

JAVIER GIL FLORES

ANÁLISIS DE DATOS CUALITATIVOS

APLICACIONES A LA INVESTIGACIÓN EDUCATIVA

LCT-91



PPU

Barcelona, 1994

0039-0514

Colección: LCT-91
(Letras, Ciencias, Técnica)

Serie: IIE-Investigación e Innovación Educativa
Dirigida por: Benito Echeverría Samanes

Primera edición, 1994

Quedan rigurosamente prohibidas, sin la autorización escrita de los titulares del *Copyright*, bajo las sanciones establecidas en las leyes, la reproducción parcial o total de esta obra por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, y la distribución de ejemplares de ella mediante alquiler o préstamo públicos.

© Javier Gil Flores

© PPU, S.A.
Promociones y Publicaciones Universitarias, S.A.
Marqués de Campo Sagrado, 16. 08015 Barcelona
Tfno. (93) 442-03-91 Fax (93) 442-14-01

Diseño cubierta: C. Riu Espart

I.S.B.N.: 84-477-0326-6
D.L.: L-240-1994

Imprime: Poblgràfic, S.A. Av. Estació, s/n. La Pobla de Segur (Lleida).

INDICE

INTRODUCCION	7
CAPITULO I	
Acercamiento al concepto de dato	13
1. ¿Qué entendemos por datos?	13
2. El dato en la investigación educativa	22
2.1. El dato en los programas de investigación positivistas	23
2.2. El dato en los programas de investigación interpretativos	24
2.3. El dato en los programas de investigación críticos	26
3. El dato cualitativo	27
CAPITULO II	
El análisis de datos cualitativos	31
1. Análisis de datos	31
1.1. Concepto	31
1.2. Lugar del análisis en el proceso de investigación	33
1.3. El continuo recogida de datos-análisis de datos	37
2. El problema del análisis de datos cualitativos	38
2.1. Los esfuerzos de sistematización	40
2.2. El recurso a la cuantificación	41
2.3. Principios del análisis de datos cualitativos	43

EL ANÁLISIS DE DATOS CUALITATIVOS

Datos cualitativos pueden ser las cadenas verbales producidas en una entrevista, en una reunión, en un contacto interpersonal o en documentos escritos, las conductas y los sucesos recogidos en las notas de campo. También son datos cualitativos las fotografías, las filmaciones, las grabaciones realizadas en los contextos estudiados y que aportan al investigador información sobre ellos. Los objetos, su distribución, su uso, etc. constituyen fuentes de datos cualitativos.

El hecho de que las interacciones, las opiniones expresadas por los sujetos o la información que se transmite en los contextos educativos se produzca en códigos verbales es una importante razón para que la mayor parte de los datos elaborados en tales contextos se apoyen en la expresión verbal escrita. La información extraída por el investigador es expresada a través del lenguaje, y generalmente plasmada en registros escritos (notas de campo redactadas, transcripción de entrevistas, descripciones de artefactos, etc.). Por esta razón, se habla de datos narrativos o textuales en lugar de datos cualitativos. En realidad, el dato textual es un tipo particular de dato cualitativo. Pero dado que es posible elaborar datos textuales a partir de datos plasmados inicialmente por medio de otro tipo de registros, la mayor parte de las ocasiones el investigador se enfrenta al análisis de datos cualitativos textuales. El rasgo distintivo de los datos cualitativos textuales no es sino su expresión en forma verbal escrita, es decir, en forma de textos.

Los datos cualitativos son elaborados por procedimientos y técnicas tales como la entrevista en profundidad, la observación participante, el trabajo de campo, la revisión de documentos personales u oficiales. Para recoger tales datos, los investigadores suelen permanecer prolongadamente en el campo, implicándose en las situaciones estudiadas hasta llegar a obtener datos que se ajustan a lo que la gente dice y/o hace.

1. ANÁLISIS DE DATOS

1.1. CONCEPTO

Los datos que elabora el investigador no suelen ofrecer suficiente información para resolver los complejos problemas planteados en la investigación educativa, si no son organizados y manipulados de alguna forma. En este sentido, suele hablarse de análisis de datos, momento presente en cualquier proceso de investigación. Como afirma Ander-Egg (1990), los datos tienen una importancia limitada; es necesario encontrarles significación si queremos que resulten útiles para el investigador. La tarea del análisis consiste precisamente en interpretar y extraer significado de los datos recogidos.

Podemos hacernos una idea de lo que supone analizar datos si examinamos el propio concepto de análisis. El análisis es un proceso aplicado a la realidad que nos permite discriminar sus componentes a algún nivel determinado, describir las relaciones entre tales componentes y utilizar esa primera visión conceptual del todo como instrumento para realizar nuevos análisis con mayor profundidad, que lleven a síntesis más adecuadas (Bunge, 1985). La idea de examen de las partes de algo por separado y de las relaciones entre ellas, inherente al concepto de análisis, queda también recogida por Spradley (1979:92 y 1980:85), cuando define el análisis como un proceso de pensamiento:

"El análisis de cualquier clase supone un proceso de pensamiento. Se refiere al examen sistemático de algo para determinar sus partes, las relaciones entre las partes, y sus relaciones con el todo".

El proceso de análisis es susceptible de ser aplicado a objetos, hechos o fenómenos muy diversos, los cuales, admitiendo que la realidad es pluridimensional, podrían ser analizados en más de una forma. Así, el número de unidades que comprende un fenómeno real puede variar según el criterio o la dimensión considerados al dividirlo, y los vínculos entre tales unidades diferirán en función del tipo de relaciones en que centremos nuestra atención. El fin de todo análisis es

lograr una mayor comprensión de la realidad analizada, sobre la cual podría llegarse a elaborar algún tipo de modelo conceptual explicativo.

En cualquier caso, el análisis de datos forma parte del proceso de investigación. La investigación cuenta con procedimientos mediante los cuales trata de dar respuesta a los problemas que se plantea. Ello nos lleva a la idea de método, definido por Bunge (1980:28) como "un procedimiento regular, explícito y repetible para lograr algo, sea material sea conceptual", que constituye el atributo característico de la actividad investigadora frente a la mera observación y especulación (Shulman, 1988).

El método de la investigación educativa no es único ni tampoco exclusivo. La existencia de diferentes programas y enfoques de investigación hace que se dé una variedad de procedimientos empleados para llevar a cabo lo que denominamos investigación. A pesar de esta diversidad, el análisis de datos está siempre presente como un elemento constitutivo o como una fase de los métodos de investigación desarrollados en el ámbito educativo (Fox, 1981; Pérez Juste, 1985; Spradley, 1979; etc.).

Del mismo modo, cuando se habla de diseño de investigación, entendido como secuencia lógica que conecta los datos empíricos a las cuestiones de investigación y a las conclusiones, es decir, un plan de acción que lleva desde las cuestiones iniciales a las conclusiones sobre ellas (Yin, 1987), el análisis de datos es considerado uno de sus componentes.

Por tanto, la planificación de los procesos indagativos tendrá que avanzar el tipo de datos que es necesario recoger para dar respuesta al problema planteado, los instrumentos y técnicas que podrían utilizarse para obtenerlos, y prever también los procedimientos de análisis que se emplearán.

El análisis de datos realizado en la investigación educativa puede ajustarse total o parcialmente al concepto general de análisis, es decir, división en unidades y examen de relaciones entre ellas y respecto al todo. A veces los datos son, por ejemplo, expresados en números, que se consideran isomorfos con una realidad fragmentable, y por tanto, divisible en elementos y variables. Los procedimientos estadísticos a los que se suele recurrir en estos casos permiten la comprobación de leyes que establecen relaciones entre las variables. También para autores enmarcados en programas de investigación cualitativa, el análisis de datos implica un proceso de varias etapas en el que un fenómeno global es dividido en unidades,

éstas son clasificadas en categorías, y a continuación ensambladas, relacionadas para conseguir un todo coherente (Goetz y Lecompte, 1988).

En el mismo sentido, el análisis de datos es concebido como un proceso consistente en aislar las unidades fundamentales de conocimiento cultural o dominios en que se organiza lo que los sujetos conocen, explorar la estructura interna de tales dominios y buscar relaciones entre los dominios y de éstos con el todo (Spradley, 1979,1980), o es visto como un conjunto de actividades tendentes a reducir los datos a unidades manejables y significativas, estructurar y presentar los datos, y extraer conclusiones que avancen en la comprensión del todo (Miles y Huberman, 1984b).

No es infrecuente que aquello que los investigadores denominan análisis de datos se aparte de la idea general de análisis que hemos presentado en párrafos anteriores. El análisis de datos puede ser visto, a veces, como un proceso espontáneo, asistemático, intuitivo, flexible, orientado a encontrar sentido a los datos, sin que para ello se pase necesariamente por las etapas de separación de elementos y de reconstrucción del todo a partir de las relaciones entre ellos. Woods (1989), ejemplificando posiciones próximas a este tipo de enfoques, reivindica el componente artístico en el análisis de datos dentro de la investigación de corte etnográfico.

En suma, podemos definir el análisis de datos como el *conjunto de manipulaciones, transformaciones, operaciones, reflexiones y comprobaciones realizadas a partir de los datos con el fin de extraer significado relevante en relación a un problema de investigación.*

1.2. LUGAR DEL ANÁLISIS EN EL PROCESO DE INVESTIGACIÓN

Tradicionalmente, en la investigación en ciencias sociales, se suele caracterizar el desarrollo del proceso investigativo por su linealidad en el tiempo. Las etapas que lo componen tienen un sentido secuencial: planteamiento de problemas, formulación de hipótesis, recogida de datos, análisis, interpretación y redacción del informe. Desde este punto de vista, el análisis de datos constituiría una fase posterior a la recogida de datos, e iría seguido de la interpretación. Esta característica se suele contraponer a los planteamientos metodológicos de los programas de investigación de tipo cualitativo, en los que es frecuente que el análisis de datos se realice de forma simultánea a otras tareas, tales como la recogida de datos o la elaboración de informes, por citar algunas.

Sin embargo, tal oposición parece derrumbarse si examinamos el momento en que se da el análisis en relación a otras tareas que comporta el proceso de investigación, con independencia de enfoques cualitativos o cuantitativos.

Las tareas que habitualmente denominamos recogida de datos y análisis de datos aparecen entrelazadas en el proceso de investigación. Podemos ver la recogida de datos, en sí misma, como una forma primitiva de análisis, puesto que representa algún tipo de elaboración sobre la realidad. Así, por ejemplo, registrar la conducta de un sujeto en una lista de control en la que hubiera que especificar numéricamente la intensidad de esa conducta o cualquier otra característica, representa una decisión del observador en el momento de registrar el dato, para clasificar lo que observa y valorar su grado. Obtener la frecuencia de un hecho o fenómeno implica una evaluación continua de cada dato empírico para determinar si puede ser clasificado dentro de la categoría de fenómenos que estamos contando.

Del mismo modo, describir lo que vemos al observar una situación educativa implica ineludiblemente una interpretación de la realidad por parte del observador y una categorización de la misma implícita en el uso del lenguaje. De alguna manera, se está llevando a cabo un análisis de lo observado.

El análisis está presente cuando el investigador recoge datos, por ejemplo, en forma de notas de campo. Es frecuente que junto a éstas aparezcan juicios, opiniones, sospechas, dudas, reflexiones, interpretaciones que el investigador añade a la mera información descriptiva sobre lo que ocurre en el campo de estudio. Así, Taylor y Bogdan (1986) incluyen en las notas de campo -forma que presentan los datos elaborados en la observación participante- no sólo los datos descriptivos sobre palabras y conductas de las personas, sino también sentimientos, interpretaciones emergentes, intuiciones, hipótesis de trabajo y preceptos del investigador, áreas futuras de indagación, etc., que aparecen como comentarios del observador. Otros autores prefieren ir anotándolos separadamente en forma de memorandos (Miles y Huberman, 1984b), que pueden ser considerados formas preliminares de análisis. Como afirma Woods (1989), al observar, entrevistar, tomar notas de campo y confeccionar el diario de investigación, el etnógrafo no se limita a registrar, pues junto a tales tareas se da una reflexión.

Dejando a un margen este modo de concebir la elaboración de datos, que implica en sí misma un primer análisis de la realidad, podemos afirmar que momentos de análisis y momentos de recogida de datos suelen alternarse o ser simultaneados en los diseños de investigación. Rechazamos, por tanto, la idea de que el análisis constituye una fase final, posterior a la recogida, pues "ello excluye

la posibilidad de recoger nuevos datos para rellenar huecos o comprobar nuevas hipótesis que emergen durante el análisis" (Miles y Huberman, 1984b:49). La interconexión de los procesos de recogida y análisis en la investigación se ha concretado metodológicamente, dentro de la tradición cualitativa, en las denominadas estrategias de selección secuencial (Goetz y Lecompte, 1988), en las que la progresiva construcción teórica determina la recogida de datos. Entre las formas de selección secuencial se encuentran la selección de casos negativos, que refutan o contradicen un concepto y la selección de casos discrepantes que permiten modificarlos; el muestreo teórico, decidiendo qué datos recoger y dónde recogerlos de cara a desarrollar la teoría emergente (Glaser y Strauss, 1967); o la selección de implicantes, que posibilita la comprobación de teorías a partir de casos seleccionados en función de las implicaciones de esas teorías (Campbell, 1986).

Análisis y recogida aparecen interconectados también en investigaciones de orientación cuantitativa. Un ejemplo podría ser el desarrollo de estudios descriptivos mediante procedimientos estadísticos, a partir de los cuales se seleccionan muestras más reducidas sobre las que recoger nuevos datos para análisis de otro tipo. Incluso la misma tradición científica nos habla de una constante comprobación -sometimiento a prueba- de las hipótesis. Cada nueva explicación verosímil de un hecho nos incita a recoger evidencia y a contrastar esa explicación.

Recogida y análisis se condicionan mutuamente, por tanto, en un proceso en el que la información recogida y la teoría que emerge de su análisis son usadas para orientar una nueva recogida de datos (Wilson, 1977). Posiblemente el modo gráfico en que Lofland y Lofland (1984:132) presentan el desarrollo simultáneo de análisis y recogida de datos en el contexto de la investigación cualitativa (ver Figura 2.1) sea suficientemente clarificador. Para estos autores, la consideración de ambos procesos como fases separadas del trabajo del investigador resulta poco productiva, inclinándose por la concomitancia de tales tareas con un progresivo incremento del análisis, que prosigue una vez finalizada la recogida de datos. Durante este último período se desarrollan y ponen en orden ideas surgidas previamente. Tal relación entre recogida y análisis es considerada productiva.

Algunos autores entienden que el análisis de datos no sólo está presente en la fase de la investigación correspondiente a la recopilación de datos sino también en la fase de la elaboración de informes (Erickson, 1989). Una prosa lúcida y un buen estilo literario pueden constituir bases adecuadas para la interpretación y el análisis, pues para muchos investigadores el proceso de escribir es el período donde el conocimiento y la penetración aparecen más espontáneamente (Wax y Wax, 1980). En este sentido, pueden recordarse las recomendaciones de Spradley (1979),

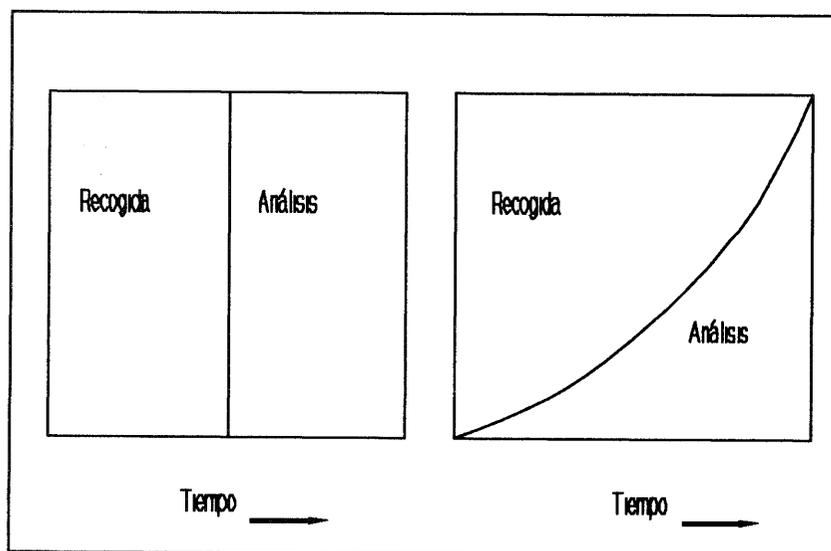


Figura 2.1: Relaciones productiva e improductiva entre recogida y análisis de datos.

quien es partidario de comenzar a escribir una etnografía pronto, pues escribir implica un proceso de pensamiento y análisis.

El hecho de que cada fase o componente del proceso de investigación pueda ser estudiado de manera aislada no debe alejarnos de su carácter unitario, del todo coherente que conforma. El tipo de problema de investigación que se plantee implica un modo de conceptualizar el objeto de estudio al que va referido. De alguna manera, en él van implícitos el tipo de datos que necesitamos y las estrategias analíticas adecuadas para su resolución.

El problema estudiado condiciona, por tanto, las decisiones acerca del tipo de datos necesarios y de técnicas o instrumentos que se emplearán en su elaboración, y avanza los procedimientos de análisis que éstos y la naturaleza del problema aconsejan. Sin embargo, la relación del problema de investigación con la elaboración y análisis de los datos no es unidireccional. Todo proceso de investigación supone, de hecho, una invitación a que sean los propios datos o los resultados de su análisis los que influyan o determinen el verdadero carácter del problema de investigación o la siguiente y más fundamental pregunta a investigar.

Para Goetz y Lecompte (1988), quienes entienden que recogida y análisis de datos están indisolublemente unidos, en la medida en que se profundiza en la cultura y en las interpretaciones que hacen los participantes se da una redefinición de las preguntas de investigación. La influencia de los resultados que va arrojando el análisis puede hacer que un problema sea focalizado en aspectos cada vez más concretos (Bogdan y Biklen, 1982). Incluso podría ser diferido el planteamiento del problema hasta que un análisis preliminar del campo de estudio haya sido realizado.

Los resultados de un análisis correlacional, por ejemplo, podrían aconsejar la extensión de un problema de investigación para abarcar no sólo la constatación de relaciones entre variables, sino la posible existencia de relaciones causa-efecto entre ellas.

Si trascendemos el nivel de un estudio concreto y nos centramos, con una perspectiva diacrónica, en el proceso de investigación sobre un determinado tema educativo, encontramos, con independencia del paradigma en que nos situemos, que los nuevos conocimientos alcanzados como consecuencia del estudio y análisis de la realidad educativa llevan a que, en los sucesivos estudios, se revisen los problemas y los supuestos iniciales. Por tanto, recogida y análisis de una parte y problema de investigación de otra pueden llegarse a condicionar mutuamente en el proceso de investigación educativa.

1.3. EL CONTINUO RECOGIDA DE DATOS-ANÁLISIS DE DATOS

Consideramos interesante detenernos brevemente en una idea apuntada en el apartado anterior. Recogida de datos y análisis no sólo ocurren de manera simultánea como se ha señalado, sino que se encuentran íntimamente interrelacionados e incluso es posible verlos como procesos indiferenciados conceptualmente. La recogida se ha venido asociando aquí a elaboración de datos. Entendemos que el dato es producto de una elaboración y que esa elaboración cristaliza en la aplicación y uso de determinados instrumentos o técnicas de recogida y registro. La elaboración del dato implica en sí misma una intervención del investigador, quien conceptualiza la información al interactuar con la realidad y determina una forma de expresarla o conservarla.

Pero es frecuente que el investigador sobrepase la elaboración que supone el mero dato descriptivo y añada a éste un componente interpretativo, reflexivo, al que podríamos atribuir una naturaleza analítica. Se llega a afirmar que los datos en sí mismos son sustancialmente interpretaciones (Angulo, 1990:41). Esta conexión

e indiferenciación que puede llegar a darse entre elaboración de datos y análisis hace que a veces resulte difícil establecer dónde acaba el dato y dónde comienza el resultado o la interpretación del análisis. Determinadas formas de elaboración de datos, tales como la aplicación de escalas de observación, suponen un esquema previo de categorías en virtud del cual las conductas o fenómenos observados son codificados directamente en la recogida, es decir, la elaboración de datos lleva implícito un componente analítico como es la reducción a categorías.

Datos y productos del análisis tendrían, por tanto, una misma naturaleza. Van Maanen (1985) atribuye a datos y teorías una misma naturaleza conceptual. Los datos son conceptos de primer orden y las teorías conceptos de segundo orden que se utilizan para explicarlos.

Elaboración -o recogida- y análisis de datos forman parte de un continuo que lleva desde los fenómenos u objetos reales que actúan de referentes hasta las conclusiones de un estudio, de forma que la información extraída de la realidad va siendo procesada con un nivel de complejidad creciente.

2. EL PROBLEMA DEL ANÁLISIS DE DATOS CUALITATIVOS

El análisis de datos cualitativos, dado el carácter polisémico de éstos, su naturaleza predominantemente verbal o su difícil reproducción, resulta un proceso de cierta complejidad. Además, estos datos aparecen generalmente en forma de textos de gran extensión, cuyo manejo durante el análisis encierra no pocas dificultades.

Para el análisis de datos cuantitativos se han desarrollado procedimientos sistemáticos y mayoritariamente aceptados por la comunidad científica. Tales métodos de análisis están basados en determinados planteamientos epistemológicos asumidos en el estudio de la realidad social, conforme a los cuales se aborda el análisis de datos siguiendo técnicas estadísticas. Para estos investigadores, la realidad es susceptible de ser medida y puede ser descrita objetivamente. Es posible encontrar leyes, a las que generalmente subyacen modelos matemáticos, capaces de explicar y predecir los fenómenos sociales.

Cuando se manejan datos de naturaleza cualitativa la situación ha venido siendo bastante diferente. El análisis de datos cualitativos representa, en cierto

modo, un problema para el investigador cualitativo en la medida en que concurren diferentes factores que vienen a complejizar esta tarea:

- Indefinición de los métodos de análisis.
- Importancia de la componente artística en el análisis.
- Pluralidad de enfoques de investigación cualitativa.
- Escaso tratamiento del análisis en la literatura sobre investigación.
- Falta de un lenguaje común al hablar del análisis.

A pesar de la riqueza y frescura de significados intrínseca a los datos cualitativos, la escasa definición de los métodos de análisis hace que muchos autores vean a aquéllos simplemente como una "molestia atractiva" (Miles, 1985). No se suele contar, pues, con vías definidas o convenciones claras que orienten el modo de analizar los datos, por lo que los procedimientos utilizados son de tal variedad y singularidad que a veces resultan genuinos, no ya de la disciplina o de la tradición donde nos situemos, sino del propio investigador que los emplea, quien ha llegado a configurarlos a partir de una experiencia acumulada de trabajo con este tipo de datos, y de acuerdo con su propio estilo como investigador.

"Todos los investigadores desarrollan sus propios modos de analizar los datos cualitativos" (Taylor y Bogdan, 1986:159).

Tal es la variedad de los procedimientos aplicados al análisis cualitativo de datos que podría afirmarse que el único punto de acuerdo entre los investigadores es la idea de que el análisis es el proceso de extraer sentido de los datos (Tesch, 1990).

Frente al carácter sistemático y formalizado de las técnicas de análisis aplicadas a los datos cuantitativos, los métodos usados con datos cualitativos se basan con frecuencia en *la intuición y la experiencia del investigador*. A veces se ha afirmado que la tarea del analista es más un arte que una técnica. Esto supone una dificultad añadida a la investigación cualitativa, pues sería necesario sumar a las habilidades y tareas propias de todo proceso analítico otras de carácter especial.

Para Woods (1989), el analista tendría que poseer actitudes y cualidades de *creatividad, imaginación* y un *"olfato de detective"* para extraer temas de interés a partir de indicios observados y llegar a producciones teóricas. Glaser y Strauss (1967) añaden a la imaginación la *ingenuidad*, y hablan de la *sensibilidad teórica* como característica que ha de poseer el investigador para conceptualizar y formular una teoría que emerge de los datos cualitativos. Otros aluden a la habilidad del

investigador para *trascender su propia perspectiva* y llegar a conocer las perspectivas de aquéllos a quienes está estudiando (Wilson, 1977). Autoras como Goetz y Lecompte (1988) hablan de la necesidad de *pensamiento divergente* y *talento artístico*, aunque no refiriéndose a todo el proceso de análisis sino a algunos de los elementos instrumentales presentes en la interpretación (metáforas, analogías, etc.).

La indefinición existente en relación a los procedimientos de análisis se ve agravada por el carácter plural de la investigación cualitativa, en la que coexisten un amplio número de tradiciones de investigación (Jacob, 1988).

Por otra parte, el análisis de datos cualitativos es objeto de una escasa atención en la literatura especializada. Una gran parte de los trabajos de investigación omiten o sólo describen vagamente -como si de una caja negra se tratase- la forma en que los datos son procesados hasta conseguir extraer conclusiones a partir de ellos. Lo mismo ocurre en los manuales de investigación, en los que se dedica un exiguo número de páginas a tratar este tema. Es significativo que en trabajos como el de Bogdan (1983), destinados a definir un modo de enseñar e introducir a principiantes en algunos métodos de investigación cualitativa, se excluya cualquier alusión al modo de analizar los datos.

Para Miles y Huberman (1984b), los métodos de análisis son raramente relatados con detalle, en parte como consecuencia del grado de incertidumbre acerca del análisis cualitativo. Pero aun cuando los investigadores llegan a explicitar sus métodos, la falta de un lenguaje común hace que resulten ambiguos. No sólo existe, por tanto, desacuerdo en el modo en que los investigadores realizan el análisis de sus datos cualitativos, sino también en la terminología que usan para referirse a los elementos y operaciones implicados en él. Ello supone una dificultad añadida para el empleo de las técnicas de análisis por otros investigadores y especialmente por los investigadores noveles.

2.1. LOS ESFUERZOS DE SISTEMATIZACIÓN

La privacidad de los métodos de análisis desarrollados desde una perspectiva cualitativa y la falta de detalle con que son descritos en los informes de investigación, concentran parte de las críticas relativas a la credibilidad en la investigación cualitativa. Esta situación ha provocado reacciones por parte de los investigadores que trabajan dentro de este enfoque, quienes han hecho esfuerzos dirigidos a clarificar y también a guiar los procesos analíticos. La indefinición de

procedimientos y pautas para llevar a cabo el análisis de datos cualitativos ha suscitado en diversos autores intentos de dar rigor y exhaustividad a las técnicas cualitativas de análisis, tratando de que se aproximen en sistematización y fiabilidad a los procedimientos cuantitativos, si bien algunos autores los rechazan por considerar que esclerotizan todo el proceso en perjuicio de su naturaleza intuitiva y creativa. Con tales intentos se sale al paso de las numerosas críticas vertidas sobre la investigación denominada cualitativa y, en particular, sobre el modo en que en esta perspectiva eran analizados los datos, al que se calificaba de oscuro, difuso, falta de rigor, poco definido.

Este tipo de técnicas analíticas aplicadas a los datos cualitativos se incluyen dentro de lo que Firestone y Dawson (1982) denominan enfoques "de procedimiento", que incluyen una serie de reglas y pasos a seguir, frente a enfoques "intuitivos" o "intersubjetivos", basados respectivamente en la intuición y en la interacción entre investigadores o entre investigadores y participantes en los contextos estudiados. En estos enfoques de procedimiento se hacen esfuerzos por describir, con mayor o menor detalle, una serie de operaciones y pasos concretos, ofrecen consejos, recomendaciones y advertencias para el manejo, la disposición o la presentación de los datos, y suelen ser, en general, menos explícitos a la hora de explicar los principios que subyacen y las reglas, mecanismos o estrategias que permitan la extracción final de conclusiones, operación para la que no se cuenta con criterios de decisión como los aportados, por ejemplo, en las pruebas estadísticas.

Algunos autores han llegado a desarrollar procesos sistemáticos de análisis que han contado con amplia difusión. Entre los de mayor tradición se encuentran la inducción analítica (Znaniecki, 1934) o el método de la comparación constante (Glaser y Strauss, 1967).

2.2. EL RECURSO A LA CUANTIFICACIÓN

El avance que representan en el análisis de datos cualitativos los denominados enfoques de procedimiento a los que nos hemos referido, no constituye óbice para que aún sean muchos los investigadores que recurren en algún momento de su proceso analítico a las técnicas estadísticas, con la intención de complementar o contrastar las conclusiones obtenidas por otras vías. Incluso este rasgo ha sido visto como característico de los enfoques de investigación interpretativos o cualitativos, los cuales recurren, sobre todo en el análisis de datos, a procedimientos cuantitativos característicos del enfoque positivista (De Miguel, 1988). Este es uno de los elementos, a nivel metodológico, que permiten hablar de una superación del

enfrentamiento entre paradigmas. El acercamiento entre técnicas cuantitativas y cualitativas suele hacerse, bien por la introducción de la cuantificación numérica en los análisis cualitativos, o bien por la complementariedad de ambas, en la medida en que producen distinto tipo de información.

La naturaleza de los datos cualitativos -palabras, y no números- constituye un problema crónico en la investigación educativa, pues aunque las palabras aportan mayor significado, los números resultan menos ambiguos y son procesados con mayor economía. Por esta razón, la mayor parte de los investigadores prefiere trabajar con números y trasladan los textos a números lo más rápidamente posible. Para los investigadores que toman este camino, el terreno está bien marcado por las teorías de la medición, reglas de decisión, niveles de confianza, algoritmos para el cálculo, etc. (Miles y Huberman, 1984a).

No obstante, sería difícil establecer una rígida división entre procedimientos exclusivamente cualitativos o exclusivamente cuantitativos. Procesos de ambos tipos suelen aparecer como momentos diferentes de una misma manipulación de datos. Incluso algunos de los métodos analíticos caracterizados como genuinamente cualitativos, dejan abierta una puerta a la posibilidad de la cuantificación. Para Miles y Huberman (1984b:215), "cuando identificamos un tema o modelo estamos aislando algo que ocurre un número de veces y que ocurre consistentemente de un modo específico". Tanto el número de veces como los juicios acerca de la consistencia se basan en un recuento. Los números, según estos autores, son útiles para tener una noción de qué tenemos ante nosotros cuando trabajamos con un amplio cuerpo de datos, para verificar una sospecha o hipótesis, y para ser honestos en el análisis tratando de evitar sesgos. Goetz y Lecompte (1988) contemplan la enumeración como uno de los procedimientos analíticos generales que usa el investigador y, al referirse a las tareas de teorización, consideran que la comparación, contrastación, agregación y ordenación constituyen un prerrequisito para determinar las frecuencias de los fenómenos.

"Una vez identificados los elementos, es posible reducirlos a forma cuantificable examinándolos, elaborando listas, codificándolos y asignándoles puntuaciones" (178-179).

A partir de las frecuencias de los fenómenos, identificaríamos las relaciones que se dan entre los elementos mediante algún tipo de tratamiento estadístico pertinente.

Aquí consideraremos que la cuantificación y el análisis estadístico son herramientas analíticas con las que cuenta el investigador en su trabajo con datos cualitativos, y pueden ser utilizados conjuntamente con otras herramientas no cuantitativas. La formación del investigador, sus objetivos y sus concepciones epistemológicas le llevan a configurar un proceso de análisis más o menos desplazado hacia uno de estos tipos de técnicas.

Dependiendo del estudio, la extensión de la cuantificación de datos varía ampliamente, desde ninguna hasta el uso de técnicas estadísticas de variado grado de complejidad (Wilcox, 1982). La cuantificación puede llegar a ser el aspecto central del análisis, e incluso a veces los datos cualitativos son, desde el primer momento, trasladados a índices numéricos y analizados cuantitativamente.

El análisis cuantitativo de datos cualitativos no está exento sin embargo de la crítica de quienes consideran que la complejidad de las situaciones sociales desaconseja el tratamiento de la información elaborada sobre ella recurriendo a la división en elementos y el cómputo de frecuencias.

2.3. PRINCIPIOS DEL ANÁLISIS DE DATOS CUALITATIVOS

A pesar de que la dispersión y variedad de métodos que se han puesto en práctica para abordar el análisis de datos cualitativos resulta una de las razones por las que éste puede ser, en cierto modo, considerado un problema, podemos encontrar ciertos elementos característicos del análisis de este tipo de datos. En general, el proceso empleado para analizar los datos y elaborar conclusiones sobre la relación entre los elementos encontrados en un estudio se apoya en una serie de *principios o criterios* básicos, algunos de los cuales quedan recogidos en la siguiente relación. Todos ellos se combinan de distinto modo configurando estrategias para el análisis de los datos.

* *Agrupamiento*. Los datos aislados no son significativos, y por esa razón es preciso unirlos a otros a fin de llegar a conclusiones. La agrupación de elementos es característica de los procesos de categorización. A su vez, las categorías son susceptibles de ser agrupadas, dando lugar a factores o dimensiones relevantes para comprender el significado de los datos. Técnicas estadísticas como los análisis de conglomerados o procedimientos taxonómicos se basan en la idea de agrupamiento.

* *Recuento.* El análisis procede separando unidades y agrupándolas en función de determinadas afinidades (identidad, similitud, equivalencia, etc.). El cómputo de las unidades encontradas de cada tipo es una operación básica en todo análisis. El recuento de unidades permite concluir respecto de la existencia e importancia de los fenómenos. Un examen de las distribuciones permite hacer rápidamente afirmaciones acerca de los datos.

* *Pasar de lo particular a lo general.* En la extracción de conclusiones es esencial moverse desde las observaciones empíricas hacia niveles más elevados de abstracción. Cuando agrupamos una serie de unidades incluyéndolas bajo una misma categoría estamos avanzando en este proceso.

* *Inclusión.* También está presente en la categorización, pues ésta implica identificar en qué tipo de entidades se incluye un hecho o circunstancia. Interrogándonos por la relación de inclusión podríamos llegar al establecimiento de clasificaciones, sistemas mediante los que se expresa el entramado de relaciones inclusivas existentes dentro de un conjunto de elementos.

* *Subordinación.* La existencia o la ocurrencia de un fenómeno depende de otro en el cual se inserta, del que constituye una etapa o faceta, etc. Las relaciones de subordinación permiten construir jerarquías entre los elementos conceptuales identificados en el corpus de datos.

* *Ordenación.* La ordenación es otra operación que nos lleva a obtener importantes conclusiones en la tarea analítica. Ordenar implica disponer elementos en determinada forma de acuerdo con algún criterio explícito (magnitudes cuantitativas, cronología, ubicación, funcionalidad, causalidad, etc.). Llegar a establecer un orden es en sí mismo una conclusión de cualquier estudio.

* *Coocurrencia.* Constatar la aparición simultánea de dos fenómenos es un resultado útil, que da pie a cuestionarse el tipo de relación que está detrás de esa coincidencia. Cuando analizamos un texto, la coocurrencia podría venir definida por la proximidad espacial dentro del mismo.

* *Covariación.* La covariación implica una vinculación mayor entre elementos, pues no sólo supone la aparición conjunta sino también que la variación en una dirección determinada de uno de los elementos está aparejada a una variación similar del otro.

* *Causalidad.* Establecer relaciones causa-efecto es el objetivo de muchos estudios. Se dan cuando puede razonablemente considerarse que la presencia, ausencia o variación de un elemento es causa de la del segundo. En el análisis de datos textuales algunos autores han sugerido la detección de este tipo de relaciones a partir de la coaparición de elementos en una secuencia consistente (Huber, 1988). Los métodos de análisis estadísticos basados en la dependencia entre variables presuponen relaciones de causalidad.

3. EL PROCESO GENERAL DE ANÁLISIS DE DATOS CUALITATIVOS

Además de señalar principios comunes, es posible establecer una serie de tareas u operaciones que constituyen el proceso analítico básico, común a la mayoría de los estudios en que es necesario trabajar con datos cualitativos textuales, con independencia de que las vías elegidas para el análisis tengan un carácter marcadamente cualitativo o cuantitativo.

En la descripción de tales tareas adoptamos la distinción establecida por Miles y Huberman (1984), para quienes en el análisis concurrirían actividades dirigidas a la reducción de datos, presentación de datos y extracción o verificación de conclusiones.

En cada uno de estos grupos describiremos manipulaciones, operaciones, transformaciones concretas que son realizadas durante el análisis de datos, aunque no necesariamente aparecen todas presentes en el trabajo de cada analista. En ocasiones, determinadas actividades pueden extenderse hasta constituir por sí mismas el proceso de análisis o, por el contrario, pueden no ser tenidas en cuenta en el tratamiento de los datos, de acuerdo con los objetivos del trabajo, el enfoque de investigación o las características del investigador. Entre ellas no siempre se establece una sucesión en el tiempo, y pueden darse de forma simultánea, o incluso hallarse dentro de una misma manipulación o tarea.

El marco general podría esquematizarse en las siguientes tareas:

- * Reducción de datos.
 - Separación de elementos.
 - Identificación y clasificación de elementos.
 - Agrupamiento.
- * Disposición de datos.
 - Transformación y disposición.
- * Obtención y verificación de conclusiones.
 - Proceso para extraer conclusiones.
 - Verificación de conclusiones.

3.1. LA REDUCCIÓN DE DATOS

La limitada capacidad humana para procesar grandes cantidades de datos hace necesario reducir la amplia información contenida en los datos textuales, diferenciando unidades e identificando los elementos de significado que soportan. Los procesos de categorización y codificación son los más representativos de la reducción de datos textuales. En momentos avanzados del análisis, nuevas reducciones de datos pueden hacerse sobre resultados de tratamientos previos efectuados con los datos "brutos".

a) Separación en unidades

De acuerdo con el significado básico del término "análisis", una de las primeras operaciones a que tendrían que ser sometidos los datos es la *separación de los segmentos* o unidades que conforman el conjunto global de datos objeto de análisis. Precisamente esa segmentación de los datos, esta división en unidades relevantes y significativas, es considerada una de las prácticas más características del análisis de datos cualitativos (Tesch, 1990:95).

La división de la información en unidades puede realizarse siguiendo diferentes criterios. Las consideraciones *espaciales* o *temporales* permiten definir unidades de carácter artificial, surgidas a partir de criterios exclusivamente físicos. De acuerdo con tales criterios, constituirían unidades los bloques de un determinado número de líneas, las páginas, los bloques de minutos, etc. Es más frecuente, y

probablemente más adecuado, considerar unidades en función del *tema abordado*. Considerando conversaciones, sucesos, actividades que ocurren en la situación observada es posible encontrar segmentos que hablan de un mismo tema. Sin embargo, el criterio temático conlleva una pequeña dificultad cuando las unidades son objeto de cómputo para posteriores análisis. Podríamos encontrarnos tanto fragmentos de gran extensión como fragmentos reducidos que hablen de un mismo tema, resultando unidades que llegan a ser consideradas equivalentes en análisis ulteriores, a pesar de que no lo son. Este problema ha sido soslayado en el análisis de contenido mediante la distinción entre *unidades de registro* y *unidades de enumeración* (Cartwright, 1978). Mientras las primeras determinan fragmentos referidos a un mismo tópico, las segundas aluden a la unidad sobre la cual se realiza la cuantificación. Las unidades pueden establecerse también a partir de *criterios gramaticales* (el párrafo o la oración podrían ser considerados las unidades básicas de un texto) o *conversacionales* (declaraciones, turnos de palabra, etc.). Con ellas no es necesario un juicio sobre el significado para establecer los límites de la unidad.

b) Identificación y clasificación de unidades

La identificación y clasificación de elementos consiste en examinar las unidades de datos para encontrar en ellas determinados componentes temáticos que nos permitan clasificarlas en una u otra categoría de contenido. Las operaciones más representativas de las actividades de identificación y clasificación son las conocidas como codificación y categorización. La *categorización*, que constituye una importante herramienta conceptual en el análisis de datos cualitativos, hace posible clasificar conceptualmente las unidades que son cubiertas por un mismo tópico. Una categoría soporta un significado o tipo de significados. Las categorías pueden referirse a situaciones y contextos, actividades y acontecimientos, relaciones entre personas, comportamientos, opiniones, perspectivas sobre un problema, métodos y estrategias, procesos. La categorización se realiza conjuntamente a la división en unidades cuando ésta se hace atendiendo a criterios temáticos. Si una unidad es separada por referirse a determinado tópico, automáticamente puede ser incluida en la categoría correspondiente a ese tópico.

El proceso de *codificación* no es más que la operación concreta por la que se asigna a cada unidad un indicativo (código) propio de la categoría en la que la consideramos incluida. Es el proceso físico, manipulativo mediante el cual se realiza la categorización. Por esta razón, no es extraño que el estudio de ambas operaciones se aborde en los manuales de investigación educativa de forma conjunta, y que incluso en algún momento lleguen a identificarse como un único proceso, aspecto

que queda reflejado en afirmaciones tales como "categorizar es asignar etiquetas a algo" (Bliss, Monk y Ogborn, 1983:8), o "los códigos son categorías" (Miles y Huberman, 1984b:56).

Los códigos, que representan a las categorías, pueden ser números con los que se marca cada unidad de datos (Bogdan y Biklen, 1982). Aunque es más frecuente utilizar abreviaturas de palabras con las que se han etiquetado las categorías. Miles y Huberman (1984b) desaconsejan usar números y, con el fin de permitir que los investigadores recuperen lo más rápidamente posible las unidades de información sin tener que trasladar el código al concepto que representa, propone la utilización de nombres estrechamente relacionados con el concepto original que describen.

Para Miles y Huberman (1984b:56), el código es "una abreviación o símbolo aplicado a un segmento de texto -muchas veces una frase o un párrafo de notas de campo transcritas- con el fin de clasificarlo". Estos autores señalan que la codificación puede llevarse a cabo en diferentes momentos de la investigación. Distinguen códigos *descriptivos* (atribuyen una unidad a una clase de fenómenos), que son utilizados en un primer momento, y códigos con un mayor contenido inferencial (*interpretativos* y *explicativos*) utilizados posteriormente.

Ambos procesos -categorización y codificación- giran en torno a una operación fundamental: la decisión sobre la asociación de cada unidad a una determinada categoría. La categoría es definida por un constructo mental al que el contenido de cada unidad puede ser comparado, de modo que quede determinada su pertenencia o no a esa categoría.

Los sistemas de categorías

Las categorías utilizadas en un estudio pueden establecerse a priori, de acuerdo con el marco teórico y conceptual previo, a las cuestiones o hipótesis que guían la investigación, a las categorías ya usadas en estudios de otros investigadores o a los instrumentos de investigación empleados (el uso de cuestionarios o guiones de entrevista, por ejemplo, permite utilizar las cuestiones planteadas como fuente de categorías útiles para organizar los datos).

Ha habido intentos de proporcionar sistemas de categorías válidos para múltiples estudios, como el ampliamente difundido sistema de categorías de Bales para el análisis de conductas lingüísticas en un grupo (ver Anguera, 1978). Los sistemas de observación sistemática, mediante los cuales el observador recoge

información sobre actividades o fenómenos usando un conjunto predeterminado de categorías, pueden ser aplicados como herramientas analíticas para la categorización de registros en audio, video y también transcripciones de discursos, elaborados a partir de las situaciones observadas (McIntyre y Maclean, 1986).

El partir de un sistema de categorías a priori no es una condición suficiente, sino un requisito previo para conseguir una misma codificación en distintos codificadores. Es preciso además una definición operativa de categorías, ofreciendo reglas que especifiquen los aspectos del contenido que deben tomarse como criterio de pertenencia a una categoría (Cartwright, 1978).

Apoyadas en este planteamiento a priori se han diseñado aplicaciones informáticas para el análisis de datos textuales capaces de realizar el proceso de codificación de modo automático, aunque con la limitación de que las categorías deben estar definidas por la aparición de un término o frase en los datos, que señalen la relevancia del fragmento para una categoría (Tesch, 1990). La codificación automática eliminaría los efectos de las características personales y circunstanciales del codificador, las cuales afectan al proceso de reducción de datos (Molly, 1984). Características físicas y mentales, motivación, mecanismos de memorización, estímulos o factores ambientales, como temperatura, iluminación, etc. están detrás de los errores cometidos por el codificador (Franzosi, 1990).

Los sistemas de categorías pueden elaborarse inductivamente a partir de los propios datos. Al examinar los datos nos preguntaríamos por el tópico capaz de cubrir cada unidad. Posteriores comparaciones entre los datos agrupados bajo un mismo tópico o entre tópicos permitirían refinar las categorías emergentes (Glaser y Strauss, 1967). En el proceso denominado "codificación abierta" (Strauss, 1987), se parte de la búsqueda de conceptos que traten de cubrir los datos. Tales conceptos son provisionales al principio. El investigador busca línea a línea o párrafo a párrafo indicadores empíricos consistentes en conductas y sucesos observados y descritos en documentos, y en las palabras de entrevistados e informantes, se pregunta qué categoría indica un determinado incidente, y asigna un nombre de código provisional, que es continuamente verificado al examinar los restantes datos.

En muchos casos, el sistema de categorías deja de ser únicamente una herramienta que utiliza el investigador para organizar sus datos y se constituye por sí mismo en un producto del análisis, un resultado de investigación. Para Spradley (1979, 1980), uno de los objetivos es precisamente llegar a determinar las categorías (dominios, en la terminología de este autor) en que se clasifican los términos de una cultura y reducir de ese modo la complejidad de la experiencia humana. Su interés

está en descubrir cómo las personas organizan su conocimiento, por lo que las categorías son un fin en sí mismas; no tiene sentido imponer categorías analíticas a los datos obtenidos durante el estudio.

Por lo general, procedimientos deductivos e inductivos se combinan en el trabajo del analista, de modo que a partir de unas categorías a priori se producen modificaciones para adaptarlas al conjunto de datos a que están siendo aplicadas.

Tanto si los códigos son definidos a partir de los datos como si son preelaborados en función del estudio o tomados de sistemas generales, codificación y categorización no se desarrollan linealmente, sino que son procesos recurrentes, durante los cuales el sistema de categorías puede ser alterado.

“En el proceso de codificación el investigador descubre normalmente que las categorías desarrolladas en principio no son siempre adecuadas. Pueden necesitar ser renombradas, modificadas en su contenido o sustituidas por otras nuevas”(Tesch, 1990:91).

A medida que el investigador posee un mayor conocimiento del material que analiza, las reglas de codificación pueden alterarse inconscientemente mientras se lleva a cabo ésta. Cuando surjan incongruencias derivadas de la aplicación inconsistente de códigos, habrá que revisar lo realizado hasta el momento. Durante la codificación es posible apercibirse de que determinados códigos podrían haber sido agrupados en uno solo o, por el contrario, que cabría dividir uno de ellos en otros diversos.

Los sistemas de categorías no son únicos. Mediante las categorías no buscamos sino extraer distinciones y nombrarlas, y esas distinciones pueden hacerse en varias dimensiones independientes (Bliss, Monk y Ogborn, 1983). Según el aspecto en que nos fijemos, un texto puede ser categorizado de distinto modo.

Autores que ven en la categorización un modo de transformar los datos textuales en datos susceptibles de medición y tratamiento cuantitativo (Ander-Egg, 1990; Cartwright, 1978; Kerlinger, 1975; Sánchez Carrión, 1985, etc.) demandan de los sistemas de categorías una corrección lógica que se alcanza con los requerimientos clásicos de:

- *Exhaustividad* de las categorías. Cualquier unidad debe poder ser ubicada en alguna de las categorías. Es decir, el sistema de categorías es capaz de cubrir todas las posibles unidades diferenciadas en los datos.

- *Exclusión mutua*. Cada unidad se incluye en una sola categoría. Un segmento de texto diferenciado no puede pertenecer simultáneamente a más de una categoría.

- *Unico principio clasificadorio*. Las categorías deben estar elaboradas desde un único criterio de ordenación y clasificación.

Algunos autores añaden otras características a los sistemas de categorías correctamente construidos (Mucchielli, 1988):

- *Objetividad*. Las categorías deben resultar inteligibles para distintos codificadores, de forma que la mala interpretación del contenido de las categorías no dé lugar a una cierta inconsistencia intercodificadores.

- *Pertinencia*. Las categorías habrían de ser relevantes en relación a los objetivos del estudio y adecuadas al propio contenido analizado.

Sin embargo, concepciones más recientes en la investigación en ciencias humanas abogan por categorías que no están claramente delimitadas unas de otras, con superposición entre las áreas temáticas que cubren. Se habla en estos casos de *membreía* en una categoría, para expresar en qué grado determinada unidad puede formar parte o ser miembro de ella, y de la noción de *prototipo*, que representa el tipo ideal con un cien por ciento de pertenencia a una sola categoría. Tal elemento no existe realmente; es un constructo mental que sirve como punto de referencia al que comparar las unidades de datos reales, para determinar su grado de *membreía* (Tesch, 1990).

c) *Agrupamiento*

La identificación y clasificación de elementos están estrechamente unidas al agrupamiento. Cuando categorizamos estamos ubicando diferentes unidades de datos bajo un mismo tópico o concepto teórico. La categorización supone en sí misma una operación conceptual de agrupamiento.

En el caso del análisis de contenido automático se ha recurrido a dos procedimientos para llegar a configurar grupos de palabras que podrían verse como categorías o conceptos: los enfoques a priori y los enfoques empíricos.

Los *sistemas a priori* se basan en la búsqueda de palabras en los textos identificadas con conceptos teóricos relevantes para el problema investigado

(Mochmann, 1985). Es necesario definir previamente la relación entre conceptos y palabras, creando lo que se denomina un diccionario de análisis de contenido. El ordenador se encargará de localizar todas las palabras que se agrupan bajo cada concepto y ofrecer sus frecuencias de aparición en las unidades consideradas para permitir cálculos estadísticos. Existen diccionarios generales, apoyados en teorías de gran alcance, y también diccionarios especializados o incluso específicos de cada estudio (Mohler, 1985).

Otros enfoques, denominados *empíricos*, clasifican automáticamente las palabras (excluyendo las palabras funcionales: artículos, conjunciones, etc.) sin utilizar categorías preestablecidas, y evitan así toda influencia del investigador en el agrupamiento resultante. Las categorías surgen mediante procedimientos estadísticos como análisis factoriales o análisis de conglomerados realizados sobre matrices de correlaciones obtenidas a partir del texto (Iker y Harway, 1969).

Las tareas revisadas hasta aquí -segmentación en unidades, codificación/categorización, agrupamiento- constituyen modos de contribuir a la reducción de datos que en el fondo presupone todo análisis (se parte de un conjunto amplio y complejo de datos para llegar a elementos más manejables que permitan establecer relaciones y sacar conclusiones).

Pero es posible también advertir la función reductora incluso antes de contar con los datos. Miles y Huberman (1984b) consideran que puede darse una reducción de datos anticipada cuando focalizamos y delimitamos la recogida de datos, aspectos constituyentes de una forma de preanálisis que comporta descartar determinadas variables y relaciones y atender a otras. De alguna manera, los datos recogidos suponen un conjunto conscientemente reducido del universo de datos que podrían haberse obtenido.

La reducción de datos es común a los procedimientos cualitativos de análisis y a las estrategias que recurren a la cuantificación. Preside procesos de análisis estadístico tales como los análisis factoriales o de conglomerados, citados más arriba, o la construcción de índices representativos de un grupo: medidas de tendencia central, frecuencias, porcentajes, y está presente también cuando el investigador resume o esquematiza sus notas de campo o cuando selecciona, entre los datos obtenidos, aquéllos que considera relevantes para sus cuestiones de investigación.

3.2. LA DISPOSICIÓN DE DATOS

Una disposición es un conjunto organizado de información, presentada en alguna forma espacial ordenada, que permite extraer conclusiones. Los datos cualitativos aparecen generalmente en forma textual y resultan difíciles de abarcar por las capacidades cognitivas humanas. Además, los textos narrativos resultan dispersos, ofrecen a veces una visión secuencial de hechos simultáneos, dificultan el examen de varias variables a la vez y están vagamente ordenados (Miles y Huberman, 1984b). Por esta razón, es necesaria una *transformación y ordenación* de los datos que permita presentarlos de manera abaricable y operativa de cara a resolver las cuestiones de investigación. La transformación supone presentar los datos en algún otro lenguaje o forma de expresión.

Los procesos de reducción de datos llevan a simplificar la información, pero ésta debe ser mostrada o dispuesta en algún modo que posibilite su procesamiento posterior y facilite la obtención de conclusiones. Los procedimientos para la disposición de los datos difieren en los análisis que recurren a la cuantificación y los análisis cualitativos. La disposición de datos en *análisis que recurren a la cuantificación* suele adoptar la forma de una matriz numérica donde los valores de cada celdilla corresponden, en la mayoría de las ocasiones, a las frecuencias alcanzadas por las diferentes categorías en cada unidad considerada para el estudio. La operación básica es el cómputo.

Modalidades de tablas numéricas serían los tres tipos diferenciados por Lebart, Morineau y Fénelon (1983):

- *Tablas de contingencia*, en las que una misma población puede ser objeto de dos particiones, que se sitúan en filas y columnas respectivamente. En la intersección de éstas aparecerá el número de sujetos o casos que pertenece simultáneamente a dos partes consideradas. Estas tablas tienen por finalidad "estudiar la relación entre dos o más variables de tipo nominal u ordinal" (Sánchez Carrión, 1989:23).

- *Tablas de incidencia*, o también denominadas tablas de presencia-ausencia. En éstas, los sujetos y variables son trasladados a filas y columnas; según un individuo u objeto presente o no determinada característica o atributo, se consigna 1 ó 0 en la celda intersección correspondiente.

- *Tablas de valores numéricos*, recogiendo en las intersecciones el valor de las variables ubicadas en las columnas para los individuos colocados en las filas.

En los estudios lexicométricos, se transforma el corpus textual en *tablas léxicas*, a partir de las cuales desarrollar posteriores análisis estadísticos (Lebart y Salem, 1988). Estas tablas recogen ordenadamente la información sobre ocurrencia de las formas gráficas -palabras-, situando en las columnas las partes consideradas en el texto (momentos diferentes, emisores diferentes, etc.) y en las filas las formas del corpus textual o sólo aquéllas que rebasan una determinada frecuencia. Los términos son colocados en orden lexicométrico: frecuencias o longitudes decrecientes, orden alfabético.

La presentación de datos en forma numérica puede también darse una vez que se ha llevado a cabo un cierto procesamiento estadístico de la información. Las tablas de frecuencia suelen emplearse para mostrar porcentajes, proporciones, desviaciones sobre valores centrales, etc. Otras tablas o matrices numéricas son empleadas para presentar ordenadamente datos tales como correlaciones, pesos y puntuaciones factoriales, distancias entre sujetos, etc. obtenidos tras diversos análisis.

Tanto si se trata de datos directos como de datos transformados, las tablas cumplen una función analítica. Su finalidad no es almacenar datos, sino comunicarlos en algún modo ordenado que permita extraer conclusiones (Wainer, 1992).

Los *procedimientos gráficos*, como medio de presentar información que ha podido ser objeto de un mayor o menor grado de elaboración, son tradicionalmente empleados en los enfoques cuantitativos del análisis de datos y, en particular, del análisis de datos cualitativos.

Los histogramas, curvas de nivel, diagramas de sectores circulares son utilizados para describir gráficamente distribuciones de frecuencias. Gráficos de caja, diagramas de dispersión y otros diversos procedimientos de presentación son inseparables de las técnicas cuantitativas. Técnicas como el análisis factorial de correspondencias, el escalamiento multidimensional o el análisis de conglomerados conllevan representaciones gráficas que muestran la situación de los elementos respecto a ejes factoriales, su posicionamiento en un espacio geométrico, o el modo en que se clasifican o agrupan (dendograma). Los gráficos permiten no sólo presentar los datos, sino advertir relaciones y descubrir su estructura profunda (Wainer, 1992).

La comprensión e interpretación de los datos y la extracción de conclusiones se ve facilitada con estos modos de transformar y presentar gráficamente la información inicial. Incluso los procedimientos gráficos pueden constituir el eje central del análisis de los datos (Bertin, 1987).

Los *procedimientos de disposición usados en análisis cualitativos* difieren algo de los presentados hasta ahora. Las vías cualitativas, por las que son a veces analizados los datos cualitativos textuales, pueden excluir el uso de índices numéricos o frecuencias y recurrir a procedimientos verbales o gráficos.

El uso de las técnicas de modelado gráfico se ha extendido en el análisis y manipulación de datos cualitativos (Dobbert, 1982). Un buen ejemplo podemos encontrarlo en Spradley (1979,1980) quien, en cada etapa del proceso de análisis que propone, emplea diversas formas de presentar la información obtenida -perfil general, diagrama de caja, diagrama en árbol, paradigmas-, facilitando la aplicación de las modalidades analíticas subsiguientes.

Los diagramas, definidos como "representaciones gráficas o imágenes visuales de las relaciones entre conceptos" (Strauss y Corbin, 1990:198) constituyen importantes instrumentos de análisis. La elaboración de *diagramas esquemáticos*, mediante los cuales se visualicen las relaciones entre taxonomías, es vista por Spradley (1979,1980) como una de las mejores estrategias para el descubrimiento de temas culturales, que conectan los diferentes dominios de una cultura.

Miles y Huberman (1984b) han dedicado especial atención a describir diferentes procedimientos para la presentación de datos cualitativos. Contemplan el diseño de *matrices* que pueden albergar diferentes tipos de información (expresada por fragmentos de texto, citas, frases, abreviaciones o figuras simbólicas) y adoptar distintos formatos.

Generalmente, las matrices consisten en tablas de doble entrada en cuyas celdas se aloja una breve información verbal, de acuerdo con los aspectos especificados por filas y columnas. Es frecuente construirlas para sintetizar los fragmentos codificados en una misma categoría (columna) para diferentes sujetos, situaciones, casos, etc. (filas). Algunas de las dimensiones a partir de las cuales se elaboran pueden implicar la ordenación de fenómenos o componentes de fenómenos a lo largo del tiempo, en función de los roles desempeñados, etc. Determinadas matrices presentan procesos y productos, relaciones de causa-efecto, agrupamientos conceptuales o empíricos de datos.

Bliss, Monk y Ogborn (1983) proponen un *sistema de redes* ("networks") mediante las cuales tratan de clarificar y presentar las distinciones que delimitan a las diversas categorías consideradas en el análisis del material cualitativo, las cuales a su vez pueden estar subdivididas. Para ello usan una notación precisa, con la que indican nombres de categorías, subdivisiones de éstas, diferencias, circunstancias, restricciones, etc.

La presentación de un mismo corpus de datos puede hacerse en múltiples formas, según el objetivo que el analista persiga con ello. A veces, el interés es exclusivamente descriptivo, tratando de ofrecer la información disponible y permitir una visión de conjunto que lleve a análisis o presentaciones más detalladas.

En la mayoría de los casos, el analista presenta junto a los datos ciertos resultados o conclusiones. La ordenación de los datos según criterios cronológicos, causales, de inclusión, etc. implica una interpretación de los mismos y, en cierto modo, constituye un producto del análisis. La elaboración de figuras y matrices supone pensar y reflexionar sobre los datos. El informe narrativo de los resultados podría ser considerado, en sí mismo, una forma de disposición.

3.3. OBTENCIÓN Y VERIFICACIÓN DE CONCLUSIONES

a) Obtención de resultados y conclusiones

Un importante grupo de actividades constituyentes de los procesos de análisis se dirige a la obtención de resultados y conclusiones, en relación a los objetivos particulares de cada investigación.

Llegar a conclusiones implicaría, de acuerdo con las connotaciones semánticas del término análisis, ensamblar de nuevo los elementos diferenciados en el proceso analítico para reconstruir un todo estructurado y significativo. Bajo la denominación de conclusiones aparecen generalmente los resultados, los productos de la investigación y la interpretación de los mismos.

Considerando que la finalidad de la investigación educativa es "identificar y entender el significado que tiene la educación y desarrollar teorías que expliquen y resuelvan los problemas educativos" (García Jiménez, 1990:56), los resultados de un estudio se dirigirán a tales fines, es decir, avanzarán en la explicación, comprensión y conocimiento de la realidad educativa y contribuirán a la teorización

sobre la misma. Un sólo estudio difícilmente produciría una teoría social completa, pero sí puede desarrollar un grupo de proposiciones teóricas.

Los procedimientos estadísticos de carácter descriptivo, exploratorio, arrojan como resultados índices, agrupamientos, representaciones gráficas que permiten profundizar en la comprensión de los datos y de la forma en que se estructuran. La investigación de corte cualitativo se preocupa especialmente por el significado que los participantes atribuyen a sus prácticas y a las situaciones educativas en las que se desarrollan. Los resultados o conclusiones de esta investigación engloban toda una serie de decisiones del investigador sobre el significado de las cosas, constatación de regularidades, modelos, explicaciones, configuraciones posibles, flujos causales y proposiciones (Miles y Huberman, 1984b) a que el investigador llega, y cuya certeza intenta comprobar.

Van Maanen (1985) denomina a las conclusiones "conceptos de segundo orden", pues se construyen a partir de los datos, o conceptos de primer orden, es decir, a partir de las propiedades estudiadas en el campo y las interpretaciones que hacen los propios participantes. Estos conceptos de segundo orden son nociones usadas por el investigador para explicar el modelo que configuran los datos, y consisten en afirmaciones acerca de las relaciones entre ciertas propiedades observadas o en interpretaciones sobre las interpretaciones que hacen los sujetos participantes.

Las conclusiones no suelen limitarse a la presentación ordenada de los datos adecuadamente reducidos, e implica mayores niveles de inferencia. Para Goetz y Lecompte (1988), las conclusiones de un estudio tendrían que superar los límites de la mera descripción y aproximarse a una interpretación e integración teórica, explicitando los vínculos contextuales y las generalizaciones encontradas por el investigador. El dato sólo adquiere significación dentro de un contexto. La importancia del contexto en la obtención de conclusiones queda reflejada cuando se habla de "relativización de los datos" (Taylor y Bogdan, 1986) para comprenderlos e interpretarlos en el propio contexto donde fueron recogidos, o de buscar el "significado contextual" (Van Maanen, 1985) de cada elemento de información para determinar qué representa.

Las conclusiones son extraídas por el investigador durante todo el proceso de recogida de datos y durante el análisis (Miles y Huberman, 1984b). Las tareas de reducción y presentación de datos le permiten hacer afirmaciones que progresivamente avanzan desde lo descriptivo a lo explicativo y desde lo concreto a lo

abstracto. Estas afirmaciones tendrán un carácter provisional; son tentativas, no definitivas.

Algunos autores opinan que establecer relaciones y conexiones importantes puede resultar favorecido si entre la finalización del trabajo de campo y la reflexión teórica del investigador media un cierto espacio de tiempo (Goetz y Lecompte, 1988). En este caso, tras un informe descriptivo preliminar, se sucederían informes con un nivel creciente de abstracción.

Las conclusiones son, por tanto, afirmaciones, *proposiciones en las que se recogen los conocimientos adquiridos por el investigador en relación al problema estudiado*. Dado que el análisis persigue separar las unidades elementales de un fenómeno para reconstruir un todo significativo, las conclusiones suelen recoger la relación, más o menos compleja, encontrada entre dos o más de estos elementos o variables.

Las categorías obtenidas en el proceso de reducción son en sí mismas conclusiones del estudio y, como se señaló más arriba, la ordenación de los datos en una matriz, un cuadro, una figura, etc. presupone algún tipo de conclusiones respecto a su estructura.

La mayoría de los estudios suelen implicar conclusiones sobre el modo en que se relacionan los elementos o variables observadas, sobre el tipo de conexiones que conforman una "configuración significativa".

Cuando se recurre a procedimientos interpretativos para el análisis de datos cualitativos, lo habitual es que el modo en que se llega a las conclusiones, es decir, el proceso lógico aplicado para extraer conocimiento sobre la realidad educativa en estudio, sea el aspecto más oscuro del análisis en los informes de investigación. Los investigadores describen detalladamente cómo recogen los datos y, tal vez, cómo proceden en el análisis reduciendo o reestructurando y presentando la información, pero no explicitan las herramientas conceptuales y lógicas que les llevan a sus conclusiones. La lectura de una gran parte de los trabajos que se publican sugeriría que las conclusiones emergen directamente a partir del corpus de datos, sobre todo una vez que éste ha sido adecuadamente reducido y presentado.

Sin duda, una de las principales herramientas intelectuales en el proceso de obtención de conclusiones es la comparación. La comparación permite destacar las semejanzas y diferencias entre las unidades incluidas en una categoría, y hace posible la formulación de sus propiedades fundamentales, a partir de las cuales

puede llegarse a una definición, ilustración y verificación de esa categoría (Fielding y Fielding, 1986). La comparación facilita llegar a conclusiones a partir de las matrices y otras figuras usadas para la presentación de los datos. La comparación de filas o columnas de una matriz, por ejemplo, constituye la principal vía para extraer conclusiones (Miles y Huberman, 1984b). Las conclusiones pueden emanar de la comparación con otros escenarios, casos, situaciones, etc. similares al estudiado. Tales comparaciones son necesarias si se desea extraer conclusiones respecto a diferentes casos o generalizar y aplicar las conclusiones obtenidas a otros casos.

Las conclusiones se dirigen igualmente hacia la contextualización del estudio. Una vez alcanzados ciertos resultados es necesario integrarlos con los de estudios anteriores, poner de manifiesto cómo se integran en unas teorías más amplias, en qué medida contribuyen al cuerpo de conocimientos sobre un tema. Goetz y Lecompte (1988) proponen algunas estrategias para la interpretación de los resultados y su integración en marcos más amplios:

- *Consolidación teórica*. Consiste en confrontar los resultados obtenidos con los marcos teóricos desarrollados originalmente, modificando éstos para intentar que se ajusten más a los datos. Este enfoque ha sido ampliamente desarrollado por Glaser y Strauss (1967), quienes han defendido la utilidad de generar teoría fundamentada en los datos.
- *Aplicación de otras teorías*. La búsqueda de otros marcos más generales en los que integrar los datos permite la generalización de los resultados.
- *Uso de metáforas y analogías*. Son medios útiles para establecer conexiones entre temas aparentemente no relacionados, o conectar los resultados a la teoría (Miles y Huberman, 1984b). Preguntarse "¿a qué me recuerda esto?" acerca de los diferentes aspectos estudiados puede ser una estrategia adecuada para encontrar analogías y metáforas (Bogdan y Biklen, 1982:153).
- *Síntesis de los resultados con los obtenidos por otros investigadores*, para lo que es necesario comparar, contrastar, agregar, ordenar, establecer relaciones y especular. Es decir, se repite el proceso analítico pero esta vez con los resultados alcanzados y, por tanto, a niveles más elevados de abstracción y generalización.

b) Verificación de conclusiones

Una vez alcanzadas las conclusiones de un estudio, es necesario verificar esas conclusiones, es decir, confirmar que los resultados corresponden a la realidad, en el caso de los enfoques positivistas, o a los significados e interpretaciones que los participantes atribuyen a la realidad, en el caso de las investigaciones cualitativas o interpretativas.

El concepto de validez ha sido usado en una variedad de formas en la literatura sobre metodología de investigación. La American Psychological Association (1954) definió cuatro tipos de validez: predictiva, concurrente, de contenido y de constructo. La subdivisión en validez interna y validez externa, introducida por Campbell (1957), ha ido seguida de diferentes formulaciones y desarrollos aparecidos en las obras de Campbell y Stanley (1963) y Cook y Campbell (1979) sobre investigación experimental y cuasi-experimental. Como indicador de la aproximación de las conclusiones de un estudio a la realidad suele tomarse la validez interna, mientras que la validez externa se refiere a las posibilidades de generalizar los resultados a otras poblaciones o contextos.

Los investigadores cualitativos han adoptado distintas posiciones (Howe y Eisenhart, 1990) en relación a los conceptos tradicionales de validez interna y validez externa surgidos en el ámbito experimental. Mientras que la subdivisión establecida por Campbell ha sido adaptada a la investigación cualitativa por autores como Denzin (1978) o Goetz y Lecompte (1988), otros han rechazado el concepto convencional de validez (Erickson, 1989; Guba y Lincoln, 1981), basándose en argumentos epistemológicos que distancian a los enfoques positivista-experimental e interpretativo.

Lecompte y Goetz (1982) adoptan los conceptos de fiabilidad interna y externa, asociados con la replicabilidad de los resultados, y validez interna y externa, relacionados con su corrección, señalando los factores por los que se ven afectados. Al igual que estos autores, Guba y Lincoln (1981) engloban la cuestión de la validez en el problema más general de la credibilidad de la investigación, pero proponen criterios alternativos, aunque análogos, a los de Campbell para valorarla. Estos autores consideran que la credibilidad está relacionada con el valor de verdad de los descubrimientos y con otras tres preocupaciones fundamentales: su aplicabilidad a otros contextos o sujetos, la consistencia de los resultados y la neutralidad de los descubrimientos frente a inclinaciones, motivaciones, intereses, concepciones teóricas del investigador.

Cualquiera que sea la posición adoptada, todos los autores suelen aludir a aspectos de calidad de la investigación cuando abordan el tema de la validez. Pero, aunque es posible considerar la validez del diseño de investigación, de la definición de variables, de los procedimientos de recogida de datos o de los datos obtenidos, lo más apropiado es hablar de ella en relación a los hallazgos o conclusiones de la investigación (Hammersley, 1991).

Verificar las conclusiones de un estudio significa, por tanto, comprobar el valor de verdad de los descubrimientos realizados, o lo que es igual, comprobar su validez.

Desde posiciones racionalistas, la validez quedaría determinada si se consigue demostrar un isomorfismo entre los datos con que trabaja el investigador y las realidades o fenómenos que representan. Puesto que ello no es posible, salvo con un conocimiento absoluto del mundo, los racionalistas recurren a extraer explicaciones alternativas, de modo que las hipótesis puedan ser refutadas mostrando que la hipótesis alternativa es, con una probabilidad alta, correcta (Guba, 1983). Los procedimientos estadísticos basados en modelos probabilísticos suelen recurrir a este tipo de comprobaciones para establecer la significación de sus resultados, aceptando comúnmente como criterio un cinco por ciento de máxima probabilidad de que la hipótesis alternativa sea cierta. Además tratan de favorecer la obtención de resultados válidos mediante procedimientos técnicos de aleatorización o el control de variables para evitar las fuentes de error. Pero como advierte Glass (1981), incluso el más cuidado diseño deja abierta una puerta a juicios del investigador en su interpretación, por lo que no quedaría completamente garantizada la validez interna de los resultados.

Cuando los datos cualitativos son analizados siguiendo procedimientos interpretativos, la estimación de la validez es menos precisa que en la investigación cuantitativa, basándose en juicios sobre la correspondencia entre los hallazgos y la realidad. La validez de los resultados puede incrementarse, o simplemente ser calibrada, de acuerdo con estrategias desarrolladas en las fases de diseño, recogida de datos, análisis de datos o redacción del informe (Dawson, 1982). Mencionaremos aquí algunas de las más comunes:

- Por el hecho de estar presente, el investigador provoca en los participantes conductas que ordinariamente no habrían ocurrido y es, al mismo tiempo, influenciado por ellos. La *presencia prolongada en el campo* hace que el investigador forme parte de la situación estudiada. lo cual le permite soslayar los efectos de su presencia, comprobar los prejuicios y percep-

ciones propios o de los participantes, y comprender lo que es esencial o característico (Guba, 1983), haciendo posible el ajuste entre las categorías científicas extraídas y la realidad.

- El *intercambio de opiniones con otros investigadores*, exponiéndose a sus preguntas y críticas puede ser útil para llegar a conclusiones válidas. La publicación de resultados constituye un modo de presentarlos a la revisión por parte de otros expertos (Goetz y Lecompte, 1988).

- La *triangulación* es un procedimiento ampliamente usado que consiste en el uso de diferentes fuentes de datos, investigadores, perspectivas o metodologías para contrastar los datos e interpretaciones (Denzin, 1978). Si diversas clases de datos conducen a la misma conclusión podemos estar algo más seguros de tal conclusión (Hammersley y Atkinson, 1983). Sin embargo, se han puesto objeciones a la triangulación en el caso en que métodos de diferentes perspectivas metodológicas son usados conjuntamente a pesar de responder a asunciones ontológicas y epistemológicas diferentes (Blaikie, 1991).

- La *comprobación con los participantes* supone contrastar los datos e interpretaciones de los mismos con los sujetos que constituyen la fuente de esos datos. En tanto que profundos conocedores de la realidad que se investiga, los participantes podrían actuar como jueces que evalúen los principales descubrimientos de un estudio (Denzin, 1978). Análogamente, pueden realizarse esas comprobaciones de los resultados frente a datos nuevos procedentes de otras fuentes (documentos, grabaciones, etc.) recogidos con tal fin durante el estudio, estrategia denominada por Guba (1983) *establecimiento de adecuación referencial*.

- Miles y Huberman (1984b) proponen la *ponderación de la evidencia* como técnica que incrementaría la validez de los resultados. Su idea es la de considerar ciertos datos mejores que otros a la hora de extraer conclusiones, ya sea por proceder de determinados informantes, por las circunstancias en que fueron recogidos (al final o al principio de la estancia en el campo, revelados en privado o en presencia de otros informantes, etc.) o por haber sido validados mediante estrategias como la triangulación o la comprobación con los participantes.

- Debe ser comprobada la *coherencia estructural* para asegurarse de que entre los datos e interpretaciones no se dan contradicciones o incoherencias

(Guba, 1983). Un modo de aproximarse a ella es la *búsqueda de evidencia negativa* (Miles y Huberman, 1984b), consistente en preguntarse, una vez alcanzadas las conclusiones, si hay datos que se oponen o son inconsistentes con esas conclusiones.