

Visual Layout of Print Questionnaires:
Effect on Responses of Middle School Students

Sheri Bauman

University of Arizona

Robert Steiner

New Mexico State University

Francesca López

University of Arizona

Abstract

A three-page questionnaire was modified to a one-page format. Questionnaires were administered in classroom groups to 300 middle school students. Classrooms were randomly divided into two groups, with one group receiving the original three-page format and the other receiving a single-page version of the same questionnaire. The visual layout of the two versions was different, and included variations in font, placement of response options, and spacing. Item non-response was greater for the single-page format. No statistically significant differences were detected between response patterns or internal consistency of the two versions of the questionnaire. Implications for school counselors are discussed.

Visual Layout of Self-Administered Surveys:
Effect on Responses of Middle School Students

With the increased emphasis on accountability in the field of education including school counseling, practitioners need to collect data about programs and interventions. Doing so involves administering surveys and using pre- and post-tests to evaluate effectiveness. However, most school counselors are master's level practitioners (Bauman, 2004), and as such have had a single course in assessment that is unlikely to cover the relationship between the visual layout of surveys and respondent answers (CACREP, 2001). Because the appropriate visual layout of a survey or questionnaire is essential for limiting measurement error (Dillman & Christian, 2005; Dillman, Redline, & Carley-Baxter, 1999; Sanchez, 1992; Schwartz, 1995; Ware, 2000), an understanding of the effect that various design elements might have on response patterns is important for school counselors seeking to utilize the best measurement techniques for collecting needed data.

There is a body of literature that examines elements of survey and questionnaire construction and the influence of these elements on participant response patterns. Studies of questionnaire format have focused on question type (Dillman & Christian, 2005), question order and placement (Schwartz, 1995; Schwarz, Strack, & Mai 1991), as well as questionnaire length (Gazel, Schwer, & Daneshvary, 1998). Fewer researchers have examined the visual layout of surveys, and the few published studies on this topic have focused on the effects of visual layout in computerized or internet surveys (Couper, 2000; Couper, Traugott, & Lamias, 2001; Dillman & Christian, 2005).

Although researchers have explored how survey questions should be asked when working with children and adolescents (Borgers, Hox, & Sikkel, 2003), no studies were located addressing the visual layout of questionnaires designed for children or adolescents. In order to assess the effectiveness of programs and interventions, children and adolescents in schools are often asked by the school counselor to complete surveys and questionnaires. It is important to consider how the visual layout of surveys and questionnaires may influence the responses of this population.

Cognitive psychologists have long recognized that the words used in verbal communication transmit only a portion of the message. Non-verbal elements such as tone, emphasis, and gestures convey a significant portion of the message to the recipient. These elements of verbal communication have analogous components in written communication (Argyle, 1972). In self-administered questionnaires, the respondent looks for information from the questionnaire on how to answer the questions by using both verbal and visual elements (Couper, Traugott, & Lamias, 2001). Tourangeau, Couper, and Conrad (2004) observed that the verbal language (content) of questionnaires and surveys is supplemented by visual elements such as symbols and graphic layout, although the influence of these elements is less well understood in the context of questionnaires. These visual elements are analogous to the non-verbal components of oral communication.

There is general agreement that the manner in which questionnaire items and response options are displayed on a page is important (Christian & Dilman, 2004; Sanchez, 1992; Smith, 1995), but there is not sufficient understanding of the specific ways in which the influence of graphical layout affect responses. For example, in a

study of Web-based surveys (Couper, Traugott, & Lamias, 2001), responses to items on the same screen were more highly correlated than those items when they appeared on separate screens. The same study also found fewer item non-responses when the items were placed on a single screen than when they were on separate screens. Long sets of items using the same response format, often used in standardized tests, are especially prone to response sets (Herzog & Bachman, 1981; Susan Losh, personal communication, July 12, 2005). Response sets include extreme checking style (selecting the extreme poles of a rating scale), response range, and acquiescence (Hui & Triandis, 1985), and are particularly susceptible to the length of the questionnaire. That is, if a questionnaire or survey is long, respondents tend to choose a response option, such as “agree” repeatedly in later parts of the questionnaire. One of the principal concerns with response sets is that measurement error increases, and subsequently inhibits accurate data analysis. It is important for those school counselors who use surveys to understand whether and how the design of surveys and questionnaires influence the responses of their participants so that they may eliminate design features as competing explainers of results.

The design experiment described in this paper was imbedded in a research project conducted with 300 middle school students in a southwestern city. In that study, four questionnaires were administered, including the Social Experience Questionnaire – Self-Report (Crick & Grotpeter, 1996). This is a questionnaire regarding experiences of bullying in school, and can be used as a pre- and post-test to evaluate an anti-bullying intervention, or as an initial survey to estimate the prevalence of these behaviors in the school.

Order of administration of the four questionnaires was counterbalanced by classroom group. Each of the four questionnaires was printed on a different color paper, and each of the other three questionnaires was on a single page. When the researcher distributed the SEQ-SR to the first classroom, several students exclaimed, “Wow! This one is really long! It’s three pages!” The researcher pointed out that the questionnaire only contained 15 questions, which was a similar number as the other questionnaires, and administration proceeded. However, the researcher was concerned that this initial negative reaction might influence the persistence of students. In addition, there were more student participants in this study than originally anticipated, and more copies of the questionnaire needed to be made for subsequent classes. The three-page version would be more costly to provide.

Thus, the researcher decided to embed a study on the influence of visual layout within the larger study. Based on a review of the literature, the hypotheses were: 1) the internal consistency of the single page version would be significantly higher than that of the three-page version, 2) the correlations for sequential items will be significantly greater when the items appear on the same page than when they are on different pages, and 3) the proportion of item non-response will be lower on the multiple-page version than on the single page version.

Method

Participants

Participants were 300 students (129 males and 171 females) in grades six ($n = 123$), seven ($n = 105$) and eight ($n = 72$) in a middle school in a mid-sized Southwestern U. S. city. Approximately 90% of students in the school qualified for free or reduced-

price lunch in the most recent year for which data are available. Race/ethnicity was determined by self-report of participants. Hispanic/Latino students comprised 89% of the sample ($n = 234$), with 2% ($n = 5$) African American, 6% Caucasian ($n = 17$), 1% Asian American ($n = 3$), and 2% Native American ($n = 5$) making up the rest of the sample. Racial/ethnic information was omitted by 36 students.

Instrument

The *Social Experience Questionnaire – Self Report* (SEQ-SR; Crick & Grotpeter, 1996) is a 15-item questionnaire with three subscales: Overt Victimization, Relational Victimization, and Recipient of Prosocial Behavior. Respondents indicate the frequency with which they have experienced the situation in each statement using a Likert scale with five response options (Never, Almost Never, Sometimes, Almost All the Time, All the Time). Internal consistency coefficients (Cronbach's alpha) were reported by Crick and Grotpeter to be .80, .78, and .77 for the Relational Victimization, Overt Victimization, and Recipient of Prosocial Behavior subscales respectively. In the original version, the appearance of the items is depicted in Figure 1. An example of the alternative format designed for this study is shown in Figure 2. The original version was printed on three pages; the alternative format was printed on a single page.

Procedure

Students with written parental permission who gave written assent to participate were administered four measures, including the SEQ-SR, as part of a research project on school bullying. Questionnaires were administered in classroom groups during one period of a required science class. Questionnaires were read aloud by the researcher. Classes in each grade were randomly divided into two groups. One group received the

DIRECTIONS: Here is a list of things that sometimes happen to kids your age at school. How often do they happen to you at school?

A. How often do you eat lunch at school?

1 NEVER	2 ALMOST NEVER	3 SOMETIMES	4 ALMOST ALL THE TIME	5 ALL THE TIME
------------	----------------------	----------------	-----------------------------	-------------------

B. How often does your class get free time?

1 NEVER	2 ALMOST NEVER	3 SOMETIMES	4 ALMOST ALL THE TIME	5 ALL THE TIME
------------	----------------------	----------------	-----------------------------	-------------------

Figure 1. Appearance of items in the original three-page version of the SEQ-SR.



DIRECTIONS: Here is a list of things that sometimes happen to kids your age at school. How often do they happen to you at school?

	Never	Almost Never	Some- times	Almost all the time	All the time
EXAMPLE A. How often do you eat lunch at school?					
EXAMPLE B. How often does your class go outside to play?					

Figure 2. Appearance of items in the alternative single-page format of the SEQ-SR.

original format (three pages), and the other received the alternative format (single page). Within each classroom group, all students received the same format.

Results

The two format groups (three-page and single page) were compared on demographic variables of gender, ethnicity, and grade to determine whether there were systematic differences between them. No differences were detected on gender or ethnicity. For grade, a chi-square analysis revealed that there were more eighth graders than expected in the group that received the alternative (single-page) format ($\chi^2 = 12.72, df = 2, p < .002$).

To test the first hypothesis, we compared the internal consistency of the two versions using the procedure described by Duhachek and Iacobucci (2004) for comparing Cronbach's alpha coefficients obtained from two independent samples. This new statistical procedure allows researchers to compare to internal consistency coefficients and determine whether observed differences are statistically significant. The procedure involves determining whether zero is included in the multi-sample confidence interval. If zero is included, then the measurement error affecting both samples is equal. If not, the conclusion is that the error is different in the two samples. In this case, $\alpha = .80$ ($n = 111$) for the alternative single-page version and $.72$ ($n = 170$) for the original three-page version. The confidence interval $(-0.00212, 0.16272)$ includes zero so we conclude that there is not a statistically significant difference in the internal consistency of the two versions of the questionnaire. Prior to Duhachek and Iacobucci's procedure, we would have observed that the alpha coefficient for the single-page version was higher, but we would not have been able to determine whether that difference was

statistically significant. In addition to this procedure, chi-square tests of homogeneity of variance across the two formats were calculated for each item. Table 1 presents the results of the item-by-item analysis. Note that none of the differences was significant at $\alpha = .05$ and therefore hypothesis 1 was not supported. In other words, the expected difference in internal consistency in favor of the one-page version was not found.

Table 1

Chi-Square Tests of Homogeneity across Two Formats, Item by Item

Question	Degrees of Freedom	Chi-Square	<i>p</i>
1	4	4.3312	.3630
2	4	3.3940	.4942
3	4	9.3790	.0523
4	4	7.6542	.1051
5	4	0.9794	.9129
6	4	2.9694	.5630
7	4	0.4572	.9775
8	4	2.3927	.6639
9	4	4.6820	.3215
10	4	1.0414	.9035
11	4	0.4575	.9775
12	4	4.6374	.3266
13	4	1.6530	.7992
14	4	1.1245	.8904
15	4	3.5197	.4749

To test the second hypothesis, correlations were calculated for the sequential items that appeared on separate pages in the original version. For items 3 and 4, $r = .271$ ($n = 128$) in the single-page format and $.433$ ($n = 171$) for the three-page format. For items 9 and 10, $r = .387$ ($n = 128$) for the single-page visual layout, and $.238$ ($n = 171$) for the three-page version. Using the procedure described in Hinkle, Wiersma, and Jurs (1994), the correlation coefficients were converted to z scores and compared. The obtained values of z were 1.61 for items 3, 4 and 1.38 for items 9, 10. As the critical value of z at $p < .05$ is 1.64, the differences were not statistically significant and hypothesis 2 was not supported. This means that the correlation coefficients for sequential items did not differ in the two versions. The responses were no more similar when they were on the same page than they were on sequential pages.

To test the missing data hypothesis, the total number of items with missing data was calculated for each form. For the single-page format, 27 item non-responses were counted; for the three-page format, 4 item non-responses were observed. Using the procedure in Hinkle, Wiersma and Jurs (1994) for comparing proportions for independent samples, $z = 5.36$, which exceeds the critical value of 2.33 at $p < .01$. This indicates that the difference in the proportion of item non-response between the two formats is statistically significant. The effect size ($h = .668$) is a medium effect, and means that about 67% of the variance in the item completion rate is explained by the format of the questionnaire. An additional chi-square analysis was conducted comparing the number of missing items (0, 1, or 2) in the two formats. The difference was also significant ($\chi^2 = 11.24$, $df = 2$, $p < .002$). Since we hypothesized that item non-responses would be significantly fewer in the single page format, and the reverse was

found, the hypothesis was not supported. We found the opposite of what we hypothesized; that is, there was more missing data on the single-page version than on the three-page version.

Discussion

Although our hypotheses regarding the influence of the visual layout on the questionnaire response patterns of middle school students were not supported, several useful conclusions can be inferred. Before discussing the practical implications, several caveats will be noted.

One concern about the current study is that the change in design was originally made for practical reasons. Thus, more than one change was involved. The condensation to one page from the original three pages was achieved by a change in placement of response options. Rather than having a separate box with response options labeled within and placed below each item, as in Figure 1, the alternative format used a column approach with the labels for the response options at the top of the column only. Response options were to the right of all items (see Figure 2). In order to improve readability of the condensed version, font was changed as well. The single-page version is more crowded (contains more text) than the multiple-page version. Thus, the alternate version differs from the original in number of pages, font, and placement and labeling of response options, and obtained results cannot isolate which of the changes may be responsible for the findings.

Several aspects of the data need to be mentioned with respect to the comparison of alphas used to address the first hypothesis. First, Likert data are never normally distributed, although a reasonable approximation can be obtained. Second, the two

covariance matrices test as unequal. Finally, although we obtained significant results regarding item non-response, those results were in the opposite direction from the prediction.

We expected the correlations for sequential items to be significantly greater when the items appear on the same page than when they are on different pages based on previous research (Couper, Traugott, & Lamias, 2001). However, it is possible that we did not detect differences in correlations between instruments because of the mode and length of the survey. It remains unclear whether respondent dynamics in a web-based survey differ substantially from a paper format. Because much of the recent research has focused on web-based surveys, more studies that explore the respondent dynamics regarding paper-based surveys are necessary. Additionally, future studies that explore visual layout in surveys designed for children and adolescents should include length as a variable of interest in order to better understand the variables that contribute to accurate data collection.

Although the findings were not what we expected, they do suggest that variations in the visual layout of the type used in this study do not appear to have a significant effect on the consistency of responses of middle school students on a short questionnaire. Interestingly, the rate of item non-response was greater in the single-page format than for the three-page format. This means that school counselors who wish to minimize student concerns about length (to reduce negative bias toward the study), or who wish to reduce expenses for duplication of questionnaires, can condense survey page length without having a large impact on the internal consistency or response patterns for sequential items. However, because item non-response was

greater on the single-page format in this study, and as school counselors typically want to minimize non-response so as to obtain the greatest amount of complete data, they might want to exercise caution before condensing a questionnaire to a single-page format. One strategy that might mitigate the negative reaction to a multi-page questionnaire or survey would be to anticipate the initial reaction and prepare them in advance by pointing out the number of items. Middle school students can think abstractly enough to understand such explanations.

The column format for response options used in the single-page version is widely used in testing for children and adolescents, and since the internal consistency was not compromised by employing this format, school counselors may conclude that students in middle school are able to utilize this format without loss of consistency of responses.

Among school faculty, school counselors are usually the individuals with the strongest training in assessment (Schafer & Lissitz, 1987). It is not surprising that both elementary and secondary school counselors have a better understanding of test selection validity, and ethical practices than administrators and teachers (Impara & Plake, 1995). However, with the increase in accountability and the need for school counselors to utilize appropriate data collection techniques, researchers need to further explore the effect various design elements might have on response patterns. Given that there is virtually no previous research that examined the question of visual layout in printed surveys or questionnaires for children and adolescents, this initial study provides exploratory findings. Further studies investigating item non-response in different formats would be useful. Although web surveys are increasingly popular, printed questionnaires and surveys are not obsolete, and an empirical basis for making decisions about visual

layout will assist researchers by ensuring the quality of their data and school counselors by ensuring that surveys and questionnaires they use will provide the maximum information.

References

- Argyle, M. (1972). Non-verbal communication in human social interaction. In R. A. Hinde (ed.), *Nonverbal communication* (pp. 243-269). Cambridge, NY: Cambridge University Press.
- Bauman, S. (2004). School counselors and research revisited. *Professional School Counseling, 7*, 141-151.
- Borgers, N., Hox, J., & Sikkel, D. (2003). Response quality in survey research with children and adolescents: The effect of labeled response options and vague quantifiers. *International Journal of Public Opinion Research, 15*, 83-94.
- Christian, L. M., & Dilman, D. A. (2004). The influence of graphical and symbolic language manipulations on responses to self-administered questions. *Public Opinion Quarterly, 68*, 57-80.
- Council for Accreditation of Counseling and Related Educational Program (CACREP). (2001). 2001 standards. Retrieved March 26, 2007 from <http://www.cacrep.org/2001Standards.html>
- Crick, N. R., & Grotpeter, J. K. (1996). Children's treatment by peers: Victims of relational and overt aggression. *Development and Psychopathology, 8*, 367-380.
- Couper, M. P. (2000). Review: Web surveys: A review of issues and approaches. *The Public Opinion Quarterly, 64*, 464-494.
- Couper, M. P., Traugott, M. W., & Lamais, M. J. (2001). Web survey design and administration. (data quality in self-administered surveys). *Public Opinion Quarterly, 65*, 230-253.

- Dillman, D. A. & Christian, L. M. (2005). Survey mode as a source of instability in responses across surveys. *Field Methods, 17*, 30-52.
- Dillman, D. A., Redline, C. D., & Carley-Baxter, L. R. (1999). Influence on type of question on skip pattern compliance in self-administered questionnaires. Paper presented at the Joint Statistical Meetings of the American Statistical Association, Indianapolis.
- Duhachek, A. & Iacobucci, D. (2004). Alpha's standard error (ASE): An accurate and precise confidence interval estimate. *Journal of Applied Psychology, 89*, 792-808.
- Herzog, A. R & Bachman, J. G. (1981). Effects of questionnaire length on response quality. *The Public Opinion Quarterly, 45*, 549-559.
- Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1994). *Applied statistics for the behavioral sciences* (3rd ed.). Boston: Houghton Mifflin.
- Hui, C. H., & Triandis, H. C. (1985). The instability of response sets. *Public Opinion Quarterly, 49*, 253-261.
- Impara, J. C. & Plake, B. S. (1995). Comparing counselors', school administrators', and teachers' knowledge in student assessment. *Measurement and Evaluation in Counseling and Development, 28*, 78-87.
- Gazel, R. C, Schwer, R. K., & Daneshvary, R. (1998). Interview mode choice by survey respondents: A methodological analysis. *Social Science Computer Review, 16*, 185-191.
- Sanchez, M. E. (1992). Effects of questionnaire design on the quality of survey data. *Public Opinion Quarterly, 56*, 206-217.

- Schafer, W. D. & Lissitz, R. W. (1987). Measurement training for school personnel: Recommendations and reality. *Journal of Teacher Education*, 38, 57-63.
- Schwartz, N. (1995). What respondents learn from questionnaires: The survey interview and the logic of conversation. *International Statistical Review*, 63, 153-177.
- Schwarz, N., Strack, F., & Mai, H. P. (1991). Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis. *Public Opinion Quarterly*, 55, 3-23.
- Smith, T. W. (1995, August). *Little things matter: A sampler of how differences in questionnaire format can affect survey responses*. Paper presented at the annual meeting of the American Statistical Association, Orlando, FL.
- Tourangeau, R., Couper, M., & Conrad, F. (2004). Spacing, position, and order: Interpretive heuristics for visual features of survey questions. *Public Opinion Quarterly*, 68, 368-393.
- Ware, C. (2000). *Information visualization: Perception for design*. San Francisco: Morgan Kaufman.

Authors Biographical Statement

Dr. Sheri Bauman is an associate professor and director of the School Counseling and Guidance master's degree program at the University of Arizona. Prior to earning her doctorate in 1999, Dr. Bauman worked in public schools for 30 years, 18 of those as a school counselor. She is a licensed psychologist, and maintains a small private practice in Las Cruces, New Mexico.

Dr. Bauman teaches in the master's program and conducts research on bullying, professional issues in school counseling, and group work. She is a member the editorial boards for the *Journal of Counseling and Development*, *Professional School Counseling*, and the *Journal for Specialists in Group Work*. In July 2007, she will become editor of the *Journal for Specialists in Group Work*. Dr. Bauman is involved in several professional organizations, with current memberships in APA, ACA, AERA, ASCA, and ASGW. She has presented at state, national, and international conferences. She will be presenting on cyberbullying at the National Coalition Against Bullying conference in Melbourne, Australia in November 2007. Her book, *Essential Topics for the Helping Professional*, is scheduled for publication by ABLongman in August, 2007. Her vita lists numerous publications in peer-reviewed journals.

Dr. Robert Steiner is a professor in the Department of Economics and International Business at New Mexico State University, Las Cruces, NM. He consults widely on statistical questions in a variety of fields.

Ms. Francesca López is a doctoral candidate in the Department of Educational Psychology at the University of Arizona. Her research interests include acculturation and the academic achievement of Mexican American students.