Applying cognitive design principles to formatting HRQOL instruments

Paul A. Mullin[1], Kathleen N. Lohr[1], Brian W. Bresnahan[2] & Pauline McNulty[2]
[1]*Research Triangle Institute, P.O. Box 12194, Research Triangle Park, NC 27709-2194;* [2]*Johnson & Johnson, ICOM Health Economics, USA*

Abstract

Researchers developing or using health-related quality of life (HRQOL) instruments can benefit from knowledge of state-of-the-art formatting methods for self-administered questionnaires. Three objectives in formatting design are: (1) to reduce errors in respondent navigation through the questionnaire that lead to item non-response and question misinterpretation; (2) to reduce respondent and administrative burden; and (3) to enhance respondent motivation in question answering and compliance with the request to participate. Based on an extensive literature review to identify techniques that have been shown to meet these objectives, we developed specific guidelines for HRQOL instruments concerning all aspects of questionnaire formatting. These guidelines represent well-motivated recommendations for improving HRQOL instruments, although their overall impact has not been empirically tested. We applied the guidelines to several HRQOL instruments that are widely used internationally, and obtained approval from the developers for all formatting changes to their instruments. Applying cognitive design principles and empirically substantiated formatting techniques produces an HRQOL instrument formatting with six critical attributes: simple, consistent, organized, natural, clear and attractive. The present paper contributes to the emerging research literature on the cognitive processes by which respondents answer HRQOL questions and demonstrates how 'cognitive aspects of survey methodology' research can improve HRQOL data collection efforts.

Key words: Quality of life, Questionnaire design, Survey methods

Introduction

Developers of health-related quality of life (HRQOL) instruments often control the content of their instruments through copyright protection and contractual arrangements. In contrast, instrument formatting usually varies depending on the individual needs of the user, although developers may provide a recommended version. In this context, HRQOL researchers may benefit from knowledge of survey methods research on the formatting design of self-administered questionnaires. The use of empirically tested formatting methods can lead to improvements in data collection efforts in the following areas: [1, 2]

- lower risk of navigational errors that can lead to item nonresponse,
- better reading performance and comprehension by respondents,
- less time and effort required by respondents to complete a questionnaire,
- greater proportion of time spent thinking about answers to questions,
- greater likelihood of respondent participation and compliance,
- less reliance on assistance from administrators by respondents, and
- less time and effort required for data entry staff.

In this paper, we introduce research and theory underlying state-of-the art formatting methods for self-administered questionnaires and apply the principles and techniques of sound design to the layout of HRQOL instruments. Our examination of HRQOL formatting began with an extensive

14

review of the research literature. We identified design principles and a collection of effective formatting techniques, developed formatting guidelines for HRQOL instruments, and applied these guidelines to several HRQOL instruments used internationally: the *Medical Outcomes Study Health Survey Short Form 36* (SF-36) [3], the *Kidney Disease Quality of Life Short Form* (KDQOL-SF) [4], the *Functional Assessment of Chronic Illness Therapy – Anemia* (FACT-An) [5], and the *Diabetic Foot Ulcer Scale* (DFS) [6]. We then presented various instrument design changes predicated on the guidelines to the developers of these instruments, HRQOL research experts, and survey research methodologists. In these discussions, we received approval from the developers to use our formatting design for their instruments in future research studies.

The present paper extends these discussions of HRQOL instrument formatting and illustrates key points with examples from these instruments. The guidelines presented here represent well-motivated recommendations for improving HRQOL instruments based on empirical studies involving other data collection instruments and documents. The potential advantages of applying these formatting methods to HRQOL instruments warrant empirical investigation.

Questionnaire design objectives

Respondents engage in several activities when completing a questionnaire. They evaluate the content of questions and generate answers, and they navigate spatially between and within questions to read the questions and mark their responses [1, 7]. Questionnaire formatting communicates to respondents about how to proceed through the instrument. Thus, one objective of formatting design is to avoid common navigation errors by respondents, such as item non-response and the failure to read instructions.

A second objective of formatting design is to increase the ease of responding to the instrument. Respondents are willing to expend only a finite amount of time and effort in completing a questionnaire. They often short-cut the amount of work they must do, but they also attempt to find ways to hide this and appear as if they are answering responsibly [8]. By minimizing the cognitive demands of navigating an instrument, a questionnaire designer can help respondents to focus on the substantive issues posed by the items of the questionnaire. Moreover, lessening the burden of the data collection process on respondents is meritorious in its own right; the Medical Outcomes Trust Scientific Advisory Committee has included administrative burden among its evaluation criteria for HRQOL instruments [9].

A third objective of formatting design is to increase the motivation of respondents to invest time and effort in the survey. Respondents clearly take cues to the apparent complexity of a questionnaire into account when determining the amount of effort they will dedicate to answering the questions. Respondent-friendly design can make a questionnaire look inviting to respondents and has been shown to significantly increase response rates to mail surveys [10].

Design principles and techniques

A limited set of design principles aptly summarizes the effectiveness of various formatting techniques in meeting the three objectives for self-administered questionnaires that are described above. In particular, a soundly designed questionnaire can be described by six characteristics: simple, consistent, organized, natural, clear, and attractive. We review the basis for each of these principles below and provide illustrations of formatting methods that help to produce an instrument with these characteristics. In some cases, specific points pertain to more than one design characteristic, and their location in the discussion is arbitrary.

Simple design

Under many circumstances, irrelevant information (noise) can interfere with the perception of relevant (signal) information [11, 12]. To reduce the degree to which respondents are visually distracted when reading questions and marking responses, it is necessary to eliminate graphical complexities and unnecessary or redundant information in the questionnaire formatting.

Using alternatives to grid lines
Numerous HRQOL instruments employ table grid lines to guide respondents through the rows and columns of certain types of questions (see Figure 1a). These grid lines add considerable complexity to the instrument. Moreover, lines are universally used to represent barriers or stopping points, and respondents must regularly cross these barriers when dealing with a question that is embedded in a table grid. In this case, the grid lines are used in a manner that is inconsistent with the meaning typically ascribed to these graphical features (and thus they also fail to create a *clear* and *natural* design). To achieve a simple design without table grid lines, dot leaders can be used to guide respondents horizontally, and response box columns marked at the top by upside-down triangles serve well to guide respondents vertically (see Figure 1b) [1].

De-emphasizing irrelevant information
A key aspect of making an instrument simple is to de-emphasize material that is not relevant to respondents [1]. This includes material that is pertinent solely to researchers, such as copyright notices and information that identifies the research study protocol. One effective practice is to print respondent-irrelevant material in a small font size and place it outside respondents' primary navigational path. This can be accomplished by locating irrelevant material in a footer and using a thin-lined border to separate the footer from relevant

a. Traditional Version

3. Because of your illness, how often have you...

	None of the time	A little of the time	Some of the time	Most of the time	All of the time
a. Had to depend on others to help you bathe and dress?	1	2	3	4	5
b. Had to depend on others to do household chores like cooking and cleaning?	1	2	3	4	5
c. Had to depend on others for getting out of the house?	1	2	3	4	5
d. Had to spend more time planning or organizing your daily life?	1	2	3	4	5

b. Reformatted Version

3. Because of your illness, how often have you...

	None of the time ▼	A little of the time ▼	Some of the time ▼	Most of the time ▼	All of the time ▼
a. Had to depend on others to help you bathe and dress?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b. Had to depend on others to do household chores like cooking and cleaning?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c. Had to depend on others for getting out of the house?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
d. Had to spend more time planning or organizing your daily life?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Figure 1. A traditionally designed HRQOL matrix question using table grid lines to guide respondents (a) and a reformatted version using dot leaders and simple graphical symbols to guide respondents (b).

material. Another solution is to print irrelevant material in a low-contrast color that is different from the color of the main text of the instrument. Respondents can efficiently segregate relevant and irrelevant information based on color [13].

Using response boxes

Having respondents circle numbers to indicate their answer selections is a common practice in HRQOL data collection efforts. The focus of this design is to facilitate the task for data entry staff, who record these codes for the selected responses. However, the numbers are typically printed with the same font characteristics as the question text and are difficult for respondents to separate visually from the question text (see Figure 1a) [14].

Because question answering relies primarily on the processing of language, or verbal information, a preferable way to guide respondents when they are marking responses is to use graphical information such as response boxes. The boxes differentiate the response area of the question from other parts of the question (see Figure 1b), and the concurrent performance of different tasks (i.e., navigation, question answering) is usually superior if the tasks rely on different forms of information (e.g., auditory versus visual, verbal versus graphical) [15, 16]. Data entry codes can then be placed adjacent to the response boxes. These codes should be printed in a smaller font than the rest of the text because they are not specifically relevant to the respondent.

Using response boxes places no additional burden on respondents. Respondents mark response boxes at the same speed as they circle or underline responses [3, 17]. In general, however, it is a good idea to ask respondents to mark crosses (e.g., '×') in the response boxes because they may be inconsistent about keeping check marks (e.g., '√') from straying into nearby boxes.

Emphasizing selected terms

One aspect of instrument design that is frequently misunderstood relates to the selective emphasis of specific words or phrases that are deemed especially important for a respondent to read. Questions in some HRQOL instruments emphasize primary subjects, objects, or action verbs. There is usually little reason to emphasize these terms because they are already the focus of the questions. Moreover, the overuse of emphasis can lead to uneven reading and may actually dilute the intended emphasis.

In general, selective emphasis should be reserved for words and phrases that either define important criteria on which respondents must base their response (e.g., a reference period) or differ between two highly similar, adjacent questions. Underlining is often the best choice for the emphasis of selected words. Such underlining increases memory performance for the emphasized textual material for both younger adults [18] and the elderly [19]. Nonetheless, underlining must be used judiciously so as not to cause visual interference with the text.

Avoiding repeated instructions

Instrument developers sometimes provide instructions about the response task (e.g., 'circle one answer on each line') and repeat these instructions within each question. Some methodologists support the idea of building such redundancy into the instrument for respondents who might be confused about what to do [20]. However, such redundancy is unnecessary for many other respondents. Because respondents tend to read only what they believe is necessary to read [21], there is some risk that respondents may begin to ignore all instructional material when they find that certain instructions are not useful to them.

Consistent design

Ideally, a formatting design should be employed so that no confusion arises about the response task for a given question and repeating instructions is unnecessary. This can be accomplished by ensuring that the response tasks are consistent for similar-looking types of questions and by giving distinctive visual characteristics to different types of questions. For example, Version 2 of the SF-36 Health Survey [22] contains multiple choice questions in which the response options are arrayed horizontally in a row (see Figure 2a), rather than vertically in a column under the question (see Figure 2b) as has been typically recommended [23, 24]. This horizontal formatting design increases the similarity of these multiple choice questions to the matrix questions, which ask respondents to

a. Horizontal Orientation

1. In general, would you say your health is:

Excellent	Very good	Good	Fair	Poor
▼	▼	▼	▼	▼
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

b. Vertical Orientation

1. In general, would you say your health is...

Excellent .... ☐ 1

Very good .... ☐ 2

Good .... ☐ 3

Fair .... ☐ 4

Poor .... ☐ 5

Figure 2. Multiple choice questions with horizontally-aligned response options (a) and vertically-aligned response options (b). Both are formatted according to the guidelines.

select an answer for each of set of related items under an overall question. It adds consistency to the respondents' task and eliminates the need to repeat response task instructions.

Consistent design minimizes the cognitive demands of completing a questionnaire by helping respondents develop mental models, or expectations, for the repetitive behaviors that they use to navigate an instrument and mark responses [25, 26]. Respondents can become skilled at completing a questionnaire only to the extent that meaning and behaviors are consistently mapped to design elements [27]. Thus, a design element should be used for only one purpose, and different design elements should not be used for the same purpose. For example, if underlining is used in one location to emphasize words, then italics or bold text should not be used in another location.

Maintaining graphical structure
The example above concerning the SF-36 Version 2 shows how consistency in the graphical layout of the questions allows respondents to use a regular set of navigation behaviors. This applies equally well to the broader graphical design of the page structure. An instrument should not be designed with a mixture of single-column and two-column formatting, or landscape and portrait page orientations. Additionally, the columns of response boxes for similar types of questions on a page should be vertically aligned to the extent possible [1]. These types of considerations also add to the attractiveness of the instrument.

Ordering scale options
In some HRQOL instruments, the order in which the response options are presented varies in different questions (e.g., 'never' to 'always' versus 'always' to 'never'). Apparently, the developers have introduced this inconsistency in the attempt to disrupt the tendency of some respondents to respond positively without evaluating the content of the items (i.e., acquiescence bias). This practice adds complexity to the respondents task, however, and unexpected changes in response scale order may go unnoticed. The more widely used approach to controlling acquiescence effects is to vary the negative or positive content of item wording [28]. With a balanced scale of positive and negative items, statistical control of the acquiescence effects is possible [29].

Organized design

The ability of respondents to comprehend questions and instructions depends on the presentation of contextual information in appropriate locations [30]. In designing an organized questionnaire, it is important to adhere to the *proximity compatibility principle*, which states that the degree to which different displays of information are relevant to

common mental tasks should guide the physical proximity of these displays [31].

Locating instructions
Some HRQOL instruments begin with a page of instructions to guide respondents in completing the questionnaire. Survey methodologists generally recommend against relying on these introductory instruction pages because such material is not read reliably by respondents, presents information out-of-context, and requires respondents to maintain the information in memory or refer back to it when it is needed [1, 20]. Instead, instrument developers should attempt to place instructions in the context of the questions where the instructions are more likely to be seen. For example, respondents can be instructed about how to mark their answers (e.g., by putting an '×' in a response box) in a sentence that appears immediately after the first question of the instrument.

Grouping concepts
When organizing a questionnaire, it is important to minimize switching between survey topics and, if possible, response tasks and layout designs. A coherent grouping of questions can enhance validity and reduce the time to complete the questionnaire [31, 32]. Research also shows that answers to questions on topics that are presented without prior warning are especially susceptible to biases [33]. Therefore, brief section headers and lead-in statements should be used to help guide respondents through the topics of a questionnaire and orient them to new topics. Respondents report a sense of accomplishment as they proceed through an instrument that is divided into labeled sections [34].

Natural design

Most respondents are highly familiar with the behaviors, such as reading, that are required to complete a questionnaire. Designers should attempt to take advantage of what comes most naturally to respondents because skilled performance is automatic, fast, and relatively effort-free in comparison to the performance of unfamiliar tasks [25, 26, 27]. Moreover, the expectations that skilled readers have about formatting can affect their ability to understand the material presented to them [35]. For example, placing 'no' before 'yes' in the response options for a question, contrary to conversational norms, can decrease the speed of question answering and may yield responses that are less valid than when the options are ordered according to convention [36].

Creating a natural reading flow
Jenkins and Dillman discuss a general strategy of formatting an instrument to create a natural reading flow, for example, in left to right reading [1]. Figure 3 shows an example of a response scale format that violates the principle of creating a natural reading flow. (This example is taken from an actual clinical report form.) To use the scale, respondents must determine which of the two navigational paths they should follow, across the rows or along the columns. The natural inclination to read from right to left leads respondents to encounter the response options out of order.

When response options are aligned in a vertical column under a multiple choice question, placing the response boxes to the right of the response options creates a natural reading flow. In this design, the respondent can read the text of an item and continue in the same rightward direction to mark their response (see Figure 2b). In contrast, respondents must reverse direction and return to the start of an item to mark their response when the boxes are located to the left of the response options.

When the response options are arrayed horizontally in a row under the question (see Figure 2a), the optimal location for the response

6. During the *Past Four Weeks*, to what extent has your *Physical Health* or *Emotional Problems* interfered with your normal social activities with family, friends, neighbors, or groups?

    ☐ *Not at all*   ☐ *Moderately*   ☐ *Extremely*

    ☐ *Slightly*   ☐ *Quite a bit*

Figure 3. An illustration of a question format that does not follow the principle of natural design. Respondents normally read from right to left, but the response scale follows a column format and there are no directional cues. Additionally, words that are not proper names are capitalized contrary to the conventions of the English language.

boxes is below the response options. Respondent errors occur when researchers position the boxes to the side of each response option. In this arrangement, some boxes appear between two response options on the row and respondents can become confused about which box corresponds to a given response.

Positioning question screens and skip instructions
The structure of most HRQOL instruments is fairly simple relative to many other questionnaires because follow-up questions and complicated skip patterns are not often used. Sometimes, screening instructions are employed in instruments to inform subgroups of respondents to skip over an item that is not relevant to them. These instructions should be located in the established flow of the text to ensure that they are visible to respondents.

When question skip instructions are used, they can be placed immediately to the right of the appropriate response boxes to maintain the natural reading flow (see Figure 4) [1]. Because these instructions are uncommon in HRQOL instruments, respondents may overlook them unless the formatting calls attention to them (e.g., with an arrow). Similarly, when violations of the natural reading flow are necessary (e.g., because of space constraints), it is appropriate to point respondents in the appropriate direction with an arrow.

Using left justification
In some HRQOL instruments (e.g., the *Nottingham Health Profile* [37]) the response boxes are aligned in a column by right-justifying the items on the page. This results in lines of text that have variable starting points, and is more difficult for respondents to read. Centered text also has variable starting points, and respondents are more likely to skip centered headers because they are not located in the natural reading flow [1]. Finally, the irregular spacing in fully-justified text decreases reading speed [38]. Although these are all relatively minor performance costs, the preferred approach is to left-justify all text in order to maintain a natural reading flow.

Clear design

Designers can reduce the cognitive demands of questionnaire navigation by structuring the instrument graphically to make important features stand out clearly and distinctly. This involves using graphical symbols with spatial meaning (e.g., arrows), figure-ground segregation effects, and perceptual grouping effects [1]. However, the effective use of graphics to create a clear design is not always straightforward, as illustrated by the matrix question shown in Figure 5. In this example, the formatting relies on dot leaders to guide respondents across the rows, but the dot leaders end rather abruptly midway through the row. This leads to a sense of discontinuity. The formatting

8. Are you sexually active?

    Yes............................ ☐ 1
    No .... ☐ 2 → Skip to Question 10

Figure 4. A question containing a skip instruction that is positioned in the natural reading flow and emphasized with an arrow.

In the past 7 days, to what extent have you felt each of the following emotions about your health?

		Not at all	A little bit	Somewhat	Quite a bit	Very much
QR1	Sad....................................	0	1	2	3	4
QR2	Satisfied	0	1	2	3	4
QR3	Nervous	0	1	2	3	4
QR4	Worried	0	1	2	3	4

Figure 5. An illustration of a formatting design that attempts to guide respondents with graphics, but has a few limitations. Dot leaders intended to guide respondents abruptly stop. A shaded box used to separate the item codes from the question instead draws attention to the codes.

also includes a shaded box to separate the respondent-irrelevant item codes from the question material. This design highlights the item codes and draw attention to them, and this interferes somewhat with the reading of the items.

Using graphical guides
Respondents frequently do not take the time to read written instructions [17]. However, many graphical symbols, such as arrows, automatically attract and direct attention [39], and can be used to efficiently guide respondent navigation behaviors. Jenkins and Dillman recommend using distinctive visual guides such as upside-down triangles underneath the response scale options in matrix questions to point respondents down the column of response boxes corresponding to each scale option (see Figure 1b) [1]. In addition, they recommend the use of an upside-down open (or three-sided) box over the response options. This helps respondents to refer back to the response options as they need to do so.

Applying figure-ground segregation
Printing the response boxes in reverse white on a gray or colored page background can draw attention to the response boxes and enhance the appearance of the instrument [1]. A minimum of a 60 percent contrast level between text and background is necessary for readability, so one should avoid using lightly colored text on a white background or black text on a medium to dark colored background [40].

Exploiting perceptual grouping effects
Perception research demonstrates that humans automatically organize visual information prior to interpreting its meaning [41]. By manipulating the spacing in an instrument, it is possible to help respondents to perceive distinctions rapidly and appropriately between different questions or different parts of a question [1]. For example, different questions should be separated by slightly more space than the space between the different parts of a single question. This is a key aspect of the layout of the questions on a page.

For instruments composed entirely of multiple choice questions with response options of limited length, two-column formatting will reduce printing costs. However, two-column formatting can increase question density and make pages more complex [42, 43]. Therefore, the space between columns should be adjusted carefully. Designers must strike an important balance between avoiding crowding that reduces the readability of the text and avoiding a disjointed appearance [1].

Selecting print styles and font sizes
Designers can improve respondent reading performance by heeding research on the differential efficiency of reading text printed in different styles and with different characteristics. In general, a Times or Times Roman font facilitates reading because serifs in this font aid in word identification by making letter and word forms more distinctive [44]. Italicized text and words printed in all upper-case letters are more difficult to read than regular text [45, 46, 47, 48].

Text needs to be at least 8-point size for easy reading by normal adults, but a 12-point font should generally be used [45, 46]. However, reading performance with 14-point and smaller fonts declines from ages 60 to 90 years [49], and significant reductions in data quality can be expected if the needs of the elderly and persons with impaired vision are not accommodated [50]. Conservatively, this means using a 16- or 18-point font [51], but compromise is usually necessary to avoid crowding due to space constraints and because younger readers have difficulty reading such large print sizes [52].

Identifying pages clearly
Elderly individuals may experience reduced sensitivity in their fingertips, and sometimes they may turn more than one page without realizing this [53]. Using heavier bond paper is cost-prohibitive, so highly visible page and question numbers are usually the best strategy for ensuring that elderly respondents and others do not skip pages accidentally. Adding the word 'Page' before each page number helps to identify the meaning of these numbers for respondents.

Attractive design

One should not underestimate the influence of an attractive and respondent-friendly questionnaire design on respondent motivation and data quality. By using the techniques described above, an

uncluttered, well-structured, natural-looking, and graphically differentiated instrument can be created which is visually appealing. Nonetheless, an additional element of artistry is also helpful.

Using asymmetric composition
One important element of graphic design is the technique of asymmetric composition [54]. This technique may be effectively employed in questionnaires by placing the main question text in a bolded and slightly larger font than the response options (or the sub-question items in a matrix question) [1]. The distinctiveness of the question text also serves as a guide for respondents as they proceed through the questionnaire. This use of asymmetric composition is shown in Figure 1b, whereas an undifferentiated design is shown in Figure 1a.

Enhancing interest with a title page
When inserting an HRQOL instrument in a clinical report form (CRF), some investigators may place the research protocol information at the top of the first page (see Figure 6). This information is not relevant or meaningful to respondents, and such formatting fails to take advantage of an opportunity to motivate respondents. An attractive cover page that is meaningful to respondents can help increase respondent interest in completing an instrument [1].

Figure 7 provides an example of a respondent-friendly cover page. Investigators should include a title for the instrument that is meaningful and understandable to respondents, not just to other researchers. The cover page should include a brief introductory statement about the general topic of the questions and its importance because this can have significant effects on both respondent participation rates and data quality [55]. The cover page is also a good place to thank respondents in advance for completing the questions. Finally, researchers may place one or more relevant graphical images on the cover page. Respondents find this very appealing, and the practice enhances response rates in mail surveys [34, 56]. However, finding an image that is culturally inclusive for all respondents or that represents abstract concepts such as health and quality of life can be difficult.

Discussion

This paper presents a review of the survey methods research literature and espouses the use of empirically-supported principles and techniques to the formatting of HRQOL instruments. One goal of the paper is to increase knowledge of effective formatting methods. Another goal is to increase awareness among HRQOL researchers about the potentially detrimental effects of ignoring a developer's recommended formatting.

Overall, the original formats of the DFS, FACT-An, KDQOL-SF, SF-36 instruments examined in the present research are reasonable. Nevertheless, the developers of these instruments have indicated that they have made or plan to make adjustments to their instruments based on the findings of the present research. In developing a new instrument, the decision to use the formatting techniques suggested by research in survey

The HRQOL Research Institute
Protocol HR-QOL-00000: Diabetic Foot Ulcer Scale, CAREGIVER   Page 1

Subject Number: ____ ____ ____ Number: 1	Subject Initials: ____ ____ ____	Visit
Visit Date (m/d/y): ____ / ____ / ____		Screening Visit

Diabetic Foot Ulcer Scale: Caregiver

Instructions

These questions ask about the effect...

Figure 6. An example of the type of the material a respondent might see first in an HRQOL instrument that has been insert into a clinical report form. Neither the response protocol information nor the instrument title are relevant or meaningful to respondents.

Your Health

– and –

Well-Being

SF-36 Health Survey

This survey asks for your views about your health. This information will help keep track of how you feel and how well you are able to do your usual activities.

Thank you for completing these questions!

Figure 7. An illustration of a respondent-friendly cover page.

methods may be obvious because the cost of implementation is minimal. By contrast, the decision to *re*format existing instruments according to these methods brings in other factors. Instrument reformatting may be especially beneficial when multiple instruments are to be used in a research study because it can increase the consistency of the tasks for respondents and data entry staff. Large research studies should also consider how reformatting may reduce the costs associated with having research staff answer clarification questions by respondents or assist respondents who have difficulty reading instruments printed in small font sizes.

Instrument psychometrics

Good formatting practices may also benefit data quality. Based on the literature we have cited, we expect that formatting HRQOL instruments according to the guidelines presented here should facilitate the collection of valid and reliable HRQOL data. Clearly, the advantages of reformatting a given instrument will depend on the characteristics of the instrument's current formatting. We have not completed any empirical investigations comparing the psychometrics of reformatted HRQOL instruments to their original developer-produced versions.

The instrument developers, HRQOL researchers, and survey methodologists who we consulted in developing the formatting guidelines have not expressed concerns about potential negative implications of the formatting on instrument psychometrics. A recent study, in which five questions from the SF-36 were embedded in a broader survey and reformatted using some of the techniques we have recommended, obtained reliability statistics well within the range that is expected for these questions [57]. In general, most widely used HRQOL instruments have robust psychometrics that are undisturbed by minor variations in formatting. This robustness appears to hold even when the instruments are converted to computerized touch-screen versions [58]. Nevertheless, care is always required in the reformatting of an instrument.

While the instrument developers who were involved in the present research gave their approval to use reformatted versions of their instruments, they suggested that HRQOL domain mean scores (norms) may change as the result of major formatting revisions. These mean score adjustments are especially likely to occur when the physical structure of a response scale is altered. For example, adjustments may occur when reformatting changes the orientation of the response options in a multiple choice question from vertical to horizontal.

Consider a vertically aligned response scale that is ordered from high to low (e.g., 'all of the time' to 'none of the time'). The high end of this scale is positioned at the top of the column of response options and is the first to be read (see Figure 8a). This position seems natural. When changing the orientation of this scale, it is unclear whether the appropriate place for the high end of the scale is on the left, where it will be the first to be read, or on the right, which seems to be the more natural manner of arranging the order (see Figure 8b). Such reformatting may systematically affect the magnitude of scores (either lowering them or raising them depending on the direction of the response scale); however, if universally applied it is unlikely to affect the relationship among scores.

We strongly encourage research to examine the impact of the recommended formatting on the psychometrics and mean scores (norms) for individual HRQOL instruments. Because of the large sample sizes required for this research, empirical studies might be best completed by instrument developers who are testing new versions of their instruments.

Effects on instrument length

Instrument length is a crucial consideration, especially if certain respondent-friendly formatting techniques, including larger font sizes and a cover page, are used. Common concerns about instrument length include the risk that it will inflate respondent expectations of the burden involved and the monetary and environmental costs of the additional paper. These concerns need to be weighed against costs attributable to poor formatting in the areas of nonresponse (when survey participation is an issue), questionnaire administration, data entry, and data quality. Respondents do not react simply to the length of an instrument when considering

a. Vertical Orientation

1. During the past four weeks, how much of the time has your health interfered with your daily life?

Excellent .... ☐ 1

Very good .... ☐ 2

Good .... ☐ 3

Fair .... ☐ 4

Poor .... ☐ 5

b. Horizontal Orientation

1. During the past four weeks, how much of the time has your health interfered with your daily life?

All of the time	Most of the time	Some of the time	A little of the time	None of the time
▼	▼	▼	▼	▼
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

OR

None of the time	A little of the time	Some of the time	Most of the time	All of the time
▼	▼	▼	▼	▼
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Figure 8. Reformatting the response options in a multiple choice questions from a vertical (a) to a horizontal (b) orientation may have effects on mean scores (norms), and the effects may depend on the order in which the scale in presented in the horizontal orientation.

the burden of responding, they also consider the density of the questions and the manner in which the materials are presented.

Interestingly, our application of the formatting guidelines to the four HRQOL instruments did not lead to a net increase in the number of pages. Some instruments increased in length by a small amount owing mainly to the addition of a cover page. Other instruments decreased in length, which was attributable in part to reductions in the size of respondent-irrelevant material.

Extensions of the present research

Cross-cultural applicability
The formatting guidelines discussed in the present paper were developed to apply for U.S., Western European, and Scandinavian versions of the HRQOL instruments. Care will need to be taken in extending the use of these formatting guidelines to other languages and countries. Of particular concern are potential violations of survey design and

reading conventions, which may lead to respondent confusion.

Interviewer-administered forms
Research investigations can benefit from the application of the sound formatting methods in the preparation of all materials provided to participants, including consent forms. This also includes forms that are used by interviewers. These instruments need to be designed so that the interviewers' task of administering them quickly becomes automatic. Interviewers must divide their attention between the questionnaire and the respondents, and respondents can become distracted and lose interest during pauses in the flow of questions.

Computerized instruments
A host of new concerns arise with the increasing use of computerized and Web-based administrations, such as how to fit large questions within the available space on a screen and how to avoid requiring respondents to switch frequently between

mouse-based and keyboard-based actions. Regardless, the formatting principles described here apply equally well for computerized instruments, although implementing some of the recommended graphical elements may be difficult, depending on the software used.

Special populations
In developing the formatting guidelines, the special needs of the elderly and individuals with poor vision were considered by providing, for example, recommendations for enlarged font sizes. Moreover, a simple and consistent instrument that contains graphical features to automatically direct attention should facilitate the administration of HRQOL instruments to other special populations such as children and cultural groups not familiar with questionnaires. Further research is needed, however, to examine how best to tailor the formatting of HRQOL questionnaires for psychiatric patients and other people with perceptual or attentional disorders. There may be special formatting methods that will specifically benefit these individuals, but applied investigations on the design of forms and the collection of responses from special populations are extremely limited.

Conclusions

This paper demonstrates how HRQOL questionnaire formatting can be informed by scientific research. We have reasoned that the application of practices supported by survey methodological research will improve respondent navigation through HRQOL questionnaires; reduce respondent burden and, possibly, administrative costs; and decrease data errors and nonresponse. To advance the field of HRQOL measurement, these potential advantages to data quality and respondent satisfaction deserve to be explored more fully in well-designed empirical investigations involving HRQOL instruments.

Questionnaire formatting will likely retain some of the characteristics of a graphic art, requiring a practiced eye. Firm prescriptions about some design issues are difficult because they depend on many factors, and designers must frequently weigh tradeoffs. An effective approach to resolving design conflicts is to pretest the questionnaire formatting using cognitive interviewing methods [59, 60].

In the past several years, studies investigating the cognitive processes by which respondents answer HRQOL questions have begun to appear [61]. These studies have the potential to improve greatly the efficiency by which developers craft effective new instruments and the quality of HRQOL data. The present application of established formatting methods and principles to HRQOL instrument development contributes to this emerging research literature, and attempts to further extend the influence of the Cognitive Aspects of Survey Methodology (CASM) movement in survey research [62].

Acknowledgements

References

1. Jenkins CR, Dillman DA. Towards a theory of self-administered questionnaire design. In: Lyberg L, Biemer P, Collins M, et al. (eds), Survey Measurement and Process Quality. New York: John Wiley & Sons, 1997.
2. Wright P. Strategy and tactics in the design of forms. Vis Lang 1980; 14: 151–193.
3. Ware JE, Snow KK, Kosinski M, Gandek B. SF-36 Health Survey Manual and Interpretation Guide. Boston, MA: New England Medical Center, The Health Institute, 1993.
4. Hays RD, Kallich J, Mapes DL, et al. Kidney Disease Quality of Life Short Form (KDQOL-SF), Version 1.3: A Manual for Use and Scoring. Santa Monica, CA: RAND, 1997.

5. Cella D. The Functional Assessment of Cancer Therapy-Anemia (FACT-An) scale: A new tool for the assessment of outcomes in cancer anemia and fatigue. Seminars in Hematol 1997; 34: 13–19.

6. McGrath C, Marquis P, Turner R. Development of a quality of life instrument for use in clinical trials of diabetic foot ulcer treatment. Drug Inform J 1996; 30: 856–857.

7. Wright P, Barnard P. Asking multiple questions about several items: the design of matrix structures on application forms. Appl Ergon 1978; 9: 7–14.

8. Krosnick JA. Response strategies for copying with the cognitive demands of attitude measures in surveys. Appl Cognitive Psychol 1991; 5: 213–236.

9. Lohr KN, Aaronson NK, Alonso J, et al. Evaluating quality-of-life and health status instruments: development of scientific review criteria. Clin Therap 1996; 18: 979–992.

10. Dillman DA, Sinclair MD, Clark JR. Effects of questionnaire length, respondent-friendly design, and a difficult question on response rates for occupant-addressed Census mail surveys. Public Opin Quart 1993; 57: 289–304.

11. Broadbent DE. Perception and Communication. London: Pergamon, 1958.

12. Kahneman D. Attention and Effort. Englewood Cliffs, NJ: Prentice Hall, 1973.

13. Treisman A, Gelade G. A feature integration theory of attention. Cognitive Psychol 1980; 12: 97–136.

14. Eriksen B, Eriksen CW. Effects of noise letters upon the identification of a target letter in a nonsearch task. Perception & Psychophys 1974; 12: 201–204.

15. Baddeley AD, Hitch GJ. Working memory. In: Bower GA (ed), Recent Advances in Learning and Motivation (Vol. 8). New York: Academic Press, 1974; 47–90.

16. Kinsbourne M, Hicks R. Functional cerebral space. In: Reguin J (ed), Attention and Performance VII. Hillsdale, NJ: Erlbaum, 1978.

17. Barnard P, Wright P, Wilcox P. Effects of response instructions and question style on the ease of completing forms. J Occup Psychol 1979; 52: 209–226.

18. Glynn SM, Muth KD. Text learning capabilities of older adults. Educational Gerontol 1979; 4: 253–269.

19. Taub HA, Sturr JF, Monty RA. The effect of underlining cues upon memory of older adults. Exper Aging Res 1985; 11: 225–226.

20. Fowler FJ. Survey Research Methods. Newbury Park, CA: Sage Publications, 1993.

21. Gower AR, Dibbs R. Cognitive research: Designing a 'respondent friendly' questionnaire for the 1991 Census. Proceedings of the Bureau of the Census Fifth Annual Research Conference 1989: 257–266.

22. Ware JE, Kosinski M. SF-36 Health Survey (Version 2.0) Technical Note. Boston, MA: The Health Institute, 1996.

23. Sudman S, Bradburn NM. Asking Questions: A Practical Guide to Questionnaire Design. San Francisco, CA: Jossey-Bass Publishers, 1982.

24. Berdie DR, Anderson JF, Niebuhr MA. Questionnaires: Design and Use. Netuchen, NJ: The Scarecrow Press, 1986.

25. Chandler P, Sweller J. Cognitive load theory and the format of instruction. Cognition and Instruct 1991; 8: 293–332.

26. Sweller J. Cognitive load theory, learning difficulty and instructional design. Learning and Instruct 1994; 4: 295–312.

27. Shiffrin RM, Schneider W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychol Rev 1977; 84: 127–190.

28. Schuman H, Presser S. Questions and answers in attitude surveys. New York: Academic Press, 1981.

29. Mirowsky J, Ross CE. Eliminating defense and agreement bias from measures of the sense of control: A 2×2 index. Social Psychol Quart 1991; 54: 127–145.

30. Bradford JD, Johnson MK. Contextual prerequisites for understanding: some investigations of comprehension and recall. J Verbal Learning and Verbal Behav 1972; 11: 717–726.

31. Wickens C, Carswell CM. The proximity compatibility principle: its psychological foundation and relevance to display design. Human Factors 1995; 37: 473–494.

31. Schriesheim CA, Denisi AS. Item presentations as an influence on questionnaire validity: A field experiment. Educational Psychol Measur 1980; 40: 175–182.

32. Mullin PA, Cashman ER, Straub HR. Answering question sequences: attention switching and memory organization. In: Herrmann D, Johnson M, McEvoy C, Hertzog C, Hertel P (eds), Basic and Applied Memory Research: Practical Aspects (Vol. 2). Hillsdale, NJ: Erlbaum, 1996: 179–192.

33. Hastie R, Park B. The relationship between memory and judgment depends on whether the judgment is memory-based or on-line. Psychol Rev 1986; 93: 258–268.

34. Dillman DA. Mail and Telephone Surveys: The Total Design Method. New York: John Wiley & Sons, 1978.

35. Wright P, Threlfall SM. Readers' expectation about format influence the usability of an index. J Res Commun Studies 1980; 2: 99–106.

36. Holbrook AL, Krosnick JA, Carson RT, Mitchell RC. Violating conversational conventions disrupts cognitive processing of attitude questions. J Exper Social Psychol (in press).

37. Hunt S, McKenna SP, McEwen J, et al. The Nottingham Health Profile: subjective health status and medical consultations. Soc Sci Med 1981; 15A: 221–229.

38. Gregory M, Poulton EC. Even versus uneven right-handed margins and the rate of comprehension in reading. Ergonomics 1970; 13: 427–434.

39. Posner M. Chronometric Explorations of Mind. Hillsdale, NJ: Erlbaum, 1978.

40. Poulton EC. Skimming lists of food ingredients printed in different brightness contrasts. J Appl Psychol 1969; 53: 498–500.

41. Rock I. The Logic of Perception. Cambridge, MA: MIT Press, 1983.

42. Hartley J, Burnhill P, Fraser S. Typographic problems of journal design. Appl Ergon 1974; 5: 15–20.

43. Whalley PC, Fleming RW. An experiment with a simple recorder of reading behavior. Programmed Learning Educational Technol 1975; 12: 120–123.

44. Albers J. Interaction of Color. New Haven, CT, 1975.

45. Tinker MA. Bases for Effective Reading. Minneapolis, MN: University of Minnesota Press, 1965.

46. Spenser H. The Visible Word. London: Lund Humphries, 1969.

47. Poulton EC. Searching for newspaper headlines printed in capitals or lower-case letters. J Appl Psychol 1967; 51: 417–425.

48. Poulton EC. Rate of comprehension of existing teleprinter output of possible alternatives. J Appl Psychol 1968; 52: 16–21.

49. Taub HA, Baker MT, Kline GE, Sturr JF. Comprehension of informed consent information by young-old and old-old volunteers. Exper Aging Res 1987; 13: 173–178.

50. Vanderplas JH, Vanderplas JM. Effects of legibility on verbal test performance of older adults. Perceptual and Motor Skills 1981; 53: 183–186.

51. Charness N, Bosman EA. Age-related changes in perceptual and psychomotor performance: Implications for engineering design. Exper Aging Res 1994; 20: 45–59.

52. Aberson DH, Bouwhuis L. Silent reading as determined by age and visual acuity. J Res Reading 1997; 20: 184–204.

53. Tim Smith. Research Triangle Institute, personal communication, 1997.

54. Edwards B. Drawing on the Right Side of the Brain. Los Angeles, CA: Jeremy P. Tarcher, 1989.

55. Couper MP. Survey introductions and data quality. Public Opin Quart 1997; 61: 317–338.

56. Nederhof AJ. Effects of a final telephone reminder and questionnaire cover design in mail surveys. Social Sci Res 1988; 19: 353–361.

57. McBean AM, Turner CF, Fitterman LK, et al. Monitoring the health status and impact of treatment on older Americans: The Medicare Beneficiary Health Status Registry. Med Care 1999; 37: 189–203.

58. Velikova G, Wright EP, Smith AB, et al. Automated collection of quality-of-life data: A comparison of paper and computer touch-screen questionnaires. J Clin Oncol 1999; 17: 998–1007.

59. Willis GB, Royston P, Bercini D. The use of verbal report methods in the development and testing of survey questions. Appl Cognitive Psychol 1991; 5: 251–267.

60. Redline C, Smiley R, Lee M, DeMaio T, Dillman D. Beyond concurrent interviews: An evaluation of cognitive interviewing techniques for self-administered questionnaires. Proceedings of the American Statistical Association: Section on Survey Research Methods. Alexandria, VA: American Statistical Association, 1998, (in press).

61. Schecter S, Beatty P, Willis G. Asking survey respondents about health status: Judgement and response issues. In: Schwarz N, Park D, Knauper B, Sudman S (eds), Cognition, Aging, and Self-Reports. Philadelphia, PA: Psychology Press 1998: 265–284.

62. Jobe JB, Mingay DJ. Cognition and survey measurement: history and overview. Appl Cognitive Psychol 1991; 5: 175–192.

Address for correspondence: Paul A. Mullin, Research Triangle Institute, P.O. Box 12194, Research Triangle Park, NC 27709-2194, USA
Phone: 919 541 6013; Fax: 919 541 1261
E-mail: pmullin@rti.org