THE MANAGER AS AN INTUITIVE STATISTICIAN

(Forthcoming in Journal of Marketing Behavior)

Bart de Langhe

Bart de Langhe (bart.delanghe@colorado.edu) is an assistant professor of marketing at the Leeds School of Business, University of Colorado, 419 UCB, Boulder, CO 80309. The author acknowledges the helpful input of the editor and two anonymous reviewers, and thanks Stefano Puntoni, Nick Reinholtz, and Stijn van Osselaer for comments on an earlier draft.

Abstract (max. 150 words; now 118)

Business decisions are increasingly based on data and statistical analyses. Managerial intuition plays an important role at various stages of the analytics process. It is thus important to understand how managers intuitively think about data and statistics. This article reviews a wide range of empirical results from almost a century of research on intuitive statistics. The results support four key insights: (1) Variance is not intuitive, (2) Perfect correlation is the intuitive reference point, (3) People conflate correlation with slope; and (4) Nonlinear functions and interaction effects are not intuitive. These insights have implications for the development, implementation, and evaluation of statistical models in marketing and beyond. I provide several such examples and offer suggestions for future research.

The ability to make accurate inferences from customer and market data is an increasingly important driver of firm performance (Brynjolfsson, Hitt, and Kim 2011; Germann, Lilien, and Rangaswamy 2013). Statistical models can help managers predict customer lifetime value based on transaction data (Gupta et al. 2006), understand how sales vary as a function of advertising expenditures (Assmus, Farley, and Lehmann 1984), or forecast the adoption of new products (Mahajan, Muller, and Bass 1990).

Managerial intuition plays an important role at various stages of the analytics process. For instance, variable selection and model specification depend on managerial intuitions about empirical relationships in the marketplace (Einhorn and Hogarth 1982), forecasting accuracy can be improved by averaging model-based forecasts with managerial judgment (Blattberg and Hoch 1990), and good models may fail to improve firm performance if managers reject them (Dietvorst, Simmons, and Massey 2015). A profound understanding of managers' statistical intuitions is thus necessary for marketing analytics initiatives to reach their full potential (Bijmolt et al. 2010).

Half a century ago, Peterson and Beach (1967) summarized three decades of research on human statistical reasoning in an influential article titled "Man as an Intuitive Statistician." The research stream on intuitive statistics is now much more mature and novel insights have emerged. My goal in writing this article is *not* to provide a comprehensive review of all research on intuitive statistical reasoning that was published since then. I will omit, for instance, the large and influential literature on cognitive heuristics (e.g., availability, representativeness, take-the-best, recognition) and the ongoing debate regarding whether these heuristics yield statistical inferences that are inappropriate and normatively irrational versus adaptive and ecologically rational (Gigerenzer and Gaissmaier 2011; Kahneman 2003). Many of these heuristics have received a lot of attention in the academic and popular press. Instead, my goal is to identify a select number of insights that have garnered relatively less attention, although they are well-supported empirically and relevant for understanding managerial decisions in today's data-driven world. Specifically, I'll review research that supports four major ideas: (1) Variance is not intuitive, (2) Perfect correlation is the intuitive reference point, (3) People conflate correlation with slope; and (4) Nonlinear functions and interaction effects are not intuitive.

I. VARIANCE IS NOT INTUITIVE

Absolute versus Squared Deviations

Several statistical techniques commonly used in marketing analytics involve squared deviations. For instance, the parameters of a linear regression model are typically estimated by minimizing the sum of squared errors of prediction and *k*-means clustering determines the location of centroids by minimizing their squared (Euclidean) distances from observations. From a statistical point of view, there are various reasons to use squared instead of absolute deviations (Judd, McClelland, and Ryan 2009). Minimizing squared deviations yields an estimate of the mean while minimizing absolute deviations yields an estimate of the median.[1] When variables are normally distributed, the sampling distribution for the mean has a lower variance than the sampling distribution for the median and thus fewer observations are required to precisely estimate the mean than the median. Moreover, squared deviations have mathematical properties

[1] Consider for instance that the numbers 1, 3, 5, 9, and 14 have a mean of 6.4 and a median of 5. The sum of squared deviations from the mean is lower than the sum of squared deviations from the median (107.2 < 117), but the sum of absolute deviations from the mean is higher than the sum of absolute deviations from the median (20.4 > 19). See Judd et al. (2009) for a formal proof.

that are simply more convenient in mathematical proofs. However, studies using a number of different paradigms suggest that people find absolute deviations more natural than squared deviations.

Judgments of Variability. In one paradigm, participants directly judge the variance of a series of numbers. Beach and Scopp (1968) presented participants with samples of numbers from two different populations and asked them to judge which population had the larger variance. The use of squared deviations in statistics implies that large deviations are weighted more than small deviations. However, people weighted small deviations slightly more than large deviations when they intuitively compared variances. Goldstein and Taleb (2007) asked Ivy League graduate students preparing for a career in financial engineering and professional portfolio managers to estimate the standard deviation of returns for a stock. Although the standard deviation is a common measure of risk in finance, the modal response corresponded to the mean absolute deviation. Thus, financially- and mathematically-savvy decision makers that deal with variance on a daily basis are inclined to use absolute instead of squared errors.

Judgments of Correlations. In another paradigm, participants estimate correlations for bivariate datasets. If people are sensitive to absolute errors, outliers should affect subjective estimates less than objective correlations. This is because outliers have large squared errors which affect objective correlations disproportionally. Erlick and Mills (1967) presented participants with bivariate data in scatterplots. They manipulated whether a scatterplot featured a few large deviations from the theoretical regression line (i.e., outliers) versus many small deviations. This manipulation held constant the sum of squared deviations (and thus the Pearson correlation), but the mean absolute deviation was smaller for scatterplots with few large deviations than for scatterplots with many small deviations. Participants thought the relationship depicted in the graphs was stronger when deviations were few and large than when deviations were many and small. Thus, correlation estimates correspond more closely to absolute deviations than to squared deviations. Several other studies support this finding (Bobko and Karren 1979; Meyer, Taieb, and Flasher 1997).

Point Predictions. In a final paradigm, participants make predictions with the goal of minimizing a given loss function. Peterson and Miller (1964) asked participants to predict the driving speeds in miles per hour for 400 hypothetical cars passing a stop sign. Participants received feedback about the actual driving speed after each prediction. The driving speeds were randomly sampled from a J-shaped distribution and prediction errors were penalized according to three different schedules. The cost for making a prediction error was (1) $2 for each error, irrespective of the magnitude of the error, (2) 20 cents times the absolute amount of prediction error in miles per hour, or (3) one cent times the square of the prediction error. The optimal prediction strategy under the first schedule was thus to predict the mode of the distribution, under the second schedule to predict the median, and under the third schedule to predict the mean. Although participants were able to minimize costs when the optimal prediction was the mode or the median, they were unable to minimize costs when the optimal prediction was the mean. Instead of minimizing squared errors and making predictions close to the mean, participants minimized absolute errors and made predictions close to the median. Evidently, it is more intuitive for people to minimize loss functions based on the count of errors or absolute errors than it is to minimize a loss function based on squared errors. Also when the distribution is normal instead of skewed, people fail to discriminate between loss functions based on absolute versus squared errors (Winkler 1970).

Forecasters in the field of marketing and sales often use asymmetric loss functions that penalize underestimation more than overestimation or overestimation more than underestimation (Franses, Legerstee, and Paap 2011; Lawrence, O'Connor, and Edmundson 2000). For example, if overestimating sales results in excessive stocks that can easily be cleared in future time periods while underestimating sales leads to delays in delivery and customer dissatisfaction, it makes sense to penalize underestimation more than overestimation (Goodwin 1996). In a lab experiment, Lawrence and O'Connor (2005) found that people can intuitively learn to minimize asymmetric absolute loss functions that favor under- or overestimation of the target outcome. Thus, people's inability to minimize squared prediction errors does not appear to reflect a general inability to minimize more complex loss functions. Rather, it seems specifically related to the nonlinearity of the loss function based on squared errors. I elaborate on the intuitive complexity of nonlinearity in Section 4 below.

It is interesting to note that, even though people have difficulty making predictions that minimize squared error (which is equivalent to predicting the mean), they do learn about the mean and they are able to reproduce it when asked about it directly. Mean estimates tend to be highly accurate, regardless of whether the distribution of numbers is symmetric or skewed, although judgments are biased by sample size and accuracy tends to decrease as the variance of the numbers increases (Beach and Swenson 1966; Duffy et al. 2010; Laestadius 1970; Malmi and Samson 1983; Smith and Price 2010; Spencer 1961; Wolfe 1975). Evidently, people have fairly good intuitions about the mean, but they fail to associate it with a loss function based on squared deviations.

Summary. In sum, there is a mismatch between the intuitive and formal definitions of variance and this may harm the communication between managers and analysts. For instance, managers may intuitively evaluate regression models that minimize squared errors as less accurate than regression models that minimize absolute errors. On the other hand, judgmental forecasts following instructions to minimize squared errors may in fact minimize absolute errors and could be seen as useless by analysts. Miscommunications are more likely when squared deviations differ more from absolute deviations, for instance when variables are heavy-tailed or skewed.

There are at least two ways to resolve the discrepancy between intuitive and formal statistics. First, it may be possible to adapt formal models such that they are based on minimizing absolute instead of squared deviations. In fact, the use of squared deviations in formal statistics is socially constructed and a product of history (Gorard 2005). Squares may be more convenient to work with when solving equations analytically, but absolute deviations may be more meaningful from a business perspective. Advances in computing power have made other approaches that do not rely on squared deviations more accessible. Second, it may be possible to teach managers heuristics that allow approximating the formal definition of variance based on other measures of dispersion that have more intuitive appeal. For instance, it seems easier to estimate the range of a variable and it can be shown that dividing a variable's range by four provides a decent approximation of its standard deviation (Browne 2001). Another welcome advantage of this range-based heuristic is that it allows people to estimate variance even when information is sparse. Consider a store manager who only knows that in the past year daily sales were minimally $1,000 and maximally $10,000.

The Mean-Dependence of Variance Perception

Research suggests that the mean is the most salient moment of a distribution. Consider for instance a consumer contemplating which of two products is of higher quality based on observing the distributions of user ratings for both products. From a statistical point of view, the consumer's confidence about relative quality should be based on the difference in mean user ratings, the variances of the distributions of ratings, and the sample sizes. However, consumers tend to give a lot of weight to the average user ratings of the products, and they give little or no weight to the variability of the distribution of ratings and sample size (de Langhe, Fernbach, and Lichtenstein 2016; Obrecht, Chapman, and Gelman 2007). The dominance of central tendency over dispersion is also reflected in customer satisfaction reports. These tend to emphasize the mean or other measures of central tendency, although the variance of customer ratings might be more important for successful brand management (Luo, Wiles, and Raithel 2013).

From a statistical point of view, the mean and variance of a continuous variable can vary orthogonally. Adding a constant to a continuous variable increases the mean but does not change the variance. However, mean and variance are intuitively related. Reinholtz (2015) presented participants with severable durable goods with different mean prices and asked them to estimate the minimum and maximum prices that other consumers paid for these goods. All participants expected higher price dispersion for higher mean prices. It may be ecologically rational to expect a higher variance for a higher mean because mean and variance are indeed positively correlated in many domains. For instance, price dispersion is higher for more expensive product categories (Grewal and Marmorstein 1994), the variance in population density is positively related to the mean population density across space and time (Taylor 1961), and the spread of response time distributions increases with the mean (Wagenmakers and Brown 2007).

People also judge the same variance as smaller when the mean is high than when the mean is low, perhaps because they expect higher variance for higher means and evaluate observed variance relative to expected variance. Hofstätter (1939) found that judgments about the variability of stick lengths decreased as mean stick lengths increased. Judgments were thus inconsistent with the statistical definition of variance. Instead, they were consistent with the coefficient of variation, which is the standard deviation divided by the mean. Several other studies using similar paradigms support this result (Beach and Scopp 1968; Lathrop 1967).

The studies above suggest that people's intuition about the variance of a distribution is biased by their perception of the mean. More recent studies started examining downstream consequences of this mean-dependence of variance perception for decision making under risk and consumer search (Reinholtz 2015; Weber, Shafir, and Blais 2004).

According to normative utility models, people should maximize expected returns while minimizing risk, with risk defined as outcome variance (Levy and Markowitz 1979; Sharpe 1964). However, variance per unit of return (that is, the coefficient of variation) tends to be a better predictor of risk sensitivity in both humans and animals (Weber et al. 2004). According to normative models of consumer search, a consumer's decision to search for better prices should be based on the variance of previously observed prices (Stigler 1961). If the variance of previously seen prices is low, it is unlikely that continued search will reveal a much lower price, and thus the economic value of further search is limited. Instead, if the variance of previously seen prices is high, it is more likely that continued search will reveal a much lower price, and thus the economic value of further search is higher. However, Reinholtz (2015) showed that consumers learn about price dispersion very slowly and inaccurately, and that consumers' decisions to search for better prices are based mostly on the mean of previously seen prices, instead of the variance.

More research needs to examine implications for managerial judgments and decisions. A manager's decision to search for a better model should depend on the variance in the outcome that is not accounted for by the model. However, people's standards for prediction accuracy might depend on the mean of the outcome such that they accept greater prediction errors for higher means (Beach and Solak 1969; Laestadius 1970). It is easy to see how the mean-dependence of variance perception can lead to suboptimal judgments and decisions. Imagine for instance that a company currently has 10,000 consumers visiting its website each day. Based on statistical analyses the company decides to adapt its marketing mix. According to the model, this should lead to 12,000 daily visitors, but the actual number of visitors turns out to be only 11,000. Now suppose that instead of predicting the overall number of visitors (i.e., 12,000), the model is set up such that it predicts the change in number of visitors (i.e., 2,000). Now the change in the marketing mix leads to only 1,000 new visitors, which feels like a greater deviation from the prediction than the difference between 11,000 and 12,000. However, regardless of whether the model predicts the overall number of visitors or the change in number of visitors, model performance in terms of absolute error (1,000 visitors) or squared error (1,000,000 visitors^2) remains the same. As this example illustrates, the mean absolute error and mean squared error of prediction remain the same regardless of whether a constant is added to or subtracted from the outcome variable of interest. However, if people evaluate variance relative to the mean, they may perceive the prediction error as smaller if the model predicts the overall number of visitors instead of the change in number of visitors.

If judgments about unexplained variance are systematically biased in this way, judgments about correlation may be biased as well. In terms of linear regression, the correlation coefficient depends on the regression slope, the variance of the predictor variable, and the variance in the outcome variable that is not explained by the predictor variable. These components are not affected by whether a constant is added or subtracted from the outcome variable. However, if people perceive unexplained variance relative to the mean level of the outcome variable, they may perceive the correlation between a predictor variable and an outcome variable as stronger when a constant is added to the outcome variable. But people may also evaluate slope relative to the mean level of the outcome variable. The slope may be perceived as less steep when a constant is added to the outcome variable which may in turn have a negative effect on perceived correlation. Although perceived error variance has a stronger influence on correlation judgments that perceived slope (Lane et al. 1985), this may attenuate the positive effect of adding a constant on correlation judgments. Future research should examine the extent to which adding constants affects perceptions of error variance, slope, and correlation.

In sum, although mean and variance can vary orthogonally from a statistical point of view, the mean is a major driver of variance perceptions. More research is needed to develop a deeper understanding of how the mean affects perceptions of variance and correlation, and ultimately variable selection in model development and the intuitive evaluation of model performance.

Variance of Linear Combinations

Mean and variance are central to financial portfolio analysis (Markowitz 1952). Recently, marketing researchers started applying financial portfolio theory to customer portfolio management (Ryals 2003; Tarasi et al. 2011). An individual customer is more attractive when his or her historical mean return is higher and the variability of his or her returns over time is lower. Similarly, portfolios of customers with higher mean returns and lower variance are more attractive. The goal of customer portfolio management is to allocate resources across customers such that the mean return is maximized for a given variance of returns, or equivalently, the variance of returns is minimized for a given mean return. In general, diversifying investments over more customers is desirable because it decreases the variability of returns without compromising mean returns. Understanding the benefit of diversification requires an understanding of linear combinations of random variables (here, individual customer returns).

Calculating the mean return of a portfolio is fairly straightforward. It is simply the weighted average of the expected mean returns of its components. For instance, if the mean annual return is 10% for customer A and 20% for customer B, the mean annual return of a portfolio that invests 50% in customer A and 50% in customer B is 15%. Early research in a non-investment context suggests that people have reasonable intuitions about the mean of linear combinations of variables. Levin (1974) presented participants with two random samples of IQ scores from students in a school and manipulated the size of the samples. Participants' estimates of the mean IQ score in the school were well described by a weighted average model where weight varies as a function of sample size. However, more recent research in a stock-investing context suggests that people's intuitions about the mean are biased (Reinholtz, Fernbach, and de Langhe 2015). Suppose two stocks have expected annual returns of 10% and 20% respectively. A person who invests everything in one randomly selected company can expect the same return as a person who splits her investment evenly across both companies (i.e., 15%). However, many people believe that a portfolio with more stocks is associated with higher returns than a portfolio with fewer stocks. This belief is more common among people high in financial literacy. These people know that diversification conveys a benefit, but mistakenly associate the benefit with the central tendency rather than the dispersion of the outcome distribution.

Calculating the variance of portfolio returns is more complex. It is based on the variances of the components and on the correlations between the components. For instance, if the variance is 4% for customer A and 100% for customer B, and the correlation between the returns of the two customers is 0.75, then the expected variance of a portfolio that invests 50% in customer A and 50% in customer B is 26.01%. This is lower than a portfolio that invests everything in one randomly selected customer. However, in the context of stock investing, many people believe that a portfolio with more stocks is associated with higher variance than a portfolio with fewer stocks. This belief is more common among people low in financial literacy (Reinholtz et al. 2015). Given that the accuracy of judgments about the mean and variance of linear combinations appears to be context-dependent, future research should examine managerial intuitions about mean and variance in the context of product and customer portfolio management.

II. PERFECT CORRELATION IS THE INTUITIVE REFERENCE POINT

Consumer behavior is determined by the interrelated influences of a multitude of variables (such as individual differences, price, product attributes, brand availability and image, the availability and image of competing brands, group influences, etc.). Even if consumer behavior were deterministic in theory, it is stochastic in practice because these variables cannot be perfectly measured or controlled (Bass 1974). Noise puts an upper limit on the predictive accuracy of any statistical model related to consumers and markets (Black 1986). Marketers should thus expect and appreciate fairly weak correlations between explanatory variables and outcomes of interest. For instance, personality and socioeconomic variables can only account for a very small fraction of the variance in consumer behavior, but the potential increase in explanatory power is nevertheless valuable to marketers (Kassarjian 1971).

Unlike variances, which do not have an upper bound, correlations range between -1 and 1. Variables with a clear lower and upper bound tend to be more evaluable and exert a larger influence on judgments and decisions (Slovic et al. 2004). One might therefore expect that people have good intuitions about correlations. However, many studies suggest that people are insufficiently sensitive to weak correlations (Baumgartner 1995; Beach and Scopp 1966; Bobko and Karren 1979; Boynton 2000; Boynton, Smith, and Stubbs 1997; Cleveland, Diaconis, and McGill 1982; Jennings, Amabile, and Ross 1982; Meyer et al. 1997; Erlick 1966; Pollack 1960). When people intuitively evaluate correlations, they take perfect correlation ($|r| = 1$) as the reference point and show decreasing sensitivity to correlations further away from that reference point. People thus discriminate better between strong correlations than between weak correlations. Moreover, people systematically underestimate objective correlations such that judgments for negative correlations are not sufficiently negative and judgments for positive correlations are not sufficiently positive. Estimates for correlations between -0.5 and +0.5 are virtually zero. This pattern of results is supported by studies using a variety of data presentation formats (e.g., data presented sequentially, in a table, or in a scatterplot).

Thus, a dissociation exists between business reality, in which recognizing weak correlations can yield large gains, and the manager's phenomenology, in which even relatively large correlations are likely to go unnoticed. This dissociation may be compounded by managerial distrust of statistical models. Recent research suggests that people are more sensitive to deviations from perfect correlation when predictions are based on an algorithm than when they come from a human forecaster (Dietvorst, Simmons, and Massey 2015). In one study, participants were presented with the actual rank of individual U.S. states in terms of the number of airline passengers that departed from that state and with forecasted ranks based on a statistical model. Although the correlation between actual and forecasted ranks was .92, most participants decided to reject the model's forecasts and rely on their own intuitive forecasts instead, even though intuitive forecasts were less accurate than model-based forecasts. Interestingly, people have a strong tendency to reject good models when they are rewarded for making accurate predictions (Arkes, Dawes, and Christensen 1986).

More research is needed to understand why people are so intolerant of a model's prediction errors. Perhaps statistical models create unreasonable expectations about the environment's predictability. Hogarth and Soyer (2011) presented graduate students in economics with regression output (parameter estimates and their standard errors, statistical significance, R-squared) and then asked questions about the distribution of the outcome at different levels of the predictor variable. Participants believed that outcomes were much more predictable than they actually were. They incorrectly inferred unexplained variation in the outcome from the standard errors of the parameter estimates and failed to take into account the R-squared statistic. Academic economists at leading universities worldwide made identical mistakes (Soyer and Hogarth 2012). Similarly, managers may have unrealistic expectations about the prediction accuracy of marketing models and this may lead to disappointment and frustration.

It is also not clear why people tend to underestimate objective correlations. Psychophysics does not offer an explanation. Theoretically, it is possible to observe decreasing marginal sensitivity but no underestimation of objective correlations. One potential explanation is that people have a tendency to mentally discretize continuous variables in an attempt to simplify and structure their environment. For instance, clinical psychologists discretize continuous measures of psychopathology when diagnosing patients and marketing researchers frequently perform median splits. Correlations computed over dichotomized variables are lower than correlations computed over continuous variables (MacCallum et al. 2002). Discretizing continuous measures of psychopathology thus hurts reliability and validity (Markon, Chmielewski, and Miller 2011) and median splits substantially reduce statistical power (Fitzsimons 2008; Irwin and McClelland 2003). Assessing correlations based on dichotomized variables may feel more natural to people than assessing correlations based on continuous variables because it relies on frequencies instead of squared errors. As I discussed in Section 1 above, squared errors are not intuitive, but a rich literature in psychology and consumer behavior suggests that frequencies are. Hasher and Zacks (1984) argued that a few fundamental aspects of an experience are automatically coded in memory and that frequency of occurrence is one such aspect. In line with this, Alba and Marmorstein (1987) found that the mere number of positive and negative attributes is very salient to consumers and a major driver of choice. A potentially interesting avenue for future research is thus to examine how managerial tendencies to discretize-and-count affect the detection and interpretation of statistical correlations.

III. PEOPLE CONFLATE CORRELATION WITH SLOPE

Correlation versus Slope in Linear Regression

If a predictor variable x is linearly related to an outcome variable y, then the relation between x and y can be represented as $y = a + bx + e$, where a is the intercept, b is the slope of the functional relationship, and e is a random error term. The squared correlation between x and y indicates the proportion of variability in the outcome that can be attributed to variability in the predictor. It is determined by the slope together with the variance of the predictor (S^2_x) and the variance in the outcome that is not accounted for by the predictor (S^2_e): $r^2 = b^2S^2_x / (b^2S^2_x + S^2_e)$. The denominator in this equation, $b^2S^2_x + S^2_e$, is equal to the total variability in the outcome, while the numerator, $b^2S^2_x$, is equal to the variability in the outcome that can be attributed to variability in the predictor. Thus, while slope and correlation are related, correlation can vary holding constant slope, and *vice versa*, slope can vary holding constant correlation.

Knowledge about correlation is important to evaluate the relative importance of predictor variables. Knowledge about slope is important to predict outcome values given predictor values; that is, to form conditional expectations. However, people tend to conflate correlation with slope. As a result, assessments of relative importance are biased by slope and conditional expectations are biased by correlation.

The Relative Importance of Predictors: Correlation versus Slope

The extent to which a predictor x is good or important for predicting an outcome y is measured by the Pearson correlation coefficient (r). A predictor that correlates more strongly with an outcome accounts for more variation in the outcome, and is thus more important. The slope of the functional relationship is just one determinant of correlation. A predictor variable with a steeper slope is not necessarily more important. Consider for instance that the slope of the function relating unit sales to prices expressed in British Pounds is much steeper than the slope of the function relating unit sales to prices expressed in US Dollars (because 1 British Pound equals more than one US Dollar). That is, a one-unit increase in British Pounds implies a much larger decrease in unit sales compared to a one-unit increase in US Dollars. This does not imply that British Pounds are more important than US Dollars for predicting sales.

However, people intuitively understand the importance of a predictor in terms of the slope of the function (instead of the correlation). Brehmer and Qvarnstrom (1976) asked participants to predict an outcome based on two predictors and instructed participants to give one predictor twice the weight of the other predictor. To examine how participants interpreted this instruction, they derived slopes and correlations from participants' predictions. The ratio of the slopes for both predictors was equal to two, suggesting that people's intuitive interpretation of relative importance corresponds to the slope of the function that relates the predictors to the outcome. Slovic (1969) asked stockbrokers to rate the growth potential of stocks on the basis of several predictors and to subjectively rank the importance of these predictors. Subjective impressions of relative importance corresponded more closely to the slopes inferred from their predictions than to the correlations inferred from their predictions. Several other studies suggest that people intuitively equate importance with slope (Brehmer, Hagafors, and Johansson 1980; Summers, Taliaferro, and Fletcher 1970). Managers are oftentimes asked to evaluate the importance of predictors when deliberating which variables to include in marketing research or statistical analyses. The findings above suggest that communications between managers and analysts about the relative importance of predictors can easily go wrong (Goldstein and Beattie 1991).

Conditional Expectations: Slope versus Correlation

The conditional expectation of *y* given *x* depends on the slope of the function together with the intercept. Holding constant intercept and slope, the extremity of conditional expectations should not vary as a function of the correlation. For instance, the intercept and slope of the functions relating price to quality ratings for restaurants are the same in Figures 1A and 1B. However, because quality ratings are less noisy for lower-priced restaurants, the correlation between price and quality is higher in Figure 1B than in Figure 1A. If people form conditional expectations based on intercept and slope (as they should do from a statistical point of view), the predicted level of quality for a high-priced restaurant should be the same in both situations. However, people tend to predict higher quality for high-priced restaurants in the second situation, suggesting that conditional expectations are driven (at least partly) by correlation (De Langhe et al. 2014).

[Insert Figure 1 here]

Variables are continuous in the restaurant example above, but research using dichotomous variables finds a similar conflation between conditional expectation and correlation. When people judge the probability that an outcome is present given that the predictor is present, their answers reflect not only the normative representation of conditional probability, but also cue-outcome contingency, or correlation. Consider a situation where two managers are exposed to the same probability that an online advertising campaign on Facebook is successful (say, .50) but different probabilities of success when the same campaign is done through another channel (say, .50 vs. .10). See Table 1. Although the correlation between Facebook (yes vs. no) and success (yes vs. no) for the first manager is zero and the correlation between Facebook and success for the second manager is positive, the conditional probability of success given that the campaign was done on Facebook is the same for both managers (i.e, .50), and thus both managers should provide equal conditional probability judgments of success for campaigns launched on Facebook. However, research suggests that the second manager who learned that the correlation between Facebook and success is positive will judge the conditional probability as higher than the first manager who learned that there is no correlation between Facebook and success (Mitchell et al. 2013; Price and Yates 1995; Lagnado and Shanks 2002). Conditional probability judgments are thus biased by correlations.

[Insert Table 1 here]

Although the correlation between a cue and an outcome is symmetric, this is not the case for conditional probabilities. Consider for instance the comparison between model forecasts and actual customer churn in Table 2. If the model predicts that a customer will churn, there is a 90% chance that the customer will actually churn. This does not imply that 90% of churning customers are correctly classified by the model. In fact, only 50% of churning customers are correctly classified by the model. More generally, the probability of A given B (here, the probability that a customer churns given that the model classifies the customer as churning) is different from the probability of B given A (here, the probability that the model classifies a customer as churning given that the customer actually churns). However, people tend to equate a given conditional probability with its inverse (Koehler 1996; Villejoubert and Mandel 2002; Wolfe 1995). In other words, they inappropriately extrapolate the property of symmetry associated with correlations to conditional probabilities.

[Insert Table 2 here]

IV. NONLINEAR FUNCTIONS AND CONFIGURAL EFFECTS ARE NOT INTUITIVE

Functional relationships in marketing are not only noisy, predictors often have nonlinear and configural effects on outcomes (Green 1973; Rust and Bornman 1982). Nonlinearity refers to the relationship between an individual predictor and an outcome. For instance, the relationship between advertising expenditures and sales can be concave such that there are decreasing returns to advertising (Lambin 1969), the relationship between customer satisfaction and willingness to pay can be inverse S-shaped such that willingness to pay increases steeply at low and high levels of satisfaction but is virtually flat at intermediate levels of satisfaction (Homburg, Koschate, and Hoyer 2005), and the relationship between age and coffee consumption can be inversely U-shaped such that coffee consumption peaks at about 50 years of age (Wheatley, Chiu, and Stevens 1980). Nonlinearity can be modeled by applying nonlinear transformations to predictors before they are combined with other predictors (e.g., a logarithmic or quadratic transformation).

Configurality refers to the relationship between two or more predictors and an outcome. The simplest case is when the effect of a predictor on an outcome depends on the level of another predictor. For instance, the negative effect of price on sales may be stronger when advertising expenditures are higher (Eskin and Baron 1977), the relationship between customer satisfaction and repurchase likelihood may depend on gender or age (Mittal and Kamakura 2001), or the effectiveness of celebrity endorsement may depend on customers' involvement with a product (Petty, Cacioppo, and Schumann 1983). Configurality can be modeled by including interaction effects between predictors in the model, but marketing researchers often have wrong intuitions about how to interpret interaction effects in formal regression analyses (Irwin and McClelland 2001; Spiller et al. 2013).

Psychologists have used function-learning paradigms to compare people's ability to learn linear versus nonlinear effects and additive versus configural effects. Participants in such studies are asked to predict an outcome that is related to one or more predictors through a function, most often with a random component. Participants receive feedback about the actual outcome values after each prediction which allows them to learn about the function and make more accurate predictions over time.

Deane, Hammond, and Summers (1972) found that people made more accurate predictions when the functional relationship between the predictor and the outcome was positive linear than when it was quadratic inversely U-shaped. Brehmer (1974) added a negative linear and a quadratic U-shaped condition and found that prediction accuracy was highest when the functional relationship was positive linear, followed by negative linear, quadratic inversely U-shaped, and finally quadratic U-shaped. He also asked people to rate how frequently they encountered these functions and how difficult they found it to come up with real-life examples for them. The positive linear function was rated as most frequent and easiest, followed by the negative linear function, the quadratic inversely U-shaped function, and finally the quadratic U-shaped function. DeLosh et al. (1997) found higher prediction accuracy for a positive linear function than for a quadratic inversely U-shaped function and that prediction accuracy for a concave exponential function was in between these two. The intuitive primacy of linear functions is supported by many other studies (Brehmer 1971; Brehmer, Kuylenstierna, and Liljergren 1974; Brehmer and Qvarnstrom 1976; Hammond and Summers 1965; Mellers 1980; Rothstein 1986; Sawyer 1991; Sheets and Miller 1974; Steinmann 1976).

People even have difficulty dealing with nonlinear relationships that are deterministic. They tend to linearize deterministic decreasing functions, such as the relationship between miles per gallon and gas savings (Larrick and Soll 2008) or the relationship between Megabits per second and download time (de Langhe and Puntoni 2016). Similarly, they linearize deterministic increasing functions, like the growth of populations over time (Wagenaar and Sagaria 1975) or the growth of savings over time (McKenzie and Liersch 2011).

Studies also show that configural effects are less intuitive than additive effects. Mellers (1980) asked participants to learn a function with additive effects of predictors on an outcome: $y = 2[(x_1 - 6) + (x_2 - 6)] + 20 + e$. Other participants learned a function where predictors had an interaction effect on the outcome: $y = (x_1 - 6)(x_2 - 6) + 20 + e$. Learning was much quicker in the condition with additive effects than in the condition with interaction effects. Several other studies using similar paradigms support this result (Brehmer 1969; Brehmer and Qvarnstrom 1976; de Langhe, van Osselaer, and Wierenga 2011; Juslin, Karlsson, and Olsson 2007; Olsson, Enkvist, and Juslin 2006; Pachur and Olsson 2012; Soyer and Hogarth 2015). De Langhe et al. (2011) examined how outcome versus process accountability affects learning of additive versus configural functions. Participants that were held outcome accountable were evaluated solely based on the accuracy of their judgments, regardless of whether they came to their judgment based on solid understanding and analysis or not. Participants that were held process accountable were evaluated based on whether they could provide a sound justification for the process that underlay those judgments, regardless of the accuracy of their judgments. Process accountability enhanced learning of additive effects, but not configural effects.

Although people do eventually learn to accurately predict outcomes for most types of nonlinearity and configurality, they do not seem to actually abstract the underlying function (Juslin, Karlsson, and Olsson 2007; Olsson, Enkvist, and Juslin 2006). Instead, people store previously seen pairs of predictor and outcome values in long term memory. These memory traces allow accurate predictions for previously seen predictor values but not for new predictor values. In other words, people fail to extrapolate nonlinear and configural functions to novel predictor values. For linear and additive functions, people do learn the underlying function and this allows accurate predictions even in extrapolation tests.

The errors made by naïve participants in laboratory tasks like the ones reviewed above are also common among professionals (Wagenaar and Sagaria 1975) and performance in function learning tasks predicts learning of complex skills in natural settings (e.g., pilot performance; Matton, Raufaste, and Vautier 2013). Farrell, Luft, and Shield (2007) presented MBA's with data about quarterly spending on employee training and the financial performance for various plants within a firm. Participants were familiar with the notion of decreasing returns but nevertheless tended to linearize the quadratic functional relationship between spending and performance, consistent with the results above. It is likely that managerial intuitions are incorrect in simpler deterministic scenarios as well. Consider for instance the relationship between annual customer retention rates and customer lifetime duration (and thus customer lifetime value). It is compelling to believe that increasing the customer retention rate from 20% to 60% has a larger effect on relationship duration than increasing the retention rate further from 60% to 80%. In fact, the first increase extends the average duration of customer relationships by 1.25 years, while the second increase extends it by 2.5 years. Many people find this counterintuitive and believe that larger absolute differences (40 vs. 20) or larger proportional differences (3 vs. 1.33) in the predictor variable imply larger differences in the outcome variable (de Langhe and Puntoni 2016).

The intuitive complexity of nonlinearity and configurality supports calls for simpler models in analytics. Brighton and Gigerenzer (2015) argue that simple marketing models often make better predictions than more complex models because simple specifications generalize better to different samples. For instance, Wübben and Wangenheim (2008) showed that a simple recency-of-last purchase analysis often discriminates better between active and inactive customers than more sophisticated stochastic customer base analysis models, such as the Pareto/NBD (Schmittlein, Morrison, and Columbo 1987) and BG/NBD models (Fader, Hardie, and Lee 2005). Simple models do not always perform better and in those cases the analyst needs to make a tradeoff between accuracy and simplicity. Given the difficulty people have to understand nonlinearity and configurality, the analyst should expect a high burden of proof or risk managerial rejection.

CONCLUDING REMARKS

Marketing managers are large consumers of data, but remarkably little research has examined their statistical intuitions. This article reviewed a broad set of empirical studies on intuitive statistical reasoning that support four key insights: (1) Variance is not intuitive; (2)

Perfect correlation is the intuitive reference point; (3) People conflate correlation with slope; and (4) Nonlinear functions and interaction effects are not intuitive. Although these insights are mostly based on studies using students as participants, they are clearly relevant in a variety of managerial settings, as a few examples illustrate throughout the paper.

This article just scratched the surface. Research on managerial intuitions about statistics is rife with opportunities, and as the use of data in business accelerates, I expect it will become an influential area of academic research. In particular, we need more research along (at least) two lines. First, we need better tools to measure individual differences in the understanding of statistics. While several measures of statistical literacy exist, most scales focus on people's understanding of probability (Lipkus, Samsa, and Rimer 2001; Peters et al. 2006). But statistical literacy involves a much broader range of skills (Rothman et al. 2009). Future research should examine which skills are most important in marketing and develop new tools to measure these skills. Second, research should explore ways to alleviate biases and improve judgment quality. For instance, data visualization is increasingly hailed as a tool to improve data-driven decision-making in business. Procter & Gamble's managers now meet in "business spheres" that immerse the team in data visualizations (Davenport 2013). Tableau, a leading data visualization software company, envisions future business meetings where managers work together with their teams, creating visuals on the fly, as they try to understand their data (Clancy 2014). While data visualization clearly conveys some benefits, we need to remain cautious when visuals are used in data analysis and increase our understanding of the psychology used to interpret visuals.

REFERENCES

Alba, Joseph W. and Howard Marmorstein (1987), "The Effects of Frequency Knowledge On Consumer Decision Making," *Journal of Consumer Research,* 14, 14-25.

Arkes, Hal R., Robyn M. Dawes, and Caryn Christensen (1986), "Factors Influencing the Use of a Decision Rule in a Probabilistic Task," *Organizational Behavior and Human Decision Processes,* 37 (1), 93-110.

Assmus, Gert, John U. Farley, and Donald R. Lehmann (1984), "How Advertising Affects Sales: Meta-Analysis of Econometric Results," *Journal of Marketing Research*, 21 (1), 65-74.

Bass, Frank M. (1974), "The Theory of Stochastic Preference and Brand Switching," *Journal of Marketing Research*, 11 (1), 1-20.

Baumgartner, Hans (1995), "On the Utility of Consumers' Theories in Judgments of Covariation," *Journal of Consumer Research*, 21 (4), 634-643.

Beach, Lee Roy and Thomas S. Scopp (1966), "Inferences about Correlations," *Psychonomic Science*, 6 (6), 253-254.

Beach, Lee Roy and Thomas S. Scopp (1968), "Intuitive Statistical Inferences about Variances," *Organizational Behavior and Human Performance,* 3, 109-123.

Beach, Lee Roy and Frances Solak (1969), "Subjective Judgments of Acceptable Error," *Organizational Behavior and Human Performance,* 4, 242-251.

Beach, Lee Roy and Richard G. Swenson (1966), "Intuitive Estimation of Means," *Psychonomic Science,* 5 (4), 161-162.

Bijmolt, Tammo H. A., Peter S. H. Leeflang, Frank Block, Maik Eisenbeiss, Bruce G. S. Hardie, Aurelie Lemmens, and Peter Saffert (2010), "Analytics for Customer Engagement," *Journal of Service Research*, 13 (3), 341-356.

Black, Fischer (1986), "Noise," *The Journal of Finance*, 41 (3), 529-543.

Blattberg, Robert C. and Stephen J. Hoch (1990), "Database Models and Managerial Intuition: 50% Model + 50% Manager," *Management Science*, 36 (8), 887-899.

Bobko, Philip and Ronald Karren (1979), "The Perception of Pearson Product Moment Correlations from Bivariate Scatterplots," *Personnel Psychology*, 32, 313-325.

Boynton, David M. (2000), "The Psychophysics of Informal Covariation Assessment: Perceiving Relatedness Against a Background of Disperion," *Journal of Experimental Psychology: Human Perception and Performance*, 26 (3), 867-876.

Boynton, David M., Laurence D. Smith, and D. Alan Stubbs (1997), "Sensitivity and Bias in Covariation Detection: A Direct Approach to a Tangled Issue," *Organizational Behavior and Human Decision Processes*, 72 (1), 79-98.

Brehmer, Berndt (1969), "Cognitive Dependence on Additive and Configural Cue-Criterion Relations," *American Journal of Psychology*, 82 (4), 490-503.

Brehmer, Berndt (1971), "Subjects' Ability To Use Functional Rules," *Psychonomic Science*, 24 (6), 259-260.

Brehmer, Berndt (1974), "Hypotheses About Relations Between Scaled Variables in the Learning of Probabilistic Inference Tasks," *Organizational Behavior and Human Performance*, 11 (1), 1-27.

Brehmer, Berndt, Roger Hagafors, and Roger Johanson (1980), "Cognitive Skills in Judgment: Subjects' Ability to Use Information About Weights, Function Forms, and Organizing Principles," *Organizational Behavior and Human Performance,* 26 (3), 373-385.

Brehmer, Berndt, Jan Kuylenstierna, and Jan-Erik Liljergren (1974), "Effects of Function Form and Cue Validity on the Subjects' Hypotheses in Probabilistic Inference Tasks," *Organizational Behavior and Human Performance,* 11. 338-354.

Brehmer, Berndt and Göran Qvarnström (1976), "Information Integration and Subjective Weights in Multiple-Cue Judgments," 17 (1), 118-126.

Brighton, Henry and Gerd Gigerenzer (2015), " The Bias Bias," *Journal of Business research,* 68 (8), 1772-1784.

Browne, Richard H. (2001), "Using the Sample Range as a Basis for Calculating Sample Size in Power Calculations," *The American Statistician,* 55 (4), 293-298.

Brynjolfsson, Erik, Lorin M. Hitt, and Heekyung Hellen Kim (2011), "Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance?" working paper, Social Science Research Network, (accessed on June 18, 2015), [available at http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=1819486].

Clancy, Heather (2014), "Why this business intelligence CEO spends oodles on research," (accessed December 16, 2014), [available at http://fortune.com/2014/12/02/tableau-software-ceo-research/].

Cleveland, William S., Persi Diaconis, and Robert McGill (1982), "Variables on Scatterplots Look More Highly Correlated When the Scales are Increased," *Science,* 216 (4550), 1138-1141.

Davenport, Tom (2013), "How P&G Presents Data to Decision-Makers," *Harvard Business Review,* (accessed June 16, 2015), [available at https://hbr.org/2013/04/how-p-and-g-presents-data/].

Deane, Donald H., Kenneth R. Hammond, and David A. Summers (1972), "Acquisition and Application of Knowledge in Complex Inference Tasks," *Journal of Experimental Psychology*, 92 (1) 20-26.

Dietvorst, Berkeley J., Joseph P. Simmons, and Cade Massey (2015), "Algorithm Aversion: People Erroneously Avoid Algorithms after Seeing Them Err," *Journal of Experimental Psychology: General*, 144 (1), 114-126.

de Langhe, Bart, Philip M. Fernbach and Donald R. Lichtenstein (2016), "Navigating by the Stars: Investigating the Actual and Perceived Validity of Online User Ratings," *Journal of Consumer Research*, 42, 817-833.

de Langhe, Bart and Stefano Puntoni (2016), "Productivity Metrics and Consumers' Misunderstanding of Time Savings," *Journal of Marketing Research,* 53, 396-406.

de Langhe, Bart, Stijn M. J. van Osselaer, Stefano Puntoni, and Ann L. McGill (2014), "Fooled by Heteroscedastic Randomness: Local Consistency Breeds Extremity in Price-Based Quality Inferences," *Journal of Consumer Research*, 41 (4), 978-994.

de Langhe, Bart, Stijn M. J. van Osselaer, and Berend Wierenga (2011), "The Effects of Process and Outcome Accountability on Judgment Process and Performance," *Organizational Behavior and Human Decision Processes,* 115 (2), 238-252.

DeLosh, Edward L., Jerome R. Busemeyer, and Mark A. McDaniel (1997), "Extrapolation: The Sine Qua Non for Abstraction in Function Learning," *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 23 (4), 968-986.

Dietvorst, Berkeley, Joseph Simmons, and Cade Massey (2015), "Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err," *Journal of Experimental Psychology,* 144(1), 114-126.

Duffy, Sean, Huttenlocher Janellen, Hedges Larry V., and L. Elizabeth Crawford (2010), "Category Effects on Stimulus Estimation: Shifting and Skewed Frequency Distributions," *Psychonomic Bulletin & Review*, 17 (2), 224-230.

Einhorn, Hillel J. and Robin M. Hogarth (1982), "Prediction, Diagnosis, and Causal Thinking in Forecasting," *Journal of Forecasting*, 1 (1), 23-36.

Erlick, Dwight E. (1966), "Human Estimates of Statistical Relatedness," *Psychonomic Science*, 5 (10), 365-366.

Erlick, Dwight E. and Robert G. Mills (1967), "Perceptual Quantification of Conditional Dependency," *Journal of Experimental Psychology*, 73 (1), 9-14.

Eskin, Gerald J. and Penny H. Baron (1977), "Effects of Price and Advertising in Test-Market Experiments," *Journal of Marketing Research*, 14 (4), 499-508.

Fader, Peter S., Bruce G. S. Hardie, and Ka Lok Lee (2005), "Counting Your Customers the Easy Way: An Alternative to the Pareto/NBD Model," *Marketing Science*, 24 (2), 275-285.

Farrell, Anne M., Joan Luft, and Michael D. Shields (2007), "Accuracy in Judging the Nonlinear Effects of Cost and Profit Drivers," *Contemporary Accounting Research*, 24 (4), 1139-1169.

Fitzsimons, Gavan J. (2008), "Death to dichotomizing," *Journal of Consumer Research*, 35(1), 5-8.

Franses, Philip Hans, Rianne Legerstee, Richard Paap (2011), "Estimating Loss Functions of Experts," Tinbergen Institute Discussion Paper, TI2011-177/4.

Germann Frank, Gary L. Lilien, and Arvind Rangaswamy (2013), "Performance Implications of Deploying Marketing Analytics," *International Journal of Research in Marketing*, 30 (2), 114-128.

Gigerenzer, Gerda and Wolfgang Gaissmaier (2011), "Heuristic Decision Making," *Annual Review of Psychology*, 62, 451-482.

Goldstein, Daniel G. and Nassim Nicholas Taleb (2007), "We Don't Quite Know What We are Talking About When We Talk About Volatility," *Journal of Portfolio Management*, 33 (4), 84-86.

Goldstein, William M. and Jane Beattie (1991), "Judgments of Relative Importance in Decision Making: The Importance of Interpretation and the Interpretation of Importance," in *Frontiers of Mathematical Psychology: Essays in Honor of Clyde Coombs*, eds. Donald R. Brown and J. E. Keith Smith, Springer-Verlag.

Goodwin, Paul (1996), "Statistical Correction of Judgmental Point Forecasting and Decisions," *Omega*, 24 (5), 551-560.

Gorard, Stephen (2005), "Revisiting a 90-Year-Old Debate: The advantages of the Mean Deviation," *British Journal of Educational Studies*, 53 (4), 417-430.

Green, Paul E. (1973), "On the Analysis of Interactions in Marketing Research Data," *Journal of Marketing Research*, 10 (4), 410-420.

Grewal, Dhruv and Howard Marmorstein (1994), "Market Price Variation, Perceived Price Variation, and Consumers' price Search Decisions for Durable Goods," *Journal of Consumer Research*, 21 (3), 453-460.

Gupta, Sunil, Dominique Hanssens, Bruce Hardie, William Kahn, V. Kumar, Nathaniel Lin, Nalini Ravishanker, and S. Sriram (2006), "Modeling Customer Lifetime Value," *Journal of Service Research*, 9 (2), 139-155.

Hasher, Lynn and Rose T. Zacks (1984), "Automatic Processing of Fundamental Information," *American Psychologist,* 39 (12), 1372-1388.

Hammond, Kenneth R. and David Summers (1965), "Cognitive Dependence on Linear and Nonlinear Cues," *Psychological Review,* 72 (5), 215-224.

Hofstätter, Peter R. (1939), "Über die Schätzung von Grupeneigenschaften," *Zeitschrift für Psychologie,* 145, 1-44.

Hogarth, Robin M. and Emre Soyer (2011), "Sequentially Simulated Outcomes: Kind Experience Versus Nontransparent Description," *Journal of Experimental Psychology: General,* 140 (3), 434-463.

Homburg, Christian, Nicole Koschate, and Wayne D. Hoyer (2005), "Do Satisfied Customers Really Pay More? A Study of the Relationship Between Customer Satisfaction and Willingness to Pay," *Journal of Marketing,* 69 (2), 84-96.

Irwin, Julie R. and Gary H. McClelland (2001), "Misleading Heuristics and Moderated Multiple Regression Models," *Journal of Marketing Research*, 38 (1), 100-109.

Irwin, Julie R. and Gary H. McClelland (2003), "Negative consequences of dichotomizing continuous predictor variables," *Journal of Marketing Research*, 40 (3), 366-371.

Jennings, Dennis L., Teresa M. Amabile, and Lee Ross (1982), "Informal Covariation Assessment: Data-based versus Theory-based judgments," in *Judgment under Uncertainty: Heuristics and Biases,* eds. Daniel Kahneman, Paul Slovic, and Amos Tversky, Cambridge University Press, 211-230.

Judd, Charles M., Gary H. McClelland, and Carey S. Ryan (2011), *Data Analysis: A Model Comparison Approach*, Routledge.

Juslin, Peter, Linnea Karlsson, and Henrik Olsson (2007), "Information Integration in Multiple Cue Judgment: A Division of Labor Hypothesis," *Cognition,* 106, 259-298.

Kahneman, Daniel (2003), "A Perspective on Judgment and Choice," *American Psychologist*, 58 (9), 697-720.

Kassarjian (1971), "Personality and Consumer Behavior: A Review," *Journal of Marketing Research*, 8 (4), 409-418.

Koehler, Jonathan J. (1996), "The Base Rate Fallacy Reconsidered: Descriptive, Normative, and Methodological Challenges," *Behavioral and Brain Science,* 19, 1-53.

Laestadius, John E., Jr. (1970), "Tolerance for Errors in Intuitive Mean Estimations," *Organizational Behavior and Human Performance,* 5, 121-124.

Lagnado, David A. and David R. Shanks (2002), "Probability Judgment in Hierarchical Learning: A Conflict between Predictiveness and Coherence," *Cognition*, 83 (1), 81-112.

Lambin, Jean-Jacques (1969), "Measuring the Profitability of Advertising: An Empirical Study," *The Journal of Industrial Economics*, 17 (2), 86-103.

Lane, David M., Craig A. Anderson, and Kathryn L. Kellam (1985), "Judging the Relatedness of Variables: The Psychophysics of Covariation Perception," *Journal of Experimental Psychology: Human Perception and Performance,* 11 (5), 640-649.

Larrick Richard P. and Jack B. Soll (2008), "The MPG Illusion," *Science,* 320 (5883), 1593-1594.

Lathrop, Richard G. (1967), "Perceived Variability," *Journal of Experimental Psychology*, 73 (4), 498-502.

Lawrence, Michael and Marcus O'Connor (2005), "Judgmental Forecasting in the Presence of Loss Functions," *International Journal of Forecasting*, 21 (1), 3-14.

Lawrence, Michael, Marcus O'Connor, and Bob Edmundson (2000), "A Field Study of Sales Forecasting Accuracy and Processes," *European Journal of Operational Research*, 122 (1), 151-160.

Levin, Irwin P. (1974), "Averaging Processes and Intuitive Statistical Judgments," *Organizational Behavior and Human Performance,* 12 (1), 83-91.

Levy, Haim and Harry M. Markowitz (1979), "Approximating Expected Utility by a Function of Mean and Variance," *American Economic Review,* 69 (3), 308-317.

Lipkus, Isaac M., Greg Samsa, and Barbara K. Rimer (2001), "General Performance on a Numeracy Scale among Highly Educated Samples," *Medical Decision Making*, 21, 37-44.

Luo, Xueming, Michael Wiles, and Sascha Raithel (2013), "Make the Most of a Polarizing Brand," *Harvard Business Review,* November, 29-31.

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

Mahajan, Vijay, Eitan Muller, and Frank M. Bass (1990), "New Product Diffusion Models in Marketing: A Review and Directions for Research," *Journal of Marketing*, 54 (1), 1-26.

Malmi, Robert A. and David J. Samson (1983), "Intuitive Averaging of Categorized numerical Stimuli," *Journal of Verbal Learning and Verbal behavior,* 22 (5), 547-559.

Markon, Kristian E., Michael Chmielewski, and Christopher J. Miller (2011), "The reliability and validity of discrete and continuous measures of psychopathology: a quantitative review," *Psychological bulletin,* 137(5), 856-879.

Markowitz, Harry (1952), "Portfolio Selection," *Journal of Finance,* 7 (1), 77-91.

Matton, Nadine, Eric Raufaste, and Stephane Vaurier (2013), "External Validity of Individual Diffrences in Multiple Cue probability Learning: The Case of Pilot Training," *Judgment and Decision Making*, 8 (5), 589-602.

McKenzie, Craig R. M. and Michael J. Liersch (2011), "Misunderstanding Savings Growth: Implications for Retirement Savings Behavior," *Journal of Marketing Research*, 48, S1-S13.

Mellers, Barbara A. (1980), "Configurality in Multiple-Cue Probability Learning," *American Journal of Psychology*, 93 (3), 429-443.

Meyer, Joachim, Meirav Taieb, and Ittai Flascher (1997), "Correlation Estimates as Perceptual Judgments," *Journal of Experimental Psychology,* 3 (1), 3-20.

Mitchell, Chris J., Oren Griffiths, Pranjal More, and Peter F. Lovibond (2013), "Contingency Bias in Probability Judgment May Arise from Ambiguity Regarding Additional Causes," *Quarterly Journal of Experimental Psychology*, 66 (9), 1675-1686.

Mittal, Vikas and Wagner A. Kamakura (2001), "Satisfaction, Repurchase Intent, and Repurchase Behavior: Investigating the Moderating Effect of Customer Characteristics," *Journal of Marketing Research*, 38 (1), 131-142.

Obrecht, Natalie A., Gretchen Chapman, and Rochel Gelman (2007), "Intuitive *T* Tests: Lay Use of Statistical Information," *Psychonomic Bulletin & Review,* 14 (6), 1147-1152.

Olsson, Anna-Carin, Tommy Enkvist, and Peter Juslin (2006), "Go With the Flow: How to Master a Nonlinear Multiple-Cue Judgment Task," *Journal of Experimental Psychology: Learning, memory, and Cognition,* 32 (6), 1371-1384.

Pachur, Thorsten and Henrik Olsson (2012), "Type of Learning Impacts Performance and Strategy Selection in Decision Making," *Cognitive Psychology*, 65 (2), 207-240.

Peters, Ellen, Daniel Vastfjall, Paul Slovic, C. K. Mertz, Ketti Mazzocco, and Stephen Dickert (2006), “Numeracy and Decision Making,” *Psychological Science*, 17 (5), 407-413.

Peterson, Cameron R. and Lee Roy Beach (1967), “Man as an Intuitive Statistician,” *Psychological Bulletin*, 68 (1), 29-46.

Peterson, Cameron R. and Alan Miller (1964), “Mode, Median, and Mean as Optimal Strategies,” *Journal of Experimental Psychology*, 68 (4), 363-367.

Petty, Richard E., John T. Cacioppo, and David Schumann (1983), “Central and Peripheral Routes to Advertising Effectiveness: The Moderating role of Involvement,” *Journal of Consumer Research*, 10 (2), 135-146.

Pollack, Irwin (1960), “Identification of Visual Correlational Scatterplots,” *Journal of Experimental Psychology*, 59 (6), 351-360.

Price, Paul C. and Frank Yates (1995), “Associative and Rule-Based Accounts of Cue Interaction in Contingency Judgment,” *Journal of Experimental Psychology: Learning, memory, and Cognition*, 21 (6), 1639-1655.

Reinholtz, Nicholas (2015), “Persistence in Consumer Search,” Columbia University, unpublished dissertation.

Reinholtz, Nicholas, Philip M. Fernbach, and Bart de Langhe (2015), “Do People Understand the Benefit of Diversification?” working paper.

Rothman, Russell L., Victor M. Montori, Andrea Cherrington, and Michael P. Pignone (2008), “Perspective: The Role of Numeracy in Health Care,” *Journal of Health Communication*, 13 (6), 583-595.

Rothstein, Howard G. (1986), “The Effects of Time Pressure on Judgment in Multiple Cue Probability Learning,” *Organizational Behavior and Human Decision Processes*, 37 (1), 83-92.

Rust, Roland T. and Elizabeth O. Bornman (1982), “Distribution-Free Methods of Approximating Nonlinear Marketing Relationships,” *Journal of Marketing Research*, 19 (3), 372-374.

Ryals, Lynette (2003), “Making Customers Pay: Measuring and Managing Customer Risk and Returns,” *Journal of Strategic Marketing*, 11, 165-175.

Sawyer, John E. (1991), “Hypothesis Sampling, Construction, or Adjustment: How Are Inferences About Nonlinear Monotonic Contingencies Developed?,” *Organizational Behavior and Human Decision Processes*, 49 (1), 124-150.

Schmittlein, David C., Donald G. Morrison, and Richard Columbo (1987), “Counting Your Customers: Who Are They and What Will They Do Next?” *Management Science*, 33 (1), 1-24.

Sharpe, William F. (1964), “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk,” *Journal of Finance*, 19 (3), 425-442.

Sheets, Charles A. Jr. and Monroe J. Miller (1974), “The Effect of Cue-Criterion Function Form on Multi[ple-Cue Probability Learning,” *American Journal of Psychology*, 87 (4), 629-641.

Slovic, Paul (1969), “Analyzing the Expert Judge: A Descriptive Study of Stockbroker’s Decision Processes,” *Journal of Applied Psychology*, 53 (3), 255-263.

Slovic, Paul, Melissa L. Finucane, Ellen Peters, and Donald G. MacGregor (2004), “Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason, Risk, and Rationality,” *Risk Analysis*, 24 (2), 311-322.

Smith, Andrew R. and Paul C. Price (2010), "Sample Size Bias in the Estimation of Means," *Psychonomic Bulletin & Review*, 17 (4), 499-503.

Soyer, Emre and Robin M. Hogarth (2012), "The Illusion of Predictability: How Regression Statistics Mislead Experts," *International Journal of Forecasting*, 28 (3), 695-711.

Soyer, Emre and Robin M. Hogarth (2015), "Learning From Experience in Nonlinear Environments: Evidence From a Competition Scenario," *Cognitive Psychology*, 81, 48-73.

Spencer, J. (1961), "Estimating Averages," *Ergonomics*, 4, 317-328.

Spiller, Stephen A., Gavan J. Fitzsimons, John G. Lynch Jr., and Gary H. McClelland (2013), "Spotlights, Floodlights, and the Magic Number Zero: Simple Effects Tests in Moderated Regression," *Journal of Marketing Research*, 50 (2), 277-288.

Steinmann, Derick O. (1976), "The Effects of Cognitive Feedback and Task Complexity in Multiple-Cue Probability Learning," *Organizational Behavior and Human Performance*, 15, 168-179.

Stigler, George J. (1961), "The Economics of Information," *Journal of Political Economy*, 69 (3), 213-225.

Summers, David A., J. Dale Taliaferro, and Donna J. Fletcher (1970), "Subjective versus Objective Description of Judgment Policy," *Psychonomic Science*, 18 (4), 249-250.

Tarasi, Crina O., Ruth N. Bolton, Michael D. Hutt, and Beth A. Walker (2011), "Balancing Risk and Return in a Customer Portfolio," *Journal of Markting*, 75 (3), 1-17.

Taylor, L. R. (1961), "Aggregation, Variance and the Mean," *Nature*, 189, 732-735.

Villejoubert, Gaelle and David R. Mandel (2002), "The Inverse Fallacy: An Account of Deviations from Bayes's Theorem and the Additivity Principle, *Memory & Cognition*, 30 (2), 171-178.

Wagenaar, Willem A. and Sabato D. Sagaria (1975), "Misperception of Exponential Growth," *Perception & Psychophysics*, 18 (6), 416-422.

Wagenmakers, Eric-Jan and Scott Brown (2007), "On the Linear Relation Between the Mean and the Standard Deviation of a Response Time Distribution," *Psychological Review*, 114 (3), 830-841.

Weber, Elke U., Sharoni Shafir, and Anne-Renee Blais (2004), "Predicting Risk Sensitivity in Humans and Lower Animals: Risk as Variance or Coefficient of Variation," *Psychological Review*, 111 (2), 430-445.

Wheatley, John J., John S. Y. Chiu, and Andrew C. Stevens (1980), "Demographics to Predict Consumption," *Journal of Advertising Research*, 20 (6), 31-38.

Winkler, Robert L. (1970), "Intuitive Bayesian Point Estimation," *Organizational Behavior and Human Performance*, 5 (5), 417-429.

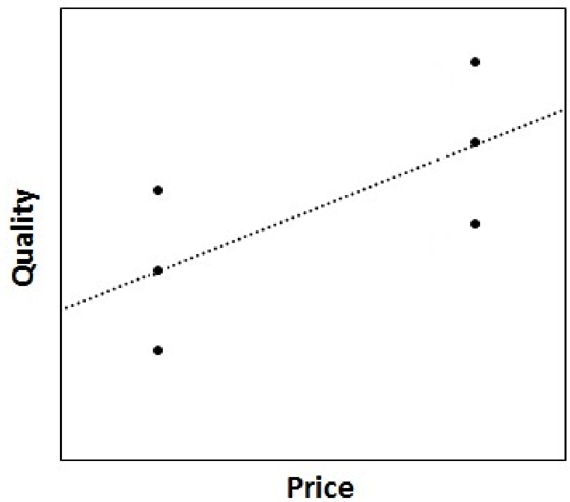
Wolfe, Christopher R. (1995), "Information Seeking on Bayesian Conditional Probability Problems: A Fuzzy-trace Theory Account," *Journal of Behavioral Decision Making*, 8 (2), 85-108.

Wolfe, Mary L. (1975), "Distribution Characteristics as Predictors of Error in Intuitive Estimation of Means," *Psychological Reports*, 36, 367-370.

Wübben, Markus and Florian von Wangenheim (2008), "Instant Customer Base Analysis: Managerial Heuristics Often "Get It Right"," *Journal of Marketing*, 72 (3), 82-93.

FIGURE 1.
The conditional expectation of quality given price is the same in panels A and B, but the correlation is higher in panel B.

PANEL A



PANEL B

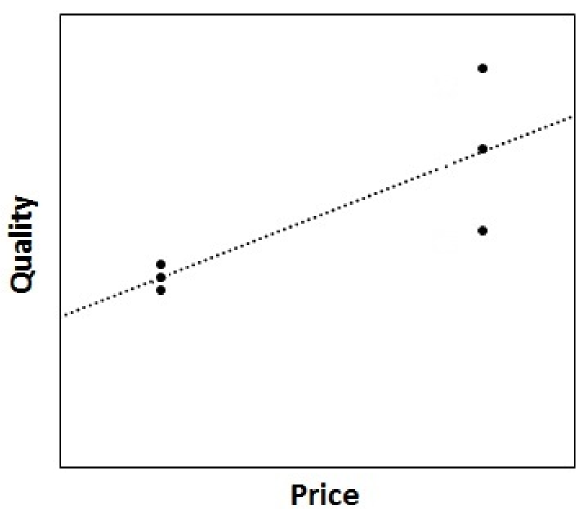


TABLE 1.
The conditional probability of success given Facebook is the same in panels A and B, but the correlation is higher in panel B.

PANEL A

		Success	
		Yes	**No**
Channel	**Facebook**	50	50
	Other	50	50

PANEL B

		Success	
		Yes	**No**
Channel	**Facebook**	50	50
	Other	10	90

TABLE 2.

Although correlations are symmetric, conditional probabilities are not. The conditional probability that a customer churns given that the model predicts churn is .90, but the conditional probability that the model predicts churn given that the customer churns is .50

		Actual	
		Churn	**Retain**
Model	**Churn**	90	10
	Retain	90	10