

Moderated Regression Analysis and Likert Scales: Too Coarse for Comfort

Craig J. Russell

Department of Organizational Behavior and
Human Resource Management
Purdue University

Philip Bobko

Department of Management
Rutgers University

One of the most commonly accepted models of relationships among three variables in applied industrial and organizational psychology is the simple moderator effect. However, many authors have expressed concern over the general lack of empirical support for interaction effects reported in the literature. We demonstrate in the current sample that use of a continuous, dependent-response scale instead of a discrete, Likert-type scale, causes moderated regression analysis effect sizes to increase an average of 93%. We suggest that use of relatively coarse Likert scales to measure fine dependent responses causes information loss that, although varying widely across subjects, greatly reduces the probability of detecting true interaction effects. Specific recommendations for alternate research strategies are made.

Saunders (1955, 1956) was the first to describe stepwise or hierarchical moderated regression analysis as a means of empirically detecting how a variable "moderates" or influences the nature of a relationship between two other variables. Statistical textbooks (e.g., Cohen & Cohen, 1983) have introduced later generations of investigators to the procedure. A simple count of the number of studies examining moderator effects in major applied psychology journals indicates that moderated regression analysis is the preferred statistical procedure for detecting interaction effects. Most applications involve random-effects designs in field settings where surveys are used to measure individual and organizational characteristics of interest. Furthermore, many theories in psychology and organizational settings postulate moderator or interactive relationships. Unfortunately, many authors have noted how rare it is for investigations to report strong, unambiguous results in support of a moderator effect (Bobko, 1986; Cronbach, 1987; Drazin & Van de Ven, 1985; Sockloff, 1976a, 1976b; Venkatraman, 1989; Zedeck, 1971). For example, one of the oldest and almost universally accepted models of work performance involves an interactive function of motivation and ability (Maier, 1955). Terborg (1977) reviewed 14 articles containing 20 tests of this interaction, finding only five results supportive of the interaction effect. Cronbach (1987) suggested that investigators redirect their attention to basic research design issues if they wish to detect true interaction effects.

The current study focused on how characteristics of the response scale affect the power of moderated regression. Specifi-

cally, a basic assumption in field studies is that the relationship between the "true" or latent variable of interest and the observed questionnaire response (for both independent and dependent variables) is linear. Busemeyer and Jones (1983) examined this assumption in "observational" or random-effects designs typically found in applied organizational research. They demonstrated that when relationships between the latent and observed variables follow some unknown, nonlinear monotonic function, moderated regression results are uninterpretable. Assumptions of linear relationships between latent constructs and observed scale scores are so common that they are rarely noted in even the most empirically oriented journals.

Pursuant to Cronbach's (1987) suggestion, Russell, Pinto, and Bobko (1991) investigated how a basic assessment design issue may be forcing subjects to operationalize latent dependent responses in a nonlinear manner. Specifically, Russell et al. considered the possibility that discrete Likert-type scales used to obtain subjects' dependent responses in interactive models may result in information loss. If five levels of the predictor (x) and moderator (z) are presented to subjects in an orthogonal fixed-effects design, the dependent response (y) produced by a "true" moderator effect will contain $5 \times 5 = 25$ conceptually distinct latent responses. (These "latent responses" do not constitute an observable random variable [y] but instead represent psychological representations of the construct of interest.) Russell et al. used a 5×5 fixed-effects design for purposes of exposition. Use of fixed- versus random-effects designs is irrelevant to conclusions drawn concerning the effects of Likert scales on measurement of the dependent variable, because the concern is with the effect of Likert scales on dependent responses.

Now, suppose subjects are provided with a 5-point Likert scale with which to portray their dependent response (y ; an observable random variable that can assume only five values). Then, the relatively coarse 5-point Likert scale will be associated with information loss, because the latent dependent response has 25 possible distinct values. Russell et al. (1991) speculated that the Likert scale requires subjects to somehow squeeze or otherwise reduce their latent response in order to

An earlier version of this article was presented at the 30th Annual Meeting of the Southern Management Association, November 1991, Atlanta, Georgia.

We would like to thank our colleagues around the United States who participated in this study (especially Dale Rude and Dirk Steiner), and we would like to thank Larry James and three anonymous reviewers for their comments.

Correspondence concerning this article should be addressed to Craig J. Russell, who is now at Department of Management, Louisiana State University, Baton Rouge, Louisiana 70803-6312.

generate an answer on the overt Likert scale. They simulated two alternate means by which subjects might reduce their latent response. These stimulated results suggested that information loss due to coarseness of the dependent scale can cause spurious increases or decreases in moderated regression effect sizes, depending on how the reduction takes place.

The effect of information loss on moderated regression analysis is not surprising. Peters and Van Voorhis (1940) and many others have demonstrated the impact of information loss in applications of correlational analysis (Cohen, 1983; Olsson, Drasgow, & Dorans, 1982) and structural equation modeling (Muthén, 1984). Russell et al. (1991) suggested that the decision to use Likert scales in operationalizing the dependent variable causes information loss that results in unknown systematic error. This systematic error can have an extreme impact on the ability to detect true interaction effects. Indeed, within-subject examinations of Vroom's (1964) expectancy theory using a coarse Likert-type response scale (Stahl & Harrell, 1981) versus a continuous response scale (Arnold, 1981) resulted in conflicting findings. (Specific procedures used by Arnold, 1981, and Stahl & Harrell, 1981, are described in the Methods section.)

The purpose of the current study was to examine the impact of a Likert-type dependent-response scale on moderated regression results when subjects were known to be providing a "true" interaction effect. This would provide an empirical extension of Russell et al.'s (1991) limited simulation and more conclusive evidence about the impact of Likert-type scales in moderated regression analyses. Russell et al. demonstrated such an impact by using a computer simulation that made assumptions about how subjects reduced or transformed their responses. The current study extended Russell et al.'s findings with actual subjects responding to a common interactive model.

In the current study, approximately one half of the subjects responded to a dependent scale measure using a 5-point Likert scale. The other half responded by placing a mark on a graphic line segment. The distance in millimeters from the left side of the line segment was obtained, resulting in a nearly continuous dependent-response measure. Hence, we were able to directly test whether an assessment design decision (to use a Likert-type scale vs. a continuous dependent scale) causes information loss and spuriously affects moderated regression results.

Method

Subjects

The subject pool was chosen to ensure that participants could reasonably be expected to provide an interaction effect when instructed to do so by the investigators. Because interaction effects tested in applied research settings are typically embedded in some theory or model, the subject pool also had to be familiar with the content of the model that predicted the interaction effect. Hence, 96 advanced doctoral students; assistant professors; or recently promoted associate professors in business schools, psychology departments, and industrial relations centers were asked to respond to a decision simulation designed to capture the interaction effect described in Vroom's (1964) expectancy theory of motivation. We selected subjects who we knew were assistant professors or who had published an article in the last 2 years in an Academy of Management publication that listed their rank as assistant profes-

sor. Subjects were purposely selected at this rank in order to maximize their ability to understand the decision scenario context (described later) and provide an interaction effect when requested to.

Procedures

Expectancy theory was chosen as the focus of this study on the basis of results reported by Stahl and Harrell (1981) and Arnold (1981). Stahl and Harrell used 11 levels of valence (v) and three levels of expectancy (e) in a within-subject design to test Vroom's (1964) multiplicative formulation of motivation ($f = v \times e$). An 11-point Likert scale was used to capture subjects' dependent responses. Hence, if Stahl and Harrell's subjects were following Vroom's multiplicative formulation, they were faced with portraying a 33-point latent-response space (3 levels of expectancy \times 11 levels of valence) on an 11-point Likert scale. Although Stahl and Harrell found some evidence of a multiplicative effect, the majority of the within-subject moderated regression analyses did not yield evidence of a significant interaction effect.

Arnold (1981) used a within-subject design with five levels of expectancy and valence in a test of the same model. However, in contrast to Stahl and Harrell (1981), Arnold used a nearly continuous graphic rating scale to capture subjects' dependent responses. Subjects were asked to place a mark on a 150-mm line segment to represent their dependent response. Arnold then measured the distance in millimeters from the left end of the line segment and recorded this as the dependent value. Hence, subjects were faced with portraying a 25-point latent-response space (five levels of expectancy \times five levels of valence) on a 150-point line. Arnold's results strongly supported the multiplicative formulation of the expectancy model.

We used similar procedures to test the effect of scale coarseness on moderated regression analysis. A decision simulation was constructed using an orthogonal fixed-effects design with five levels each of expectancy and valence. Also, the first 5 decision scenarios were repeated at the end of each questionnaire to permit an assessment of subjects' consistency reliability in their judgments. Hence, each subject was asked to respond to 25 different decision scenarios and 5 decision scenarios that duplicated the first 5 to which they had responded.

The simulation asked subjects to imagine how motivated they would be to revise a manuscript returned by an editor of a major scholarly journal. Each page of the simulation was a distinct decision situation. Expectancy was manipulated by differences in the editor's stated likelihood that a revision would be accepted for publication (10%, 30%, 50%, 70%, or 90% probability). Valence was manipulated by a senior professor's statement concerning how much impact an additional publication in that journal would have on a promotion review committee ("exceptionally strong impact," "strong impact," "moderate impact," "minor impact," or "almost no impact"). The instructions asked subjects to place themselves in the position of being 1 year away from mandatory promotion and tenure review. Subjects were asked to indicate their motivation to complete and submit a revision of their manuscript on the scale provided. Examples of scenarios using the Likert scale response format and the 150-millimeter line segment are presented in the Appendix.

The instructions clearly indicated that the investigators' goal was to gather baseline data needed to explore different ways of detecting "true" interaction effects. Subjects were explicitly instructed that the editor's percentage estimate constituted our expectancy manipulation and the senior professor's comment constituted the valence manipulation. They were then asked to respond to each scenario *in a way that supported Vroom's (1964) multiplicative formulation of expectancy theory* (which was reviewed in a brief paragraph). Finally, the instructions also indicated that the purpose of this study was not to learn how junior faculty are motivated in response to feedback from journal editors and senior colleagues.

Questionnaires were reviewed by junior faculty colleagues to ensure clarity of instructions and materials. Forty-eight copies of the questionnaire containing the discrete Likert-type response format (from *very unmotivated*[1] to *very motivated*[5]) and 48 copies containing the continuous graphical response format were mailed to subjects along with a cover letter and postage-paid return envelope. Fifty-nine responses were received, for a response rate of 61%. Three subjects included notes to us indicating that they thought our intent was to investigate how assistant professors actually made decisions to revise manuscripts. These respondents' questionnaires were dropped from the analysis, resulting in a final response rate of 58%.

Results

Ten of the subjects did not respond to the last five scenarios of the questionnaire, indicating in notes to us that these scenarios were repeats. Hence, correlations between responses to the five duplicate scenarios were derived on the remaining 46 respondents. Estimates of consistency reliability (Slovic & Lichtenstein, 1971) ranged from .667 to 1.00, for an average of .915 (all but one of the reliabilities were between .840 and 1.00). No significant difference was found between consistency reliabilities of subjects responding to Likert versus line segment scales.

Effect size in moderated regression analysis is represented by the difference between coefficients of determination ($R^2_{\text{mult}} - R^2_{\text{add}}$) obtained from the following two equations (Evans, 1991):

$$\text{Additive: } y = \beta_0 + \beta_1 x + \beta_2 z + \epsilon. \quad (1)$$

$$\text{Multiplicative: } y = \beta_0^* + \beta_1^* x + \beta_2^* z + \beta_3^* xz + \epsilon^*. \quad (2)$$

For purposes of analysis, levels of expectancy were coded as .10, .30, .50, .70, and .90, corresponding with the stated probabilities in the manipulations. Levels of valence were coded from 1 to 5, with 5 representing the most valent condition. To generate a baseline effect size, we entered all possible combinations of expectancy and valence into a statistical software package and multiplied to create a third (y) variable. That is, each level of valence (v) was crossed with each level of expectancy (e), resulting in 25 data points. The dependent variable (y) was created by multiplying e by v . Moderated regression analysis applied to this data set indicated that $R^2_{\text{mult}} = 1.00$ and $R^2_{\text{add}} = .884$. Hence, if subjects responded to the questionnaire according to Vroom's (1964) model without measurement error and generated overt responses that were linear functions of their latent responses, the expected effect size of moderated regression analysis would be $1.00 - .884 = .116$. Again, this is the maximum effect size one would expect if subjects were perfectly reliable in their use of the expectancy and valence "cues" and responded to the dependent scale without error. This figure was used as the point against which the current subjects' effect sizes were compared.

Results for the 29 subjects who returned questionnaires with Likert-type response scales and the 27 subjects who returned questionnaires with the line segment response scales are presented in Tables 1 and 2. Two results are of particular interest. First, the average effect size for subjects responding to the line segment scale was 0.058, an 93% increase in effect size relative to the average effect size for subjects responding to the Likert scales (0.030). A t test of this difference indicated that it is not likely to have occurred by chance, $t(54) = 1.852$, $p < .05$, one-

tailed. This suggests that, on average, F statistics derived to test the significance of moderator effects could be substantially higher when subjects respond to a fine scale as opposed to the relatively coarse Likert scale. Note further that this effect size was approximately half the "true" effect size of .116 that would be expected under conditions of no measurement error and a linear relationship between the latent and overt dependent responses.

The second result of interest is that there was substantial variation across individuals in moderated regression effect size. Busemeyer and Jones (1983) demonstrated that when an unknown, nonlinear transformation is made on the dependent variable, moderated regression effect sizes can spuriously increase or decrease. Our results empirically confirm that there are substantial differences in how subjects "reduce" their latent responses, causing some subjects' effect sizes to be spuriously increased or decreased.

Figures 1 and 2 contain graphs of the marginal means of subjects' motivation to revise the hypothetical manuscript for the various combinations of expectancy and valence. The clear fan-shaped pattern for subjects who responded on a line segment scale (Figure 2) is indicative of an interaction effect, whereas there is less evidence of such a pattern for subjects who responded on a Likert scale (Figure 1). Finally, the slight conver-

Table 1
Moderated Regression Results for Subjects Who Responded to a Discrete (Likert) Response Scale

| Subject | R^2_{mult} | R^2_{add} | $R^2_{\text{mult-add}}$ |
|---------|---------------------|--------------------|-------------------------|
| 1 | .834 | .723 | .111** |
| 2 | .851 | .847 | .004 |
| 3 | .737 | .735 | .002 |
| 4 | .892 | .878 | .014 |
| 5 | .830 | .824 | .006 |
| 6 | .849 | .846 | .003 |
| 7 | .818 | .811 | .007 |
| 8 | .974 | .807 | .167** |
| 9 | .864 | .864 | .000 |
| 10 | .881 | .822 | .059** |
| 11 | .824 | .770 | .054** |
| 12 | .974 | .807 | .167** |
| 13 | .919 | .864 | .055** |
| 14 | .738 | .733 | .005 |
| 15 | .832 | .811 | .021 |
| 16 | .892 | .839 | .053** |
| 17 | .895 | .893 | .002 |
| 18 | .859 | .814 | .045** |
| 19 | .855 | .806 | .049** |
| 20 | .903 | .903 | .000 |
| 21 | .909 | .898 | .011 |
| 22 | .789 | .789 | .000 |
| 23 | .838 | .834 | .004 |
| 24 | .865 | .865 | .000 |
| 25 | .857 | .851 | .006 |
| 26 | .986 | .986 | .000 |
| 27 | .808 | .788 | .028 |
| 28 | .876 | .875 | .001 |
| 29 | .922 | .907 | .015* |
| Average | .864 | .834 | .030 |
| SD | .060 | .057 | .046 |

* $p \leq .05$. ** $p \leq .01$.

Table 2
Moderated Regression Results for Subjects Who Responded to a Continuous (Line Segment) Response Scale

| Subject | R^2_{mult} | R^2_{add} | $R^2_{\text{mult-add}}$ |
|---------|---------------------|--------------------|-------------------------|
| 1 | .917 | .845 | .072** |
| 2 | .894 | .892 | .002 |
| 3 | .958 | .957 | .001 |
| 4 | .895 | .894 | .001 |
| 5 | .965 | .964 | .001 |
| 6 | .977 | .792 | .185** |
| 7 | .797 | .738 | .059* |
| 8 | .945 | .939 | .006 |
| 9 | .966 | .966 | .000 |
| 10 | .945 | .945 | .000 |
| 11 | .995 | .883 | .112** |
| 12 | .988 | .909 | .079** |
| 13 | .961 | .956 | .005 |
| 14 | .974 | .762 | .212** |
| 15 | .991 | .877 | .114** |
| 16 | .920 | .901 | .019* |
| 17 | .784 | .675 | .109** |
| 18 | .877 | .815 | .062** |
| 19 | .966 | .891 | .075** |
| 20 | .930 | .930 | .000 |
| 21 | .993 | .913 | .080** |
| 22 | .965 | .759 | .206** |
| 23 | .989 | .977 | .012** |
| 24 | .792 | .748 | .044* |
| 25 | .766 | .765 | .001 |
| 26 | .921 | .821 | .100** |
| 27 | .971 | .966 | .005* |
| Average | .927 | .870 | .058 |
| SD | .069 | .086 | .066 |

* $p \leq .05$. ** $p \leq .01$.

gence of means for high and low levels of expectancy in Figure 1 suggests that the Likert scale may also have been subject to ceiling and floor effects in the current application.

Discussion

In the current sample, use of coarse Likert response scales to capture relatively fine latent responses caused a substantial reduction in average moderated regression effect size. These results may explain many of the mixed findings in the search for moderator effects over the last 30 years. The results certainly suggest that the coarse Likert scale used by Stahl and Harrell (1981) and the fine line segment scale used by Arnold (1981) directly contributed to the differences in their support of Vroom's (1964) expectancy theory.

One implication for research designs is immediate and direct: Investigators should not attempt to discover moderator effects unless the overt measurement scale contains at least as many response options as exist in the theoretical response domain. Note that Likert (1932) and, more recently, Cicchetti, Showalter, and Tyrer (1985) demonstrated that an increase in the number of response categories to a scale does not have an attenuating effect on reliability (reliability plateaus after five to seven response options). Moreover, the consistency reliability results reported herein indicate no difference in random mea-

surement error between the two response formats. Hence, a continuous dependent-response scale will not necessarily change reliability and, our results indicate, will substantially increase the likelihood of detecting a true interaction effect. Consequently, to be safe, investigators should consider other methods of providing subjects with continuous (or nearly continuous) response scales. In this regard, the line segment method described by Arnold (1981) is an excellent beginning, although it is very cumbersome and labor intensive to employ. Methods of optical scanning and computer-assisted measurement that produce nearly continuous scales should be explored.

Note that summing responses to multiple Likert-type items on a dependent scale (as is often done in between-subjects survey designs) is not the same as providing subjects with a continuous response scale. Although the resultant "scale score" obtained by summing item responses could be considered nearly continuous, subjects are not responding with a scale score. Rather, they are responding to each item individually. Thus, if an individual responds to coarse Likert scales in a similar manner across items, the problem of reduced power to detect interaction remains. A scale formed by summing responses to Likert items may yield a significant interaction effect if the response function used by subjects is not constant across all values of the latent dependent variable. However, information loss that causes systematic error to occur at the item level would have the same effect on moderated regression effect size regardless of whether the dependent-response items were analyzed separately (as was done here in a within-subject design) or cumulated into a scale score.

One might also ask whether there are conditions in which a coarse Likert-type item might be just as capable of detecting a true interaction effect as a fine continuous response format. Depending on the nature of the interaction effect, the answer is yes. Our example asked subjects to provide responses that reflected the interaction effect postulated by expectancy theory. In this theory the interaction itself is "continuous" in that it hypothesizes that every incremental change in valence will have an influence on the relationship between expectancy and motivation. In contrast, a theory or model might hypothesize a more "discrete" interaction effect in which the relationship between x and y is constant across ranges of the moderating variable z . Such a model applied to expectancy theory would suggest a constant relationship between expectancy and motivation for an initial range of valence values. After some critical "threshold" level of valence is surpassed, the new range of valence values would dictate a different relationship between expectancy and motivation. This new expectancy-motivation relationship would hold constant until the next critical threshold level of valence is surpassed. In such a case, attenuation of moderator effects due to coarseness of the response scale would decrease. However, most theories are not definitive in their description of the nature of hypothesized interaction effects (Cronbach, 1987). Hence, a fine dependent-response scale is likely to provide the investigator with more information and increase the likelihood that any single study will shed light on the true nature of underlying interaction effects.

A second implication of the current study targets more basic measurement research. Specifically, what functions describe how subjects "reduce" their latent responses when faced with a

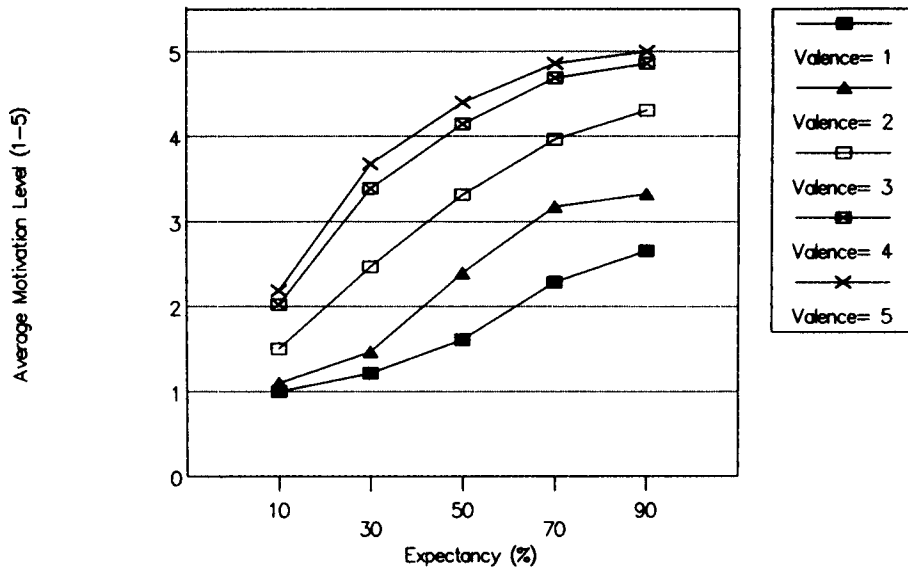


Figure 1. Likert scale marginal means.

relatively coarse Likert scale? Furthermore, some of the subjects generated moderated regression effects sizes that were greater than the expected “true” effect size of 0.116 in both the Likert scale and line segment conditions. Subject 14’s responses to the line segment scales (see Table 2) resulted in a moderated regression effect size of 0.212, almost twice as large as expected. Effect sizes greater than .116 could have been due to random measurement error (or perhaps subjects misinterpreted the task). Alternatively, the results may indicate that an assumption of linear relationships between latent and overt responses is not always reasonable, even when the subjects are given a fine response scale that permits a linear response function. That is,

there are some nonlinear functions that can spuriously increase the obtained moderated effect size (Busemeyer & Jones, 1983).

Finally, Birnbaum has demonstrated in a wide variety of applications that subjects’ judgment functions tend to be a linear function of their latent psychological representation of a construct when category or graphic rating scales are used (see Anderson, 1982; Birnbaum, 1978, 1980, 1982, & 1990, for reviews). Birnbaum used the term *difference* to describe the comparison process subjects use when using rating scales like those developed by Likert: Subjects use the psychological “difference” between their latent perceptions of a construct and their perception of the anchors used on the graphic rating scale to arrive at a

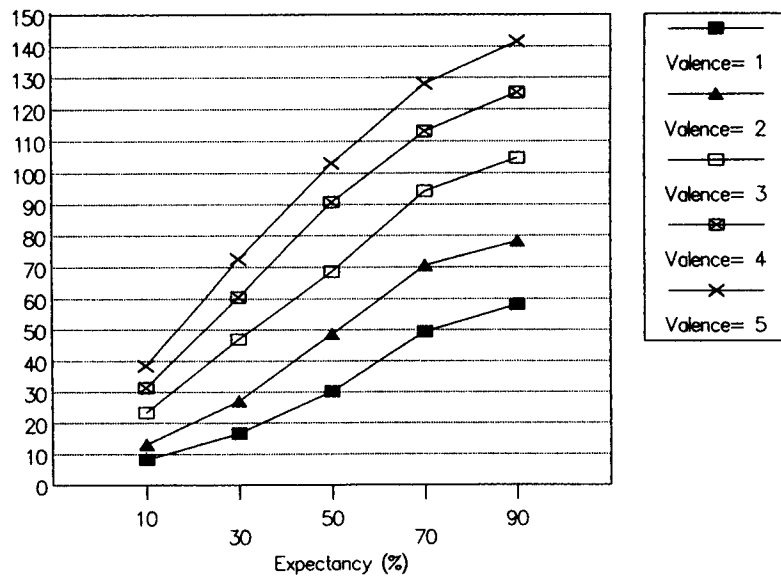


Figure 2. Line segment scale marginal means.

psychological calibration of the phenomenon of interest (the judgment function describes the relationship between the latent psychological representation and the overt response made on the graphic rating scale). Note that this literature suggests it may be possible to develop a true ratio scale of our dependent variable F_i (felt force to engage in behavior i), an approach that did not exist when Schmidt (1973) described the problems created by the absence of ratio scales in detecting interaction effects.

However, Poulton (1979) noted evidence suggesting that the judgment function relating latent psychological representations to overt responses can be influenced by contextual effects of stimulus spacing and choice of anchors (see Mellars & Birnbaum, 1982, for an example of stimulus spacing effects; see Parducci, 1983, for an example of effects due to range and frequency). The smaller than expected effect sizes for subjects responding to a Likert scale suggest an additional "contextual effect"—insufficient density in the dependent scale. However, a number of subjects (using both Likert and line segment scales) generated effect sizes that were greater than expected, suggesting that other unknown contextual effects or individual difference variables may have been distorting response functions. This is an important issue that has been largely ignored in applied settings. Busemeyer and Jones (1981) demonstrated the effect of nonlinear response functions on our ability to detect true interaction effects. The current results demonstrate how an investigator's decision to use Likert scales may force subjects to use a nonlinear response function. Assumptions regarding the relationship between latent and observed responses to questionnaires are crucial to the integrity of conclusions drawn and need to be explored more thoroughly in applications of psychological measures in organizational settings.

Regardless, our results confirm Russell et al.'s (1991) speculation that response reduction effects due to use of Likert scales can substantially influence the results of moderated regression analyses. Implications for research design are relatively straightforward and should have a considerable impact on our ability to detect interaction effects. Future research needs to address the more fundamental questions of how overt and latent responses are related.

References

- Anderson, N. H. (1982). *Methods of information integration theory*. San Diego, CA: Academic Press.
- Arnold, H. J. (1981). A test of the multiplicative hypothesis of expectancy valence theories of work motivation. *Academy of Management Journal*, 24, 128–141.
- Birnbaum, M. H. (1978). Differences and ratios in psychological measurement. In N. J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 33–74). Hillsdale, NJ: Erlbaum.
- Birnbaum, M. H. (1980). Comparison of two theories of "ratio" and "difference" judgments. *Journal of Experimental Psychology: General*, 109, 304–319.
- Birnbaum, M. H. (1982). Controversies in psychological measurement. In B. Wegener (Ed.), *Social attitudes and psychological measurement* (pp. 401–485). Hillsdale, NJ: Erlbaum.
- Birnbaum, M. H. (1990). Scale convergence and psychophysical laws. In H. G. Geissler, M. H. Mueller, & W. Prinz (Eds.), *Psychophysical explorations of mental structures* (pp. 49–57). Toronto: Hogrefe & Huber.
- Bobko, P. (1986). A solution to some dilemmas when testing hypotheses about ordinal interactions. *Journal of Applied Psychology*, 71, 323–326.
- Busemeyer, J. R., & Jones, L. E. (1983). Analysis of multiplicative combination rules when the causal variables are measured with error. *Psychological Bulletin*, 93, 549–562.
- Cicchetti, D. V., Showalter, D., & Tyrer, P. J. (1985). The effect of number of rating scale categories on levels of interrater reliability: A Monte Carlo investigation. *Applied Psychological Measurement*, 9, 31–36.
- Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, 7, 249–253.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cronbach, L. J. (1987). Statistical tests for moderator variables: Flaws in analyses recently proposed. *Psychological Bulletin*, 102, 414–417.
- Drazin, R., & Van de Ven, A. H. (1985). An examination of alternative forms of fit in contingency theory. *Administrative Science Quarterly*, 30, 514–539.
- Evans, M. G. (1991). The problem of analyzing multiplicative composites: Interactions revisited. *American Psychologist*, 46, 6–15.
- Likert, R. A. (1932). A technique for the measurement of attitudes. *Archives of Psychology, New York*, No. 140.
- Maier, N. R. F. (1955). *Psychology in industry* (2nd ed.). Boston: Houghton Mifflin.
- Mellars, B. A., & Birnbaum, M. H. (1982). Loci of contextual effects in judgment. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 582–600.
- Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 49, 115–132.
- Olsson, U., Drasgow, F., & Dorans, N. (1982). The polyserial correlation coefficient. *Psychometrika*, 47, 337–347.
- Parducci, A. (1983). Category ratings and the relational character of judgment. In H. G. Geissler & V. Sarris (Eds.), *Modern issues in perception* (pp. 5–52). Berlin: VEB Deutscher Verlag der Wissenschaften.
- Peters, C. C., & Van Voorhis, W. R. (1940). *Statistical procedures and their mathematical bases*. New York: McGraw-Hill.
- Poulton, E. C. (1979). Models for biases in judging sensory magnitude. *Psychological Bulletin*, 86, 777–803.
- Russell, C. J., Pinto, J. K., & Bobko, P. (1991). Appropriate moderated regression and inappropriate research strategy: A demonstration of information loss due to scale coarseness. *Applied Psychological Measurement*, 15, 125–135.
- Saunders, D. R. (1955). The "moderator variable" as a useful tool in prediction. In *Proceedings of the 1954 Invitational Conference on Testing Problems* (pp. 54–58). Princeton, NJ: Educational Testing Service.
- Saunders, D. R. (1956). Moderator variables in prediction. *Educational and Psychological Measurement*, 16, 209–222.
- Schmidt, F. L. (1973). Implications of a measurement problem for expectancy theory research. *Organizational Behavior and Human Performance*, 10, 243–251.
- Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance*, 6, 649–744.
- Sockloff, A. (1976a). The analysis of nonlinearity via linear regression with polynomial and product variables: An examination. *Review of Educational Research*, 46, 267–291.
- Sockloff, A. (1976b). Spurious product correlation. *Educational and Psychological Measurement*, 36, 33–44.
- Stahl, M. J., & Harrell, A. M. (1981). Modeling effort decisions with

- behavioral decision theory: Toward an individual differences model of expectancy theory. *Organizational Behavior and Human Performance*, 27, 303-325.
- Terborg, J. R. (1977). Validation and extension of an individual differences model of work performance. *Organizational Behavior and Human Performance*, 18, 188-216.
- Venkatraman, N. (1989). The concept of fit in strategy research: Toward verbal and statistical correspondence. *Academy of Management Review*, 14, 423-444.
- Vroom, V. H. (1964). *Work and motivation*. New York: Wiley.
- Zedeck, S. (1971). Problems with the use of "moderator" variables. *Psychological Bulletin*, 76, 295-310.

Appendix

Examples of Discrete (Likert) Versus Continuous (Line Segment) Decision Scenarios

Discrete Response Scenario

The editor of *The Major Scholarly Journal* has recently returned a manuscript you had previously submitted with reviewers' comments. He indicated in his cover letter that he felt substantive changes were needed in the analyses and presentation of the results. While he invited you to submit a revision, the editor estimated that even with the changes there was only a 10% chance that the manuscript would be acceptable for publication. The same day you received the reviews on the manuscript, you had an informal talk with a senior professor in your department. He indicated that, given your scholarly record to date and the recent behavior of the promotion review committee, an additional publication in *The Major Scholarly Journal* would have an *exceptionally strong impact* on how you are viewed by the promotion committee. Given that you have a number of other research projects (which may result in publication in a major scholarly journal) and commitments in the next 9 months, please circle the response on the scale below how motivated you are to revise this particular manuscript.

| | | | | |
|---------------------|---------------------------|-------------|-------------------------|-------------------|
| 1 | 2 | 3 | 4 | 5 |
| Very Unmotivated | Moderately Unmotivated | Indifferent | Moderately Motivated | Very Motivated |

Continuous Response Scenario

The editor of *The Major Scholarly Journal* has recently returned a manuscript you had previously submitted with reviewers' comments. He indicated in his cover letter that he felt substantive changes were needed in the analyses and presentation of the results. While he invited you to submit a revision, the editor estimated that even with the changes there was only a 10% chance that the manuscript would be acceptable for publication. The same day you received the reviews on the manuscript, you had an informal talk with a senior professor in your department. He indicated that, given your scholarly record to date and the recent behavior of the promotion review committee, an additional publication in *The Major Scholarly Journal* would have an *exceptionally strong impact* on how you are viewed by the promotion committee. Given that you have a number of other research projects (which may result in publication in a major scholarly journal) and commitments in the next 9 months, please place an "X" on the line below to indicate how motivated you are to revise this particular manuscript.

| | |
|---------------------|-------------------|
| 1 | 100 |
| Very Unmotivated | Very Motivated |

Received June 13, 1991
 Revision received November 21, 1991
 Accepted November 26, 1991 ■

Copyright of Journal of Applied Psychology is the property of American Psychological Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.