More Statistical and Methodological Myths and Urban Legends

Charles E. Lance

Organizational Research Methods 2011 14: 279 originally published online 30 December 2010
DOI: 10.1177/1094428110391814

The online version of this article can be found at:
http://orm.sagepub.com/content/14/2/279
Abstract
The study of urban legends represents the application of concepts developed in the academic study of traditional folktales to stories circulating in the modern world. Vandenberg adapted the notion of “urban legends” into the area of organizational research methods and coined the term statistical and methodological myths and urban legends (SMMULs) to refer to the collection of various rules of thumb and other “received doctrines” that often guide researchers’ scientific behavior. Various SMMULs have been examined over 7 years in a number of symposia at scientific conferences, in a previous Feature Topic in Organizational Research Methods (2006), and in an edited book. This Feature Topic continues this tradition by presenting five new SMMULs.

Keywords
construct validation procedures, reliability and validity, factor analysis, quantitative research, philosophy of science, meta-analysis, structural equation modeling

Have you heard this story?

A woman comes home from shopping and finds her Doberman pinscher lying in the entry of her home gagging and choking. She immediately drops her purchases, puts the dog in her car, and rushes to the veterinarian. The vet says that he will have to operate to remove the obstruction, and he advises her to return home and wait for his call. Back home, as the woman is opening her door she hears the telephone ringing, and she hurries to answer it. The vet is calling, and he urgently commands her to hang up the phone immediately and run outside; he has already called the police, who will explain. The woman does as he says, and the police soon arrive. They tell her that the vet found two fingers stuck in the Doberman’s throat. They search the house and find a burglar hiding in a closet; he is in shock and has two fingers bitten off one hand. (Brunvand, 2001, p. 71)
“The Choking Doberman” is one of the more famous “urban legends” that Jan Harold Brunvand has documented over the years (e.g., Brunvand, 1981, 1984, 1990, 2001). According to Brunvand (1984, 2001), it began circulating in the United States around 1981 and quickly spread from coast to coast and abroad, but is thought to have evolved from a much earlier Welsh legend about Prince Llewelyn and his dog Gelert in which the prince returns home to find Gelert with a bloody mouth and Lleweny’s infant’s cradle overturned and empty.

Brunvand is a professor emeritus of the University of Utah who is best known for taking concepts developed in the academic study of traditional folktales and applying them to stories circulating in the modern world (“Jan Harold Brunvand,” 2010). According to Brunvand (2001), urban legends are all those bizarre, whimsical, 99 percent apocryphal, yet believable stories that are just “too good to be true.” Often their narrative structure sets up some kind of puzzling situation that is resolved by some sudden plot twist, at which the story ends abruptly. (p. xxviii)

As such, the study of urban legends represents the application of concepts developed in the academic study of traditional folktales to stories circulating in the modern world, such as alligators teeming in New York’s sewers (see Figure 1), the $250 Neiman Marcus cookie, the nuked poodle, and the $50 Porsche (Brunvand, 2001). Many urban legends are supported by a kernel of truth (e.g., someone may once have actually flushed a baby alligator down a toilet in New York City), but largely they are myths, perpetuated in part by their (at least apparent) plausibility. The urban legend concept has become so popular as to inspire a 1998 film of that name in which a serial killer uses the methods of death described in certain urban legends. Altogether, the film incorporates no fewer than 13 popular urban legends into its plot. Although originally urban legends were passed along as part of oral tradition and folklore, today they are more often shared over the Internet (“Urban Legend,” 2010). A relatively new form of urban legend that is increasingly popular on the Internet is referred to as fauxtography and involves the circulation of retouched photographs with accompanying stories (“Fauxtography,” 2010). Fortunately, there are several Internet sites today that are devoted to sorting out truth from fiction in stories that are passed around verbally, electronically, or otherwise (e.g., snopes.com, debunker.com., straightdope.com).

In 2004, Vandenberg adapted the notion of urban legends into the realm of organizational research methods and coined the term statistical and methodological myths and urban legends (SMMULs). SMMULs are those rules of thumb, maxims, truisms, and guidelines for research
conduct, that is, “received doctrines” (Barrett, 1972, p. 1), that are taught in undergraduate and graduate classes, enforced by gatekeepers (e.g., grant panels, reviewers, editors, dissertation committee members), discussed among colleagues, and otherwise passed along among pliers of the trade far and wide and from generation to generation. They establish normative research conduct and reinforce beliefs about what’s what and what’s not. They inform us about how we should go about our research business, what constitutes adequate or exemplary research behavior, and the kinds of things we should be on the lookout for along the way. They are prescriptive and standard bearing. And like their traditional folklore counterparts, they are often founded upon a kernel of truth but are supported also by apparently plausible myth. That is, SMMULs are often based, in part, on sound rationale and justification but also, in part, on unfounded lore (Lance & Vandenberg, 2009a). For example, as Aguinis and Harden (2009, p. 273) noted,

Cohen’s [1962] definitions of small, medium, and large effect sizes are based in part on observed values as reported in the articles published in the 1960 volume of *Journal of Abnormal and Social Psychology*, and in part on his own subjective opinion. A few years later, Cohen (1988) decided to lower these values to .10 (small), .30 (medium), and .50 (large) because the originally defined values seemed a bit too high. Given the history behind the conventional values for small, medium, and large effects, it is not surprising that Cohen (1992) himself acknowledged that these definitions “were made subjectively” (p. 156).

Yet, as Aguinis and Harden document, Cohen’s small, medium, and large effect sizes have become industry standard “lines in the sand” across the social sciences. As a second example, Baron and Kenny (1986) did indeed advocate a four-step procedure for testing mediation hypotheses, a procedure that is widely applied today; this despite the facts that (a) some of the procedure’s basic requirements were subsequently rescinded (Kenny, Kashy, & Bolger, 1998) and (b) the four-step procedure has been shown to be among the least effective approaches available for testing mediation hypotheses (LeBreton, Wu, & Bing, 2009; MacKinnon, Lockwood, Hoffman, West, & Shee, 2002; Shrout & Bolger, 2002).

There now exists a 7-year-long tradition of exposés on various SMMULs, beginning with a very popular Academy of Management symposium chaired by Vandenberg in 2004 in which panel participants addressed the following questions: (a) What is the SMMUL that we (almost) all know to be true? (b) What is the kernel of truth to the SMMUL? (c) What myths have developed around this kernel of truth? and (d) What should the state of the practice actually be? With respect to several different research issues (see Vandenberg, 2006). Three papers from this symposium were subsequently published as a Feature Topic in *Organizational Research Methods (ORM)* in April 2006 (perhaps not so coincidentally, 5 years ago, this issue). One of these (Lance, Butts, & Michels, 2006) won *ORM*’s 2006 Best Paper of the Year Award, and as of October 1, 2010, these articles are the 2nd (Spector, 2006), 6th (Lance et al., 2006), and 26th (James, Mulaik, & Brett, 2006) most often cited articles in *ORM*’s history. Researchers are paying attention.

The SMMUL tradition has continued with the presentation of five symposia at meetings of the Society for Industrial and Organizational Psychology (Lance & Vandenberg, 2008, 2010, 2011; Vandenberg & Lance, 2007, 2009), which have consistently been well attended. Several of these papers were compiled in Lance and Vandenberg’s (2009b) edited volume on SMMULs and addressed topics such as determination of sample size and adequacy of response rates, misconceptions in exploratory factor analysis, the unmeasured variables problem in nonexperimental research, self-report bias, and the importance of variance accounted for in judging the importance of research findings, among other topics. In one of these papers, Highhouse and Gillespie (2009) addressed the issue of sample generalizability in applied research. For example, a study addressing an important issue in organizational behavior might be criticized because it used a student sample, and authors are
often apologetic in their discussion sections for the particular samples they used in their studies. In order to see what authors apologize for in their discussion sections, I randomly chose the November 2006 issue of the *Journal of Applied Psychology*. Out of 11 empirical articles in this issue, 9 included a “limitations” section, and in these sections were 11 apologies for study design limitations, 3 apologies for measures’ shortcomings, 1 apology for a small effect size, 5 apologies for using self-report data, and 5 apologies for sample characteristics. Of these latter apologies, authors apologized for using students (Price, Harrison, & Gavin, 2006), teachers (Trevor & Wazeter, 2006), firefighters (Neal et al., 2006), engineers (Joireman, Kamdar, Daniels, & Duell, 2006), and managers (Morgeson & Humphrey, 2006) as research participants (what kind of sample does one not have to apologize for?). Highhouse and Gillespie’s analysis of sample generalizability issues suggests that despite the repeated apologies for sample characteristics that are seen in empirical studies (and, we suspect, reviewer comments that provoke these apologies), “it is rare in applied behavioral science for the nature of the sample to be an important consideration for generalizability” (p. 250).

The intent of the SMMUL tradition is to identify pervasive beliefs that researchers hold regarding the conduct of social and organizational research; to determine the (partial?) veracity of these beliefs and the accompanying lore that also perpetuates their existence and application; and to settle, to the extent possible, what the best practices actually are with respect to these beliefs. It is this tradition that also motivates the articles in the present Feature Topic.

Some researchers I know (and I bet you know some too) routinely partial out or control for certain variables before they undertake their focal analysis, and sometimes it seems that these “control variables” are used with little thought except for fending off potential reviewer criticisms. Spector and Brannick (2011, in this issue) critique this practice, which they call the “purification principle,” that is based on the belief that the use of statistical controls produces more accurate estimates among focal variables. Among other recommendations, Spector and Brannick suggest that researchers (a) avoid using demographic variables that are mere proxies for the variables that are of real theoretical interest and (b) be more explicit about the theoretical roles that alleged control variables play in substantive models of interest.

The advent of meta-analysis some 35 or so years ago has had a profound impact on the ways that researchers understand their literatures, theories, and research findings across the sciences (Schmidt & Hunter, 2003), yet nearly from the beginning meta-analysis was controversial (e.g., Algera, Jansen, Roe, & Vijn, 1984; Eysenck, 1978). There have been many efforts at identifying and clarifying controversies associated with meta-analysis (e.g., Le, Oh, Shaffer, & Schmidt, 2007; Rosenthal & DiMatteo, 2001; Schmidt, Hunter, Pearlman, & Hirsch, 1985), but many misunderstandings remain. Aguinis, Pierce, Bosco, Dalton, and Dalton (2011, in this issue) identify and clarify seven of these, including the beliefs that a single meta-analytically derived effect size can summarize an entire literature, that meta-analysis makes lemonade out of lemons, and that a discrepancy between meta-analysis results and results of a large randomized controlled trial indicates that the meta-analysis is flawed.

Concomitant with the rise of meta-analysis has been increased emphasis on the reporting of effect sizes and confidence intervals, in addition to, or perhaps in lieu of, more traditional test statistics’ values and probability levels (APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008; Cohen, 1994). More extreme versions of this position would argue for the abandonment of null hypothesis significance testing (NHST) altogether in favor of what Cortina and Landis (2011, in this issue) call effect size testing (EST; e.g., Hunter, 1997; Schmidt, 1996). Cortina and Landis summarize arguments that have been made in favor of and against both NHST and EST and conclude that despite recent arguments in favor of EST, NHST holds considerable merit as an effective translation mechanism.

The development of modern structural equation modeling (SEM) techniques also corresponded roughly with the advent of meta-analysis (Bentler & Bonett, 1980), and one issue that has been
controversial since its inception is the assessment of model fit (e.g., Hu & Bentler, 1998, 1999; Marsh, Balla, & McDonald, 1988). As Williams and Boyle (2011, in this issue) point out, nearly all of the model fit indices that have been proposed over the years are global fit indices that simultaneously assess fit of the measurement and structural portions of latent variable models. This is problematic because it is the structural portion of the model that is usually of primary interest, but global fit indices do not allow separate assessment of the structural model’s fit. Williams and Boyle develop and provide some preliminary evaluation of two new fit indices that are designed to overcome this limitation of presently available global fit indices.

Another issue that has risen to some prominence more recently in the SEM literature is the use of formative (vs. reflective) indicators (MacKenzie, Podsakoff, & Jarvis, 2005; Podsakoff, Shen, & Podsakoff, 2006). Consistent with the common factor model, reflective indicators are caused by their corresponding latent variables, while, consistent with the principal component model, formative indicators serve as causes of their corresponding latent variables. Edwards (2011, in this issue) examines a number of purported advantages of formative measurement models but in the end concludes that (a) formative models are plagued with more problems than their proponents have acknowledged and (b) the objectives of formative measurement models can actually be achieved at least as effectively using reflective indicators.

Of course these articles do not complete the list of SMMULs that could be made of those that exist in the social and organizational sciences, but it is hoped that they will make interesting contributions to the SMMUL tradition. I certainly found them interesting and hope that you do too.

In closing, I would like to thank the former editor, Bob Vandenberg, for encouraging me to create and manage this Feature Topic, and the articles’ authors for their outstanding contributions—it really was a pleasure working with you. Finally, all of the articles in this Feature Topic were shaped, and I will say improved, as a function of the reviewers’ thoughtful and constructive comments, and so I express my sincere appreciation to them in alphabetical order: Tom Becker, Art Bedeian, Roy Howell, Ron Landis, Huy Le, George Marcoulides, Dan Newman, Raymond Nickerson, Fred Oswald, Hannah Rothstein, Marcia Simmering, Gene Stone-Romero, Dirk Timme, and Ryan Zimmerman. Couldn’t have done it without you!

Oh, and did you hear about the woman who wanted to get a fast suntan? All the tanning salons had a 20-minute limit so she visited all the salons in her town for the maximum period over the course of several days. She acquired a deep tan but began to suffer a bad odor emanating from her body despite repeated bathing, cologne etc., and she started to feel weak and sick. She finally saw her doctor, explained her situation and was told that the tanning rays have cooked her insides and that she has only a few weeks to live! (“Curses! Broiled Again!” Brunvand, 2001).

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the authorship and/or publication of this article.

Funding
The author(s) received no financial support for the research and/or authorship of this article.

Notes
2. As of October 21, 2010, this article had been cited 11,309 times in the PsycINFO database—1,038 times in 2010 alone (an average of about 3.5 citations per day!).
References


Cortina, J. M., & Landis, R. S. (2011). The Earth is not round (p = .00). Organizational Research Methods, 14, 332-349.


Lance, C. E., & Vandenberg, R. J. (2010, April). Statistical and methodological myths and urban legends: Part V. Symposium chaired at the meeting of the Society for Industrial and Organizational Psychology, Atlanta, GA.


**Bio**

Charles E. Lance is a Professor of I/O psychology at The University of Georgia. His work in the areas of performance measurement, assessment center validity, research methods, and structural equation modeling has appeared in such journals as *Psychological Methods*, *Organizational Research Methods (ORM)*, *Journal of Applied Psychology*, *Organizational Behavior and Human Decision Processes*, *Journal of Management and Multivariate Behavioral Research*. He is also co-editor of *Performance Measurement: Current Perspectives and Future Challenges* (with Wink Bennett and Dave Woehr) and *Statistical and Methodological Myths and Urban Legends: Received Doctrine, Verity, and Fable in Organizational and Social Research* (with Bob Vandenberg). Chuck is a Fellow of the Society for Industrial and Organizational Psychology (SIOP) and the American Psychological Association and is a member of the Society for Organizational Behavior. He is currently Associate Editor of *ORM*, and on the editorial board of *Personnel Psychology, Human Performance, and Group & Organization Management*. 