Reprint from:

EDUCATIONAL AND
PSYCHOLOGICAL MEASUREMENT

"Improving the Estimation of Moderating Effects by Using Computer-Administered Questionnaires"

by Herman Aguinis, William H. Bommer, and Charles A. Pierce

Books
Journals
Newsletters
University Papers
Annual Series

2455 Teller Road
Newbury Park
California 91320

Phone (805) 499-0721
Fax (805) 499-0871
Cable SagePub
Telex (510) 1000799

Sage Publications, Inc.

IMPROVING THE ESTIMATION OF MODERATING EFFECTS BY USING COMPUTER-ADMINISTERED QUESTIONNAIRES

HERMAN AGUINIS
University of Colorado at Denver

WILLIAM H. BOMMER
Southern Illinois University—Edwardsville

CHARLES A. PIERCE
Montana State University

A program designed to administer questionnaires on IBM and IBM-compatible personal computers is documented. The program prompts subjects to indicate their response by clicking on a graphic line segment displayed on the screen or by entering a numeric value using the keyboard. Features of the program include (a) easily modifiable anchors for the graphic line segment (e.g., *agree-disagree*, *satisfied-dissatisfied*), (b) no limit to the number of questions included, and (c) responses that are stored directly into an ASCII file. The program enhances the accuracy in estimating moderating effects because it overcomes two limitations of the more traditional paper-and-pencil mode of administration: (a) transcriptional errors and (b) scale coarseness.

Increasing levels of theoretical sophistication in education and psychology often lead researchers to test hypotheses regarding not only main effects but also moderating or interaction effects. A variable Z is defined as a

A previous version of this article was presented at the annual meeting of the Eastern Psychological Association, Boston, April 1995. Preparation of this article was facilitated by a Faculty Grant Award and a Summer Research and Development Stipend to Herman Aguinis from the University of Colorado at Denver. We thank John D. Lowe for his assistance in Turbo C programming. Address correspondence and reprint requests to Herman Aguinis, Department of Psychology, University of Colorado at Denver, Campus Box 173, P.O. Box 173364, Denver, CO 80217-3364, E-mail haguinis@castle.cudenver.edu.

Educational and Psychological Measurement, Vol. 56 No. 6, December 1996 1043-1047

moderator of an *X*-*Y* relationship when the nature of this *X*-*Y* relationship is contingent upon values or levels of *Z* (Aguinis, Pierce, & Stone-Romero, 1994; Zedeck, 1971).

Hypothesized effects of moderator variables are often tested with data collected using paper-and-pencil questionnaires (PPQs) (e.g., Aguinis, Nesler, Quigley, Lee, & Tedeschi, 1996; Nesler, Aguinis, Quigley, & Tedeschi, 1993). However, despite the ubiquitous use of PPQs in educational and psychological research, these questionnaires have several shortcomings that become particularly evident when estimating moderating effects. These limitations of PPQs are observed in meta-analytic as well as primary-level studies.

A limitation of PPQs that concerns meta-analysts interested in estimating moderating effects is transcriptional errors. Validity generalization researchers recognize that across-study variability observed in validity coefficients is due, in part, to computer data entry errors (Hunter & Schmidt, 1990). This artifactual across-study variability can be erroneously attributed to moderating effects, and researchers may commit a Type I statistical error regarding the null hypothesis of no moderating effect (Schmidt et al., 1993). Stated differently, variability due to transcriptional errors may be falsely attributed to the "detected" moderator variable.

A concern for primary-level researchers regarding the estimation of moderating effects with data collected using PPQs is what Russell and his colleagues labeled criterion variable "scale coarseness" (Russell & Bobko, 1992; Russell, Pinto, & Bobko, 1991). Scale coarseness has been identified as an artifact that adversely affects the statistical power of moderated multiple regression (MMR) to detect moderating effects (for descriptions of additional factors that lower the power of MMR, see Aguinis, 1995; Aguinis & Pierce, 1995; Aguinis & Stone-Romero, in press; Stone-Romero, Alliger, & Aguinis, 1994). Scale coarseness refers to operationalizing a criterion variable in such a manner that it does not include sufficient scale points, incurs in possible information loss, and thus prevents a hypothesized moderating effect from being detected.

If, for example, a predictor variable *X* and a hypothesized moderator variable *Z* are measured on 7-point Likert-type scales, the product term $X \times Z$ has a possible range of $7 \times 7 = 49$ distinct responses. However, when a PPQ is used, the criterion *Y* is typically measured on a "coarse" 7-point scale rather than on a 49-point scale. Thus information regarding the relationship between *Y* and $X \times Z$ (which carries information regarding the moderating effect) is lost and, therefore, statistical power decreases inevitably below Cohen's (1988) recommended level of .80. In an experiment supporting the aforementioned argument (Russell & Bobko, 1992), subjects were assigned to one of two conditions. They responded to a dependent variable consisting of either (a) a 5-point Likert-type scale or (b) a continuous graphic line segment on which they placed a mark indicating their response. As expected, the esti-

mated size of the moderating effect was larger when respondents used the continuous, as opposed to the "coarse," criterion scale.

The Program

A computer program was designed to address the need to improve the accuracy in estimating moderating effects by avoiding (a) transcriptional errors, which lead to increased Type I error rates (i.e., falsely discovering a moderator); and (b) scale coarseness, which leads to increased Type II error rates (i.e., erroneously dismissing a model that includes a moderator). The program was written in Turbo C and runs on IBM and IBM-compatible personal computers using CGA, EGA, or VGA monitors.

The researcher first creates an ASCII file that includes the instructions, questions, debriefing, and other statements to be displayed on the screen. Respondents then provide their answers to each of the questions by using a mouse interface and clicking on any point on the continuum or, for specific researcher-defined questions, by using numbers on the keyboard (e.g., for demographic and background information). If desired, researchers can easily adjust the mouse sensitivity by modifying the default sensitivity index (s_index).

Respondents advance through the questions sequentially by clicking a NEXT icon displayed at the bottom of each screen. Also, as is typical in computer-based testing (CBT) (e.g., Lautenschlager & Flaherty, 1990; Lee, Moreno, & Sympson, 1986), respondents are not provided the opportunity to revise a response to a previous item. Note that despite this apparent difference between CBT and PPQs, research examining the equivalence of these two types of administrations generally indicates that they yield similar results (e.g., Booth-Kewley, Edwards, & Rosenfeld, 1992; Mead & Drasgow, 1993). Finally, all responses are stored directly into an ASCII file, thereby eliminating the need to perform any additional steps before conducting subsequent analyses using statistical software packages such as SPSS, SAS, or BMDP.

Input

The instructions, questions, and debriefing information that is displayed on the screen is first entered by the researcher into an ASCII file (CAQ.TXT). In this file, each question or statement ends with the symbol | followed by a number n (where $2 \leq n \leq 999999$) or the symbol A, S, or D. The number n indicates that the answer to an item will be entered using the keyboard, and the value taken on by n indicates the highest possible score. This is a useful feature for avoiding out-of-range values. Alternatively, the symbols A and S indicate that the item will be answered using the mouse interface on the

graphic line segment; *A* indicates that the anchors *agree-disagree* will appear on the screen; *S* indicates that the anchors *satisfied-dissatisfied* will be displayed. Researchers can easily edit the source-code version of the program to append other anchors as needed (e.g., *always-never*). Finally, the symbol *D* indicates that the preceding text should only be displayed on the screen and that no input from the respondent is expected. For instance, text for the instructions and debriefing should be followed by | *D*.

Output

All responses are stored directly into an ASCII file (CAQ.OUT) with six digits per response. Responses provided on the line graph segment may range from the utmost left extreme of the scale (i.e., 000000, indicating 0% of the line) to the utmost right extreme (i.e., 100000, indicating 100.000% of the line) in increments of approximately 000160 (i.e., .16%) on EGA/VGA monitors, and 000318 (i.e., .32%) on CGA monitors. Numerical responses from the keyboard are also stored using six digits (e.g., a 2 is stored as 000002). The first question in the program, which is a default that can be altered by the researcher, inquires about the last four digits of the respondent's social security number as a method for identifying unique cases during subsequent data analyses. The participants' responses are appended to the existing CAQ.OUT file. Thus the same floppy diskette can be used for as many respondents as the available disk space permits.

Program Availability

The executable (CAQ.EXE) and source (CAQ.C) versions of the program, along with an ASCII file that includes illustrative instructions, questions, and a debriefing statement (CAQ.TXT), are available at no cost on either a 3.5-in. or 5.25-in. diskette (double or high density). Individuals who wish to obtain the program should send a blank formatted diskette and a self-addressed, stamped envelope to Herman Aguinis, Department of Psychology, University of Colorado at Denver, Campus Box 173, P.O. Box 173364, Denver, CO 80217-3364.

References

Aguinis, H. (1995). Statistical power problems with moderated multiple regression in management research. *Journal of Management, 21*, 1141-1158.

Aguinis, H., Nesler, M. S., Quigley, B. M., Lee, S., & Tedeschi, J. T. (1996). Power bases of faculty supervisors and educational outcomes for graduate students. *Journal of Higher Education, 67*, 267-297.

Aguinis, H., & Pierce, C. A. (1995, April). *Combating heterogeneity of residual variance in moderator variable detection*. Paper presented at the meeting of the Rocky Mountain Psychological Association, Boulder, CO.

Aguinis, H., Pierce, C. A., & Stone-Romero, E. F. (1994). Estimating the power to detect dichotomous moderators with moderated multiple regression. *Educational and Psychological Measurement, 54*, 690-692.

Aguinis, H., & Stone-Romero, E. F. (in press). Methodological artifacts in moderated multiple regression and their effects on statistical power. *Journal of Applied Psychology.*

Booth-Kewley, S., Edwards, J. E., & Rosenfeld, P. (1992). Impression management, social desirability, and computer administration of attitude questionnaires: Does the computer make a difference? *Journal of Applied Psychology, 77*, 562-566.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.

Lautenschlager, G. J., & Flaherty, V. L. (1990). Computer administration of questions: More desirable or more social desirability. *Journal of Applied Psychology, 75*, 310-314.

Lee, J. A., Moreno, K. E., & Sympson, J. B. (1986). The effects of mode of test administration on test performance. *Educational and Psychological Measurement, 46*, 467-474.

Mead, A. D., & Drasgow, F. (1993). Equivalence of computerized and paper-and-pencil cognitive ability tests: A meta-analysis. *Psychological Bulletin, 114*, 449-458.

Nesler, M. S., Aguinis, H., Quigley, B. M., & Tedeschi, J. T. (1993). The effect of credibility on perceived power. *Journal of Applied Social Psychology, 23*, 1407-1425.

Russell, C. J., & Bobko, P. (1992). Moderated regression analysis and Likert scales: Too coarse for comfort. *Journal of Applied Psychology, 77*, 336-342.

Russell, C. J., Pinto, J., & Bobko, P. (1991). Appropriate moderated regression and inappropriate research strategy: A demonstration of the need to give your respondents space. *Applied Psychological Measurement, 15*, 257-266.

Schmidt, F. L., Law, K., Hunter, J. E., Rothstein, H. R., Pearlman, K., & McDaniel, M. (1993). Refinements in validity generalization methods: Implications for the situational specificity hypothesis. *Journal of Applied Psychology, 78*, 3-12.

Stone-Romero, E. F., Alliger, G. M., & Aguinis, H. (1994). Type II error problems in the use of moderated multiple regression for the detection of moderating effects of dichotomous variables. *Journal of Management, 20*, 167-178.

Zedeck, S. (1971). Problems with the use of "moderator" variables. *Psychological Bulletin, 76*, 295-310.