

Reflexiones sobre el análisis a priori de los cuestionarios basado en técnicas del Análisis Estadístico Implicativo

Pablo Gregori*, Pilar Orús**, Irene Pitarch***

Departament de Matemàtiques, Universitat Jaume I de Castelló, E-12071 Castellón (España)

*gregori@mat.uji.es

<http://www3.uji.es/~gregori>

**orus@mat.uji.es

<http://www3.uji.es/~orus>

***ipitarch@mat.uji.es

Resumen. El objetivo de esta comunicación es profundizar en la utilización del Análisis Estadístico Implicativo, en los análisis de cuestionarios que consideran como “datos” tanto los obtenidos experimentalmente como los obtenidos en los estudios a priori. Para ello se retomará el cuestionario elaborado por Pitarch (2002), su análisis a priori y sus datos empíricos. Con dicho cuestionario se pretendía caracterizar el trabajo lógico-matemático y los razonamientos de los alumnos de la ESO (Enseñanza Secundaria Obligatoria) al realizar el tratamiento de datos de una tabla booleana de datos primarios.

1 Introducción

Cuando el investigador (en Didáctica de las Matemáticas) diseña un cuestionario, intenta materializar en un conjunto de preguntas (ítems) la problemática que quiere estudiar, seleccionando un conjunto de aspectos de esta problemática que serán abordados a través de las diferentes preguntas. Brousseau y colaboradores (Brousseau y Lacasta, 1995, Orús, 1986, 1992, etc.), hacen especial hincapié en este proceso de elaboración y caracterización del cuestionario -el análisis a priori- introduciendo a nivel metodológico el análisis multivariante de los datos que proporcionan las matrices a priori (características \times cuestiones) del cuestionario; estos análisis permiten poner de manifiesto relaciones a priori entre las variables (las preguntas del cuestionario) en función de sus características y contrastarlas con el resultado de la experimentación: la muestra recogida establece unas relaciones entre los ítems del cuestionario que pueden ser, o no, similares a las esperadas por el investigador y constituir, así pues, una fuente de información, a posteriori, que contraste el conocimiento anterior.

Para cada característica seleccionada, se puede interpretar que cada ítem del cuestionario posee, o no posee, la citada característica. Así se genera una matriz a priori booleana, y si el cuestionario genera también una tabla de datos binarios, cada característica se puede interpretar como un “individuo ficticio”, si consideramos conjuntamente los resultados a priori y a posteriori del cuestionario. Brousseau y Lacasta (1995) ya han mostrado esta utilización de las características de matrices a priori como “alumnos suplementarios” de la experimentación, en análisis multivariantes como el análisis factorial de correspondencias, y el análisis en componentes principales; también Orús y Gregori (2005) han utilizado la comparación de la coherencia de la estructura del cuestionario entre la matriz de datos reales y la matriz ampliada con los datos ficticios en el Análisis Estadístico Implicativo (ASI): si dicha comparación resulta positiva, la aparición de los individuos ficticios en los grupos óptimos para la tipicidad y contribución de individuos a la formación de clases o implicaciones, puede suponer una fuente de información que actualice los conocimientos a priori del investigador.

Otras contribuciones sobre la utilización del ASI en los análisis de datos que consideran la matriz a priori, se han basado en la comparación de los resultados gráficos, de los grafos de la clasificación, implicación y cohesión resultantes de los análisis con CHIC de los datos obtenidos de la experimentación (a posteriori) con los “datos ficticios” del experimentador (a priori). En esta opción, se ha comparado la coherencia de la estructura del cuestionario (los resultados del ASI), entre la matriz de datos reales y la matriz de “datos ficticios” (Orús y Pitarch, 2000, Pitarch, 2002, Pitarch y Orús, 2005).

Esta comparación directa de resultados gráficos es intuitiva para el usuario de ASI y puede resultar suficiente para interpretar mejor los resultados de los datos empíricos, considerando la matriz a priori como matriz explicativa de éstos. Pero puede ser complementada y medida, por los resultados numéricos (Spagnolo, 1997 y Spagnolo y Gras, 2007).

El objetivo de esta comunicación es profundizar en esta línea de utilización del ASI, en los análisis de cuestionarios que consideran “datos” tanto los obtenidos experimentalmente como los obtenidos en los estudios a priori.

Para ello se retomará el cuestionario elaborado por Pitarch (2002), su análisis a priori y sus datos empíricos. Con dicho cuestionario se pretendía recoger información y caracterizar el trabajo de los alumnos de la ESO (Enseñanza Secundaria Obligatoria) al realizar el tratamiento de una tabla de datos primarios.

1.1 Metodología utilizada

Queremos destacar dos aspectos diferenciados de nuestro trabajo en torno al cuestionario que denotaremos por Q (ver Anexos 1, 2 y 3) la aplicación del marco teórico de la Teoría de Situaciones (Brousseau, 1986) - particularmente las nociones de "medio", "situación a-didáctica" e "ingeniería didáctica" para la elaboración y diseño de todo el cuestionario y del estudio a priori realizado (Pitarch y Orús, 2005) y por otro lado la utilización del CHIC, que nos ha permitido utilizar el ASI, para comparar los análisis a priori y a posteriori del cuestionario, es decir su caracterización y su experimentación; aspecto que centrará esta comunicación.

Los diferentes análisis del software estadístico CHIC se han utilizado para:

- Analizar, a priori, el cuestionario diseñado. El estudio teórico de Q se plasmará en una matriz a priori del cuestionario (MAP). Esta matriz representa la caracterización de las preguntas del cuestionario según los diversos contenidos matemáticos que conllevan dichas preguntas. El análisis de similitudes de la matriz MAP, nos permitirá verificar la coherencia de los criterios elegidos para caracterizar a priori las cuestiones de Q. Si las clases resultantes nos ofrecen una estructura coherente del cuestionario, ésta se utilizará para continuar el análisis a priori, estableciendo las implicaciones que deberían verificarse en el interior de cada clase.
- Analizar, a posteriori, las respuestas de los alumnos al cuestionario. Se utiliza CHIC sobre la matriz de datos (MP) que representan las respuestas (correcta/incorrecta) de los estudiantes. De esta forma se realiza el estudio a posteriori del cuestionario, mediante técnicas de ASI. Los grafos y árboles que se obtienen permiten analizar los resultados de los estudiantes en las diferentes cuestiones con respecto a su trabajo matemático, al utilizar una tabla de valores para contestar este cuestionario; CHIC nos ofrece la estructura global del cuestionario, en función de las respuestas de los alumnos y del tipo de análisis realizado.
- Realizar un análisis comparativo entre los datos a priori y los resultados a posteriori, con el objetivo de obtener conclusiones del trabajo global. Se compararán las estructuras de Q resultantes en ambos análisis, para ver en qué medida los resultados de la experimentación, pueden explicarse en función de los criterios de las cuestiones, establecidos a priori. Se analizarán las implicaciones a posteriori entre las cuestiones, obtenidas en la experimentación, comparándolas con las implicaciones establecidas a priori entre las cuestiones, en función de sus características y conocimientos matemáticos.

La utilización del ASI y de CHIC ya ha sido presentada en otras investigaciones en el campo de la Didáctica de las Matemáticas, aportando algunos resultados, entre otros, para "medir" y modelizar situaciones didácticas a partir de la clasificación y la implicación (Bodin, 1996, Gras, 1995, Gras et al., 1996, Lerman, 1981, Lerman et al., 1981); también para facilitar la toma de decisiones didácticas en la elaboración de una ingeniería didáctica (Orús y Pitarch, 2000).

Nuestra comunicación es una nueva aportación en este campo, mostrando la potencialidad del ASI y del software CHIC (Couturier, 2000) para el análisis didáctico de un cuestionario Q, utilizando el estudio a priori de éste, previo a la experimentación.

2 Análisis a priori del cuestionario

2.1 El cuestionario Q: sobre gustos musicales

El objetivo principal del cuestionario Q (Pitarch, 2002) que vamos a analizar, era poder observar y cuantificar diversos aspectos del tratamiento de datos multivariantes en los alumnos de enseñanza secundaria, en España, a partir de una tabla con una cantidad significativa de valores (335 valores, correspondientes a las 300 respuestas dadas por 20 sujetos, sobre 15 variables).

El cuestionario Q trata sobre los resultados a una encuesta realizada a jóvenes de entre 15 y 20 años, referente a sus gustos musicales respecto a diversos tipos de música (ver Anexos 1, 2 y 3).

2.2 El cuestionario Q: adecuación del cuestionario "Pitufos"

El punto de partida de este trabajo es un cuestionario similar (Orús, 1992) elaborado para alumnos de primaria en Francia: el cuestionario de los "Pitufos" utilizado por P. Orús en su tesis (Orús, 1992, pp.259-293) para modelizar diversos aspectos del razonamiento que puede movilizar el tratamiento de datos booleanos:

lógico, numérico y clasificatorio, fundamentalmente. Este cuestionario se ha adaptado para poder ser respondido por los alumnos de Enseñanza Secundaria Obligatoria (ESO) de España, pero sigue constando de:

- La Consigna: Instrucciones que se les entrega a los alumnos de la experimentación para presentarles la situación de trabajo y enmarcar las preguntas que deben contestar.
- La Tabla de datos (T): Matriz booleana (con valores binarios 0 y 1) que representan las respuestas de los jóvenes entrevistados sobre sus gustos musicales: las columnas de la tabla representan a 20 jóvenes (de edades 15-20 años) que han sido entrevistados para obtener información sobre el gusto musical que tiene cada uno de ellos, y están codificados con letras mayúsculas, joven A, joven B, joven C, etc.; las filas de la tabla representan los 15 tipos de música diferente sobre los que se ha pedido opinión a los jóvenes, por ejemplo la fila 3 corresponde al criterio "música máquina".
- El cuestionario propiamente dicho, formado por todas las cuestiones que se les plantean a los alumnos de la experimentación. Dichas cuestiones hacen referencia a la encuesta que han respondido los jóvenes, por lo tanto los alumnos de la experimentación responden a estas preguntas del cuestionario utilizando la tabla de datos como soporte de trabajo.

Estos tres elementos de trabajo se les entrega a cada uno de los alumnos de la experimentación como material impreso tal y como se presentan en los Anexos 1—3.

La tabla T mantiene las mismas dimensiones de 15 filas por 20 columnas, por considerar que:

- Aunque los alumnos de los últimos cursos de la enseñanza primaria correspondientes a la enseñanza francesa, que utilizaron esta matriz en la experimentación realizada en Orús (1992, p.297), tuvieron un éxito en la lectura y manejo básico de la tabla correspondiente a más de un 81%, sin embargo esta tabla nos permitió observar otras dificultades en el tratamiento de datos, ligadas al tipo de trabajo propuesto.
- Los alumnos de la etapa educativa de la ESO no han trabajado habitualmente con tablas de valores con dimensiones tan grandes como la tabla (T) y, por tanto, los alumnos de esta etapa educativa deberán también trabajar este tipo de tablas, como mínimo con el mismo nivel de éxito, aunque no necesariamente haber superado las dificultades lógico-matemáticas observadas en enseñanza primaria, al no haber tenido un aprendizaje específico sobre el tema.

Identificar y cuantificar esas dificultades es nuestro objetivo.

2.3 Estudio a priori del cuestionario

El análisis a priori de nuestro instrumento didáctico fue fundamental para nuestro trabajo, puesto que éste nos iba a permitir identificar el trabajo matemático que realmente conlleva el cuestionario, bajo distintos puntos de vista matemáticos (la lógica, la teoría de conjuntos y la estadística). En los trabajos de Pitarch (2002) y Pitarch y Orús (2005) se detalla este análisis global del cuestionario y de cada cuestión.

Una primera identificación se lleva a cabo caracterizando las preguntas del cuestionario según las operaciones matemáticas que éstas suponen, bajo tres grandes grupos de operaciones: la manipulación de la tabla (sobre elementos constitutivos de la propia tabla y también mediante la comparación entre sus filas y columnas), las operaciones lógico-matemáticas y estadísticas, y por último los razonamientos que se exigen los alumnos para que justifiquen sus respuestas.

Por último, a partir del estudio a priori del cuestionario también se analizan las distintas relaciones (proximidad-distancia, implicación estadística, implicación jerárquica,...) que existen entre las preguntas del cuestionario, según la caracterización matemática realizada anteriormente, después de aplicar el programa informático C.H.I.C.

2.3.1 Análisis didáctico: Modelización del cuestionario como una situación a-didáctica.

El cuestionario de nuestra experimentación está enmarcado en la Teoría de Situaciones Didácticas de G. Brousseau (Brousseau, 1998) y por tanto está diseñado según la modelización que permite este marco teórico:

- La propia utilización de la tabla y del cuestionario, constituye una situación a-didáctica específica de los conocimientos matemáticos: tratamiento de datos, lógica y estadística elemental. La utilización del tratamiento de datos implica a su vez trabajar con el análisis multivariante, la estadística descriptiva y el análisis tipológico. Y la utilización de la lógica exige un trabajo con los predicados. En resumen, tanto la utilización del tratamiento de datos como el de la lógica conllevan un trabajo con la teoría de conjuntos.
- Las preguntas del cuestionario son las variables didácticas que permiten la gestión del enseñante sobre el trabajo del alumno. Estas variables están establecidas a priori y, por tanto, permiten modelizar el trabajo

lógico y estadístico a través del análisis a priori de las propias cuestiones realizadas a los alumnos. Como consecuencia, el cuestionario será el "medio a-didáctico" al que el alumno es confrontado.

El trabajo matemático propuesto al alumno en cada pregunta del cuestionario no depende de la intervención del profesor, sino únicamente de la naturaleza de la pregunta; aunque éstas son controladas y controlables por el enseñante.

Mediante este estudio a priori del cuestionario, el profesor puede modificar las estrategias de juego de los alumnos, permitiendo el aprendizaje del conocimiento matemático en juego en la situación a-didáctica.

A continuación exponemos un ejemplo para poder explicar cuál es el valor asignado a la variable didáctica: la pregunta "controlada" y "establecida a priori" y qué otras posibilidades se pueden aplicar para ir modificando las estrategias de juego que permitan el aprendizaje de un conocimiento matemático concreto.

Ejemplo: Q16. ¿En cuántas respuestas de la encuesta están de acuerdo los jóvenes C y J? (ésta es la pregunta nº 16 que se les formula en el cuestionario a los alumnos de la experimentación):

La modelización del trabajo a realizar para responder Q16 consiste en comparar las 20-tuplas de las columnas 3 (Joven C) y 10 (Joven J), para encontrar las coincidencias que existen en las respuestas de los jóvenes C y J: las copresencias (1,1) y las coausencias (0,0).

Otros valores posibles de la variable (es decir de la pregunta del cuestionario) podrían ser las siguientes variantes de la pregunta

Variante 1: ¿A los jóvenes C y J, qué tipo de música les gusta a ambos? Si quisiéramos que sólo trabajaran las copresencias (1,1).

Variante 2: ¿Qué tipo de música no le gusta a ninguno de los dos jóvenes, C y J? Si quisiéramos que sólo trabajaran las Coausencias (0,0).

Variante 3: ¿Cuáles son las discrepancias musicales, entre los jóvenes? Si quisiéramos que trabajaran sobre las diferencias, (0,1) y (1,0)

Variante 4: ¿En cuántas respuestas no están de acuerdo los jóvenes C y J? Si quisiéramos que trabajaran sobre las diferencias, (0,1) y (1,0)

2.3.2 La matriz a priori MAP.

El análisis a priori del cuestionario nos permite obtener la caracterización de cada una de las cuestiones, definiendo distintos criterios, según los tipos de operaciones que deben realizar los alumnos al responder al cuestionario.

Los tres grandes grupos de operaciones que hemos retenido: la manipulación de la tabla (sobre elementos constitutivos de la propia tabla y también mediante la comparación entre sus filas y columnas), las operaciones lógico-matemáticas y estadísticas y, por último, los razonamientos que se exigen los alumnos para que justifiquen sus respuestas, se pueden detallar con más precisión, a través de los criterios siguientes¹:

- Trabajo sobre la tabla (T),
- recuentos obligatorios o alternativos (R),
- comparación entre filas y/o columnas (C),
- diferencias y coincidencias positivas y/o negativas (copresencias, coausencias,...) (F),
- operaciones lógicas (implicación, conectores lógicos, cuantificadores) (L),
- operaciones básicas de estadística (frecuencias, moda...) (E)

Esta identificación de las operaciones que aportan las cuestiones indicadas, permite visualizar el tratamiento simultáneo de los datos, de la lógica y de la estadística descriptiva a partir de la tabla de datos T y del cuestionario Q.

La caracterización de dichas cuestiones se traduce en una matriz booleana que identificamos como la matriz a priori del cuestionario Q: MAP.

En esta matriz las columnas son claramente cada una de las preguntas del cuestionario (en este caso serán los sujetos MAP). En las filas aparecen los criterios de la matriz, todos ellos codificados con cuatro iniciales: la primera indica el bloque de operaciones que se identifica en la pregunta y las tres siguientes intentan clarificar la operación concreta que se realiza (por ejemplo, ROBL: significa que se debe realizar un recuento obligatorio. CSEM: realizar una comparación con interpretación semántica entre filas o columnas. LEXI: utilización de un cuantificador lógico existencial. EFEX: utilización de la frecuencia absoluta de forma explícita. EMIM: utilización del parámetro estadístico de centralización MODA, de forma implícita).

¹ La letra mayúscula indicada entre paréntesis representa la codificación del bloque de operaciones que se ha identificado en dicha pregunta del cuestionario.

	Q1	Q2	Q3	Q4	Q4.1	Q5	Q6	Q7	Q7.1	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21
TABL	1	1	1	0	0	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
TFIL	1	1	0	1	1	1	0	1	1	0	1	1	1	0	0	0	1	0	1	1	1	0	1
TCOL	1	0	1	0	0	0	1	0	0	1	0	0	0	1	1	1	0	1	0	0	0	1	0
ROBL	0	1	1	0	0	0	0	0	0	0	1	1	1	0	1	0	1	1	0	0	1	1	1
RALT	0	0	0	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0
CFIL	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	1	0	1	1	1	0	1
CCOL	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	1	0	1	0	0	0	1	0
FDIF	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0
FCOI	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0
FPRE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0
FAUS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
CSEM	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	1
LIMP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
LEXI	0	0	0	1	1	0	1	1	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0
LUNI	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
LCON	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	1	1	0	0	1	0	0
LDIS	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
LNEG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
EFIM	0	1	1	1	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	1
EFEX	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
Efre	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
EPER	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
EMIM	0	0	0	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ARGU	0	0	0	0	1	0	0	0	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1

TAB 1 – Matriz a priori de Q (MAP).

2.3.3 Clasificación de las cuestiones (a priori).

La aplicación del análisis de similitud sobre la matriz MAP –mediante el programa CHIC– nos proporciona claramente cuatro clases de cuasi-equivalencia de las preguntas del cuestionario (ver Fig. 1), en función de los criterios establecidos en la matriz a priori (MAP)

- $C_1 = \{Q1, Q2, Q10, Q9, Q11, Q3\} \equiv$ trabajo sobre la tabla y estadística
- $C_2 = \{Q15, Q19, Q17, Q18\} \equiv$ operaciones lógicas
- $C_3 = \{Q4, Q4.1, Q6, Q7, Q7.1\} \equiv$ recuentos alternativos y cuantificadores
- $C_4 = \{Q8, Q16, Q20, Q12, Q13, Q14\} \equiv$ proximidad y distancias
- $C_5 = \{Q5, Q21\} \equiv$ comparación de cardinales de variables

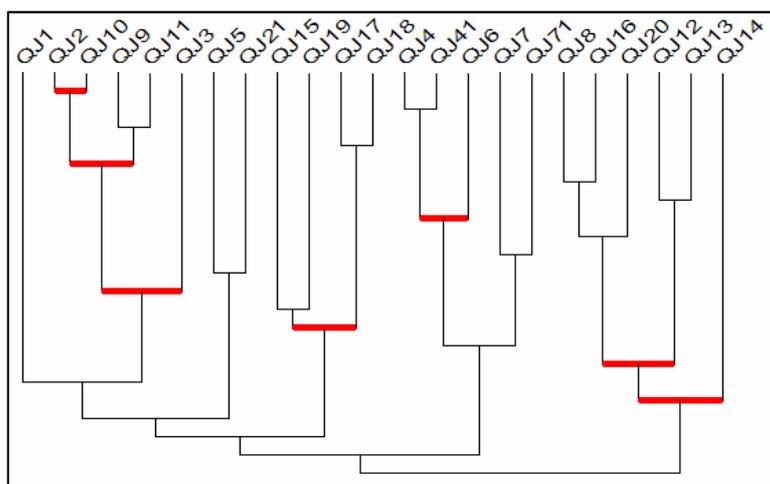


FIG. 1 – Árbol de similitudes de la MAP

2.3.4 Implicaciones a-priori.

Teniendo en cuenta los criterios atribuidos a priori a las cuestiones y la clasificación que nos ofrece la Fig. 1, vamos a analizar al interior de cada clase cuales son las implicaciones que se pueden considerar a priori entre las cuestiones que la forman. Hemos distinguido dos tipos de implicaciones diferentes que representaremos con los símbolos lógicos de *implicación* (\Rightarrow) y de *condicional* (\rightarrow), que podrían identificarse unas, como implicaciones fuertes o necesarias (\Rightarrow) entre esas cuestiones en función de los conocimientos matemáticos escolares en juego en dichas cuestiones y las otras, como relaciones que consideramos muy probables (\rightarrow) entre las cuestiones, en función de las estrategias de resolución o de conocimientos “próximos”, o incluso entre conocimientos que se impliquen a nivel matemático, pero que escolarmente no se trabaje esa relación.

Es decir, atribuiremos una relación de implicación (\Rightarrow) entre las cuestiones, si existe esta relación de implicación lógica entre los conocimientos matemáticos que caracterizan dichas cuestiones; es decir si estos conocimientos pueden ser considerados como condiciones necesarias (CN) o suficientes (CS) entre ellos. Por ejemplo diremos que $Q11 \Rightarrow Q10$, ya que el conocimiento “frecuencia relativa” que caracteriza Q11, se puede suponer como CS para el conocimiento “frecuencia absoluta” que exige Q10, o bien que este conocimiento de Q10, es CN para el conocimiento de Q11.

Mientras que atribuiremos una implicación de tipo condicional (\rightarrow) a la relación entre las cuestiones Q11 (frecuencia relativa) y Q9 (porcentaje): matemáticamente sabemos el porcentaje es una de las representaciones de la frecuencia relativa de una variable; pero en la enseñanza obligatoria ambos conceptos no se aprenden “necesariamente relacionados”. Por tanto aunque matemáticamente consideraríamos $Q11 \Rightarrow Q9$, a nivel didáctico, consideramos que se puede verificar $Q11 \rightarrow Q9$ (pero no necesariamente).

A continuación detallamos las relaciones de implicación entre las cuestiones atribuidas a priori, en función de los diversos conocimientos lógico-matemáticos del cuestionario.

i) Cuestiones sobre conceptos estadísticos

Observando el árbol de similitud de la MAP, encontramos estas cuestiones sobre conocimientos estadísticos en la clase $C_1 = \{Q1, Q2, Q3, Q10, Q9, Q11\}$, conjuntamente con las cuestiones básicas sobre el funcionamiento de la tabla, todas ellas con un marcado carácter de “actividad enumerativa”.

Las cuestiones identificadas a priori como cuestiones específicas sobre conceptos estadísticos, son Q9, Q10 y Q11, al ser las cuestiones que hacen referencia explícita a la terminología estadística (frecuencias, variable estadística, porcentaje); no obstante diferenciamos la pregunta Q9 sobre porcentajes, de las otras dos preguntas sobre frecuencias, por considerar que el término “porcentaje” está más incorporado al lenguaje natural y, por tanto, menos específico como termino estadístico, determinando así las siguientes implicaciones a priori entre estas cuestiones, Implicaciones que traducen las relaciones de los conocimientos matemáticos en juego:

$$\begin{array}{c} Q11 \Rightarrow Q10 \\ \downarrow \\ Q9 \end{array}$$

El conocimiento del concepto de frecuencia absoluta es condición necesaria para la definición de frecuencia relativa, y el conocimiento de la representación porcentual de esta frecuencia relativa, se puede suponer como una consecuencia derivada del dominio de dicho conocimiento, pero no necesariamente: el alumno puede conocer el concepto de frecuencia relativa como cociente entre la frecuencia absoluta de la variable y el cardinal de la población y desconocer la representación porcentual de esta frecuencia.

Si añadimos a estas cuestiones, las cuestiones relacionadas con ellas pero que no exigen conocimientos específicos de estadística para responderlas, sino que la interpretación semántica de la tabla permite responderlas con éxito, podríamos encontrar la siguiente cadena de implicaciones relacionadas:

$$\begin{array}{ccc} Q11 & \rightarrow & Q9 \\ \downarrow & & \downarrow \\ Q10 & & Q3 \\ \downarrow & & \downarrow \\ Q2 & \rightarrow & Q1 \end{array}$$

Teniendo en cuenta que Q1, Q2 y Q3 son las preguntas básicas que establecen el conocimiento mínimo de los alumnos sobre el sentido de una tabla de datos (lo que representa una fila, una columna y una casilla o celda, en el conjunto de los datos representados), se espera que sean respondidas correctamente prácticamente por todos los alumnos, ya que en caso contrario el resto del cuestionario carecería de sentido. El carácter básico de estas cuestiones hace previsible que puedan aparecer como consecuentes, al final de las diferentes cadenas de implicaciones posibles entre todas las cuestiones de Q. Un ejemplo es el que acabamos de presentar con las implicaciones entre las cuestiones sobre estadística.

ii) Cuestiones sobre operaciones lógicas

$$C_2 = \{Q15, Q19, Q17, Q18\} \text{ y } C_5 = \{Q5, Q21\}$$

Las cuestiones que aparecen en esta segunda clase C_2 están identificadas en la matriz a priori como cuestiones que requieren utilizar operaciones lógicas.

Por otra parte, las preguntas Q5 y Q21 hacen referencia a disyunciones entre dos y tres tipos de música, respectivamente, pero además, en ambas interviene un recuento alternativo o cardinal de los valores positivos de algunas filas (dos o tres) de la tabla del cuestionario, para poder comparar dichos recuentos y dar respuesta a estas cuestiones. Vemos que la clase de estas preguntas se forma entre las clases C_1 y C_2 y aparece aislada, no perteneciendo a ningún nodo o nivel significativo, y por ello también la hemos considerado en nuestro análisis como una clase diferenciada de las otras cuatro clases que marcan los niveles significativos.

En este caso consideramos la implicación a priori entre estas dos cuestiones como sigue: $Q21 \Rightarrow Q5$, ya que en la pregunta Q21 intervienen tres criterios distintos y, en cambio, en Q5 sólo aparece la disyunción entre dos criterios, y en la cuestión Q21 también se exige el razonamiento de la respuesta

Las cuestiones que aparecen en la clase C_2 son: Q15, que requiere el cardinal de la conjunción de dos criterios, Q19, que involucra el cardinal de la conjunción de negaciones, y Q17 y Q18, en las que se trabaja la implicación lógica (aunque como contenido matemático queda reflejado en la pregunta Q18).

Estas cuatro últimas cuestiones exigen todas ellas una comparación entre las celdas (1,1) de las filas implicadas.

Si además tenemos en cuenta que los contenidos matemáticos más complejos de tratar por los alumnos son la implicación lógica y la conjunción de negaciones -ya que el alumnado está más familiarizado con las preguntas formuladas en positivo (Q15)-, podemos establecer la siguiente cadena de implicaciones:

$$\begin{array}{ccc} Q18 & \rightarrow & Q19 \\ \downarrow & & \downarrow \\ Q17 & \rightarrow & Q15 \end{array}$$

Análisis a priori de cuestionarios basado en técnicas ASI

Por lo tanto suponemos, a priori, que si el alumno tiene asimilada la implicación lógica (como contenido matemático más fuerte) correctamente, entonces puede realizar con éxito el cardinal de la conjunción de negaciones y por tanto el cardinal de una conjunción.

iii) Cuestiones sobre cuantificadores existenciales y universales

$$C_3 = \{Q4, Q4.1, Q6, Q7, Q7.1\}$$

Esta clase está compuesta por cuestiones que, según la matriz a priori MAP, están caracterizadas por el cuantificador existencial y/o universal, a la vez, que pueden requerir un recuento alternativo por parte del estudiante.

Tanto en la pregunta Q4, como en la Q6 se le exige al alumno que realice la búsqueda de un modelo « homogéneo », en concreto de una fila y una columna, respectivamente, en la que todas las celdas tengan el valor 1. Por lo tanto establecemos a priori, la relación entre ellas con la siguiente bicondicional: $Q4 \leftrightarrow Q6$

En cambio, en la cuestión Q7 se exige una comparación entre criterios mediante un modelo variable, ya que estableciendo un criterio (tipo de música) previo, se induce a la búsqueda de otros criterios con mayor cardinal que el prefijado. Por ello, suponemos que el alumno que tenga éxito en este tipo de estrategia también lo tendrá en la respuesta de la cuestión con un modelo homogéneo (Q4 y Q6), i.e., $Q7 \Rightarrow Q4$.

Las preguntas Q4.1 y Q7.1 están directamente relacionadas con Q4 y Q7 respectivamente, con la diferencia de que las primeras son más abiertas y para ser contestadas se necesita el tratamiento global de la tabla.

Así pues, teniendo en cuenta todas estas consideraciones, podemos establecer las siguientes cadenas de implicaciones:

$$\begin{array}{ccc} Q7.1 & & Q4.1 \\ \Downarrow & & \Downarrow \\ Q7 \Rightarrow Q4 & \leftrightarrow & Q6 \end{array}$$

iv) Cuestiones sobre relaciones de proximidad y distancia

$$C_4 = \{Q8, Q16, Q20, Q12, Q13, Q14\}$$

Las cuestiones de esta última clase tratan la proximidad y distancia entre los sujetos (los jóvenes) de la tabla del cuestionario.

En las preguntas Q8, Q14, Q16 y Q20 se trabajan las coincidencias (copresencias y coausencias). Concretamente, en la pregunta Q8 se trata de comparar dos columnas para comprobar si son iguales o no, y en Q14 se pide encontrar tres pares de columnas iguales en toda la tabla. Por tanto, estas dos cuestiones quedan relacionadas de la siguiente forma: $Q14 \Rightarrow Q8$.

La operación que requieren tanto la cuestión Q16 como la Q20 es el cálculo del cardinal de las coincidencias (positivas y negativas) entre columnas concretas. Suponemos, a priori, que la pregunta Q20 es más compleja que la Q16, puesto que ésta última es más directa, tanto en su redacción (más natural para el alumnado) como en los cálculos que deben hacer en cada una de ellas: en Q16 deben contar el número de coincidencias entre dos columnas (sujetos) y, en cambio, en la Q20 se establece un sujeto previo a partir del cual se debe calcular el cardinal de coincidencias con otros dos sujetos y después comparar los resultados. De ahí que la implicación entre estas dos cuestiones queda establecida como $Q20 \Rightarrow Q16$.

También suponemos a priori la implicación $Q14 \rightarrow Q16$, ya que como la Q14 es una pregunta abierta con búsqueda en toda la tabla, la consideramos más complicada que la Q16 por ser concreta y relacionar sólo dos sujetos concretos.

Asimismo, consideramos que el alumno que responda con éxito las preguntas Q14 y Q16 también será capaz de contestar correctamente la cuestión Q13, en la cual se exige el cardinal de las diferencias existentes entre dos columnas de la tabla. Por tanto, debe verificarse el siguiente encadenamiento de las cuestiones (incluyendo la transitividad entre Q14 y Q13):

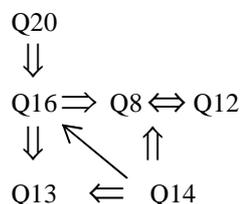
$$Q14 \rightarrow Q16 \Rightarrow Q13$$

También podemos establecer la implicación $Q16 \Rightarrow Q8$, ya que, si un alumno sabe calcular el cardinal de coincidencias entre dos jóvenes (sujetos) tal y como pide la primera cuestión, también será capaz de comprobar si dos columnas son iguales (coinciden todas las celdas) según se exige en la pregunta 8.

Por último, consideramos equivalentes las preguntas Q8 y Q12, en las que se pregunta –aunque de forma distinta– las coincidencias entre dos sujetos concretos. La Q8 se formula a partir del concepto de « preferencias entre tipos de música » y en la Q12 se pregunta utilizando « el mismo gusto musical »:

$$Q8 \Leftrightarrow Q12$$

En resumen, la cadena de implicaciones entre las cuestiones de esta clase quedaría representada de la siguiente forma:



3 Comparación de los análisis del Cuestionario Q

3.1 Resultados de la experimentación

Presentamos a continuación una aproximación a los porcentajes de respuestas justas a las preguntas del cuestionario Q, obtenidas por Pitarch (2002) en la experimentación realizada en diferentes centros de Castellón, sobre 239 estudiantes, diferenciando los resultados en los dos ciclos de la ESO y relacionándolos según la caracterización de las cuestiones realizada a priori.

Cuestión Número	Operación básica según MAP	Porcentaje éxito en la respuesta	1r ciclo (1º,2º ESO)	2º ciclo (3º 4º ESO)
Q1, Q2, Q3	Tratamiento tabla	97 %	no existen diferencias signific.	
Q4, Q4.1, Q6	Cuantificadores (\forall, \exists)	90 %	menor éxito	mayor éxito
Q5 ; Q21	Comparación cardinales	92 % ; 62%	no existen diferencias signific.	
Q7 ; Q7.1	Comparación filas	70 %	no existen diferencias signific.	
Q8	Comprobar coincidencias	87 %	no existen diferencias signific.	
Q9 ; Q10 ; Q11	Estadística elemental	39%, 15% ; 8%.	menor éxito	mayor éxito 4º
Q12; Q13	Enumerar diferencias	81 %	no existen diferencias signific.	
Q14	Coincidencias-toda Tabla	11 %	menor éxito	mayor éxito 3º
Q15, Q19	Conjunción[^]/Negación \neg	51 %	menor éxito	mayor éxito 4º
Q17, Q18	Implicación lógica y con razonamiento	13 %	menor éxito	mayor éxito 4º
Q16, Q20	Coincidencias-columnas	66 %	menor éxito	mayor éxito 4º

TAB 2 – Porcentajes de respuestas justas de la experimentación, en Castellón (Pitarch, 2002).

Estos resultados confirman los obtenidos por Orús (1992) en alumnos de último curso de enseñanza primaria en Francia (CM2), mostrando que se siguen observando las dificultades en cuestiones que exigen una operación o razonamiento lógicos, con más de un 80% de alumnos que presentan dificultades con la implicación, pese al mayor porcentaje de éxito en 4º curso (último año de la enseñanza obligatoria). Queremos señalar también el bajo porcentaje de éxito en las preguntas de estadística elemental, inferior al 20% en el cálculo de la frecuencia de una variable de la tabla T. Dificultades que no pueden ser atribuidas al manejo de la tabla, puesto que todas las cuestiones que sólo exigen estrategias de búsqueda, enumeración o comparación de elementos de información de la tabla (filas o columnas) tienen un altísimo porcentaje de éxito, superior al 80%; aunque también este porcentaje disminuya cuando la complejidad de la tarea de búsqueda aumenta.

3.2 El ASI y los análisis priori y a posteriori

3.2.1 Las implicaciones de la experimentación.

Los grafos implicativos efectuados por CHIC, versión 3.7, sobre los resultados de la experimentación (Pitarch, 2002) muestran las relaciones de implicación entre todas las cuestiones del cuestionario; en la Fig. 2 se muestran las relaciones más significativas, con un índice de implicación de 0.99 (izquierda) y también las nuevas relaciones de implicación que aparecen entre las cuestiones al rebajar el índice a 0.95 (derecha).

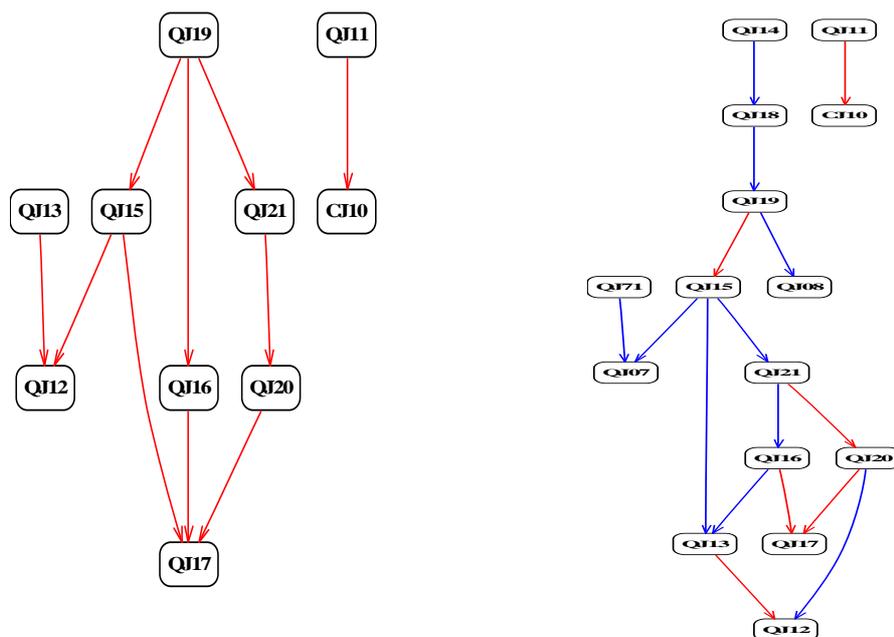


FIG. 2 - Implicaciones a posteriori: 0.99 (izquierda) y 0.95 (derecha)

Analizaremos estas implicaciones con más detalle, en función de los conocimientos de las cuestiones, retomando las clases del análisis a priori, una vez contrastadas la validez de las implicaciones que habíamos establecido a priori al interior de estas clases.

3.2.2 Validación de las implicaciones a priori.

Con la experimentación realizada vamos a evaluar la validez de nuestro análisis (implicativo) a priori, usando la metodología de Spagnolo (1997). En ella, el autor trata el tema de la validación-refutación de ciertas hipótesis didácticas a través del uso de un cuestionario y del análisis estadístico implicativo. Por una parte, el investigador traduce las hipótesis didácticas que quiere someter a examen, en una serie de implicaciones entre algunas clases de cuestiones del cuestionario que, a priori, serían lógicas (es decir, exentas de contraejemplos). Posteriormente, la muestra estructura las (cuasi-)implicaciones entre variables a partir de los contraejemplos presentes en ella. Entonces, calcula las intensidades de las implicaciones de los datos recogidos (la contingencia), y las mismas para las implicaciones a priori (que resultan fijando a cero el número de contraejemplos y manteniendo el número de ocurrencias de antecedente y consecuente), y crea un estadístico de tipo χ^2 , que describe la distancia entre cada implicación a priori y a posteriori (la diferencia, normalizada por el valor la intensidad a priori). La distancia entre las implicaciones a priori y a posteriori entre clases de cuestiones queda definida como la media aritmética de las distancias entre las implicaciones que resultan de combinar cada variable de la clase antecedente con cada variable de la clase consecuente.

Finalmente, la regla de decisión consiste en la validación de la implicación entre clases (i.e. de la hipótesis didáctica) si y solamente si la diferencia entre lo ideal y la contingencia es inferior al valor 0.25.

En nuestro caso, las hipótesis didácticas han sido trasladadas a una serie de implicaciones entre variables, y no entre clases de variables. Los cálculos de intensidades de implicación a priori y sobre los datos del

cuestionario han sido realizados usando la teoría clásica y la ley de Poisson, dando como resultados los expuestos en la Tab 2.

Implicación analizada	Intensidad a priori	Intensidad a posteriori	Diferencia	Diferencia Normalizada	Decisión
Q11 → Q10	1	0.999	0	0	CONCORDANCIA
Q10 → Q2	0.83	0.126	0.704	0.848	REFUTADA
Q2 → Q1	0.956	0.704	0.253	0.264	REFUTADA
Q6 → Q3	0.998	0.589	0.41	0.41	REFUTADA
Q3 → Q1	0.955	0.697	0.258	0.27	REFUTADA
Q11 → Q9	0.998	0.826	0.172	0.172	CONFIRMADA
Q41 → Q4	1	0.894	0.105	0.105	CONFIRMADA
Q4 → Q6	1	0.794	0.206	0.206	CONFIRMADA
Q6 → Q4	1	0.864	0.136	0.136	CONFIRMADA
Q71 → Q7	0.993	0.979	0.014	0.014	CONFIRMADA
Q7 → Q4	1	0.789	0.211	0.211	CONFIRMADA
Q18 → Q19	1	0.954	0.046	0.046	CONFIRMADA
Q18 → Q17	0.999	0.845	0.154	0.154	CONFIRMADA
Q19 → Q15	1	1	0	0	CONFIRMADA
Q21 → Q5	0.997	0.425	0.572	0.574	REFUTADA
Q17 → Q19	1	0.842	0.158	0.158	CONFIRMADA
Q17 → Q15	1	0.868	0.132	0.132	CONFIRMADA
Q20 → Q16	1	0.973	0.027	0.027	CONFIRMADA
Q16 → Q13	1	0.964	0.036	0.036	CONFIRMADA
Q16 → Q8	1	0.552	0.448	0.448	REFUTADA
Q8 → Q12	1	0.798	0.202	0.202	CONFIRMADA
Q12 → Q8	1	0.747	0.253	0.253	REFUTADA
Q14 → Q13	0.976	0.674	0.301	0.309	REFUTADA
Q14 → Q8	0.915	0.915	0	0	CONFIRMADA
Q14 → Q16	0.989	0.92	0.069	0.069	CONFIRMADA

TAB 3 – Implicaciones analizadas, valores de intensidad a priori y a posteriori, diferencias absolutas y normalizadas, y decisión sobre su validez.

3.2.3 Análisis por clases, de las implicaciones a priori y a posteriori.

Retomamos las clases de cuestiones, obtenidas en el análisis de similaridad realizado a priori, sobre el cuestionario, en función de los conceptos matemáticos en juego, para seguir analizando cada una de ellas,

Análisis a priori de cuestionarios basado en técnicas ASI

contrastando las implicaciones a priori, con toda la información obtenida a posteriori sobre las implicaciones entre las cuestiones, de los datos de la experimentación.

i) Cuestiones sobre conceptos estadísticos

La comparación entre las implicaciones establecidas a priori entre estas cuestiones y las obtenidas a posteriori, en la experimentación, han permitido confirmar estas relaciones (Ver Tab. 2).

$$\begin{array}{ccc} Q11 \Rightarrow Q10 & & \\ \downarrow & & \\ Q9 & & \text{Confirmadas} \end{array}$$

Al considerar el subgrafo de las cuestiones de la clase C_1 del análisis implicativo de la experimentación (ver Fig. 4), solamente aparecen relaciones de implicación entre las cuestiones que habíamos identificado a priori como cuestiones sobre conceptos estadísticos Q10, Q11 y Q9, presentando una fuerte intensidad (de 0.99) la implicación de la cuestión Q11 referente a la frecuencia relativa de una variable, hacia la cuestión Q10 sobre la frecuencia absoluta. Sin embargo los buenos resultados sobre el porcentaje Q9 parecen depender más directamente de los aciertos sobre la frecuencia absoluta que sobre la frecuencia relativa, como habíamos considerado a priori.

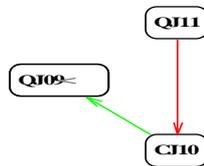


FIG. 3 - Implicación a posteriori entre las cuestiones de la clase $C_1 = \{Q1, Q2, Q3, Q10, Q9, Q11\}$

El resto de cuestiones de la clase C_1 , cuyas relaciones a priori, han sido establecidas por la interpretación semántica de dichas cuestiones y el tipo de trabajo sobre la tabla que exigen sus respuestas, no aparecen en los resultados del análisis implicativo de la experimentación (como puede apreciarse en las Figs. 2 y 3). En la comparación a priori-a posteriori, estas relaciones han sido refutadas.

$$\begin{array}{ccc} Q6 \rightarrow Q3 \rightarrow Q1 & & \\ \uparrow & & \\ Q10 \rightarrow Q2 & & \text{Refutadas} \end{array}$$

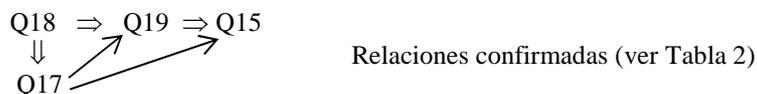
ii) Cuestiones sobre operaciones lógicas

Estas cuestiones aparecían en la clase $C_2 = \{Q15, Q19, Q17, Q18\}$ y el subgrafo implicativo del análisis a posteriori (ver Fig. 4) muestra la siguiente cadena de implicaciones, todas ellas con un índice superior a 0.95 y Q9, Q15 y Q17 con un índice de implicación de 0.99.



FIG. 4 - Implicación a posteriori entre las cuestiones de la clase $C_2 = \{Q15, Q19, Q17, Q18\}$

Las relaciones establecidas a priori, según el trabajo lógico y las estrategias de búsqueda, también han sido confirmadas.



Habíamos considerado en este apartado también las cuestiones de la clase $C_5 = \{Q5, Q21\}$, sobre el cálculo del cardinal de la disyunción y se había establecido la implicación a priori $Q21 \Rightarrow Q5$; pero esta implicación no aparece en el grafo implicativo de la experimentación (Fig. 1), y aparece como refutada en la comparación a priori/a posteriori de la Tabla 2.

iii) Cuestiones sobre cuantificadores existenciales y universales

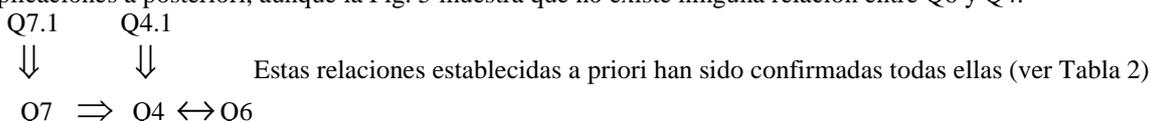
Recordemos que las cuestiones de la clase $C_3 = \{Q4, Q4.1, Q6, Q7, Q7.1\}$ están caracterizadas por utilizar un cuantificador existencial y/o universal, a la vez, que pueden requerir un recuento alternativo por parte del estudiante.

Las implicaciones a posteriori del subgrafo de estas cuestiones de la clase C_3 (ver Fig. 5), son más débiles que en las otras clases: solamente aparecen $Q7.1$ y $Q7$ con un índice de 0.95 y el resto de cuestiones aparecen relacionadas cuando se desciende el índice a 0.85.



FIG. 5 - Implicaciones a posteriori entre las cuestiones de la clase $C_3 = \{Q4, Q4.1, Q6, Q7, Q7.1\}$, con el umbral fijado en 0.95 (izquierda) y 0.85 (derecha)

Hay que señalar que todas las relaciones establecidas a priori, son validadas por la comparación con las implicaciones a posteriori, aunque la Fig. 5 muestra que no existe ninguna relación entre $Q6$ y $Q4$.



Las dos preguntas $Q6$ y $Q4$, involucran la misma actividad matemática, establecer el valor de verdad de una función proposicional generalizada en un dominio finito:

$$\zeta \exists P_i / [\forall x \in X / P_i(x)]: \text{verdadera, siendo } X = \{x_1, x_2, \dots, x_n\}?$$

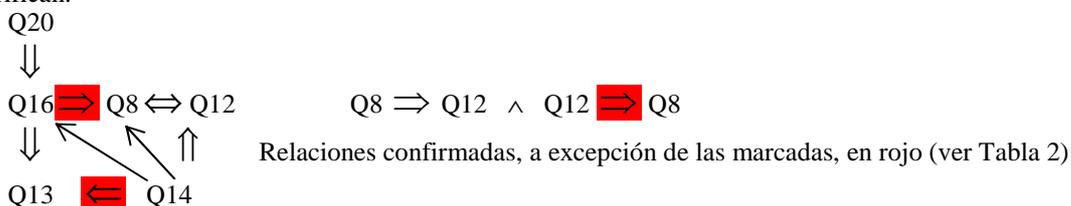
Sin embargo, la ausencia de relación a posteriori entre estas dos preguntas muestra que la interpretación semántica de la tabla marca el trabajo de los alumnos y que el sentido que atribuyen a una fila o a una columna es distinto: las respuestas de los alumnos, entre la pregunta $Q4$: *¿Hay algún tipo de música elegida por todos los jóvenes?* y la pregunta $Q6$: *¿Hay algún joven a quien le guste todos los tipos de música?* no establecen ninguna relación de implicación entre ellas.

En el análisis a-priori, ya habíamos establecido entre ellas solamente una relación $Q4 \leftrightarrow Q6$, en función del tipo de estrategia de resolución de ambas preguntas ya que tanto en la pregunta $Q4$, como en la $Q6$ se le exige al alumno que realice la búsqueda de un modelo «homogéneo», en concreto de una fila y una columna, respectivamente, en la que todas las celdas tengan el valor 1; pero no le habíamos atribuido una relación de complicación $Q4 \leftrightarrow Q6$ que le hubiera correspondido por la equivalencia del conocimiento matemático subyacente en ellas, puesto que conjeturamos que no era el modelo matemático de estas preguntas, sino la estrategia de búsqueda lo que marcaría las respuestas de los alumnos..

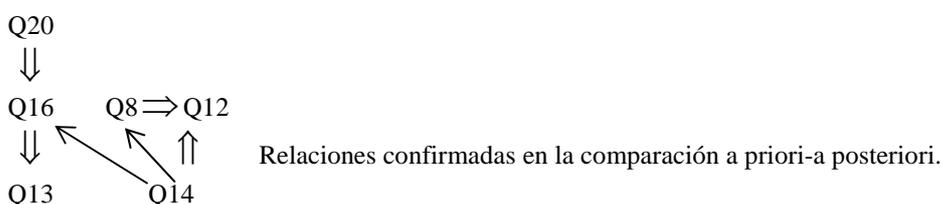
Análisis a priori de cuestionarios basado en técnicas ASI

iv) Cuestiones sobre relaciones de proximidad y distancia

No todas las relaciones de implicación que habíamos considerado a priori, entre las cuestiones de la clase $C_4 = \{Q8, Q16, Q20, Q12, Q13, Q14\}$, se han visto confirmadas en la comparación con los datos a posteriori; en el esquema que habíamos presentado para relacionar a priori estas cuestiones, hemos marcado las que no se verifican.



Por tanto la estructura de implicación que ha quedado confirmada para esta clase C_4 es la siguiente:



La coimplicación a priori que habíamos establecido entre Q8 y Q12, debida a la equivalencia del trabajo a realizar para responder ambas preguntas, ha sido refutada, mostrando que para los alumnos existe una diferencia entre estas preguntas; si nos fijamos en la formulación de la pregunta Q8 (*prefieren el mismo tipo de música*), parece que resulta más confusa a los alumnos que la pregunta Q12 (*tienen el mismo gusto por la música*), esto puede explicar que los alumnos que saben interpretar correctamente Q8 también saben responder correctamente Q12, es decir $Q8 \Rightarrow Q12$ mientras que no sucede lo contrario y por tanto no se verifica $Q12 \Rightarrow Q8$.

La confusa redacción de la pregunta Q8, también puede explicar la no validación de la implicación establecida a priori entre Q16 y Q8, al poder aumentar la posibilidad de contraejemplos, con contestaciones justas para Q16 (*¿En cuentas respuestas de la encuesta están de acuerdo los jóvenes C y J?*) y erróneas para Q8 (*Los jóvenes H y O ¿prefieren el mismo tipo de música?*).

El hecho de que la validación de la cadena de implicaciones establecida a priori entre Q20, Q16 y Q13, todas ellas cuestiones que comparan coincidencias o diferencias entre los gustos musicales de jóvenes concretos de la encuesta, mientras que la implicación atribuida a priori entre las cuestiones Q14 y Q13 no se confirma, parece corroborar la "originalidad" de la cuestión Q14 (*Encuentra, si puedes, a tres parejas de jóvenes a quienes les guste el mismo tipo de música*). Se solicita del alumno una actividad de mayor complejidad y dificultad por generalizar la comparación entre sujetos a toda la tabla, que parece ser una condición suficiente para responder correctamente a muchas otras cuestiones, apareciendo como antecedente en muchas de las relaciones de implicación (todas ellas confirmadas, como se ve en el esquema) con esas otras cuestiones Q16, Q8 y Q12, pero no con Q13. La particularidad de Q13 es que es la única pregunta explícita sobre las diferencias en las comparaciones que se cuestionan.

El siguiente grafo (Fig. 6), presenta las relaciones de implicación a posteriori en C_2 e ilustra las reflexiones anteriores sobre la validación o no de las relaciones de implicación que habíamos establecido a priori en esta clase C_2 .

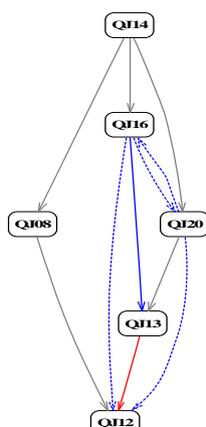


FIG. 6 - Implicación a posteriori entre las cuestiones de la clase $C4 = \{Q8, Q16, Q20, Q12, Q13, Q14\}$, considerando un índice de implicación de 0.75 y la opción de transitividad

Solamente añadir, que aparece una relación de implicación con un índice de 0.99, la implicación $Q13 \rightarrow Q12$, y ésta no la habíamos señalado en el análisis a priori, por considerar que la actividad de enumeración sobre la tabla resultaba más asequible a los alumnos que cualquier otro tipo de actividad comparativa, y por tanto que $Q13$, debería aparecer como consecuente en las cadenas de implicación, tal y como se ha visto en el esquema a priori. Esta fuerte relación de implicación a posteriori entre $Q13$ (*¿Cuántas diferencias entre los jóvenes C y J?*) y $Q12$ (*Los jóvenes R y N ¿tienen el mismo gusto por la música?*) parece mostrar que es la peculiaridad de $Q13$ cuestionando sobre las diferencias, la que la sitúa en el antecedente de la implicación

4 Conclusiones

4.1 Conclusiones didácticas sobre la experimentación de Q en la Enseñanza Secundaria Obligatoria

El trabajo presentado en esta comunicación ilustra una vez más la conveniencia -incluso necesidad- de los análisis a priori de los cuestionarios, para que éstos se adecuen a los objetivos que la investigación pretende abordar; esta necesidad ha sido reclamada constantemente en Didáctica de las Matemáticas, y especialmente desde la Teoría de Situaciones (Brousseau y colaboradores). En nuestro caso, el rol de situación a-didáctica que se pretendía jugar con el cuestionario, es el que ha marcado la necesidad del estudio a priori de dicha situación y de los conocimientos presentes en el medio a-didáctico, para determinar las variables que queríamos observar y los valores que se les iba a otorgar a dichas variables. En Pitarch y Orús (2005), se puede consultar este análisis y el interés didáctico que puede tener el cuestionario, para trabajar en clase, más allá del valor instrumental de recogida de información que ha supuesto la experimentación del cuestionario.

No obstante creemos que el análisis de los datos de la experimentación, que hemos presentado en los diferentes apartados de esta comunicación, comparando los análisis a priori y a posteriori, aporta una información más completa y matizada sobre la realidad del tratamiento de la información en la ESO y del bajo índice de éxito de algunos de los conocimientos lógico-matemáticos y estadísticos asociados, notablemente las operaciones lógicas complejas, combinando operadores lógicos (conectores o cuantificadores, y la negación ellos) o en forma de funciones proposicionales condicionales.

El bajo porcentaje de éxito en las cuestiones explícitas sobre el concepto estadístico de frecuencia de una variable estadística “materializada” en la tabla como la cabecera de una fila (el gusto por un determinado tipo de música) y de los diferentes valores que toma esa variable, en función de los sujetos, muestran las dificultades de los alumnos con la apropiación del concepto de variable estadística y no sólo con el cálculo de las frecuencias, que si son capaces de realizar sobre datos con presentación “estándar”, tipo ejercicio de manual.

Consideramos que la utilización didáctica de la tabla de datos por parte del profesor puede permitir recrear, en situación escolar, el sentido de éstos y otros conceptos estadísticos, como puede ser la clasificación de datos (Orús, 1992).

A la vez, la tabla permite así mismo dar sentido al trabajo de los alumnos con operadores lógicos que actúan sobre funciones proposicionales y no sólo sobre proposiciones; Durand-Guerrier (1996) señala la necesidad de

un trabajo de esta naturaleza con alumnos de la enseñanza obligatoria, para evitar las actuales carencias de formación lógica observada y analizada por ella misma en alumnos de primeros años de Universidad, carencias que esta autora identifica como el origen de muchas de las actuales dificultades de los alumnos universitarios de diferentes titulaciones en el estudio de las asignaturas de Matemáticas.

4.2 Conclusiones metodológicas: ASI y análisis a priori de Q

El análisis a priori del cuestionario nos condujo a buscar las variables de las cuestiones, transformándolos en unos criterios que nos pudieran permitir identificar o caracterizar esas cuestiones. De esta manera hemos obtenido la matriz a priori que hemos llamado en toda la comunicación MAP, en la que los criterios podían ser considerados los “sujetos-tipo” que caracterizaban las cuestiones. En este sentido nos pareció interesante analizar el comportamiento a priori del cuestionario, solamente en función de los criterios seleccionados y poder “comparar” estos análisis con el comportamiento de estas cuestiones en la experimentación.

El análisis de similaridad de la MAP, como todo análisis tipológico o clasificatorio, permite hacer “emerger” relaciones entre las variables (las cuestiones) que no se habían considerado inicialmente (Chandon y Pinson, 1981) y por ello quisimos realizar este análisis de la MAP. Las clases obtenidas nos han confirmado que los criterios retenidos para caracterizar las cuestiones estaban bien seleccionados, dando sentido e identificando las clases resultante. También puso de manifiesto algunas proximidades no previstas, entre cuestiones, como por ejemplo agrupar en la clase C_1 , las cuestiones sobre el manejo de la tabla, con las cuestiones explícitas sobre estadística.

Sin embargo, la naturaleza no simétrica del ASI nos puso de manifiesto las posibles dificultades que podía presentar el análisis implicativo de la MAP. Por ejemplo, aparecía que la cuestión Q10 (frecuencia absoluta) implicaba a Q11 (frecuencia relativa) y no al revés; la comparación de los valores de la MAP, para estas dos cuestiones nos muestra que son idénticos para todos los criterios, a excepción de *EFRE*, que identifica la frecuencia relativa y por tanto Q11 tenía el valor 1 y Q10 el valor 0, aportando un contraejemplo para la condicional $Q11 \rightarrow Q10$, pero sin embargo se verifica $Q10 \Rightarrow Q11$ (como implicación lógica). Pero esto se podía haber evitado simplemente cambiando el criterio, por el de frecuencia absoluta, invirtiéndose todas las relaciones anteriores, o incluso no diferenciar la frecuencia absoluta de la relativa y dar a ambas cuestiones el valor 1 y a todas las demás 0.

Hemos constatado como una misma identificación de las cuestiones puede hacerse por tanto, con criterios diferentes e incluso utilizando la formulación negativa de los criterios, se puede invertir la asignación de 1 de 0 a las cuestiones y por tanto alterar o invertir los contraejemplos de las implicaciones a priori.

Por todo ello no hemos utilizado los resultados del ASI sobre la matriz a priori MAP; sino que hemos establecido a priori las implicaciones entre las cuestiones, en función de los contenidos implicados y no de la codificación asignada; y estas implicaciones son las que han sido comparadas con las obtenidas experimentalmente.

La aplicación de este método, ya utilizado en Spagnolo (1997), ha resultado muy fructífera en nuestro trabajo:

- El establecimiento a priori de implicaciones entre las cuestiones, ha propiciado el verificar las inversiones que aparecían en las implicaciones entre esas cuestiones en el grafo implicativo de la MAP y descartar estos análisis no simétricos sobre la MAP, tanto el análisis implicativo como el cohesitivo.
- La comparación de estas implicaciones a priori, con los resultados de la experimentación nos ha permitido confirmar la mayoría de nuestras hipótesis respecto a la estructura de nuestro cuestionario y su pertinencia como instrumento de medida de los conocimientos implicados.
- Ha permitido complementar y refinar la información obtenida con la aplicación de técnicas ASI, sobre los resultados de la experimentación, al interior de cada clase de cuestiones.

En definitiva, creemos que esta comunicación muestra una vez más la pertinencia, e incluso la necesidad, de los análisis a priori de los cuestionarios para poder garantizar la coherencia de los análisis de los resultados, en función de los objetivos del propio cuestionario; nuestra constatación hecha en el ámbito de la investigación en Didáctica de las Matemáticas, la consideramos también válida y generalizable a aquellos dominios de investigación en los que la utilización de los cuestionarios juega un papel metodológico importante.

Bibliografía

Bodin A. (1996), Modèles sous-jacents à l'analyse implicative et outils complémentaires, Cahiers du séminaire de didactique de l'IRMAR de Rennes

- Brousseau G. (1986), Fondaments et méthodes de la didactique des mathématiques. Thèse d'Etat, Université de Bordeaux-I. Ed: LADIST-Université de Bordeaux II, Bordeaux.
- Brousseau G. (1998), Théorie des situations didactiques, La Pensée Sauvage, Grenoble. [Versión en lengua inglesa: Theory of didactical situations in mathematics: Didactique des mathématiques 1970-1990 (N. Balacheff, M. Cooper, R. Sutherland and V. Warfield, Eds. and Trans.). Dordrecht, The Netherlands: Kluwer (1997)].
- Brousseau G. y Lacasta E. (1995), L'analyse statistique des situations didactiques, Actes du Colloque Méthodes d'analyses statistiques multidimensionnelles en Didactique des Mathématiques, ARDM, pp 53-107.
- Couturier R. (2000), Traitement de l'analyse statistique dans CHIC, Actes des Journées sur la Fouille de Données par la Méthode d'Analyse Statistique Implicative, Caen, IUFM de Caen, pp. 33-50
- Durand-Guerrier V. (1996), Logique et raisonnement mathématique. Défense et illustration de la pertinence du calcul des prédicats pour une approche didactique des difficultés liées à l'implication. Thèse d'Etat, Université Claude Bernard - Lyon-I. Ed. CNRS de l'Université Lumière-Lyon-2.
- Gras R. (1995), Méthodes d'analyses statistiques multidimensionnelles en didactique des mathématiques. Actes du Colloque Méthodes d'analyses statistiques multidimensionnelles en Didactique des Mathématiques. Ed. ARDM, pp 53-107.
- Gras, R et al. (1996), L'implication statistique. Nouvelle méthode exploratoire de données. Ed. La Pensée Sauvage.
- Lerman I.C. (1981), Classification et analyse ordinaire des données, Dunod, Paris
- Lerman I.C., Gras R. y Rostam H. (1981), Elaboration et évaluation d'un indice d'implication pour des données binaires, I et II, Mathématiques et Sciences Humaines n° 74, p. 5-35 y n° 75, p. 5-47.
- Orús P. (1986), L'enseignement des méthodes de classification. Proposition d'une ingénierie pour le cours moyen. Collection Études en Didactique des Mathématiques. IREM de Bordeaux.
- Orús P. (1992), La place du raisonnement naturel des élèves dans la négociation didactique: rôle possible d'une initiation à l'analyse classificatoire dans la scolarité obligatoire. Thèse, U.E.R. Scientifique. Université de Bordeaux-I, Ed. IREM de Bordeaux.
- Orús P. y Gregori P. (2005), Des variables supplémentaires et des élèves "fictifs" dans la fouille des données avec CHIC. Quaderni di Ricerca in Didattica. Supplemento n.2 al N. 15, Ed. GRIM, Palermo (pp. 279- 293)
- Orús P. y Pitarch I. (2000), Utilisation didactique des tableaux des données et du logiciel CHIC à l'école élémentaire, Actes des Journées sur La fouille dans les données par la méthode d'analyse statistique implicative, École Polytechnique de l'Université de Nantes 2000, pp 85-98.
- Pitarch I. (2002), Estadística y Análisis de datos en la ESO. DEA presentado en el Departamento de Matemáticas de la Universidad Jaume-I de Castellón.
- Pitarch I. y Orús P. (2005), Análisis didáctico de un cuestionario en el marco de la Teoría de las Situaciones Didácticas. Actas del VIII Simposio de Investigación en Educación Matemática. A Coruña 9-1. Septiembre de 2004. Ed. U.Coruña, pp 1-20.
- Spagnolo F. (1997), L'analisi a-priori e l'indice di implicazione statistica di Gras, Quaderni di Ricerca in Didattica, n.7, pp 110-125.
- Spagnolo F. y Gras R. (2007), Une mesure comparative en didactique des mathématiques entre une analyse a priori et la contingence, manuscript.

Anexo 1

CONSIGNA:

Se ha hecho una encuesta, escogiendo como muestra jóvenes entre 15 y 20 años, para valorar qué tipo de música es la más elegida entre la población de estas últimas décadas.

A continuación se presenta una tabla de valores que representa las respuestas de los jóvenes a la encuesta musical y dicha información viene expresada de la siguiente manera:

- A la izquierda está el listado de los tipos de música, de los cuales se pide opinión a cada uno de los jóvenes. Por ejemplo: "¿te gusta la música heavy?"

- La letra mayúscula de cada una de las columnas, representa el nombre del joven. Por ejemplo: " el joven B"

- Los jóvenes entrevistados sólo podrán contestar con un SI o un NO. Y en la tabla de valores viene expresado de la siguiente manera: 1 quiere decir SI, y 0 equivale a NO.

Anexo 2

TABLA T

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1.¿Te gusta la música hip-hop?	1	1	1	0	1	0	1	0	1	0	0	1	1	0	0	1	1	0	1	0
2.¿Te gusta la ópera?	1	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0
3.¿Te gusta la música máquina?	1	1	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1
4.¿Te gusta el jazz?	0	1	0	1	1	0	1	1	0	0	1	1	0	0	1	1	0	0	1	0
5.¿Te gusta el flamenco?	0	0	0	1	0	1	1	1	1	0	1	0	0	1	1	0	1	0	1	0
6.¿Te gustan las rumbas?	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7.¿Te gusta la música heavy?	0	1	0	1	1	0	1	0	0	0	1	1	1	0	0	1	0	0	1	1
8.¿Te gusta la música pop?	1	1	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1
9.¿Te gusta la música clásica?	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
10.¿Te gusta la música dance?	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11.¿Te gusta la música reggae?	0	1	0	1	1	0	1	0	1	0	1	1	0	0	0	1	1	0	1	0
12.¿Te gusta el hardcore?	0	1	1	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1
13.¿Te gusta el rap?	1	0	1	1	0	1	0	1	1	0	1	0	0	1	1	0	1	0	1	1
14.¿Te gusta la música punk?	1	1	1	0	1	0	1	0	1	0	1	1	1	0	0	1	0	0	1	0
15.¿Te gusta la música ska?	1	1	1	1	1	0	0	1	0	0	1	1	1	1	1	1	1	0	1	1

TAB 4 – Tabla sobre la que se pregunta a los alumnos, en el cuestionario Q de la experimentación.

Anexo 3

EL CUESTIONARIO (Questiones para responder los alumnos)

- Q1. ¿Al joven H le gusta el jazz?
- Q2. ¿A cuántos jóvenes les gusta el flamenco?
- Q3. ¿Cuántos tipos de música le gustan al joven Q?
- Q4. ¿Hay algún tipo de música elegida por todos los jóvenes? En caso afirmativo, indica cuáles:
- Q4.1. ¿Hay otras? En caso afirmativo, indica cuáles
- Q5. Los jóvenes entrevistados ¿prefieren la música heavy o la música ska?
- Q6. ¿Hay algún joven a quien le guste todos los tipos de música? En caso afirmativo, indica su nombre:
- Q7. ¿Existe alguna música más preferida que la música punk?
- Q7.1. Si hay alguna, indícala y razona tu respuesta.
- Q8. Los jóvenes H y O, ¿prefieren el mismo tipo de música?
- Q9. ¿A qué porcentaje de los jóvenes entrevistados les gusta la música rap?
- Q10. Si consideramos - escoger la música máquina- como una variable estadística, ¿qué frecuencia absoluta dirías que tiene en esta distribución estadística?
- Q11. ¿Y qué frecuencia relativa tendría la variable estadística de la pregunta anterior?

Contesta, **razonando todas tu respuestas:**

Q12. Los jóvenes R y N, ¿tienen el mismo gusto por la música? ¿Por qué?

Q13. ¿Cuántas diferencias hay entre las respuestas de los jóvenes C y E? ¿Por qué?

Q14. Encuentra, si puedes, tres parejas de jóvenes a quienes les guste el mismo tipo de música.

Q15. ¿A cuántos jóvenes les gusta a la vez el flamenco y la música pop? ¿Por qué?

Q16. ¿En cuántas respuestas de la encuesta están de acuerdo los jóvenes C y J?

Q17. A todos los jóvenes que les gusta la música dance ¿también les gusta el hardcore?

Q18. A todos los jóvenes que les gusta el hardcore ¿también les gusta la música dance? ¿Por qué?

Q19. ¿A cuántos jóvenes no les gusta ni el flamenco ni el pop? ¿Por qué?

Q20. El joven I ¿a quien se aproxima más en su gusto musical, al joven F o al joven D?

Q21. A los jóvenes entrevistados ¿les gusta más la música heavy, el jazz o el reggae? ¿Por qué?

Summary

The goal of this contribution is to extend the use of the Statistical Implicative Analysis to the framework of questionnaire analyses, specially in questionnaires whose data are the one obtained through the sample as well as data created in *a priori* studies. The questionnaire performed in Pitarch (2002), its *a priori* analysis and empirical data is used for this purpose. That questionnaire was designed in order to characterise the logical and mathematical skills and reasoning of Compulsory Secondary School students, when performing data processing of a Boolean data table.

Résumé

L'objectif de cette communication est d'approfondir dans la ligne d'utilisation de l'Analyse Statistique Implicative dans les analyses de questionnaires que considèrent des "données", aussi bien les obtenus expérimentalement comme les obtenus dans les études *a priori*. Pour ce but, on reprendra le questionnaire élaboré par Pitarch (2002), son analyse *a priori* et ses données empiriques. Ce questionnaire a l'objectif de rassembler des informations et caractériser le travail des élèves de l'Enseignement Secondaire Obligatoire en effectuant le traitement d'un tableau booléen de données primaires.