STEPHEN A. SPILLER, GAVAN J. FITZSIMONS, JOHN G. LYNCH JR., and GARY H. McCLELLAND*

It is common for researchers discovering a significant interaction of a measured variable X with a manipulated variable Z to examine simple effects of Z at different levels of X. These "spotlight" tests are often misunderstood even in the simplest cases, and it appears that consumer researchers are unsure how to extend them to more complex designs. The authors explain the general principles of spotlight tests, show that they rely on familiar regression techniques, and provide a tutorial demonstrating how to apply these tests across an array of experimental designs. Rather than following the common practice of reporting spotlight tests at one standard deviation above and below the mean of X, it is recommended that when X has focal values, researchers should report spotlight tests at those focal values. When X does not have focal values, it is recommended that researchers report ranges of significance using a version of Johnson and Neyman's test the authors term a "floodlight."

Keywords: moderated regression, spotlight analysis, simple effects tests

Spotlights, Floodlights, and the Magic Number Zero: Simple Effects Tests in Moderated Regression

Most marketing and consumer behavior articles reporting experiments test for interactions between two or more variables. Authors may follow up an interaction of two variables with "simple effects" tests (called "conditional effects" tests by econometricians) of the effect of one variable at different levels of another. They may follow up an interaction of three variables with tests of "simple interactions" of two variables at a level of a third variable or "simple–simple" effects of one variable at chosen levels of the other two (Keppel and Wickens 2004).

This article presents a tutorial on the analysis of simple effects tests in designs in which one or more of the interacting variables are continuous and quantitative rather than categorical. Researchers primarily trained in using analysis of variance (ANOVA) frameworks for experimental designs often struggle when following up interactions in which a continuous variable interacts with one or more categorical variables and the appropriate analysis takes place in the framework of a moderated regression. In a review of Volume 48 of *Journal of Marketing Research* and Volume 38 of *Journal of Consumer Research*, we found that the reported moderated regression analyses were often "correct" but not optimally performed and, in many other cases, were simply incorrect. We observe similar small and large errors in other social sciences. In this article, we identify the most common misunderstandings and provide a simple framework for conducting these analyses.

Consider a fictional extension of McFerran et al.'s (2010) study of the effect of social influence on consumption. The

*Stephen A. Spiller is Assistant Professor of Marketing, Anderson School of Management, University of California, Los Angeles (e-mail: stephen.spiller@anderson.ucla.edu). Gavan J. Fitzsimons is R. David Thomas Professor of Marketing and Psychology, Fuqua School of Business, Duke University (e-mail: gavan@duke.edu). John G. Lynch Jr. is Ted Anderson Professor, Leeds School of Business, and Director, Center for Research on Consumers' Financial Decision Making, University of Colorado Boulder (e-mail: john.g.lynch@colorado.edu). Gary H. McClelland is Professor of Psychology and Faculty Fellow, Institute of Cognitive Science, University of Colorado Boulder (e-mail: gary.mcclelland@colorado.edu). Fitzsimons, Lynch, and McClelland are listed alphabetically and contributed equally. The authors thank Rick Staelin, Carl Mela, and Wagner Kamakura for asking questions that motivated this article and Ajay Abraham, Philip Fernbach, Yoosun Hann, Ji Hoon Jhang, Christina Kan, Peggy Liu, Matthew Philp, Adriana Samper, Julie Schiro, Scott Wallace, Elizabeth Webb, Hillary Wiener, and the review team for constructive comments, as well as seminar participants at the University of Colorado. Any errors are the authors'. Don Lehmann served as associate editor for this article. This article was invited by Robert Meyer.

ISSN: 0022-2437 (print), 1547-7193 (electronic)

authors propose that people's own consumption behavior is anchored on the quantity taken by others in their environment, but they adjust their consumption on the basis of whether others around them belong to an aspirational or dissociative group. That is, the authors propose an interaction between quantity taken by others and type of others on the amount of consumption. They find an interaction such that consumers modeled the behavior of a thin confederate more than they modeled the behavior of an obese confederate. Consumers took more candy when the confederate took 30 pieces than when she took 2 pieces, but this difference (which might reflect imitation of the model's behavior) was stronger when the model was thin than when she was obese.

Suppose that rather than manipulating the weight of the confederate over two levels, McFerran et al. (2010) had a yoked design in which pairs of undergraduate students participated in the study, and one was cast in the role of confederate and instructed to take 2 or 30 candies, testing the effect on the behavior of the other participant in the pair. Over 100 pairs, suppose they measured the body mass index (BMI) of the 100 confederate models.

How could the authors analyze the interaction and simple effects? They could choose to perform a median split and divide the undergraduate students into groups with large and small confederate models (or small, medium, and large to allow for nonlinearity). This is not a viable solution, because the problems with median splits are well documented: there is a substantial loss of statistical power from dichotomizing a single predictor variable (e.g., Irwin and McClelland 2001, 2003; Jaccard et al. 2006; MacCallum et al. 2002), and dichotomizing in multiple predictor models creates spurious effects (Maxwell and Delaney 1993; Vargha et al. 1996). Instead, the authors should use moderated multiple regression and test the model

$$Y = a + bZ + cX + dZX, \quad (1)$$

where Y is number of candies the participant takes, X is the BMI of the model, and Z is an indicator variable for number of candies the confederate model takes. That indicator variable could be dummy coded (0 = 2 candies, 1 = 30 candies), or it could be contrast coded (–1 = 2 candies, +1 = 30 candies).

A significant coefficient d in Equation 1 implies that BMI moderates the effect of number of candies taken or, equivalently, that the number of candies taken moderates the effect of BMI. Following detection of a significant interaction, the authors may want to estimate and test the simple effect of the manipulated variable Z at different levels of X, the BMI of the model. Tests of simple effects of a manipulated or categorical variable at a level of a continuous variable are often called "spotlight" tests: they shine the spotlight on the effect of the manipulated Z at a particular value of X. Spotlight analysis is a technique using basic statistics from regression analysis to analyze the simple effect of one variable at a particular level of another variable, continuous or categorical. The purpose of this article is to help authors conduct spotlight analyses in various types of experimental and correlational designs and convey their findings more effectively. We show the following:

1. Regression terms that authors sometimes interpret as "main effects" are actually simple effects of an interacting variable in a product term (ZX) when other variables in that product (interaction) term are coded as 0.
2. Researchers can shine the spotlight for the simple effect of Z on a particular value of X by adding or subtracting a constant from the original X variable to make the focal value the zero point on the recoded scale.
3. Authors in marketing and allied social sciences have been following a convention of testing simple effects of Z at plus and minus one standard deviation from the mean of X. This one standard deviation from the mean spotlight level is arbitrary and hinders generalization across studies.
4. If there are values of X that are particularly meaningful or relevant for theoretical or substantive reasons, simple effects spotlight tests should be reported at those values rather than at plus and minus one standard deviation from the mean value of X.
5. If there are no values of X that are particularly meaningful—in other words, if all values of X are relevant and interesting values for considering simple effects of the manipulated Z—authors should abandon spotlight tests and report what we call a "floodlight" test of simple effects of Z at all possible values of X. This floodlight test from Johnson and Neyman (1936) identifies regions along the X continuum where the simple effect of Z is significant and regions where it is not. It is simple to compute those regions.
6. These same principles can be applied to more complex designs about which marketing and consumer researchers have been treading with trepidation. One can readily apply these principles to multiple levels of Z, to within-participant manipulations of Z, and to higher-order factorial designs including one or more measured variables. The principles involve nothing more than basic regression techniques. We discuss certain statistical subtleties in the Appendix and explain the applications to more complex designs in Web Appendix A (www.marketingpower.com/jmr_webappendix). Table 1 covers the contents of Web Appendix A.

SIMPLE EFFECTS TESTS AND THE MAGIC NUMBER ZERO

Spotlight analysis provides an estimate and statistical test of the simple effect of one variable at specified values of another continuous variable. Aiken and West (1991), Irwin and McClelland (2001), and Jaccard, Turrisi, and Wan (1990) discuss how to conduct spotlight analyses. We reiterate the key points here to aid understanding of the general principles underlying spotlight analyses (we explain specific examples subsequently) and how this relates to our proposed floodlight analysis.

Table 1
INDEX OF WHERE TO FIND BUILDING-BLOCK DESIGNS FOR SPOTLIGHT AND FLOODLIGHT ANALYSES

Case Number	*Design*	*Covered*
0 (base case)	2 × continuous	Main text p. 279 and Table 2, Web Appendix A Table W1
1	2 (within) × continuous	Main text p. 285, Web Appendix A p. 1 and Table W2
2	2 × 2 × continuous	Main text p. 286, Web Appendix A p. 2 and Table W3
3	3 × continuous	Web Appendix A p. 4 and Table W5
4	Continuous × continuous	Web Appendix A p. 6
5	Quadratic	Web Appendix A p. 7

Take the basic moderated multiple regression model in Panel A of Table 2 for the hypothetical version of McFerran et al. (2010) we described previously. We analyze the dependent variable (Y) as a function of a two-level manipulated variable (Z), a continuous measured variable (X), and their interaction. We code Z as 0 for the group in which the model takes 2 candies and as 1 for the group in which the model takes 30 candies. The model is given by Equation 1.

In Figure 1, Panel A, we plot hypothetical data for such a model, with the continuous variable (X) plotted on the x-axis and two regression lines relating X to the dependent variable Y: one regression line for the Z = 0 group in which the model takes 2 candies and one for the Z = 1 group in which the model takes 30 candies. We discuss the specific estimates in the next section.

Some authors use the continuous value of X when testing the interaction in Equation 1 (i.e., for "analysis" of the interaction). However, when performing simple effects tests to "explicate" the interaction, they revert to using median splits, testing the simple effect of Z at different levels of the now-dichotomized X. This is incorrect. The correct test of simple effects of Z at different levels of X uses the continuous X and spotlight tests.

The simple effect of Z at a given value of X is equivalent to the distance between the regression line for the treatment group and the regression line for the control group. We find the regression line for the group in which the model takes 30 candies, where Z = 1, by replacing Z with 1:

$$Y = a + b + cX + dX = (a + b) + (c + d)X. \quad (1a)$$

The intercept, where X = 0, is given by (a + b), and the slope is given by (c + d). We found the regression line for the group in which the model takes 2 candies, where Z = 0, by replacing Z with 0:

$$Y = a + cX. \quad (1b)$$

The intercept, where X = 0, is given by a, and the slope is given by c. Therefore, the simple effect of the manipulation, Z, given by the difference between the lines,[1] is

$$\Delta Y = b + dX. \quad (1c)$$

Equation 1 and Equation 1c make clear that b is the simple effect of Z when X = 0, even though X = 0 may well be outside the range of the data or an impossible value. Equation 1 simplifies to Y = a + bZ where X = 0. Equation 1c, which estimates the simple effect as the difference between two regression lines, simplifies to ΔY = b where X = 0.

Zero is a "magic number" in moderated regression. It is "magic" because Equation 1 simplifies when either variable has a value of zero. A constant can be added to or subtracted

[1]More generally, the simple effect of Z on Y is given by the derivative of Y with respect to Z.

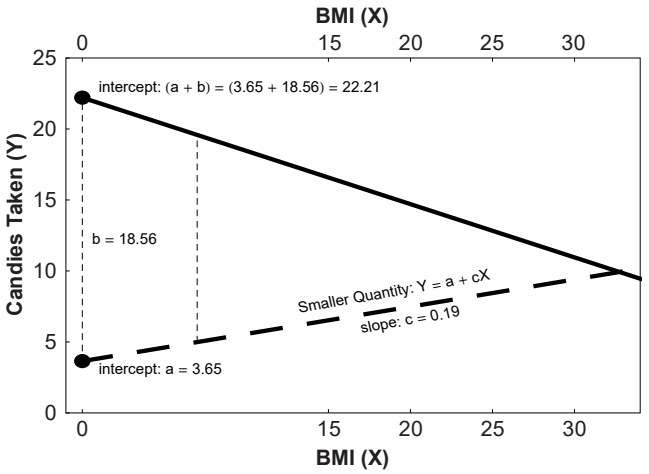
Table 2
SIMPLE EFFECTS IN A 2 × CONTINUOUS DESIGN

	Intercept	*Manipulation Z*	*Measured Variable X*	*Manipulation × Measured ZX*
A. Baseline Analysis				
Coding		0 = control 1 = treatment	Raw scale	
Coefficient	a	b	c	d
Interpretation	Estimate of Y when Z = 0 and X = 0 (i.e., for control group when X = 0)	Simple effect of treatment vs. control when X = 0	Simple slope of measured variable on Y when Z = 0 (i.e., for control group)	Change in effect of treatment vs. control when measured variable increases by one unit
B. Test the Simple Effect of Treatment Versus Control at Focal Value $X = X_{Focal}$ by Recoding X so That It Drops Out of the Equation				
		Z	*X′*	*ZX′*
Coding		0 = control 1 = treatment	$X' = X - X_{Focal}$	
Coefficient	a′	b′	c′	d′
Equivalent to	$a + cX_{Focal}$	$b + dX_{Focal}$	c	d
Interpretation	Estimate of Y when Z = 0 and X′ = 0 (i.e., for control group when $X = X_{Focal}$)	Simple effect of treatment vs. control when X′ = 0 (i.e., when $X = X_{Focal}$)	Simple slope of measured variable on Y when Z = 0 (i.e., for control group)	Change in effect of treatment vs. control when measured variable increases by one unit
C. Test the Simple Slope of X in Treatment Group by Recoding Z so That It Drops Out of the Equation				
		Z″	*X*	*Z″X*
Coding		1 = control 0 = treatment	Raw scale	
Coefficient	a″	b″	c″	d″
Equivalent to	a + b	–b	c + d	–d
Interpretation	Estimate of Y when Z″ = 0 and X = 0 (i.e., for treatment group when X = 0)	Simple effect of control vs. treatment when X = 0	Simple slope of measured variable on Y when Z″ = 0 (i.e., for treatment group)	Difference in slope of measured variable between control (Z″ = 1) and treatment (Z″ = 0)

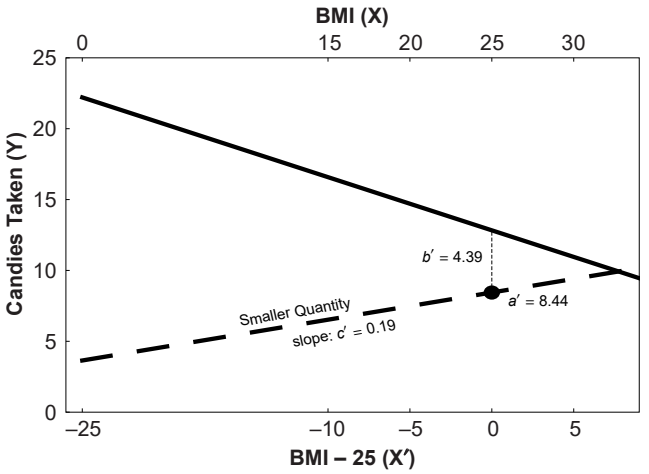
Notes: Y = a + bZ + cX + dZX.

Figure 1
GRAPHICAL INTERPRETATION OF REGRESSION PARAMETERS FROM TABLE 3

A: Model Parameters for Raw Data



B: Model Parameters for Transformed Data



Notes: The dashed lines represent the smaller quantity group where Z = 0; the solid lines represent the larger quantity group where Z = 1. Panel A shows the regression results using the untransformed data. The coefficient on quantity, b, reflects the effect of quantity for a BMI of 0, an impossible value that lies well outside the range of the data. Panel B shows the regression results using the transformed data, recoded such that the definition of borderline overweight (a BMI of 25) lies at 0. Everything about the graph is exactly the same, other than the recoded x-axis. The statistical test still is a test at 0, but now 0 corresponds to a substantively meaningful value.

from a moderating variable to make the coefficients on the other variables reflect simple effects of those variables at particular values of the moderator.

Simple Effect of Categorical Variable Z at a Given Level of Continuous Variable X

Understanding that the coefficient b reflects the simple effect of Z when X = 0 and that the coefficient c reflects the simple effect of X when Z = 0, we can recode X to examine the effect and statistical significance of Z at some value X_{Focal} other than the original X = 0. Simply subtract X_{Focal} from X to create a new variable ($X' = X - X_{Focal}$). Rerun the regression using X′ instead of X. The estimate, standard error, and significance test of b′ (the new coefficient on Z) are equivalent to those of b + dX at the focal value because when X = X_{Focal}, X′ = 0 (see Table 2, Panel B).

Simple Slope of Continuous Variable X at a Given Level of Categorical Variable Z

We can use the same "magic number zero" principles if we want to know the simple effect of the quantitative variable X at a given level of the manipulated Z. We can use the same principle to examine the estimate, standard error, and significance test of the slope of either line. Because the line for the group in which the model takes 2 candies, where Z = 0, is given by Equation 1b, the estimate, standard error, and significance test of c represent the estimate, standard error, and significance test of the slope of X for that group. To test the slope of X for the group in which the model takes 30 candies, recode Z such that Z = 0 for the group in which the model takes 30 candies and Z = 1 for the 2-candy group (see Table 2, Panel C).

Note that this is only the case when Z is dummy coded (i.e., one group is coded as 0 and the other group is coded as 1). If Z is contrast coded such that one group is coded as –1 and the other is coded as 1, the magic number zero principle still holds such that the coefficient on X still represents the relationship between X and Y when Z = 0, but this no longer represents the simple slope for either group. Instead, if Z is contrast coded, the coefficient on X represents the unweighted average of the two simple slopes (a "main effect" in ANOVA terms). The simple slope for the group represented by Z = –1 is given by (c – d), and the simple slope for the group represented by Z = 1 is given by (c + d).

These two examples, testing both the difference between two regression lines and the slope of a single line by recoding interacting variables, are examples of a broader principle: In linear models with interaction terms, the estimate, standard error, and significance test of a coefficient on a variable represent the estimate, standard error, and significance test of the simple effect of that variable when all variables it interacts with are equal to 0. As Irwin and McClelland (2001) note, many scholars incorrectly interpret these parameters as main effects rather than as simple effects. This mistake persists in recent marketing research.

Because coefficients b and c in Equation 1 are interpretable as simple effects when interacting variables are set equal to 0, strategic recoding enables a researcher to examine effect sizes and significance tests at other values of interest.[2] This is not limited to the familiar 2 × continuous design. We describe other cases subsequently in this article and in Web Appendix A (www.marketingpower.com/jmr_webappendix). In the next section, we illustrate this point and emphasize the role of focal values in the simple case of two interacting variables in a 2 × continuous design.

SPOTLIGHT ANALYSIS AT MEANINGFUL FOCAL VALUES

We begin illustrating spotlight analysis in a simple common design. We have two purposes in discussing this

[2]Rather than recoding X′ and redoing the regression analysis for a given X_{Focal}, we can directly estimate the coefficient b and its standard error using components of the variance–covariance matrix.

design. First, we establish the basic paradigm used in all extensions of the magic number zero in more complex designs. Second, we emphasize the suboptimal nature of the spotlight tests that marketing researchers most often report for these designs, in which they examine simple effects of a manipulated variable at plus and minus one standard deviation from the mean of a measured variable.

For this and subsequent examples, we generated fictitious illustrative data (N = 100) to showcase various analysis methods and results; we generated all of these data to be consistent with plausible predictions made from McFerran et al.'s (2010) results discussed previously, but we collected no real data for these examples.[3] Again, the dependent variable (Y) is the number of candies the participant takes. The dichotomous independent variable (Z) is the number of candies the confederate takes (2 candies, coded as 0, vs. 30 candies, coded as 1). The continuous variable (X) is the confederate's BMI (M = 21.97, SD = 2.90). We are interested in the effect on quantity taken by the nonconfederate participant. We estimate the parameters of the moderated regression model given by Equation 1; these appear in Table 3, Panel A, and Figure 1, Panel A.

The regression results using the untransformed data are not readily interpretable. The significant value of the coefficient d tells us that the interaction is significant. That is, the two experimental groups have different slopes relating BMI of the confederate to number of candies the participant takes. In the 2-candy condition in which Z = 0, the slope is c. In the 30-candy condition in which Z = 1, the slope is c + d. However, because of the magic number zero, a (the intercept) and b (the coefficient for Z) pertain only to when BMI = 0, an impossible value.

Given that we find a significant interaction, at what values of X should we test for simple effects of Z? The convention is to test at one standard deviation above and below the mean, though we argue that these arbitrary values are not very informative and researchers should instead test at meaningful focal values. There are commonly agreed cutoffs for BMI between underweight and normal weight, normal weight and overweight, and overweight and obese. These cutoffs represent meaningful values, and we argue that readers should be more interested in knowing the effect of Z at these focal meaningful values that are not sample dependent than they should be in tests at plus and minus one standard deviation from the sample mean for an idiosyncratic sample. Furthermore, effect sizes for sample-dependent values of X are less likely to generalize than those for meaningful focal values that are the same across studies.

Consequently, we might want to know the effect of Z when X = 25, the cutoff between being normal weight and being overweight; we could equally apply these procedures to any of the other focal cutoffs. To observe the effect of choice of large versus small quantity for a borderline overweight confederate, we would simply define a new variable $X' = X - 25$. We want to set the X value of interest equal to 0: this is the key. We set $X' = X - 25$ and reran the resulting model:

$$Y = a' + b'Z + c'X' + d'ZX'. \quad (2)$$

Table 3, Panel B, shows the parameter estimates and tests for an analysis of this model using X′ = BMI – 25 instead of X = BMI. It is important to note that 25 is subtracted from raw, not mean-centered, BMI.

Figure 1, Panel B, depicts the parameters for the model in Table 3, Panel B. In the figure, a value of X = 25 corresponds to X′ = 0. Note that the underlying models in Figure 1, Panels A and B, are identical. In particular, the slopes for the two groups are unchanged by the transformation of X, and the estimates and tests are unchanged for the coefficient c = c′, the slope for the group that observed the confederate take the smaller quantity. Furthermore, there is no change in the interaction coefficient d = d′, the difference between the slopes in the two conditions. However, now the coefficient b′ (≠ b) estimates the difference between the two groups when X = 25 (i.e., the effect of the number of candies the observed confederate takes, estimated for confederates with BMI = 25, the lower bound of the overweight range). In this case, there is a significant difference between the two groups when X′ = 0 or equivalently, when X = 25. The intercept a′ also changes because it now reflects the forecasted value when X′ = 0 (i.e., X = 25) and Z = 0; that is, the model predicts that participants who observe a borderline over-

[3]The data sets for the 2 × continuous and 3 × continuous examples are available in the Web Appendix (www.marketingpower.com/jmr_webappendix) so readers may replicate the analyses reported here or do any further examination of these illustrative, hypothetical data.

Table 3
REGRESSION RESULTS OF FICTITIOUS ILLUSTRATIVE DATA

A. Results in Raw Metric (X = BMI)					
Variable	*Coefficient*	*Estimate*	*Standard Error*	*t*	p
Intercept	a	3.65	3.93	.93	.356
Coded number of candies taken by confederate	b	18.56	5.54	3.35	.001
Confederate BMI	c	.19	.17	1.10	.273
Coded number of candies taken by confederate × confederate BMI	d	–.57	.25	–2.27	.026
B. Results After Transformation (X′ = BMI – 25) to Examine the Simple Effect for People with a BMI of 25					
Variable	*Coefficient*	*Estimate*	*Standard Error*	*t*	p
Intercept	a′	8.44	.68	12.39	<.001
Coded number of candies taken by confederate	b′	4.39	1.05	4.18	<.001
Confederate BMI – 25	c′	.19	.17	1.10	.273
Coded number of candies taken by confederate × (confederate BMI – 25)	d′	–.57	.25	–2.27	.026

weight model (BMI = 25) take 2 candies will themselves take approximately 8.4 candies. In summary, when the model includes all terms, adding or subtracting a constant to a variable X leaves the underlying model unchanged and only changes the coefficients for the intercept and other variables with which that variable is multiplied. It does not affect the coefficient on terms that include that variable, contrary to what many first expect when learning spotlight tests.

This example illustrates two major points. First, by recoding variables in a moderated regression to change what is coded as zero, we can derive simple effect spotlight tests for the effect of a variable in the model in which other variables are set to zero. Second, there are many cases in which authors should break from the convention of conducting spotlight tests at plus and minus one standard deviation from the mean. Indeed, we would argue that this convention is almost never the best approach, notwithstanding that we have both used and advocated the approach in previous studies.

There are three main problems of testing at plus and minus one standard deviation. First, if the distribution of the moderator X is skewed, one of those values can be outside the range of the data. Second, if the moderator X is on a coarse scale, it may be impossible to have a value of X exactly equal to plus or minus one standard deviation. Third, if two researchers replicate the same study with samples of very different mean levels of the moderator, they appear to fail to replicate one another even when they find exactly the same regression equation in raw score units. The tendency of authors to fail to report the mean and standard deviation of X only exacerbates this problem. Fernbach et al. (2012) face these potential problems using Frederick's (2005) cognitive reflection test, a coarse scale (four points ranging from 0 to 3) that can take on skewed distributions that vary substantially across populations. For example, one standard deviation above the mean of Frederick's Massachusetts Institute of Technology (MIT) sample (M = 2.18, SD = .94) would be an impossibly high value, one standard deviation below the mean of his University of Toledo sample (M = .57, SD = .87) would be an impossibly low value, and a "low" MIT score would be similar to a "high" University of Toledo score. Fernbach et al. (2012) successfully handle these problems by not using sample-dependent and potentially impossible values of "high" and "low" but rather by testing at the scale endpoints. We expand on these problems and provide further details in Web Appendix B (www.marketingpower.com/jmr_webappendix).

In cases in which there are meaningful focal values, we recommend a spotlight test focusing on simple effects of the manipulated variable at judiciously chosen values of the moderator rather than at an arbitrary number of standard deviations from the mean. In other cases, no particular value of the moderating variable is particularly focal. In those cases, we recommend reporting a floodlight analysis of the simple effect of the manipulated variable across the entire range of the moderator, reporting regions where that simple effect is significant. This approach is appropriate when the scale of measurement is "arbitrary"—that is, when it is an interval scale of some underlying construct with an unknown zero point.

FLOODLIGHT ANALYSES: SPOTLIGHT ANALYSES FOR ALL VALUES OF X

The Johnson–Neyman Point

Spotlight analysis provides a test of the significance of one coefficient at a specific value of another continuous variable, so it is most useful when there is some meaningful value to test. When it is not the case that some values are more meaningful than others or when the researcher anticipates that readers might be interested in other spotlight values, we recommend using an analysis introduced by Johnson and Neyman (1936) that we dub "floodlight" analysis. Whereas the spotlight illuminates one particular value of X to test, the floodlight illuminates the entire range of X to show where the simple effect is significant and where it is not; the border between these regions is known as the Johnson–Neyman point. In essence, this test reveals the results of a spotlight analysis for every value of the continuous variable. As Preacher, Curran, and Bauer (2006) note, this eliminates the arbitrariness of choosing high and low values such as one standard deviation above and below the mean.

Johnson and Neyman (1936) introduce the concept and statistical underpinnings of floodlight analysis. Rogosa (1980, 1981), and Preacher et al. (2006) contribute important later developments. It is not necessary to delve into the underlying mathematics to understand the basic concept. The Johnson–Neyman point (or points: there are always two such points that could either straddle a crossover or both be on the same side; see McClelland and Lynch 2012) is the value of X at which a spotlight test would reveal a p-value of exactly .05 (or whichever alpha one is using). In the case of a 2 × continuous interaction, it is the value of X for which the simple effect of Z is just statistically significant. Values of X on one side of the Johnson–Neyman point yield significant differences between the two groups, whereas values on the other side do not. In this way, a floodlight shines on the range of values of the continuous predictor X for which the group differences are statistically significant.

Mohr, Lichtenstein, and Janiszewski (2012) provide a recent example of such an analysis and presentation. In their research, they were interested in the effect of the interaction of dietary concern (assessed as a continuous measure averaging items rated on "arbitrary" seven-point scales, yielding, at most, an interval scale of the underlying construct) and health frame on guilt and purchase intention. Because they assess the continuous moderator on an arbitrary scale without focal values, they present the results showing the range over which the simple effect is significant rather than picking sample-dependent points without real meaning to the reader.

Conducting a Floodlight Analysis in the 2 × Continuous Case

Floodlight analysis remained obscure for years because of its apparent computational complexity. Now macros for SPSS, SAS, and R (Hayes 2012; Hayes and Matthes 2009; Preacher et al. 2006)[4] make computing Johnson–Neyman points feasible for any researcher. Rather than testing a par-

[4]The macros and instructions for using them are available at http://afhayes.com/spss-sas-and-mplus-macros-and-code.html and http://quantpsy.org/interact/index.html.

ticular value of X as in spotlight analysis, these macros solve for values of X for which the t-value is exactly equal to the critical value—in other words, values for which a spotlight analysis would give significant results on one side and nonsignificant results on the other side. Rather than spotlighting a single point, this floodlights the entire range of the data to reveal where differences are and are not significant rather than focusing on one or two arbitrary points. As Potthoff (1964) and Hayes and Matthes (2009) note, these regions do not adjust to account for multiple comparisons across the entire range. However, they do enable researchers to claim that any spotlight test within that range would be significant.

Researchers preferring not to download and learn new macros can readily perform a floodlight analysis by performing a spotlight analysis for a grid of interesting values, being sure to include the minimum and maximum plausible values of the continuous predictor. Often, there is only a discrete set of interesting or plausible values (e.g., points on a seven-point rating scale). For example, Nickerson et al. (2003) provide the spotlight values of the coefficient for a list of income ranges that were of interest. If finer precision is desired, it is easy to observe between which grid values the spotlight switches from being significant to nonsignificant; the Johnson–Neyman value must lie within that interval. Iterative spotlight analyses using numbers between those two grid values will quickly determine a fairly exact Johnson–Neyman value. Importantly, this iterative process generalizes to more complex designs for which macros often do not exist. The general strategy is to do spotlight analyses on a grid of values for one or more variables and note the regions in which the spotlight values switches from significant to insignificant. Then iterate between those values if more precision is desired.

As an illustration of performing floodlight analysis by an iteration of selected spotlight values, Table 4 displays the difference between the two example groups in number of candies taken (i.e., the vertical distance between the two regression lines in Figure 1) for spotlighted values of BMI between 16 and 32 in steps of 2, along with the associated statistical information for the coefficient describing the difference at each spotlighted value of BMI. We constructed the table by performing nine spotlight regressions. For example, we obtained the values in the first row by computing a new variable $X' = X - 16$ and then estimating the regression model (Equation 2):

$$Y = a' + b'Z + c'X' + d'ZX'.$$

The tabled values are the statistical values for the coefficient b′.

Table 4

SPOTLIGHT ANALYSES OF THE DIFFERENCE BETWEEN GROUPS IN NUMBER OF CANDIES TAKEN FOR A SYSTEMATIC SELECTION OF BMI VALUES BETWEEN THE MINIMUM AND MAXIMUM

BMI	*Group Difference (Candies Taken)*	*Lower 95% Confidence Interval*	*Upper 95% Confidence Interval*	*t(96)*	p
16	9.49	6.20	12.78	5.73	<.0001
18	8.36	5.92	10.79	6.82	<.0001
20	7.22	5.49	8.95	8.28	<.0001
22	6.09	4.65	7.52	8.43	<.0001
24	4.95	3.20	6.71	5.60	<.0001
26	3.82	1.35	6.29	3.07	.003
28	2.69	–.64	6.01	1.60	.11
30	1.55	–2.69	5.80	.72	.47
32	.42	–4.77	5.61	.16	.87

Note in Table 4 that the difference between groups in number of candies taken switches from being significantly greater than zero for BMI = 26 to not being significantly greater than zero for BMI = 28. Thus, the Johnson–Neyman point must be between BMI = 26 and BMI = 28. Further iteration within the range between 26 and 28 (not presented here) locates the Johnson–Neyman point more precisely at BMI = 27.4. Figure 2, Panel A, displays the regression lines for both model groups with the filled region (the floodlight) indicating for which values of BMI a spotlight analysis would reveal a significant difference in the number of candies taken between groups. That is, there is a significant difference between groups for values of BMI between 16 and 27.4 and not a significant difference for BMI values above 27.4. Figure 2, Panel B, graphs the simple effect of the manipulation as it varies across X, showing that the Johnson–Neyman point is located where the 95% confidence band around the simple effect intersects the x-axis.

To report a floodlight analysis, report the Johnson–Neyman point and range(s) of significance, or if using the grid search method, report the range(s) of significance and intervals tested. When graphing the results, show the Johnson–Neyman point or grid points tested and report range(s) of significance. For example,

> Regressing candies taken on the manipulation (2 candies = 0, 30 candies = 1), BMI (M = 21.97, SD = 2.90, min = 16.5, max = 29.0), and their interaction revealed a significant interaction ($t(96) = -2.27$, $p < .05$). To decompose this interaction, we used the Johnson–Neyman technique to identify the range(s) of BMI for which the simple effect of the manipulation was significant. This analysis revealed that there was a significant positive effect of candies taken by the model on candies taken by the participant for any model BMI less than 27.4 ($B_{JN} = 3.03$, SE = 1.54, $p = .05$), but not for any model BMI greater than 27.4.

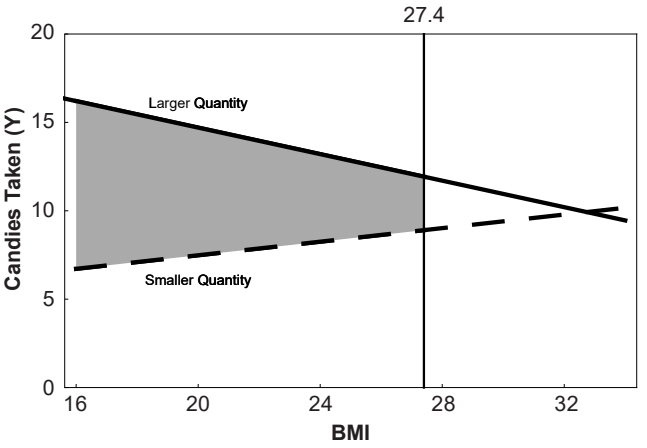
EXAMPLES OF WHEN TO USE SPOTLIGHT AND WHEN TO USE FLOODLIGHT

Sometimes spotlight analysis is more appropriate and sometimes floodlight analysis is more informative. We lay out the relevant considerations here. First, is the scale meaningful with a known correspondence to the underlying construct, or is it arbitrary with an unknown linear mapping from numbers on the scale to levels of the underlying construct? Second, do readers understand certain focal values to have meaningful referents even if not all values have meaningful referents? Third, are some values of the variable impossible due to coarseness of the scale? For a simple decision tree to determine whether to use spotlight or floodlight, see Figure 3.

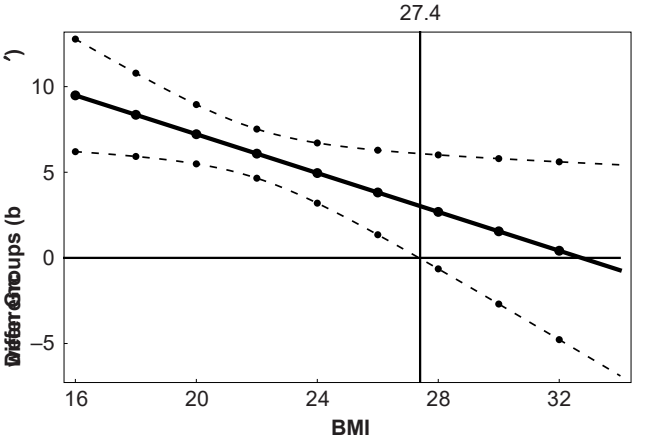
Blanton and Jaccard (2006) decry misuse of "arbitrary metrics" in psychology, wherein researchers interpret values of some interval scale as "low" or "high." A scale is "arbitrary" when the parameters of the function linking a

Figure 2
CORRESPONDENCE BETWEEN THE JOHNSON–NEYMAN POINT AND THE CONFIDENCE BANDS AROUND THE SIMPLE EFFECT OF Z

A: Regression Lines with Johnson–Neyman Point



B: Estimated Simple Effect of Z with Confidence Bands



Notes: Panel A shows a floodlight of the region of BMI values (filled area below 27.4) for which a spotlight test would reveal significant differences between the two model groups. Panel B shows a graph of the estimated simple effect (the distance between the two regression lines in Panel A) with confidence bands. Confidence bands are narrowest at mean BMI (M = 21.97). The Johnson–Neyman point in Panel A aligns with the intersection of the confidence band and the x-axis in Panel B. The crossover point in Panel A aligns with the intersection of the estimated simple effect and the x-axis in Panel B.

person's true score on a latent construct to observed scores is unknown or not transparent. Blanton and Jaccard are particularly critical of the use of scores from the Implicit Association Test. These scores are based on reaction times, which as a measure of time have ratio scale properties. However, when researchers interpret the reaction time (or difference of reaction times) as a measure of latent prejudice, it becomes an arbitrary scale, and values of zero are no longer particularly meaningful. (Difference scores computed from interval scale ratings of two objects can be meaningful values if zero truly represents no difference in perceptions/ratings of the objects. For an example, see Spiller 2011, Appendix Study 4.)

Most individual difference variables used in marketing and consumer research are arbitrary in that they are interval scales of the underlying constructs with unknown units and origins: propensity to plan, involvement, need for cognition, tightwad–spendthrift, need for uniqueness, and so on. We argue that floodlight tests are likely to be more appropriate than spotlight tests at chosen values of these scales.

When values are clearly nonarbitrary and focal values are meaningful, spotlight analysis can be particularly illuminating. Consider Study 1 from Leclerc and Little (1997), cited by Irwin and McClelland (2001) as an early example of clever rescaling of X to generate a meaningful spotlight test to examine the effect of advertising for people who were maximally brand loyal. The authors examine how the effect of advertising content type (ad type: picture vs. information) on brand attitude varied as a function of brand loyalty.

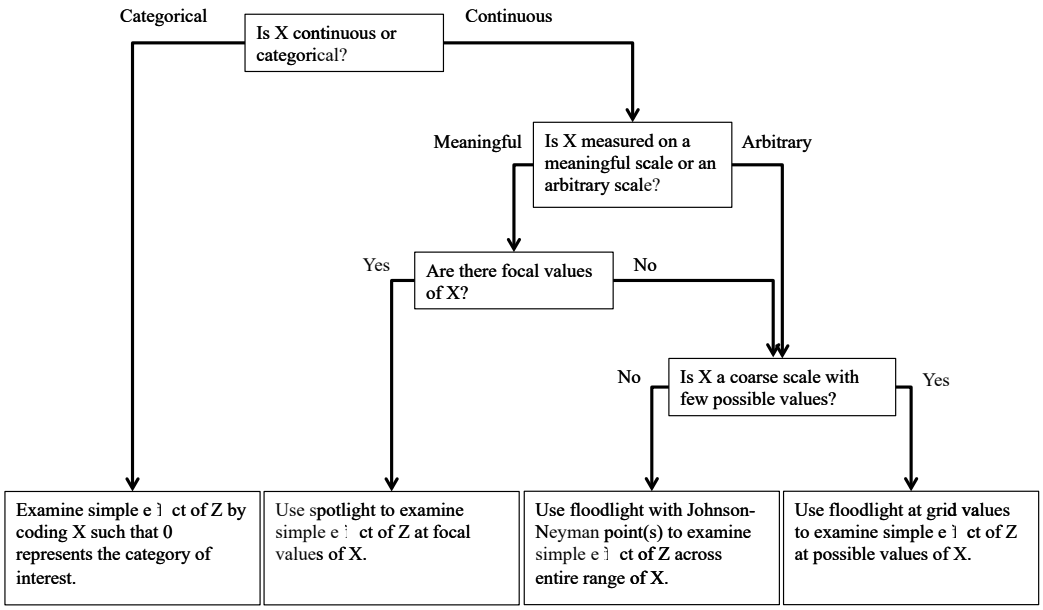
The authors operationalized brand loyalty as a function of the number of brands purchased in the product category the previous year: fewer brands indicate greater loyalty. Had they simply used the raw number of brands, the coefficient on advertising content type would have represented the effect of advertising content type for people who purchased zero brands the previous year. This would have been problematic for two reasons. First, they excluded from analysis nonusers who did not purchase any brands the previous year, so this point was outside the range of the data. Second, this was not a substantively meaningful value to test: the theory made predictions regarding brand loyalty, not brand usage.

However, because the theory made predictions for people who were brand loyal, it was meaningful to test the simple effect for people who purchased a single brand the previous year (i.e., those who were completely brand loyal). Thus, Leclerc and Little (1997) created a new variable, switching, calculated as number of brands purchased minus one. The authors regressed brand attitude on ad type, switching, and ad type × switching. They could therefore interpret the simple effect of advertising content type as the simple effect when switching was equal to zero—in other words, the simple effect for brand loyalists. Transforming one variable such that zero took on a substantively meaningful value provided the reader with information about an easily interpretable simple effect. Spotlight was particularly useful in this case; moreover, because loyalty can take on only integer values, it would be better to present spotlights at any integer values likely to be of interest to readers rather than arbitrary and impossible values one standard deviation above and below the mean.

The same point applies to our previous hypothetical extension of McFerran et al. (2010). Had relative weight been measured using a subjective seven-point scale rather than BMI, we would advocate using a floodlight analysis to consider ranges of significance rather than meaningless values one standard deviation above and below the mean.

Interval scales can become nonarbitrary when researchers develop norms for where a particular score in the distribution lies across some reasonably representative sample in a population of consumers. Churchill (1979) advocates this development of norms as the last step of scale development. However, this norming step has not been a part of practice in most marketing and consumer research on scale development, including our own.

Figure 3
FLOODLIGHT DECISION TREE



Notes: This decision tree helps researchers determine when to use spotlight analysis and when to use floodlight analysis to examine the simple effect of Z across a moderating variable X in a model of the form $Y = a + bZ + cX + dZX$.

We argue that in cases such as the McFerran et al. (2010) extension, it is more meaningful to report spotlight tests at specific values of an interval or ratio scale—which facilitates comparisons across studies using the same scale with samples drawn from different populations—than to bury the metric of the original scale by reporting plus and minus some sample-dependent standard deviation. This enables accumulation of findings over time about the range of values of the scale in which the simple effect of some independent variable is substantively and statistically significant. Edwards and Berry (2010) argue that moving beyond hypotheses that merely postulate the sign of some effect to specifying the range of values where the effect holds can increase theoretical precision.

MAGIC NUMBER ZERO FOR SPOTLIGHTS IN OTHER COMMON DESIGNS

Thus far, we have discussed spotlight and floodlight analyses in the simple 2 × continuous case. In this section, we present a simple, easy-to-implement method for accomplishing spotlight analysis in other common designs. It should be apparent that the iterative grid approach in Table 4 can derive floodlight tests in these designs as well.

From the literature, it is clear that when designs vary from the standard 2 × continuous design, authors take a variety of inappropriate strategies to analyze simple effects, including median-splitting one or more continuous variables, breaking down a higher-order interaction into separate subsamples and running piecewise analyses on each, and misinterpreting simple effects at one level of a variable as main effects across all levels of a variable. The following cases provide a better way of conducting such analyses.

In the following subsections, we show how to extend these principles to (1) the case of a 2 × continuous design when Z is manipulated within participants and (2) the case of a 2 × 2 × continuous design when all factors are between participants. Web Appendix A (www.marketingpower.com/jmr_webappendix) extends these principles to two other cases, the 3 × continuous design and the case in which X and Z are both continuous. We also consider models with quadratic terms. We emphasize that by no means are these the only designs for which we can use spotlight. Instead, these are representative examples of common designs, and the basic principle we use in these designs (recognition of "the magic number zero") can be applied to every other design that uses a linear model (including logistic regression) for analysis. For each design, we build on the basic extension of McFerran et al. (2010) described previously. Web Appendix A gives analysis templates for Cases 1, 2, and 3 along with the aforementioned extensions.

Case 1: 2 × Continuous When the Manipulation of Z Is Within Subject

Imagine a version of our original example with two levels of the manipulated factor Z (0 = confederate took 2 candies, 1 = confederate took 30 candies) and a continuous measure of X = BMI. This time, however, let Z be a repeated measures factor. In this case, we simply create a contrast score for each subject showing the effect of the manipulation for that subject: $Z_{contrast} = Y_{30} - Y_2$ (see Judd, McClelland, and Ryan 2009; Keppel and Wickens 2004); we could similarly create contrast scores for within-subject designs with more than two levels. We then analyze the $Z_{contrast}$ scores as a function of X = BMI:

$$Z_{contrast} = a + bX. \quad (3)$$

To extend the principle of the magic number zero, the test of the intercept a in this analysis is the predicted $Z_{contrast}$ score when X = 0. The coefficient b now is equivalent to a test of the interaction of X with Z in the original design. To create a spotlight test of the effect of the repeated factor Z at the borderline between normal and overweight, create X′ = X – 25. Rerun the regression $Z_{contrast} = a' + b'X'$. Now the test of the intercept a′ is the effect of the repeated factor Z at the new zero point associated with the chosen level of X.

Case 2: 2 × 2 × Continuous

Often, researchers may be interested in how a continuous variable moderates a 2 × 2 interaction, resulting in a three-way interaction. For example, in addition to manipulating quantity taken, we might also manipulate the perceived healthfulness of the item being considered (candy vs. granola, as in Study 1 of McFerran et al. 2010). The prediction might be that attenuation of assimilation only occurs for unhealthy food because participants are cued to be more vigilant when food is unhealthy than when it is healthy. (McFerran et al. find this not to be the case.) The model for this design is as follows:

$$(4)\quad Y = a + bZ + cW + dX + eZW + fZX + gWX + hZWX.$$

Here, Z and X are coded the same as they were in the opening example, and W is coded 0 for candy and 1 for granola. If the parameter h testing the three-way interaction is significant, it becomes relevant to test the simple interaction of two of the variables at different levels of the third variable. The coefficient e tests the simple ZW interaction when X = 0. (It does *not* test the ZW interaction that would be evident in plotting the ZW cell means, collapsing over levels of X.) The coefficient f tests the simple ZX interaction when W = 0. The coefficient g tests the simple WX interaction when Z = 0. To follow up a simple two-way interaction, we test the simple–simple effect of one of the variables holding constant the other two. In this model, b represents the simple–simple effect of Z when W = 0 and X = 0, c represents the simple–simple effect of W when X = 0 and Z = 0, and d represents the simple–simple effect of X when Z = 0 and W = 0.

Zero is a magic number in this analysis as well. We interpret every coefficient as the effect of that variable (or interaction) when all variables with which that term interacts are set to 0, causing them to drop out of the model.

Suppose that we obtained a significant three-way interaction ZWX and wanted to follow up with tests of the simple ZW interaction at meaningful levels of X. We would recode X′ = X – 25 just as in each of the previous examples. When X′ = 0, the new coefficient on ZW, e′, represents the simple interaction between quantity and type of snack taken (granola vs. candy) when X′ = 0, which corresponds to a BMI of 25. Spotlight analysis in a 2 × 2 × continuous design requires application of the same principle we used in the previous cases: recoding variables such that 0 represents the value of a variable at which we are interested in the simple effect of the other variables.

Web Appendix A (www.marketingpower.com/jmr_webappendix) includes detailed explanations of how to do simple effects tests in Cases 1 and 2. It also covers three additional cases (Case 3: 3 × continuous; Case 4: continuous × continuous; and Case 5: quadratic). It should be evident for each of these cases that researchers can easily accomplish floodlight analyses by iterating to redefine the value of X when X′ = 0 as in Table 4 within that particular design. Of course, they can extend these analyses to other designs not described here; for example, they could examine 2 × 3 × continuous by combining the strategies used in the 3 × continuous and 2 × 2 × continuous cases.

CONCLUSION

Spotlight tests reflect the simple effect of a variable Z at different levels of an interacting variable X. Aiken and West (1991), Jaccard et al. (1990), and Irwin and McClelland (2001) popularized these tests, but they remain misunderstood. We have shown that these tests rely on basic multiple regression principles; by changing the coding of variables to alter the zero point, tests of the parameters of a moderated regression model can provide the simple effects tests of interest.

Some researchers, reviewers, and editors seem wary or uncertain of using these tests in anything but the simple case of a dichotomous manipulated variable Z and a continuous measured variable X. For that reason, they fall back on the flawed practice of dichotomizing continuous variables when faced with more complex designs, or they use the continuous variable to test the significance of the interaction but dichotomize to graph the interaction or do simple effects tests. We show how we can apply the general principle of the magic number zero to derive ready tests of simple interactions and simple–simple effects in an array of more complex designs. When interacting variables are coded such that zero represents focal values, those interacting variables drop out of the model at their focal values. We then interpret the remaining terms in the model as simple effects at those focal values of interacting variables.

We criticize the common practice of reporting spotlight tests of the simple effect of Z at plus and minus one standard deviation from the mean on an interacting variable X. We argue that this is almost never optimal because those tests and estimates are sample dependent in defining high and low values of X, because it is possible that these estimates refer to impossible values of X and because readers are not inherently more interested in the effect of Z at plus and minus one standard deviation than at values of X somewhat higher or lower. We argue that in some cases, researchers overlook that there may be "focal" values that are of particular interest, and we encourage use of these more judiciously chosen levels of the continuous variable for spotlight tests.

There are many cases in which researchers apply spotlight analysis where no particular value of the continuous variable is focal. In these cases, we recommend abandoning the convention of testing spotlights at plus and minus one standard deviation from the mean. Instead, we recommend use of a related test that shows ranges of the continuous variable where the simple effect of a second variable is significant and where it is not. Johnson and Neyman (1936) originally reported this technique. We dub this a floodlight analysis, as it illuminates the entire range of the data rather than spotlighting a single point. Reporting floodlight analyses provides readers an efficient way to infer whether two groups differ at any given point of interest and facilitates

comparing and integrating findings across multiple samples with different sample distributions.

APPENDIX: STATISTICAL AND POWER CONSIDERATIONS IN SPOTLIGHT AND FLOODLIGHT TESTS

In this Appendix, we summarize statistical parameter estimation and power considerations in spotlight and floodlight analysis in a design with a manipulated Z with two levels (dummy coded 0 = control, 1 = treatment) and X coded as a continuous variable in its raw metric:

$$Y = a + bZ + cX + dZX. \quad \text{(A1)}$$

1. Power to detect the simple effect of Z varies with X. This is true because both the numerator and the denominator of the F test for the simple effect of Z change with X. A nonzero interaction of X and Z implies that there is some value of X where the regression lines for the two levels of Z cross over, although this crossover may occur outside the range of data. Johnson and Neyman (1936) prove that one can find two values of X where the effect of Z is exactly significant. McClelland and Lynch (2012) demonstrate that it is possible to have real data in which the effect of Z is significant to the right or left of the crossover point of the interaction, but not significant as one moves further away from the crossover. To guard against this, be sure to conduct a spotlight test at minimum and maximum values of X.
2. In the main text, we discuss how two researchers replicating the same experiment with a manipulated Z and a measured X might find exactly the same regression equation but perceive that they had failed to replicate one another's findings if the two studies used samples with high versus low average values of the measured variable X. Now consider that the researchers analyzed the two exact replicates using spotlights at the same focal value of the moderator, X_{Focal} as expressed in raw score units. The statistical tests on the simple effect of the manipulated variable Z at X_{Focal} will not match in the two replicates, because the standard error of the coefficient is smaller when the focal value X_{Focal} is closer to the sample mean (McClelland and Lynch 2012).
3. Some methodologically sophisticated colleagues have expressed skepticism when told that the simple effect of a manipulated variable Z is significant at a value of X two standard deviations from the mean. Their statistical intuition is that the analysis relies on a small subset of cases far from the mean. They are missing that the solution to the moderated regression uses all of the data from the study. Like any regression, the spotlight statistical tests of the simple effect of Z at a given level of X reflect that there are wider confidence intervals for the predicted value of Y when X is far from the mean of the data than when it is close to the mean.
4. The use of spotlight tests at each value of an arbitrary but coarse scale is different from treating each value as discrete levels of a categorical factor in an ANOVA. In that latter case, the power of the test of the simple effect of a manipulated variable at a level of the moderator variable is affected only by the number of cases at that level of the moderator variable. In contrast, the spotlight test is a regression parameter estimate that treats the moderator as a continuous variable. In this case, the power of the spotlight test of the simple effect of the manipulated factor is affected by the entire data set including all possible values of the moderator. In this case, there is no special danger of testing for extreme values of the moderator that are inside the range of the data. Like any regression, the spotlight statistical tests of the simple effect of Z at a given level of X reflect that there are wider confidence intervals for the predicted value of Y when X is far from the mean of the data than when it is close to the mean.
5. Even when the true interaction has a coefficient d exactly equal to zero, merely *including* the interaction term also causes the standard error of b to change with a rescaling of X, because now b is explicitly testing the effect of Z at the value of X coded as 0. For example, consider a model in which X takes on values from 1 to 7 with a mean of 4 and Z is dummy coded. Assume that the true interaction is zero, and when one estimates Equation 1, the coefficient d on the interaction is exactly zero. The standard error on the coefficient b is the standard error when X = 0, outside the range of the data. One will get a smaller estimate of the standard error and a larger t test on the parameter b if one mean centers $X' = X - 4$, so that now 0 is coded at the mean. The standard error of b, the distance between the two lines, is smallest at the mean of X; as one moves farther away from the mean of X, the standard error increases. Thus, even if the estimate of the interaction is exactly equal to 0, the test of the simple effect of Z will be less powerful if tested far from the mean than if tested at the mean. This has practical implications because if X ranges from 1 to 7 and the effects of Z, X, and ZX on Y are modeled using Equation 1, even if the effects of X and ZX are exactly 0, the significance test of Z will be misleading or at least misunderstood if X is analyzed in its raw metric. Note that this effect of rescaling X on the estimate and standard error of the effect of Z does *not* occur if no interaction term is included in the model; that is, for a "main effects"–only model $Y = a + bZ + cX$. In that model, neither coefficient b nor c changes, nor do the standard errors on b and c change when a constant is added or subtracted from Z or X, although the estimate of a changes.

REFERENCES

Aiken, Leona S. and Stephen G. West (1991), *Multiple Regression: Testing and Interpreting Interactions.* Newbury Park, CA: Sage Publications.

Blanton, Hart and James Jaccard (2006), "Arbitrary Metrics in Psychology," *American Psychologist*, 61 (1), 27–41.

Churchill, Gilbert A., Jr. (1979), "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research*, 16 (February), 64–73.

Edwards, Jeffrey R. and James W. Berry (2010), "The Presence of Something or the Absence of Nothing: Increasing Theoretical Precision in Management Research," *Organizational Research Methods*, 13 (4), 668–89.

Fernbach, Philip M., Steven A. Sloman, Robert St. Louis, and Julia N. Schube (2012), "Explanation Fiends and Foes: How Mechanistic Detail Determines Understanding and Preference," *Journal of Consumer Research*, 40, (published electronically December 14), [DOI:10.1086/667782].

Frederick, Shane (2005), "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 19 (4), 25–42.

Hayes, Andrew F. (2012), "PROCESS: A Versatile Computational Tool for Observed Variable Mediation, Moderation, and Conditional Process Modeling," white paper, The Ohio State University, (accessed January 18, 2013), [available at http://www.afhayes.com/public/process2012.pdf].

——— and Jörg Matthes (2009), "Computational Procedures for Probing Interactions in OLS and Logistic Regression: SPSS and SAS Implementations," *Behavior Research Methods*, 41 (3), 924–36.

Irwin, Julie R. and Gary H. McClelland (2001), "Misleading Heuristics and Moderated Multiple Regression Models," *Journal of Marketing Research*, 38 (February), 100–109.

——— and ——— (2003), "Negative Effects of Dichotomizing Continuous Predictor Variables," *Journal of Marketing Research*, 40 (August), 366–71.

Jaccard, James, Vincent Guilamo-Ramos, Margaret Johansson, and Alida Bouris (2006), "Multiple Regression Analyses in Clinical Child and Adolescent Psychology," *Journal of Clinical Child & Adolescent Psychology*, 35 (3), 456–79.

———, Robert Turrisi, and Choi K. Wan (1990), *Interaction Effects in Multiple Regression*. Thousand Oaks, CA: Sage Publications.

Johnson, Palmer O. and Jerzy Neyman (1936), "Tests of Certain Linear Hypotheses and Their Application to Some Educational Problems," *Statistical Research Memoirs*, 1, 57–93.

Judd, Charles M., Gary H. McClelland, and Carey S. Ryan (2009), *Data Analysis: A Model Comparison Approach*, 2d ed. New York: Routledge.

Keppel, Geoffrey and Thomas D. Wickens (2004), *Design and Analysis: A Researcher's Handbook*. New York: Pearson.

Leclerc, France and John D.C. Little (1997), "Can Advertising Copy Make FSI Coupons More Effective?" *Journal of Consumer Research*, 34 (4), 473–84.

MacCallum, Robert C., Shaobo Zhang, Kristopher J. Preacher, and Derek D. Rucker (2002), "On the Practice of Dichotomization of Quantitative Variables," *Psychological Methods*, 7 (1), 19–40.

Maxwell, Scott E. and Harold D. Delaney (1993), "Bivariate Median Splits and Spurious Statistical Significance," *Psychological Bulletin*, 113 (1), 181–90.

McClelland, Gary and John G. Lynch Jr. (2012), "Power Considerations in Simple Effects Tests in Moderated Regression," working paper, University of Colorado Boulder.

McFerran, Brent, Darren W. Dahl, Gavan J. Fitzsimons, and Andrea C. Morales (2010), "I'll Have What She's Having: Effects of Social Influence and Body Type on the Food Choices of Others," *Journal of Consumer Research*, 36 (6), 915–29.

Mohr, Gina S., Donald R. Lichtenstein, and Chris Janiszewski (2012), "The Effect of Marketer-Suggested Serving Size on Consumer Responses: The Unintended Consequences of Consumer Attention to Calorie Information," *Journal of Marketing*, 76 (January), 59–75.

Nickerson, Carol, Norbert Schwarz, Ed Diener, and Daniel Kahneman (2003), "Zeroing in on the Dark Side of the American Dream: A Closer Look at the Negative Consequences of the Goal for Financial Success," *Psychological Science*, 14 (6), 531–36.

Potthoff, Richard F. (1964), "On the Johnson–Neyman Technique and Some Extensions Thereof," *Psychometrika*, 29 (3), 241–56.

Preacher, Kristopher J., Patrick J. Curran, and Daniel J. Bauer (2006), "Computational Tools for Probing Interactions in Multiple Linear Regression, Multilevel Modeling, and Latent Curve Analysis," *Journal of Educational and Behavioral Statistics*, 31 (3), 437–48.

Rogosa, David (1980), "Comparing Nonparallel Regression Lines," *Psychological Bulletin*, 88 (2), 307–321.

——— (1981), "On the Relation Between the Johnson–Neyman Region of Significance and the Statistical Test of Nonparallel Within-Group Regressions," *Educational and Psychological Measurement*, 41 (1), 73–84.

Spiller, Stephen A. (2011), "Opportunity Cost Consideration," *Journal of Consumer Research*, 38 (4), 595–610.

Vargha, András, Tamás Rudas, Harold D. Delaney, and Scott E. Maxwell (1996), "Dichotomization, Partial Correlation, and Conditional Independence," *Journal of Educational and Behavioral Statistics*, 21 (3), 264–82.

市场营销研究杂志》第 L 卷（2013 年 4 月），277288

Stephen A.Spiller 是加州大学洛杉矶分校安德森管理学院市场营销学助理教授（电子邮件：
stephen.spiller@anderson.ucla.edu）。Gavan J.Fitzsimons 是杜克大学福库商学院 R.David
Thomas 市场营销和心理学教授（电子邮件：gavan@duke.edu）。John G.Lynch Jr.是利兹商学院 Ted Anderson 教授，科罗拉多大学博尔德分校消费者财务决策研究中心主任（电子邮件：john.g.lynch@colorado.edu）。Gary H.McClelland 是科罗拉多大学博尔德分校心理学教授和认知科学研究所教职研究员(电子邮件：gary.mcclelland@colorado.
edu）。Fitzsimons、Lynch 和 McClelland 按字母顺序列出，贡献相同。作者感谢 Rick Staelin、Carl Mela 和 Wagner
Kamakura 提出的问题激发了本文的写作，感谢 Ajay Abra-
ham、Philip Fernbach、Yoosun Hann、Ji Hoon Jhang、Christina Kan、Peggy
Liu、Matthew Philp、Adriana Samper、Julie Schiro、Scott Wallace、Eliza-
beth Webb、Hillary Wiener 和评审团队的建设性意见，以及科罗拉多大学研讨会的参与者。任何错误均由作者负责。Don Lehmann 担任本文的副主编。本文由 Robert Meyer 邀请撰写。STEPHEN A.SPILLER、GAVAN J.FITZSIMONS、JOHN G.LYNCH JR.
和 GARY H.MCCLELLAND*

研究人员通常会发现测量变量 X 与操纵变量 Z 之间存在显著的相互作用，以检查 Z 在 X 的不同水平上的简单影响。即使在最简单的情况下，这些聚光灯测试也常常被误解，而且消费者研究人员似乎不确定如何将它们扩展到更复杂的设计中。

作者解释了聚光灯检验的一般原理，表明它们依赖于熟悉的回归技术，并提供了一个教程，演示如何在一系列实验设计中应用这些检验。与通常的做法不同，即在 X 的平均值上下一个标准差的情况下报告聚光灯检验，建议当 X 具有焦点值时，研究人员应报告这些焦点值的聚光灯检验。当 X 没有焦点值时，建议研究人员使用作者称之为“泛光灯”的 Johnson 和 Neyman 检验版本来报告重要性范围。关键词：调节回归、聚光灯分析、简单效应检验聚光灯、泛光灯和神奇的数字零：调节回归中的简单效应检验

2013，美国营销协会
ISSN：0022-2437（印刷版），1547-7193（电子版）

大多数营销和消费者行为文章报告了两个或多个变量之间相互作用的实验测试。作者可能会使用简单效应测试（计量经济学家称为条件效应测试）来跟踪两个变量的相互作用，以测试一个变量在另一个变量的不同水平上的影响。他们可能会使用测试两个变量在第三个变量水平上简单相互作用或测试一个变量在另外两个变量的选定水平上简单效应来跟踪三个变量的相互作用（Keppel 和 Wickens 2004）。

本文介绍了一种简单效应检验分析方法，该方法适用于一个或多个相互作用变量是连续和定量的，而非分类的。主要接受过方差分析 (ANOVA)框架的实验设计培训的研究人员，在跟踪连续变量与一个或多个分类变量相互作用的交互时，经常会遇到困难，而且适当的分析是在有调节的回归框架中进行的。在对《市场营销研究杂志》第 48 卷和《消费者研究杂志》第 38 卷的回顾中，我们发现报告的有调节的回归分析通常是正确的，但执行得不是最佳的，在许多其他情况下，完全是不正确的。我们在其他社会科学中也观察到了类似的大大小小的错误。在本文中，我们找出了最常见的误解，并提供了进行这些分析的简单框架。