latent class and discrete latent trait models. Similarities and differences*. Thousand Oaks, Ca. [etc.]: Sage.
- Kaplan, D. 2009. *Structural equation modeling. Foundations and extensions*, 2.^a ed. Los Angeles, Ca.: Sage.
- Keil, T.J. y G.F. Vito. 1991. «Fear of crime and attitudes toward capital punishment: a structural equations model». *Justice Quarterly*, 8, 447-464.
- Kline, R.B. 2005. *Principles and practice of structural equation modeling*, 2.^a ed. New York y London: The Guilford Press.
- Kruttschnitt, C. y R. Macmillan. 2006. «The violent victimization of women. A life course

- perspective». Pp. 139-170 en *Gender and crime. Patterns of victimization and offending* (K. Heimer y C. Kruttschnitt eds.). New York y London: New York University Press.
- Kuhn, A. 1993. «Attitudes towards punishment». Pp. 271-288 en *Understanding crime experiences of crime and crime control* (A. Alvazzi del Frate et al. eds.). Roma: UNICRI.
- 2002. «Public and judicial attitudes to punishment in Switzerland». Pp. 115-127 en *Changing attitudes to punishment. Public opinion, crime and justice* (J.V. Roberts y M. Hough eds.). Cullompton: Willan Publishing.
- Kurki, L. 2001. «International standards for sentencing and punishment». Pp. 333-374 en *Sentencing and sanctions in Western countries* (M. Tonry y R.S. Frase eds.). Oxford y New York: Oxford University Press.
- Kury, H. y J. Obergfell-Fuchs. 2008. «Methodological problems in measuring attitudes to punishment (punitivity)». Pp. 277-302 en *Fear of crime – Punitivity. New developments in theory and research* (H. Kury ed.). Bochum: Universitätsverlag Dr. N. Brockmeyer.
- Kury, H. y E. Putkaradze. 2009. «Attitudes to punishment in Georgia. Results from a student survey». Ponencia presentada al Congreso Anual de la *European Society of Criminology*. Inédito.
- Kury, H., J. Obergfell-Fuchs y M. Würger. 2002. *Strafeinstellungen. Ein Vergleich zwischen Ost- und Westdeutschland*. Freiburg i.Br.: MPI.
- Kury, H., J. Obergfell-Fuchs y U. Smartt. 2002. «The evolution of public attitudes to punishment in Western and Eastern Europe». Pp. 93-114 en *Changing attitudes to punishment. Public opinion, crime and justice* (J.V. Roberts y M. Hough eds.). Cullompton: Willan Publishing.
- Kury, H., H. Kania y J. Obergfell-Fuchs. 2004. «Worüber sprechen wir, wenn wir über Punitivität sprechen? Versuch einer konzeptionellen und empirischen Begriffsbestimmung». *Kriminologisches Journal*, 36, 51-88.
- Kutateladze, B. 2011. «Measuring State punitiveness in the United States». Pp. 151-179 en *Punitivity. International developments, 1 - Punitiveness – a global phenomenon?* (H. Kury y E. Shea eds.). Bochum: Universitätsverlag Dr. Brockmeyer.
- Langeheime, R., J. Pannekoek y F. van de Pol. 1996. «Bootstrapping goodness-of-fit measures in categorical data analysis». *Sociological Methods and Research*, 24, 492-516.
- Mayhew, P. y J. van Kesteren. 2002. «Cross-national attitudes to punishment». Pp. 63-92 en *Changing attitudes to punishment. Public opinion, crime and justice* (J.V. Roberts y M. Hough eds.). Cullompton: Willan Publishing.
- McCutcheon, A.L. 1987. *Latent class analysis*. Newbury Park, Ca. [etc.]: Sage.
- 2002. «Basic concepts and procedures in single- and multiple-group latent class analysis». Pp. 56-85 en *Applied latent class analysis* (J.A. Hagenaars y A.L. McCutcheon eds.). Cambridge [etc.]: Cambridge University Press.
- McCutcheon, A.L. y J.A. Hagenaars. 1997. «Comparative social research with multi-sample latent class models». Pp. 266-277 en *Applications of latent trait and latent class models in the social sciences* (J. Rost y R. Langeheine eds.). Münster [etc.]: Waxman.
- McCutcheon, A.L. y C. Mills. 1998. «Categorical data analysis: log-linear and latent class models». Pp. 71-94 en *Research strategies in the social sciences. A guide to new approaches* (E. Scarbrough y E. Tanenbaum eds.). Oxford: Oxford University Press.
- Long, J.S. 1997. *Regression models for categorical and limited dependent variables*. Thousand Oaks, Ca. [etc.]: Sage.

- Powers, D.A. y Y. Xie. 2000. *Statistical methods for categorical data analysis*. San Diego, Ca. [etc.]: Academic Press.
- Rabe-Hesketh, S. y A. Skrondal. 2001. «Parameterization of multivariate random effects models for categorical data». *Biometrics*, 57, 1256-1264.
- Rabe-Hesketh, S., A. Skrondal y A. Pickles. 2004a. *GLLAMM Manual*. Berkeley, Ca.: University of California. Accesible en <<http://www.bepress.com/ucbbiostat/paper160>>.
- 2004b. «Generalized multilevel structural equation modeling». *Psychometrika*, 69, 167-190.
- Reeve, B.B. y L.C. Mâsse. 2004. «Item response theory modeling of questionnaire evaluation». Pp. 247-273 en *Methods for testing and evaluating survey questionnaires* (S. Presser et al. eds.). Hoboken, NJ: John Wiley and sons.
- Roberts, C. 2010. «Measuring social attitudes». Pp. 194-214 en *Social measurement through social surveys. An applied approach* (M. Bulmer et al. eds.). Farnham y Burlington, VT: Ashgate.
- Rudas, T. 2002. «A latent class approach to measuring the fit of a statistical model». Pp. 345-365 en *Applied latent class analysis* (J.A. Hagenaars y A.L. McCutcheon eds.). Cambridge [etc.]: Cambridge University Press.
- Sacco, P., K.K. Bucholz, E.L. Spitznagel. 2009. «Alcohol use among older adults in the National Epidemiologic Survey on alcohol and related conditions: a latent class analysis». *Journal of Studies on Alcohol and Drugs*, 70, 829-838.
- Serrano Maíllo, A. 2006. «Punitivität und Gesetzgebung: Die Situation in Spanien». Pp. 245-252 en *Kriminalität und Kriminalprävention in Ländern des Umbruchs* (H. Kury y E. Karimov eds.). Bochum: Universitätsverlag Dr. N. Brockmeyer.
- Serrano Maíllo, A. y H. Kury. 2008. «Insecurity feelings and punitivity: relationship in a national sample of adolescents and young adults in Spain». Pp. 321-348 en *Fear of crime – Punitivity. New developments in theory and research* (H. Kury ed.). Bochum: Universitätsverlag Dr. N. Brockmeyer.
- Serrano Maíllo, A. y A. Serrano Gómez. 2009. «El aumento de la punitividad a nivel legislativo en España». Pp. 293-316 en *Estudios sobre punitividad y victimación* (H. Kury y A. Serrano Maíllo eds.). Madrid: Dykinson.
- Skrondal, A. y S. Rabe-Hesketh. 2004. *Generalized latent variable modeling. Multilevel, longitudinal, and structural equation models*. Boca Raton, FL [etc.]: Chapman and Hall/CRC.
- 2007. «Latent variable modelling: a survey». *Scandinavian Journal of Statistics*, 34, 712-745.
- Stalans, L.J. 2002. «Measuring attitudes to sentencing». Pp. 15-32 en *Changing attitudes to punishment. Public opinion, crime and justice* (J.V. Roberts y M. Hough eds.). Cullompton: Willan Publishing.
- Timberlake, T.S. 2008. «A latent class analysis of nicotine-dependence criteria and use of alternate tobacco». *Journal of Studies on Alcohol and Drugs*, 69, 709-717.
- Tonry, M. y D.P. Farrington. 2005. «Punishment and crime across space and time». *Crime and Justice*, 33 – *Crime and punishment in Western countries, 1980-1999*, 1-39.
- Uebersax, J. 1993. «Dimension reduction and latent class analysis. A simple method for interpretation of latent class analysis parameters, with possible implications for factor analysis of dichotomous and ordered-category measures». Ponencia presentada al Congreso Anual de la *Classification Society of North America*. Inédito.
- 1997. «Analysis of student problem behaviors with latent trait, latent class, and related probit mixture models». Pp. 188-195 *Applications of latent trait and latent*

- class models in the social Sciences* (J. Rost y R. Langeheine eds.). New York: Waxman.
- 1999. «Probit latent class analysis with dichotomous or ordered category measures: conditional independence/dependence models». *Applied Psychological Measurement*, 23, 283-297.
- 2009. *A practical guide to conditional dependence in latent class analysis*, última versión. Accesible en <<http://www.john-uebersax.com/stat/condep.htm>>.
- Uebersax, J. y W.M. Grove. 1990. «Latent class analysis of diagnostic agreement». *Statistics in Medicine*, 9, 559-572.
- Varona Gómez, D. 2008. «Ciudadanos y actitudes punitivas: un estudio piloto de población universitaria española». *Revista Española de Investigación Criminológica*, 6-1, 1-38.
- Vaughn, M.G., M. DeLisi, K.M. Beaver y J.P. Wright. 2009. «Identifying latent classes of behavioral risk based on early childhood manifestations of self-control». *Youth Violence and Juvenile Justice*, 7, 16-31.
- Vermunt, J.K. 1997a. *Log-linear models for event histories*. Thousand Oaks, Ca. [etc.]: Sage.
- 1997b. *LEM: a general program for the analysis of categorical data*. Tilburg: Department of Methodology and Statistics, Tilburg University.
- 1999. «A general class of non-parametric models for ordinal categorical data». *Sociological Methodology*, 29, 187-223.
- 2003. «Applications of latent class analysis in social science research». *Lecture Notes in Artificial Intelligence*, 2711, 22-36.
- Vermunt, J.K. y J. Magidson. 2002. «Latent class cluster analysis». Pp. 89-106 en *Applied latent class analysis* (J.A. Hagenars y A.L. McCutcheon eds.). Cambridge [etc.]: Cambridge University Press.
- 2003. «Latent class models for classification». *Computational Statistics and Data Analysis*, 41, 531-537.
- 2007. «Latent class analysis with sampling weights. A maximum likelihood approach». *Sociological Methods and Research*, 36, 87-111.
- Vermunt, J.K., J.R. van Ginkel, L.A. van der Ark y K. Sijtsma. 2008. «Multiple imputation of incomplete categorical data using latent class analysis». *Sociological Methodology*, 33, 369-297.
- Viki, G.T. y G. Bohner. 2009. «Achieving accurate assessment of attitudes toward the criminal justice system: methodological issues». Pp. 96-119 en *Public opinion and criminal justice* (J. Wood y T. Gannon eds.). Cullompton: Willan.
- Winship, C., R.D. Mare y J.R. Warren. 2002. «Latent class models for contingency tables with missing data». Pp. 408-432 en *Applied latent class analysis* (J.A. Hagenars y A.L. McCutcheon eds.). Cambridge [etc.]: Cambridge University Press.
- Yang, X., J. Shaftel, D. Glasnapp y J. Poggio. 2005. «Qualitative or quantitative differences? Latent class analysis of mathematical ability for special education students». *Journal of Special Education*, 38, 194-207.
- Zugaldía Espinar, J.M. 2006. «Contrarreforma penal (el annus horribilis de 2003) y el anteproyecto de reforma del código penal de 2006». Pp. 1347-1382 en *Derecho penal y Criminología como fundamento de la Política criminal. Estudios en homenaje al Profesor Alfonso Serrano Gómez* (F. Bueno et al. directores). Madrid: Dykinson.