Bar coding—a functional tool in survey research

Matti Siivola, Arja Rimpelä, Sakari Karvonen, Matti Rimpelä

Abstract

Study objective—The aim was to assess the value of bar coding, together with developed microcomputer programs, as a tool to speed up mailing and handling of return responses of a survey where over 4000 adolescents received a mailed questionnaire.

Design—Bar codes were used for displaying the identification number of a questionnaire, as well as each response alternative and number of the question. These bar codes were read by a wand, a handheld scanner, connected to a microcomputer and the data were automatically entered into a computer data file.

Results and conclusions—Compared with previous similar surveys, the use of bar codes did not seem to affect a manner of responding. Bar coding saved human steps in handling of return responses, reduced an error rate, and helped to control the process of work. Bar coding is recommended as a new technical tool for preparations of future large scale surveys.

The (Finnish) Adolescent Health and Lifestyle Survey has been biannually mailed to several thousands of young people since 1977. In the 1989 survey, approximately 4000 adolescents received a 12 page, mainly structured, questionnaire containing 87 questions, and bar coding combined with assisting microcomputer programs was adopted as a tool to speed up mailing and handling of return responses. Experience with this method has been entirely positive, and we would like to explain an adaptation of this technique here. Based on the success, we recommend that researchers acquaint themselves with this technical tool before beginning preparations for future large scale surveys.

Methods

DATA ENTRY

Bar codes were used for displaying the identification numbers of questionnaires (fig 1), and within each question a bar code displaying a number of a question as well as of a response alternative (fig 2).

As a completed questionnaire was received, an identification bar code was read by a wand, a handheld scanner, connected to a microcomputer. An "approval" beep sound made it clear that the scanning had been successful. The bar codes of response alternatives circled by a

respondent were then scanned question by question in a similar manner. No subprogram for the scanning of bar codes was necessary, because the microcomputer sees the bar code as if it were a keyboard entry. This also provided for the possibility of a traditional keyboard data entry when damaged, torn, or otherwise unscannable bar codes were found.

The scanned identification numbers were automatically entered into a mailing data file by a program which kept a control file as well as updated information on the mailing process. Similarly, response data were automatically entered into a data file by a program which identified the question number and the number of a response alternative from the bar code, and located the data into an appropriate variable. Bar codes exceeding the minimum and maximum limits of each variable were non-existent and could not be scanned.

In order to make the questionnaire reasonable in length and price, it was not possible to print all response alternatives on the questionnaire (eg, illustrating age in years at which alcohol was first tried). In open questions, the response alternatives and their bar codes were presented on a separate paper from which the appropriate code was scanned. Most open questions could be coded as the data were entered.

BAR CODES

Bar codes may be found in different formats. The more dense a bar code is, the more likely it is that even the smallest printing, colouring, or wear error will make it unscannable. Bar codes used for the questions in this survey were selected in such a way that the smallest possible space was used without reducing the scannability. The bar code used for the questions included four digits and was called "interleaved 2/5" (fig 2). The bar code type used for questionnaire identification was code 39. All colours are not scanned evenly well and some of them not at all. Black, dark blue, and dark green colours with white background might be the safest alternatives.

Bar code stickers are sold by several commercial companies, or they may be self produced with an appropriate printer and programs. In 1989, we used purchased bar code stickers and had them attached by a clerk to an otherwise print ready questionnaire. New technology makes it possible to produce a print ready questionnaire including bar codes. With appropriate technology, printing offices can also print a bar code with a consecutive identification number onto a questionnaire.

ATTITUDE TOWARD BAR CODES

Instructions for filling out the questionnaire

Department of Public Health, University of Helsinki,
Töölöntullinkatu 8,
SF-00250 Helsinki 25,
Finland
M Siivola
A Rimpelä
S Karvonen
Department of Public Health, University of Turku, Finland
M Rimpelä

Correspondence to:
Dr A Rimpelä

Accepted for publication
March 1991

Figure 1 Bar code indicating identification number

Figure 2 Example of a question with bar codes set

36 Oletko koskaan polttanut (kokeillut) tupakkaa?

1 en (voit siirtyä kysymykseen 46)

2 olen

mentioned that it had been equipped with bar codes and numbered for data processing. A return percentage appeared to be a little better in this study using bar codes than in a similar survey two years earlier.

Respondents were given instructions to circle the number of an appropriate question response alternative, not the bar code (see fig 2). The number had generally been circled.

A few respondents (approximately 0.5%) had damaged or torn the identification bar code. When the identification number had also been printed onto the questionnaire, the manual keyboard entry could be used in most cases. However, often these questionnaires were also generally messy and answered incompletely, and had to be omitted from the analysis.

Compared with our previous mailings of survey questionnaires, the use of bar codes did not appear to affect the manner in which the questionnaires had been answered.

Discussion

In comparison with the traditional key to disk data entry, bar coding, together with the assisting computer programs, saved human work and reduced error rate.

Since manufacturers suggest that bar code scanning error rate is extremely small (one in two million), no rescanning was necessary. Some values exceeding the minimum or maximum limits of variables were found, but providing the manufacturer's reports on error rates are correct, these would mainly have originated from the manual keyboard entry used in some instances.

Less editing of data was found necessary and less rechecking of questionnaires due to data entry errors was needed. Two kinds of errors were now non-existent. First, no erroneous identification numbers were found. Second, a common earlier error where an entered value got shifted to a previous or next variable, thus causing a

dislocation of following variables, was also avoided. Such a mistake was now prevented by incorporating a question number into the bar code. If question numbers had been manually entered, data entry work would have doubled while in bar coding several numbers can be entered by one hand movement. Bar coding does not need any typing experience and is, therefore, easy to start. A research assistant who checks the questionnaires and does coding may also do data entry. Each questionnaire may be handled just once, whereas traditionally coding and entering data into a computer have been separate steps.

Unlike previous surveys, this one needed no outside people and companies for data entry. The total data collecting and editing processes were controlled by the research group which minimised the chance of delay and guaranteed confidentiality since the questionnaires containing sensitive information could be held in one place under the supervision of the group.

A comparison of costs using scanning of bar codes *v* traditional data entry by experienced processors is not totally possible due to variation in the appearance of questionnaires. Reading an individual bar code may be slower than the traditional entering of data, but the smaller chances for errors and savings in human work level out expenses. It is known, however, that cost per questionnaire is not any higher when using bar codes. In the future, developments in bar coding together with appropriate computer programs may prove to its further advantage.

The monotonous "approval" beep sound produced by the scanner is a disadvantage. However, compared to traditional methods an easier coordination between eyes and fingers reduces errors and makes work easy to start. Also, especially when assisted with computer programs, bar coding made work more attractive and assisted in finding workers interested in their assignments.

Bar coding combined with assisting mailing and data entry programs speeded up the total work process when less work was needed for editing the data. We were able to have the first press conference with a 16 page report within a week after the last questionnaire had been received. Thus experiences have been positive and our next survey will use bar coding.

The researchers wish to thank Dr Ansa Ojanlatva, Senior Researcher at the Adolescent Health and Life Style Survey, University of Turku, for her suggestions and for having translated this material into English.