INCENTIVES, MOTIVES, AND RESPONSE BIAS

BY CHARLES F. CANNELL AND RAMON HENSON
(WITH THE ASSISTANCE OF JUDITH HYBELS)

The effectiveness of financial incentives is considered in relation to findings on response bias. Three general patterns of response bias are discussed: effects of elapsed time between the event and the interview, the importance or salience of the event for the person, and social desirability effects. The major interpretation is that the biases are associated with patterns of respondent motivation. A model of response problems and three types of respondent motivation are discussed. Incentives, especially monetary rewards, are described in this framework. It is concluded that payment to respondents can be effective but only if used properly and an appropriate amount chosen. Remuneration presents more complex issues than are usually anticipated by the researcher.

Rather than discuss specific effects of financial incentives on survey responses, we propose to consider the issue of incentives in somewhat broader terms. We start by summarizing research findings on response bias and discuss some of the concepts of motivation relevant to these findings.

In the late 1950's the Inter-University Committee for Research on Consumer Behavior was established to investigate problems of obtaining financial data with survey research methods. Publications by Ferber and his colleagues at the Bureau of Economic and Business Research at the University of Illinois and Katona, Lansing, and associates at the Institute for Social Research, The University of Michigan, report several investigations on response bias in surveys of financial information. Somewhat later, a program of studies into problems of reporting health information was undertaken by the present authors and their colleagues.

Ferber (1959) finds that non-mortgage debt, installment debt, and liquid asset holdings are substantially underreported in surveys. Lansing, et al. (1961) report similar findings for personal debts. Health data demonstrate underreporting of hospitalizations (Cannell, et al., 1965a), physician visits (Cannell and Fowler, 1963), and the presence of chronic illness (Madow, 1967).

The results of the investigations by these and other researchers provide a basis for some generalizations on response problems, especially on the validity of report, and some correlates of invalidity. While not all of these studies show comparable results, generally there is a high level of consistency among them, from which the following generalizations can be derived.

As the time between an event and the interview increases, underreporting of information about that event becomes progressively greater. This generalization will, of course, surprise no one, but what is unexpected is the rapidity with which the failure to report the event increases with time. For reporting of visits to physicians (Cannell and Fowler, 1963) the reporting rate drops over a two-week period from 85 percent for one week preceding the interview to 70 percent for the second preceding week. Samples drawn from hospital records show that 95 percent of hospitalizations occurring within one month of the interview are reported, while only about 70 percent of the hospitalizations from 10–12 months prior to the interview are reported (Cannell, et al., 1965a). Neter and Waksberg (1965, p. 13)
found that lengthening the recall period from one to six months led to a substantial increase in underreports of residential alternations and repairs.

Events which are important to the individual are reported more completely and accurately than those of less importance. The concept of importance is variously defined. As used here, “relevance” or “salience” are synonymous terms for importance. Studies show that hospitalizations of longer duration and those involving surgery are more likely to be reported (Cannell, et al., 1965a). The larger the number of visits to a physician which the respondent has made for a chronic condition, the more likely it is that the condition will be reported to a survey interviewer. In the previously mentioned Neter and Waksberg study of underreporting of repairs, the decline in reporting was found to be considerably greater over time for jobs costing less than $20 than for more expensive ones.

We have avoided referring to the phenomenon of underreporting in terms of “forgetting” because the term has the implication that “forgotten” information is forever lost and inaccessible. Studies show that much unreported information is not irretrievable, but can be reported, given adequate incentives. Interference theory of forgetting is relevant to these findings (Manis, 1971; Underwood and Postman, 1960). Forgetting is not absolute; that is, information does not disappear from memory, but some items are more difficult than others to retrieve, because of competing associations or interferences from intervening events. As elapsed time increases, there is greater opportunity for interference to occur. The importance or salience of the event for the person show the same effects. The probability of interference is greater for unimportant events than for more important ones. This theory implies that underreporting is not that information is truly forgotten in an absolute sense but is a failure of the information retrieval process. This conceptualization has an important implication: it suggests that reporting can be improved by manipulating conditions which facilitate recall.

The conclusion is that much of the information which respondents are asked to report is not purposely withheld, but some items are more difficult to retrieve from memory than others. Recall tasks, as presented to respondents, vary in difficulty, and the level of difficulty is related to how well the task is performed. If the investigator were content to ask only for readily available information, reporting error would be much less of an issue. As the task becomes more difficult, the respondent must exert greater effort for adequate performance. Stated in terms of motivation, it suggests that the more difficult the task, the higher the level of motivation which is required for accurate reporting.

Reporting of an event is likely to be distorted in a socially desirable direction. If the event is perceived as embarrassing, sensitive in nature, threatening, or divergent from one’s self-image, it is likely not to be reported. Lansing, et al. (1961) found that primary car loans are reported well, but secondary loans are greatly underreported. Conversely, behavior perceived as desirable tends to be overreported; e.g., voting is reported when it did not occur (Parry and Crossley, 1950), respondents overstate the size of contributions to charity (Parry and Crossley, 1950), and the size of small savings accounts is over-stated (Ferber, et al., 1969). When asked the reason for their hospitalizations, respondents overreported diagnoses of benign neoplasms, stomach ulcers, and diseases of the gall bladder, and greatly underreported diagnoses such as female breast and genital disorders and diseases of the nervous system (Cannell, et al., 1965a).
Parfitt (1967) interviewed a sample of panel housewives concerning their purchases of twelve consumer products. Overreporting was found to be related to average frequency of purchase: the lower the frequency of purchase, the higher the overreporting. The author suggests, "There may be product fields that, in a face-to-face interview, a housewife may be reluctant to admit she does not buy....." (p. 19). Purchase of butter, for example, may be overreported with a concomitant underreport of margarine purchases.

In these data we see another effect of respondent motivation: the unwillingness to reveal potentially embarrassing, threatening, or undesirable information; or, conversely, the overreporting of information perceived as ego-enhancing. Much of the information which a respondent fails to report, however, is not repressed or deeply threatening to his ego; it is only mildly socially desirable or undesirable. The failure to report accurately reveals a lack of willingness to accept the respondent task and, as in the issue to retrieving material from memory storage, the solution is either to ask nothing which he perceives as embarrassing or ego-enhancing, or to increase his willingness to take some risks in revealing information. The problem of respondent motivation is again apparent.

A Model of Reporting Behavior

Our interest in these principles of response bias is not to demonstrate the frailty of survey data but to provide a basis for hypotheses regarding factors underlying reporting patterns, and to suggest some approaches for improving reporting completeness and accuracy. The following model illustrates the major processes involved in obtaining a response to a single question. (See Figure 1.) To simplify the model, we assume that the question asked refers to some past event or behavior. (A comparable model could be used for attitudinal information.) A further assumption is that the question communicates the objective perfectly and the respondent has the required information accessible.

The process begins with the interviewer asking the question and the respondent perceiving it. The steps required for adequate response performance are: (1) the respondent's consideration of what information is relevant for an adequate answer, and a request for clarification if this is unclear; (2) memory search to retrieve relevant information. The result of this memory scan is evaluated by the respondent on the basis of his understanding of how adequately the retrieved information meets the question objective. The result may be further memory search. (3) Next is an evaluation of the social desirability of the retrieved information. (4) Then the information may be reported. The report may be inaccurate because of faulty or inadequate memory search, or it may be incomplete or inaccurate because of the perceived consequences of its disclosure even though correct information was available. The interviewer then assesses the response, and evaluates its completeness and its adequacy to meet the question objective. If he considers the response to be inadequate, he uses a probe which creates a feedback loop and recycles the process. If the response is seen as adequate, it is recorded, and the cycle is complete.

As the model indicates, adequate performance requires that the respondent be thoroughly competent and complete in his memory searching activities and be
This model was developed by Charles Cannel, Kent Marquis, and André Laurent.
willing to report accurately even though the information may be evaluated as socially undesirable. Adequate task performance rests on the two components: ability and motivation, and neither by itself is sufficient (Fowler, 1973; Vroom, 1964). Data cited earlier indicate that respondents tend to fail in one or both of these tasks. An inference to be drawn from this model is that the apparently simple job of adequate reporting requires the respondent to carry out successfully several rather complex tasks. The respondent who is not adequately motivated is unlikely to perform well on any but the simplest of requests for information. We can be more specific. For a respondent to perform well he must be sufficiently motivated to be conscientious in the complex job of memory search and evaluation of the information. He must also be sufficiently motivated to report information even though it may be threatening or sensitive.

Before discussing respondent motivation in more detail, we must mention respondents' memory-searching skill and ability. That people vary in their ability to recall information is common knowledge. Recall ability is one of the requisites for educational achievement. Recall tasks are included in many intelligence tests, pointing to the relevance of memory in intellectual life. The implication of this is two-fold. First, memory retrieval is a skill or an ability in which some people are more competent than others. Second, memory searching is a task which may require a significant level of effort to perform adequately. Evidence for both of these assumptions abounds. Both Ferber and Lansing have found that higher educated respondents report savings accounts, cash loans, and life insurance policies more accurately than lower educated respondents (Lansing, et al., 1961, p. 180). Our data show similar relationships in the reporting of health events.

To achieve an adequate motivational level is not easy; researchers have been woefully unable to obtain the level of performance necessary for complete and accurate reporting, and there is very little research on the motivational variables in the interview. However, the theory relevant to motive patterns can be described, as can some studies aimed at improving report by manipulating both motivational and cognitive variables.

Most concepts of motivation involve some "path-goal" concept, consisting basically of four components: (1) a need, or some psychological or physical force on the individual, (2) a goal, something which he perceives will, if attained, partially or wholly satisfy a need. When a need is linked to a specific goal perceived as satisfying it, forces are generated to move toward that goal. (3) a path: behavior takes place only when the individual perceives that behavior as a path leading toward a goal. (4) barriers: forces which make a path less desirable or attainable, sometimes to the point where the negative values of the forces are greater than the positive values of the goal. The path, then, is no longer perceived as desirable.

The relevance of this conceptualization for respondent behavior is readily seen. Three types of respondent motive patterns may be postulated as most pertinent to an interview. One pattern assumes that the respondent is motivated by a perception that his participation in the interview will enable him to achieve certain personal goals. He evaluates the interview in terms of the attractiveness of the purposes and objectives of the survey and their compatibility with his personal goals, and he perceives the interview as an effective path to attaining those goals.
In introducing the survey to the respondent, it is usual practice for the interviewer to make statements about the importance of the survey, and about the research objectives. This statement of purpose and value hopefully demonstrates to the respondent that the interview is a path for achieving some personal or social goal.

For example, the interviewer may mention that most people look to financial security (a need), and one way they attempt to become more secure is through accumulated savings and retirement accumulations (goal). The interview is to find out about people’s situations in these areas in order to supply information to policy makers to help safeguard these assets (path). Some barriers in reporting may include the respondent’s unwillingness to report his financial situation, and the difficulty in recalling such information. The interviewer assumes that the respondent will view his contribution to the survey as instrumental towards achieving some desired personal or social goal.

In practice it is often difficult to predict respondents’ personal goals, and respondents themselves may have different and even conflicting goals. Furthermore, it is the respondent’s perceptions of how the interview will meet his goals that is important. Lengthy explanations may have little or no effect, especially if the respondent has negative attitudes towards the interview to begin with.

Based on this analysis, one can make some guesses about the effectiveness of payments to respondents. The money may be offered as a goal for the respondent where it is felt no goal exists. To be effective, the amount of money offered must be large enough to be worth working for. A crucial issue is the determination of how large the payment should be to achieve the desired effect. This is a more complex problem than it appears on the surface if one can extrapolate from psychological experiments. Studies in equity, conformity, and compliance suggest that if individuals perceive that they are overpaid, the effects on task performance may in fact be negative (Kaufmann, 1971). How to determine the appropriate payment amount is a major research project in itself.

While such payment may be particularly useful in persuading the respondent to grant the interview, it seems unlikely to be an effective motive to accurate reporting unless a link can be made clearly between the money and good role performance. In fact, the payment may lead to poor behavior if the respondent feels that to be a good respondent he should give the answers the interviewer wants. If payment is to be used, it should be made clear that it is given to encourage the respondent to work hard in recalling relevant information and to report accurately. More experiments on the use of payment to respondents are needed to provide some answers to these issues.

A second type of motive pattern is also based on a path-goal model, although the instrumental relationship is less direct and the respondent’s goals less specific. The major characteristic of this type of motivational pattern is that the respondent reacts to the interview and the interviewer according to his habitual mode of response toward requests made on him by legitimate agencies or organizations in the society. He has developed norms of good citizenship, politeness, acquiescence to requests for information, etc. The respondent may react positively to cooperating with a governmental agency or a university. He may be flattered to be selected as a respondent. He may see his participation as good citizenship, or he may merely...
be exhibiting acquiescence or politeness to the interviewer's request. Some researchers (Berkowitz, 1969; Darley and Latane, 1969) have pointed out that for many individuals a norm of social responsibility and the introduction of certain cues in the interview may activate and motivate respondents. People with attitudes favorable toward research often are better respondents (Lansing, et al., 1961).

Unfortunately, both research evidence and experience suggest strongly that these two types of motivations are not very effective in increasing respondent activity. Respondents do not share the researcher's goals, or, if they do, they fail to see the interview as an effective way of achieving that goal. While economic well-being is important, reporting one's income and savings to an interviewer is not seen as related to this goal, even though the interviewer may make a valiant attempt to demonstrate the connection. Similarly, the respondent role may be seen as related to citizen responsibility, but this is not usually a sufficiently salient or strong motive to induce a high level of activity or a willingness to report embarrassing information.

In connection with a health survey, experimental letters and brochures were sent to respondents prior to the interview (Cannell, et al., 1965b). One brochure attempted to attract the reader by emphasizing the relationship between the interview and "good health." The other stressed the citizenship role in participation in the survey. Neither received much attention, and readership appeared to have no significant effect on the response rate or on the caliber of reporting.

In another study, Dommermuth and Cateora (1963) showed that respondents who were sent individually typed letters a few days before the interview had a response rate identical with a respondent group that was not contacted prior to the interview (94.8 percent).

Ferber (1959) reports that respondents receiving a letter requesting help in evaluating the interview reported debt information more accurately than those receiving a customary introductory letter, but the former group showed a higher refusal rate. Apparently, the evaluative letter weeded out the reluctant respondents, leaving those better motivated. Generally, advance letters are of little aid either in improving the response rate or eliciting better respondent performance.

There is a third type of motivation which differs from the two extrinsic goal types mentioned above. This is the motivation which is aroused by interpersonal interaction which occurs between the interviewer and respondent during the interview. It is a common finding that respondents enjoy the interview, even though they may be ignorant of its purpose, and the goal seems remote. Cannell and Axelrod (1956), for example, found that 50 to 60 percent of respondents on four surveys assessed the interview as very interesting, and almost none found it uninteresting. But what did they enjoy? And why? And, did it affect the quality of their responses? This third type of motivation is based on somewhat more fundamental personal needs than the other two. Need for positive interpersonal relationships (or need for affiliation as Atkinson [1958] and McClelland et al. [1953] refer to it) is a source of motivation. The interactions of the interview provide an opportunity for these needs to become activated. The goal is achieved through the interactive process itself.

To learn more about these interactions in interviews, we developed a coding system for classifying each activity of the respondent and interviewer (Marquis and
Cannell, 1969; Lansing, et al., 1971). One consistent finding of the interaction studies was that the interview was characterized by a balance in activity level between the number of units of activity of the interviewer and the respondent. Not only was there an overall balance or matching of activity level, but major subclasses of behaviors tended to balance, especially behaviors which were task-oriented. Another significant finding was with regard to the amount of interviewer feedback. In these studies, feedback constituted nearly one-quarter of an interviewer’s activities. (This term includes both short interjections: “OK—I see—Good” or more extensive ones: “That’s the kind of information we want.” “You are doing a good job.” etc.) This finding (Marquis, et al., 1972) is important, particularly since in training no efforts are usually made to instruct interviewers as to how or when to provide feedback or even what feedback to use. Since findings in other fields demonstrate the marked effect of feedback on performance (Ammons, 1956), it was disturbing to find that one-quarter of behaviors, potentially very significant for influencing response behavior, was uncontrolled.

This interaction analysis led to a series of experiments designed to control and use feedback as a positive force to improve reporting—both to increase validity of response and to increase the amount of information reported. The first of these was a study of reporting of chronic and acute conditions and symptoms (Marquis, et al., 1972). From this it was concluded that it was feasible to change respondent performance in a desired direction as a response to changes in the interviewer’s behavior.

A second series of studies was developed based on the finding that respondents and interviewers tended to show a balance in the level of behavioral interactions in the interview, and that this level correlated highly with the amount of information reported (Marquis and Cannell, 1969). This suggested that if the interviewer was programmed for a high verbal output, the respondent might model his behavior accordingly, and that if this occurred, more information might be reported. Matarazzo and his colleagues (1972) have demonstrated this type of modeling in other settings.

To investigate modeling effects, we designed two questionnaires, again using health variables. In one, the questions were the standard type while the other contained the same questions lengthened considerably, but in a way which did not change the nature or amount of information requested (Laurent, 1972). This series of studies suggests that increasing the length of the question would increase the amount of information reported. Three alternative rationales for the question length effect were proposed.

1. The length of the question has cueing and motivational effects upon reporting behavior. A longer question may convey the idea that the task is important and deserves serious effort. The long questions may also seem to set an unhurried frame of reference.

2. Question length of interviewer speech duration is a proxy for another variable: namely, the time given for recall activity. A longer question increases the time available to the respondent for search activity and thus improves the information retrieval process.

3. Finally, it may be that redundancy improves the clarity of the question and leads to better understanding of what is wanted.
The next experiment undertaken was to combine question length and reinforcement with the prediction that this combination would yield even more complete and accurate information. The results did not support the prediction (Marquis et al., 1972). While both verbal reinforcement and long questions showed main effects, that is, increased the reporting of acute and chronic conditions, the combination of the two resulted in lower reporting rates than either technique by itself. In analyzing the data by educational level, some interesting findings were revealed. For lower educated respondents, reinforcement improved reporting, while long questions did not. For higher educated respondents, on the other hand, the use of long questions improved their reporting while the use of reinforcement did not.

Explanations for the differential effectiveness of the two procedures are tentative at this point. It may be that higher educated respondents find verbal reinforcement aversive, while lower educated respondents interpret it as encouragement and meaningful feedback. On the other hand, higher educated respondents may find the long questions stimulating, while lower educated respondents find them confusing. Education may also be a proxy variable for status, and the social distance between the interviewer and the respondent may be a significant mediating variable. The work of Hyman et al. (1954), Back and Gergen (1963), Dohrenwend, et al. (1968), Weiss (1968), and Williams (1968) all suggest this.

The modeling effect referred to earlier can also be accomplished another way. The literature on imitative behavior by Bandura (1972) and others (Bandura and Huston, 1964; Hicks, 1965) suggests that social learning is an important phenomenon in interpersonal settings, and many laboratory studies have been conducted to demonstrate the effectiveness of modeling procedures. To further test such effectiveness it was proposed that verbal modeling might facilitate and enhance interview performance. A recent study (Henson, 1973) played a recorded model interview to the respondents before the actual interview began. Preliminary findings suggest that modeling is effective primarily with lower educated respondents. The number of reported health conditions increased considerably among this group. For higher educated respondents modeling makes no difference in reporting. Thus, education has again shown up as a significant mediating variable.

CONCLUSIONS

Based on our own work and on our interpretation of the findings of others, we conclude that the task given to respondents in the usual sample survey is more complex and makes greater demands than is usually considered. We conclude further that respondents are generally not particularly motivated to perform the task if it makes them work hard or report embarrassing information. They are polite, willing to respond, but unmotivated to work.

For most surveys, respondents do not share the research goals and they fail to see how their own goals are served through the interview. Attempts to lead respondents to make the connection between the interview purposes and their own goals have not been successful in general. The use of payments to respondents may be useful in obtaining the interview originally, but it is unclear as to its effectiveness in motivating accurate reporting. The most potent source of motivation
appears to us to be found in the interaction of the interview itself. Manipulating interviewer behavior has shown potential for increasing good role performance. The different interviewing techniques that have been implemented, however, suggest that their effectiveness may depend on the educational level of the respondent. Whether education is a proxy variable for social distance between interviewer and respondent has not yet been determined. Much further experimentation is necessary before definite answers to the problem of response bias are possible.

References


