You can’t believe all that you’re told: the issue of unvalidated questionnaires

Ian Scott

This column arose out of discussions about a mailed questionnaire to parents on home hazards and safety practices (see Evans et al, p29). The issue being discussed was the confidence that could be placed on the results in the absence of independent validation of questions and answers. One view was that to require formal validation is unrealistic. It was felt that such validation is onerous, and, even when resources are available, it is often difficult to know how best to validate such questionnaires. Those sharing this view also note that, while it is right to be wary of non-validated surveys, if scientific journals always required such data, few papers would be accepted. Mailed questionnaires are used because they are manageable, cost effective tools, but some are of little real value. Is it unfair to say that the direction and structure of some questionnaires are influenced more by what is convenient than by what is good technique? The pressure to ask questions that can be easily coded and expressed as numbers in a good looking table is very real. As one researcher suggested in conversation, research reports that contain plenty of percentages look meaningful, lend themselves to numerical quotation, and appeal to funding agencies. It is because there is almost always a compromise between science and resources that there are no solid answers about this issue. Accordingly, this opinion was commissioned to generate debate on what is good practice when conducting this type of research.

How frequently have you been a bad parent?

Asking parents about what they do, without trying to check on their responses is problematic. What do we learn when two out of three parents say they have taken some action in the past year to reduce the likelihood of their child being injured? What parent could easily respond that they have taken no action to prevent their child having an accident? Many replies will be influenced by what respondents think a ‘good parent’ should do. For example, Nelson found that the less people actually used seat belts, the more likely they were to over-report their use: ‘self reports [of seat belt use] were highest in the [areas] with the lowest levels of observed use.’ It is not necessary to think that parents deliberately lie. Among other factors, their recall of events is influenced by their need to reinforce their view of the world. In a sample of parents, two thirds of those who used baby walkers thought they were safe, compared with 5% of non-users. None of the sample of baby walker users saw safety as a reason to stop using them, but 45% of non-users did. A national US phone survey revealed other aspects of the same issue when researchers investigated parental knowledge and attitudes. They concluded that ‘parental knowledge is a necessary, but insufficient, factor in the prevention of injuries to children’.

‘We learned something useful, everyone gave the same answers’

When there is not much variation in responses researchers must wonder whether the right questions are being asked. After all, what is learned if two out of three parents respond that they think their home is ‘as safe as I can make it’? Such results are doubly wasteful if they fail to investigate new areas or new interventions. For example, it has clearly been established that parents think their houses are ‘safe’ at least partly because they don’t understand what is hazardous.’ This is not a finding that needs to be constantly checked.

Questionnaires can highlight significant aspects of behaviour and attitude. For example, a study that found that a high level of Victorian football players wore mouthguards at matches (89%) but relatively few (13%) did so at training, provided a reason for this discrepancy. The knowledge that players perceive that physical contact is much less likely to occur at training gave investigators a direction for action. Any concentration of answers must raise concern that the tool being used is too crude to distinguish between subsets of views, or that the respondents are giving ‘socially acceptable’ answers rather than reflecting what they actually think or do.

‘I know the answer to this one. . .’

What have we learned if 199 out of 200 respondents say that medicines are ‘never’ or ‘rarely’ left within reach of their child, except that the question is not working? The most striking feature that Ann Williamson and her colleagues found in developing a workplace safety questionnaire was the lack of variation between answers on many issues. She and her team concluded that there are well known or socially acceptable beliefs in the working community about safety matters. They went on to develop a meaningful, testable, valid questionnaire.
Knowing that some issues elicit socially desirable responses can be helpful. If only one out of 200 parents reports that 'sometimes' medicines are within reach of their child, it seems a fair implication that these parents know that medicines are dangerous. Therefore, information to shape future interventions might come more effectively from opened ended questions about exactly where products are stored; questions about what places parents consider to be safe; about whether parents have had any difficulty with children getting at medicines, or from direct inspections of even a small number of homes. There are a wide range of techniques to get this information.*

**Answer this question as though you know what it means**

In order for the answers to any question to be useful, the interpretation of the question by most respondents needs to be broadly similar and unambiguous. Answering questions about how often medicine is left within reach means that respondents must have a view of what 'out of reach' means. It is highly likely that any question involving any form of judgment will be the subject of significant differences in understanding.

**Is the presence of a hazard the key variable?**

The presence of injury hazards has been the focus of much work and has been linked to a host of variables. While it is reasonable to assume that there is an association between injury rate and the presence of hazards in the home, Jolly and her colleagues found that 'while rates of household poisoning in young children are higher in disadvantaged groups... this is not reflected in hazard prevalence differences'. They concluded that household practices or exposure to the hazards must be different.*

Experience and intuition says that the relation between hazard and injury is complex, and it is reasonable to assume that injury may well relate to the frequency and way in which 'exposure' to the hazard occurs, rather than to the simple presence of a hazard in the home.* In support of this argument, one researcher cited an interview with a young mother who said that medicines were not a problem in her home because her 3 year old knew not to touch them!

*Example 1: An Australian Bureau of Statistics Home Safety Survey used 4000 interviewer administered questionnaires in the respondent's home to ask about the specific location of all medicines and household chemicals. A cue card was used with eight different possibilities—cabinet or cabinet above/below 1.5 metres, shelf above/below 1.5 metres, drawer, handbag, etc. Responses showed that 84% do not have a child resistant cupboard for medicines and 79% do not have them for household cleaners. Of those with cupboards, between 10 and 20% said they did not store specific items in them.*

**Example 2: A UK home visitor program that asked about attitudes to risk, awareness of hazards, and descriptive data was combined with a tour of the house noting 10 specific hazards.**

**There's nothing wrong with my memory...**

There are good reasons to be cautious about how reliable memory is. In developing guidelines for health related surveys, the Australian Bureau of Statistics found that an individual's unprompted recall of episodes of ill health were significantly different from their recall during detailed interviews. This finding applied to even major health events and resulted in health surveys being limited to experience in the "last 30 days" (J Harrison, personal communication).

A related problem is that common frequency descriptors, such as 'always' and 'never', are imprecise and present problems for both researchers and respondents. 'Of greatest concern is that one person's 'nearly always' can be 'always' to another and 'sometimes' to yet another.'

**Conclusions**

The argument that it is impractical and too expensive to ask questions that return meaningful answers does not hold water. In some areas detailed validation is necessary, but there are also plenty of examples of useful data coming from simple tools. There are a range of tools to check the validity of questions and answers. Careful consideration of what it is we really want to know, about whether people can and will answer questions, and about what we are likely to be able to say once the questions have been answered, would go a long way toward ensuring that any questionnaire survey provides useful and trustworthy information.

On these matters the literature and colleagues are great resources.