La recogida de datos de la EPA 2005

Carlos Ballano
subdirector General de Recogida de Datos. INE

El 1 de enero de 2005 se han introducido dos cambios importantes en la recogida de datos de la Encuesta de Población Activa; uno es la modificación del cuestionario y el otro la creación de centros CATI (computer assisted telephone interviewing) para la recogida de las entrevistas sucesivas.
En este artículo describiremos estos cambios y expondremos las razones de los mismos.

Modificación del cuestionario

Se consideró que el primer trimestre de 2005, coincidiendo con las otras modificaciones que era preciso introducir en la encuesta, entre las que se encontraba la necesidad de incluir preguntas sobre seis nuevas variables, era el momento propicio para modificar el cuestionario.

La finalidad perseguida era adecuarlo mejor al proceso de la entrevista y hacer más fluida la relación entre el entrevistador y el entrevistado, de forma que mejorara la comprensión por parte del informante de las preguntas y, como consecuencia, mejorara la estandarización de la encuesta, al requerirse con menos frecuencia a los entrevistadores explicaciones sobre el significado de las preguntas.

El trabajo desarrollado con el cuestionario se organizó en tres fases que vamos a describir brevemente a continuación.

Primera fase

El primer paso en una tarea de este tipo consiste en una revisión en profundidad del cuestionario existente, realizada en colaboración por personal con experiencia en la elaboración de cuestionarios y en el contenido y explotación de la EPA. También se tuvieron en cuenta las sugerencias recibidas desde las Delegaciones Provinciales del INE. Las principales modificaciones introducidas en esta fase han consistido en:

- Simplificar la formulación de las preguntas: esto, en algunos casos, ha conducido a desglosar una pregunta antigua en varias.
  Un ejemplo de este tipo de actuación lo ofrece la pregunta inicial del módulo de relación con la actividad económica. En el antiguo cuestionario la pregunta con la que se abría dicho módulo era:
  La semana pasada, de lunes a domingo, ¿ha realizado un trabajo remunerado (en metálico o en especie) como asalariado o por su cuenta, aunque sólo haya sido por una hora o de forma esporádica u ocasional? (Opciones de respuesta: sí, no)
  En el cuestionario actual se ha desglosado en dos preguntas: (no se indica el flujo)
  B1. *La semana pasada, de lunes a domingo, ¿trabajó aunque sólo fuera una hora?* (Opciones de respuesta: si, no, no sabe)
  B2. *¿Se le paga por ese trabajo u obtiene un beneficio económico con él?* (opciones de respuesta: si, no, no sabe).
  Al realizarse la recogida de datos asistida por ordenador, el texto de la pregunta cambia la expresión "la semana pasada, de lunes a domingo" por las fechas exactas de la semana a la que está referida la información.

> **"Se pretende no demorar en demasía el tiempo que se tarda en preguntar por los aspectos más centrales de la encuesta"**

- Además se ha modificado el orden de los distintos grupos de preguntas que componen el cuestionario y se ha reestructurado el de bastantes otras. Así, por ejemplo, el módulo sobre enseñanza y formación se pasa después de los correspondientes a la relación con la actividad, características del empleo y búsqueda de empleo, cuando en el cuestionario antiguo iba al principio de la entrevista. De esta forma se pretende no demorar en demasía el tiempo que se tarda en preguntar por los aspectos más centrales de la encuesta.

• Otras modificaciones han sido:
  – que una misma pregunta se formule de distintas maneras en función de respuestas anteriores, lo que es posible gracias a la utilización de cuestionarios informatizados en la Encuesta.
  – la introducción de opciones de respuesta “no sabe”, como se ha podido ver en el ejemplo anterior.
  – se han mejorado y completado las opciones de respuesta a las preguntas y se ha redefinido cuales de ellas son preguntas abiertas y cuales cerradas, etc.

Segunda fase

Es necesario que quede claro que las modificaciones del cuestionario no han implicado un cambio en las definiciones de la encuesta. Para ser conscientes de la posibilidad de modificar el cuestionario sin que cambien los conceptos asociados a la encuesta hay que considerar que la relación que existe entre una variable y la pregunta (o preguntas) que trata de medirla es en cierto modo análoga a la que existe entre la población y la muestra. En efecto dos formulaciones diferentes de una misma pregunta pueden proporcionar diferentes estimaciones del valor de la variable a medir, al igual que dos muestras pueden proporcionar también estimaciones diferentes.

Por tanto, la cuestión a resolver es cómo asegurar que la pregunta formulada mida con una aproximación suficiente el verdadero valor de la variable. Para alcanzar este propósito se desarrolló la segunda fase del trabajo con el cuestionario: se trata de la aplicación de técnicas cognitivas para la evaluación de las preguntas.

Este proceso consiste en aplicar a preguntas previamente seleccionadas por su interés para la encuesta distintas técnicas como entrevistas en profundidad, codificación del comportamiento o grupos de discusión, en las que se analiza, según el tipo de técnica empleada, el proceso que sigue el entrevistado desde que oye la pregunta hasta que verbaliza una respuesta (comprensión e interpretación de los conceptos incluidos en la pregunta, recuperación de la información que se le demanda, juicio interno sobre la conveniencia de proporcionar la información, etc.)

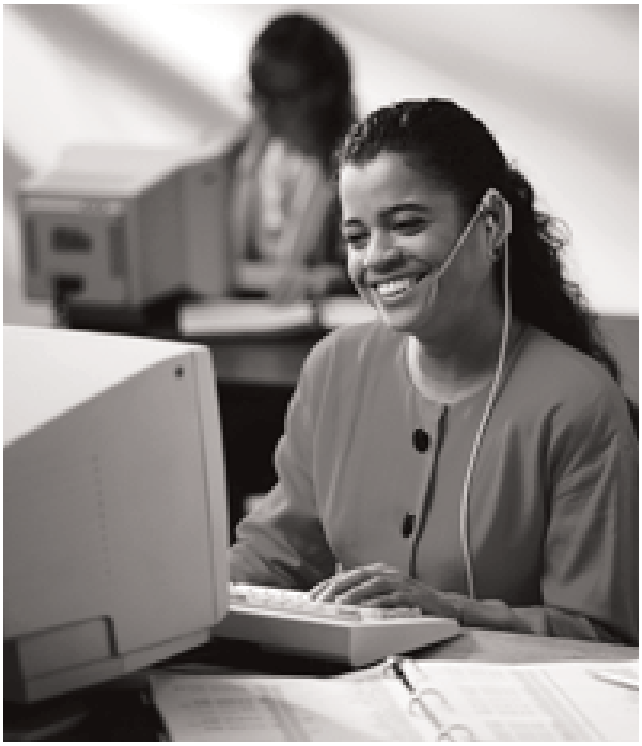
Identificamos dos grupos de preguntas como aquellos en los que era más complejo diseñar el cuestionario a la vista de los conceptos teóricos. Uno era el referido al módulo sobre búsqueda de empleo y el otro el referido al estudio de las horas extraordinarias. A continuación nos referiremos brevemente a ambos.

La definición de la categoría ‘parado’ en la EPA se recoge en el anexo I del Reglamento (CE) Nº 1897/2000. Para acotar el concepto son precisas, en la traducción en castellano del Reglamento, 520 palabras; la mayor parte dedicadas a aclarar lo que significa buscar activamente empleo. En el antiguo cuestionario la pregunta correspondiente a indagar sobre los métodos de búsqueda de empleo contemplaba diecisiete opciones de respuesta, por lo que consideramos necesario replantear el módulo. Para el diseño del nuevo módulo se elaboraron tres versiones alternativas

También se diseñaron dos alternativas de formulación de las variables referidas a horas trabajadas con el objetivo de analizar la mejor manera de medirlas (de preguntar por ellas) y especialmente cómo preguntar por las horas extraordinarias y discernir si los informantes incluyen todas o sólo las pagadas

Todas estas opciones de ambos módulos fueron evaluadas cualitativamente mediante la realización de entrevistas en profundidad y grupos de discusión en las que participaron un total de 200 personas de diversos grupos sociales relevantes para la encuesta por sus características o por el papel de informantes que juegan durante la entrevista (inmigrantes, personas jóvenes iniciando su relación con el mercado laboral, amas de casa, etc.).

Un ejemplo de resultado que se obtiene con estas técnicas es el relativo a la pregunta *¿Cuántas horas semanales trabaja habitualmente en este empleo?* Al analizar las entrevistas en profundidad se comprobó que, a pesar de la redacción de la pregunta, determinados grupos de informantes explicaban que habían contestado las horas que tenían en su contrato o acuerdo con el empresario. La solución adoptada fue la de incluir una pregunta previa, precisamente, sobre las horas pactadas o que figuran en el contrato que, además de completar la información sobre este aspecto de la encuesta, fija este concepto en los informantes de manera que cuando a continuación se les pregunta por las horas habituales hace que contesten con mayor precisión a esta segunda cuestión.



Tercera fase

En paralelo a la aplicación de estas técnicas se desarrolló una encuesta piloto dirigida a más de 3.300 viviendas, llevada a cabo en los meses de julio y septiembre de 2004, que constituyó la tercera fase del trabajo con el cuestionario. En este piloto se trabajó con hasta seis versiones de cuestionario aplicadas aleatoriamente a los hogares bajo las dos modalidades de aplicación (entrevistas personales y telefónicas).

> **“Es necesario que quede claro que las modificaciones del cuestionario no han implicado un cambio en las definiciones de la encuesta”**

Con todo el volumen de información recopilado de las fases que se acaban de describir se decidió el formato final del actual cuestionario de la EPA.

Paralelo a todo el proceso descrito se creó un grupo de trabajo con usuarios de la encuesta, en el que estaban presentes representantes de los agentes sociales (empresarios y sindicatos), universidades y organismos de la administración, a los que se ha informado de los cambios previstos y que han realizado aportaciones sobre los mismos.

Creación de centros CATI

La otra modificación de la recogida de datos que ha tenido lugar en el primer trimestre de 2005 es la creación de siete centros CATI (Computer Assisted Telephone Interviewing) desde los que se ha centralizado la recogida de las entrevistas sucesivas en aquellos hogares que disponen de teléfono. Para las primeras entrevistas no ha habido modificación del procedimiento de recogida; continúan, por tanto, recogiéndose por visita personal asistida mediante un ordenador portátil (método CAPI: Computer Assisted Personal Interviewing) y en ellas se solicita el teléfono, necesario para la realización de las entrevistas sucesivas desde los centros CATI.

Las entrevistas sucesivas ya se hacían por teléfono desde cada una de las 52 delegaciones desde el año 1999. Los entrevistadores utilizaban para la administración de estas encuestas el mismo portátil con el que se realizaban las entrevistas personales. Con el nuevo sistema se ha producido, por tanto, una especialización de los entrevistadores según que las entrevistas sean por visita personal o por teléfono.

Los principales cambios debidos a la introducción del sistema CATI son:

- Es el sistema informático el que gestiona la asignación de las entrevistas a los entrevistadores de cada centro CATI. Para ello se tiene en cuenta que se realicen distintos intentos de contacto en distintas bandas horarias y días. El número medio de intentos de contacto por vivienda finalmente ausente en el primer trimestre de 2005 ha sido de alrededor de 200. Las viviendas no se asignan a entrevistadores concretos, excepto en el caso de entrevistas interrumpidas.
- Los centros están abiertos de lunes a viernes desde las nueve de la mañana hasta las nueve de la noche, de forma ininterrumpida.
- Hay una mayor cercanía entre los entrevistadores y los responsables del diseño de la operación y del control de la recogida. La posibilidad de escuchar directamente las entrevistas proporciona mucha información que se utiliza para la mejora y estandarización del proceso.

Para la realización de la prueba piloto ya descrita se creó un centro CATI que permitió, entre otras cosas, la estimación de las tasas de no respuesta y tiempos de entrevista asociados al nuevo método.

Conclusión

Como cualquier otra organización de nuestro tiempo el Instituto Nacional de Estadística debe adaptarse a las nuevas técnicas y circunstancias sociales para cumplir con el objetivo de siempre: dar datos de calidad, que permitan la comparación entre territorios y grupos sociales en términos homogéneos. La aplicación de los nuevos métodos, tanto en el desarrollo de los cuestionarios como en la recogida mediante CATI, se van extendiendo en las encuestas dirigidas a los hogares, en la medida que los recursos y las características de las encuestas lo permiten.