# QUANTITATIVE ATTITUDES QUESTIONNAIRE: INSTRUMENT DEVELOPMENT AND VALIDATION

LEI CHANG
Chinese University of Hong Kong

This article reports on the development and validation of the Quantitative Attitudes Questionnaire (QAQ), a 20-item instrument measuring graduate students' attitudes toward quantitative research methodology. The QAQ was developed in a sample of 90 graduate students in education and was cross validated in 3 additional samples of similar size. The QAQ was found to correlate positively with quantitative exams and self-reported GRE-Math scores, and to differentiate doctoral versus master's students and students from research- versus teaching-oriented universities.

Assessing social science graduate students' attitudes toward quantitative research methodology is an important step in fostering appreciation of research and its methodology among these students. Such assessment results are also useful for curriculum development and instructional evaluation, because attitude change can be an important indicator of instructional success. However, there is currently no instrument assessing quantitative attitudes of graduate students in the social sciences. Existing attitude instruments that are most closely related to quantitative methodology include the Attitude Toward Statistics Scale developed by Wise (1985), the Statistics Attitude Survey by Roberts and Bilderback (1980), the Attitude Toward Education Research Scale (Napier, 1978), and the Mathematics Attitude Scale (McGallon & Brown, 1971). Most of these instruments target the undergraduate population. This article reports on the development of the Quantitative Attitudes Questionnaire (QAQ), a short survey prepared for graduate students in education and other social science disciplines.

The author thanks Robert Baker, Cecilia Barron, Susan Cohen, Gerry Ledford, Angela McBride, and Catherine McBride-Chang for their help with instrument development, and Larry Daniel and two anonymous reviewers for their helpful comments on the manuscript. Correspondence regarding this article should be addressed to Lei Chang, Department of Educational Psychology, Chinese University of Hong Kong, Shatin, N.T., Hong Kong. This study was conducted while the author was at the University of Central Florida..

Educational and Psychological Measurement, Vol. 56 No. 6, December 1996 1037-1042

## Development of the QAQ

For this instrument, quantitative attitudes were defined as consisting of four components: (a) utility (the extent to which one finds quantitative methodology useful in his or her work), (b) value (how important one thinks quantitative methodology is in social science research), (c) efficacy (one's belief in one's quantitative abilities), and (d) knowledge (how much one thinks or feels he or she knows about quantitative methodology).

These four related attitudes were believed to have an impact on graduate students' performance in research methodology and other quantitative areas. They were derived by six faculty judges representing the fields of education, psychology, nursing, and business administration from four universities. These four constructs were subsequently translated into test items that were intended to duplicate the constructs. An initial pool of 65 items was developed and was reviewed by the same judges for content validity. Forty-five items receiving consensus endorsements from the judges were retained.

These 45 items, together with nine filler items measuring acquiescence or overall positive opinion, were administered to a convenience sample of 90 master's students in education from three sections of a research methods course and a statistics course taught by the author. The items were administered at the beginning of the semester using a 6-point Likert-type scale anchored by the commonly used *disagree* and *agree* designations. Students were asked to comment on the items. Based on their comments, some items were eliminated or revised. The revised items were administered 1 week later to the same students using a 4-point scale. Students were again asked to comment on the items and to indicate which of the two scales, 4-point versus 6-point, they preferred.

Five criteria were used in the subsequent item analyses. Specifically, satisfactory items were expected to have (a) few negative comments from the students, (b) high test-retest correlations over the two administrations, (c) high loadings on respective factors, (d) low correlations with filler items, and (e) positive correlations with a composite score made up of the midterm and final exams of the two quantitative courses the students were taking at the time. Not all criteria were equally adhered to. For instance, for some items that were not retested, the test-retest criterion did not apply.

These item analyses resulted in two test forms. The long form contained 30 items and the short form contained 20 out of 30 items. Both forms used the 6-point scale, which most students preferred. The following validation studies were based on the 20-item QAQ.

## Validation Studies

The 20-item QAQ was subsequently administered to three convenience samples. The first sample contained 120 master's students in education who

were enrolled in different sections of a research methods course and a statistics course taught by the author. The second sample included 92 Ed.D. and counseling specialist students. Both samples were taken from a regional teaching university in the United States. The third sample contained 60 Ph.D. students and 14 faculty in psychology from a major research university in the United States. Gender composition of the sample of master's students was 80% female. There was no significant gender difference on any of the four QAQ factors. Gender information was not gathered for the other samples. However, the two departments from which the samples were taken were about 70% female.

Confirmatory factor analysis (CFA) using LISREL-7 was conducted within each of the three samples. In all three samples, the four-factor QAQ structure was identified and shown to be superior to two alternative (one-factor and five-factor) structures. Although chi-square tests of goodness of fit were statistically significant, other goodness-of-fit indexes indicated acceptable fit. For example, the average chi-square to degree of freedom ratio was 1.7. Interfactor correlations were .30 on average. The average standardized factor loadings were above .70. Internal consistency reliability estimates were similarly high. (CFA results are available from the author.) Reliability and validity information is included in Table 1.

Criterion-related validity was investigated in Sample 1 by correlating the four QAQ scores with a composite made up of the midterm and final exams of the two quantitative courses the students were enrolled in. The midterms for the two courses each had 76 multiple-choice items. The finals for the two courses each had 40 multiple-choice items. The four exams had internal consistency reliabilities above .80. Because of the different item difficulties, $z$ scores were used to compute the criterion composite. The average correlation between the four QAQ scores and the criterion was .34. Among the four factors—utility, value, efficacy, and knowledge—utility correlated lowest at .22 and knowledge correlated the highest at .46. It seems that quantitative performance is more reflective of how much one thinks one knows about quantitative methodology than how important one thinks the methodology is. Similar but stronger validity evidence was observed in Sample 2, where self-reported GRE-Math scores were used as the criteria. The average correlation was .41. Efficacy and knowledge had the highest correlations with GRE-Math at .57 and .59, respectively. Overall, these correlations showed that this attitude instrument could differentiate graduate students of high versus low quantitative performance. These results support a criterion-related validity inference of the QAQ's application.

Mean differences among the three samples on four scales were investigated using MANOVA. Hotelling's test was statistically significant, approximate $F(8, 558) = 36.68, p < .01$. As expected, Sample 3, which included Ph.D. psychology students and faculty from a major research university, had significantly higher means than either of the other two samples on all QAQ

Table 1
*Descriptive Statistics and Reliability and Validity Estimates for QAQ Scores Across Samples*

| Scale | *M* | *SD* | α | Validity Estimate[a] |
|---|---|---|---|---|
| Sample 1 ($n = 120$) | | | | |
| Utility | 4.07 | 0.87 | .81 | .35 |
| Value | 4.69 | 0.72 | .77 | .22 |
| Efficacy | 3.33 | 1.12 | .84 | .33 |
| Knowledge | 3.18 | 0.81 | .80 | .46 |
| Total score | 3.82 | 0.55 | .85 | .35 |
| Sample 2 ($n = 92$) | | | | |
| Utility | 5.18[b] | 1.11 | .77 | .29 |
| Value | 4.13[c] | 0.97 | .86 | .20 |
| Efficacy | 3.61 | 1.13 | .75 | .57 |
| Knowledge | 3.60[c] | 0.99 | .78 | .59 |
| Total score | 4.14 | 0.67 | .80 | .39 |
| Sample 3 ($n = 74$) | | | | |
| Utility | 5.35[d] | 0.68 | .82 | |
| Value | 5.00[e] | 0.79 | .74 | |
| Efficacy | 4.17[e] | 1.01 | .87 | |
| Knowledge | 4.70[e] | 0.76 | .80 | |
| Total score | 4.83 | 0.62 | .89 | |

*Note.* Sample 1 contained master's students in education enrolled in two quantitative courses. Sample 2 contained Ed.D. and specialist students. Both samples were drawn from a teaching university. Sample 3 contained faculty and Ph.D. students in psychology from a major research university. Mean comparisons were made on the four scales but not on the total score; familywise error rate exceeds the reported $p$ value.

a. Correlations with quantitative course exam for sample 1 and GRE-Math for sample 2.

b. Higher than Sample 1, $p < .01$.

c. Lower than Sample 1, $p < .01$.

d. Higher than Sample 1, $p < .01$.

e. Higher than Samples 1 and 2, $p < .01$.

factors except for utility, where the difference between Sample 3 and Sample 2 was not statistically significant. These results strengthen the validity inference of the use of QAQ. Doctoral students from a research university, who were heavily engaged in research, were expected to show more positive attitudes toward quantitative methodology than students from a teaching university, who had almost no research experience. Between the two samples of Ed.D./counseling specialist students (Sample 2) and master's students (sample 1), both of which were taken from the teaching university, the former was expected to have higher means than the latter, because there was more methodological training in the Ed.D. and specialist programs than in the master's program. On utility and knowledge, Sample 2 had significantly higher scores than Sample 1. However, Sample 1 had a higher mean on the value factor. This last finding could be due to the fact that, because Sample 1 students were enrolled in the quantitative courses at the time, they could be more positive than usual about the research value of quantitative methodology.

## Conclusion

The present study reports on 2 years of properly planned and implemented instrument development and validation procedures. Data presented here provide evidence that the 20-item QAQ is an efficient instrument to assess graduate students' attitudes toward quantitative methodology. However, there are two limitations. First, one of the criterion measures—the GRE-Math score—was based on self-reports that could contain errors. Second, gender information was incomplete. Although one of the degree programs showed no gender difference, and a similar gender composition was expected in the three programs, the mean differences found across programs could still have been confounded by potential gender differences. Validation is an ongoing process. Further research is needed to draw a stronger validity inference in the use of this new instrument.

## Appendix
## The Quantitative Attitudes Questionnaire Items

1. Knowledge of quantitative research methods is useful for my job.
2. The "truth" or falsity of a research question has to be tested by empirical data.
3. I need to know research methodology in order to do my own research.
4. I'm confident in my quantitative ability.
5. Any theory "worth its salt" has to be subjected to data-based quantitative tests.
6. I understand the basic principles of hypothesis testing and statistical inference.
7. A sound methodology is essential for quality research.
8. I see the usefulness of quantitative research methodology in my life.
9. I enjoy working with numbers.
10. I have a thorough understanding of quantitative research methods.
11. I understand the basic principles of classical test theory.
12. I'm good with numbers.
13. Quantitative research methodology is useful for my career.
14. I understand the interrelations among measurement, statistics, and research design.
15. Math has been one of my favorite subjects in school.
16. I need to keep up with quantitative development to do my job well.
17. I know which statistical procedure to use to test my hypothesis.
18. A good researcher must have a strong background in quantitative methodology.
19. Compared to others I know, I'm very good in quantitative subjects.
20. Statistical tools are invaluable for understanding and interpreting one's data.

*Note.* The scale Perceived Utility for Oneself comprises items 1, 3, 8, 13, and 16; the scale Perceived Value in Research comprises items 2, 5, 7, 18, and 20; the scale Quantitative Efficacy comprises items 4, 9, 12, 15, and 19; the scale Perceived Quantitative Knowledge comprises items 6, 10, 11, 14, and 17.

## References

McGallon, E. L., & Brown, J. D. (1971). A semantic differential instrument for measuring attitude toward mathematics. *Journal of Experimental Education, 39*(4), 69-72.

Napier, J. (1978). An experimental study of the relationship between attitude toward and knowledge of education research. *Journal of Experimental Education, 47*(2), 131-134.

Roberts, D. M., & Bilderback, E. W. (1980). Reliability and validity of a statistics attitude survey. *Educational and Psychological Measurement, 40*, 235-238.

Wise, S. L. (1985). The development and validation of a scale measuring attitudes toward statistics. *Educational and Psychological Measurement, 45*, 401-405.