

Fifty Lost Years: Why International Business Scholars Must Not Emulate the US Social-science Research Model

John L. Kmetz*

US business schools imported many social scientists into their faculties in the 1960s, in response to sharp criticism from the 1959 Ford and Carnegie Foundations reports that business research lacked rigor. While these faculty brought many benefits, a serious problem was inherited with them—the incorrect interpretation of results from Null Hypothesis Significance Testing (NHST). In the 50 years since, a mythology has grown around NHST results which renders the overwhelming majority of published business-school research an unsound, unscientific, typically incorrect body of findings. The author will show that (1) the flaws and errors of this methodology are well-documented and virtually ignored by academic researchers; (2) no study of business-school research supports the contention that it is used or useful in practice; and (3) that our research is virtually ignored even by academic authors of best-selling business books (of 3,162 total references in my 30-book sample, only 131 were from the social sciences, and of these only one from international business journals). If international business research is to avoid falling into the status of “junk science,” it must take active and aggressive measures to ensure scientific validity immediately. The most crucial of these, all of which can be readily implemented, are suggested by the author in his conclusions.

Field of Research: Research Methodology

1. Introduction

From the founding of the eponymous business school by Joseph Wharton at the University of Pennsylvania in 1881, business schools have become a nearly-universal element of the typical United States college or university. Wharton’s objectives in forming a business school were highly pragmatic, and he wanted the business school to be a source of learning about subjects as diverse as proper elocution, the nature of securities, the causes of financial panics, business law, and many other topics related to the conduct of business. Among an impressive list of Wharton “firsts” is the 1921 creation of the Industrial Research Unit, the first dedicated business research center (Wharton history, 2010). Practitioners, economists, and scholars had been writing their ideas about management and organization for centuries (George 1968), and many of these early contributions became the basis for business-school teaching; but from that point on, programmatic research and publication became established elements of the mission of most business schools, with many faculty and experienced managers publishing materials that summarized experience and emphasized practical application or “management philosophy.”

*Associate Professor of Management, Dept. of Business Administration; Faculty Director, Project Management Certificate Program, University of Delaware, Newark, DE 19716 USA, +1 302 831 1773 kmetz@udel.edu, <http://www.buec.udel.edu/kmetzj>

1959 was a watershed year in the development of the US business school. That year, two reports from the Ford Foundation (Gordon and Howell 1959) and Carnegie Foundation (Pierson 1959) were published; both were highly critical of the unscientific nature of the cases, histories, and anecdotes comprising much b-school research, and strongly recommended the adoption of more rigorous and scientific research methods. Business schools began rapidly changing the composition of their faculties to meet these challenges (Bass 1965), and one key feature of this response was the recruitment of faculty trained in the research methods and practices of the behavioral sciences (Webber *et al.* 1971). Part of this was supported directly by the Ford Foundation, which made USD 35 million in grants to support research strengthening. Faculty recruitment from the social sciences seemed logical for many reasons—many early authors had long recognized the importance of the human factor in business and had written extensively on that subject; moreover, these faculty came from a background of being trained to do the seemingly more scientific research the two foundation reports emphasized.

2. The Payoff

In the more than half-century since these reports, research has become a primary mission for many US business schools and their faculties. “Tier 1” universities are defined by their research output and the emphasis placed on research productivity by faculty. “Productivity” is measured in terms of not just the number of publications or the total page count produced, but in terms of the number of publications in the “top” journals of the authors’ fields. That productivity has boomed—from a handful of academic research journals in the early 1960’s, the number of journals and subdisciplines has grown to the point where the premier US business-school accrediting body, AACSB International, estimated that by 2005 there were over 1900 English-language journals in the field, which published 15,000 to 20,000 articles per year at an estimated annual cost of USD 320 million (AACSB 2007, p. 6). AACSB also reported that over 40 percent of b-schools placed emphasis on research at a level equal to that of teaching.

One would expect such an outpouring of research to have major positive impacts on all the disciplines of business and economics, and after half a century evidence of that would be clear through advances in both management practice and the theoretical underpinnings of practice. But such evidence is rare, and often unconvincing—despite there being a body of scholars who believe that behavioral science research now justifies a movement toward “evidence-based management” as the basis for teaching in b-schools (Rousseau 2006, 2008; Rousseau & McCarthy 2007), there is no consensus on what that “evidence” supports. One finds unsubstantiated claims that research helps practicing decision-makers (Tahai & Meyer 1999), or assessments of research “impact” that examines only what other academics value (Podsakoff *et al.* 2005, 2008). There are broader inquiries into the limited impact of the social sciences in general (*e.g.*, Beyer & Trice 1982) that basically underscore the extent to which the social sciences are ignored. A recent claim that research adds value to MBA programs in the form of higher salaries for graduates is more likely a tautology for US “Tier 1” research status (O’Brien *et al.* 2010).

Kmetz

Instead, the literature on evidence of research impact is a long record of studies showing that academic research in management is neither used by practitioners nor perceived as relevant to their interests (Behrman & Levin 1984; Bennis & O'Toole 2005; Buckley *et al.* 1998; Chia & Holt 2008; Cohen 2007; Deadrick & Gibson 2007; Duncan 1974a, 1974b; Dunnette & Brown 1968; Gopinath & Hoffman 1995; Guest 2007; Hambrick 1993, 1994; Hoffman & Gopinath 1994; Kilmann *et al.*, 1983; Lyles 1990; Miller 1999; Pfeffer & Fong 2002; Rynes *et al.* 2007; Starbuck 2006; Van de Ven 2000; Van de Ven & Johnson 2006). In their extensive first-hand study of business education, Porter & McKibbin (1988) concluded that managers “ignore academic research with impunity” (p. 304). In reviewing “key management ideas,” Crainer (1998) cites not a single top journal, and the only “key idea” based on empirical research is Herzberg’s Two-Factor Theory, which most motivation researchers consider to be a discredited idea. Similarly, Mol and Birkinshaw (2008) trace not a single one of their “giant steps in management” to empirical research, and they also cite not a single top journal; this is especially ironic because their focus is on management innovations that have had significant impact. The business press has consistently noted the lack of relevance of academic research over the years (Panning for gold 2004; Practically irrelevant 2007; Signifying nothing? 2004; Byrne 1990, 1997; Oviatt & Miller 1989; Skapinker 2008).

Compared to 1959, many academic b-school researchers perceive themselves as scientists (Cummings 2007; Rousseau 2006). But many in the academy question whether our research is truly scientific (Hambrick 1994, 2007), or forms a sound foundation for teaching (Pearce 2004; Pfeffer & Fong 2002; Starbuck 2006). Our research appears to have no impact on those who recruit our MBA graduates (Safón 2007). AACSB International itself has opened questions about the value and “impact” of b-school research (AACSB 2007, with a call for greater attention to its impact on practice). There has been expression of concern over the loss of professionalism in management education (Trank & Rynes 2003) and faddishness in both education and practice (David & Strang 2006; Greatbatch & Clark 2005; Huczynski 2005).

A recent study into research misconduct on the part of management researchers among 104 US PhD-granting, AACSB-accredited universities was chilling (Bedeian *et al.* 2010). Within the year prior to the study, nearly 80 percent of the 384 respondents reported knowledge of researchers who withheld methodological results or details; or selected only supporting data for a hypothesis and withheld the rest, among other issues. Nearly 92 percent knew of instances where hypotheses were developed after results were known (an absolute breach of scientific integrity). Nearly 27 percent knew of cases where results were fabricated!

Rather than direct impact, some have posited that there is a “food chain” through which research is processed and converted into information with practical value. McKelvey (2006) described the concept of a “knowledge food chain” not unlike a biological food chain, in which b-school research followed a path something like this:

Disciplines > Management Research > PhD/MBA Students >
Consultants > Practitioners.

Kmetz

I investigated one possible “food chain” pathway by analyzing the references from a group of 30 business books which made the best-sellers list between 1996 and 2005 (Kmetz 2011, in review). Since authors with academic backgrounds would be more likely than non-academics to know the research literature, I deliberately selected authors with academic backgrounds in greater proportion than non-academics, such that 13 of the 30 authors (43 percent) had academic affiliations. Table 1 shows the results of analysis of the 3,162 references cited in these books—as it shows, academic research in general is largely ignored, with only 361 academic journal citations in total. Of these, 154 are from economics, by far the most frequently cited discipline, and 152 of the 154 are cited in only three books, by authors with academic affiliations. Of the social-science literature, which is the focal point of this paper, there are only 131 citations, and 90 of these are accounted for by four books alone. I examined the sources of these, and only 40 of the 131 came from “top” journals; half of these 40 came from one book with an academic author, who cited the *Strategic Management Journal* 20 times.

Thus, empirical evidence on the payoff of research consistently reports a bleak and discouraging conclusion—b-school research appears to have little or no relevance to anyone other than the closed circle of academics who produce it, and even they almost entirely ignore it when addressing practitioners. The social-science model that was the great hope for the future in 1960 appears to have produced very little of value after 50 years.

Kmetz

TABLE 1
Breakdown of Best-Selling Book References by Title and Category

Title Codes (see Appendix)

	8H	101	AL	BDB*BOS*	CFF*	CR	CSH	CST*	CWP*	DC	EEQ	FP	FRK*	G2G	GBL*	GN*	HOW	IS*	M6S	MFO*	PZ	RC*	SFO*	TC	TM	WD	WMI	WNG	WRW*	Totals	PCT	
Book	53	0	61	29	68	41	3	88	15	15	36	270	143	53	52	88	29	5	87	89	33	12	24	41	14	0	136	38	0	67	1590	50.28
broadcasT	0	0	0	0	1	0	0	0	1	0	0	0	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	7	0.22
Case	0	0	0	1	0	0	0	1	0	9	2	1	0	0	2	0	6	0	4	0	22	1	1	1	0	0	0	7	0	0	58	1.83
Interview	0	0	0	0	0	0	0	0	0	0	13	2	0	0	1	0	7	0	0	3	0	6	0	0	1	0	0	0	0	0	33	1.04
Journal	3	0	1	10	33	19	0	5	59	19	9	62	7	50	10	19	21	0	62	21	87	2	84	25	1	0	2	12	0	2	625	19.77
<i>Jml brkdn</i>																																
Prac	3	0	1	3	15	11	0	2	1	4	3	22	6	1	10	1	9	0	20	10	5	2	16	22	1	0	2	12	0	2	184	5.82
Acad	0	0	0	5	17	8	0	3	54	15	4	23	0	32	0	8	6	0	41	11	72	0	59	3	0	0	0	0	0	0	361	11.42
Legal/ethx	0	0	0	2	0	0	0	0	4	0	0	0	1	9	0	0	6	0	1	0	10	0	8	0	0	0	0	0	0	0	41	1.30
Other	0	0	0	0	1	0	0	0	0	0	2	17	0	8	0	10	0	0	0	0	0	0	1	0	0	0	0	0	0	0	39	1.23
<i>Acad brkdn</i>																																
Acctg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	2	0.06
Econ	0	0	0	2	7	1	0	0	50	1	2	0	0	27	0	0	1	0	3	0	52	0	8	0	0	0	0	0	0	0	154	4.87
Finance	0	0	0	3	0	0	0	0	2	0	0	0	0	0	0	0	1	0	2	0	8	0	40	0	0	0	0	0	0	0	56	1.77
Opns	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2	0	0	0	1	0	0	0	0	0	0	4	0.13
GASSSPP	0	0	0	0	10	7	0	0	2	12	2	22	0	5	0	8	4	0	35	3	9	0	11	1	0	0	0	0	0	0	131	4.14
Other	0	0	0	0	0	0	0	3	0	1	0	1	0	0	0	0	0	0	1	6	1	0	0	1	0	0	0	0	0	0	14	0.44
<i>GASSPP brkdn</i>																																
Intl bus	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0.03
Mgmt	0	0	0	0	3	0	0	0	0	3	1	4	0	0	0	1	1	0	10	2	0	0	5	0	0	0	0	0	0	0	30	0.95
Org	0	0	0	0	0	0	0	0	0	5	1	1	0	0	0	0	0	0	9	0	3	0	1	0	0	0	0	0	0	0	20	0.63
Psych	0	0	0	0	3	0	0	0	1	0	0	13	0	1	0	6	0	0	1	1	0	0	0	0	0	0	0	0	0	0	26	0.82
Sociology	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	2	0	0	0	0	0	0	0	0	0	5	0.16
sStrategy	0	0	0	0	3	6	0	0	0	2	0	0	0	0	0	0	0	0	13	0	0	0	5	0	0	0	0	0	0	0	29	0.92
markETING	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	4	0.13
Other	0	0	0	0	1	1	0	0	1	2	0	4	0	2	0	0	1	0	1	0	2	0	0	1	0	0	0	0	0	0	16	0.51
Magazine	6	0	2	4	11	25	11	6	6	3	1	22	1	16	139	4	4	0	3	22	0	6	9	10	3	0	0	1	0	1	316	9.99
Newspaper	1	0	0	4	7	11	6	3	5	1	0	6	0	38	82	0	3	1	4	2	1	1	15	1	14	0	0	8	0	0	214	6.77
Paper	4	0	0	9	1	2	1	6	4	0	4	4	5	21	5	2	5	0	5	11	6	0	25	7	2	0	1	0	0	130	4.11	
Speech	1	0	1	1	0	0	0	0	1	1	3	0	1	0	1	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	15	0.47
Website	3	0	0	11	9	0	0	0	2	2	0	0	0	0	1	0	0	1	3	5	0	0	1	0	0	1	0	0	0	0	39	1.23
Z other	7	0	0	6	3	3	3	0	4	2	4	4	2	3	15	3	6	0	3	5	0	5	51	3	1	0	1	0	0	1	135	4.27
Corrected Totals**	78	0	65	75	133	101	24	109	97	52	72	371	159	183	310	116	81	7	171	158	149	33	215	88	36	1	140	67	0	71	3162	

* Author with academic affiliation ** Column totals corrected for double counting of journals, academic journals, and social science journals.
GASSSPP = Generally Accepted Soft Social Science Publishing Process, the US social-science research model (see text)

3. What Went Wrong?

Unfortunately, one of the traditions that came into business schools along with the new behavioral-science faculty was their approach to evaluating the outcomes of statistical research. As Hubbard & Ryan (2000) demonstrate, Null Hypothesis Significance Testing (NHST) grew explosively in psychology between 1940 and 1955, and had essentially become “synonymous with empirical analysis” (Hubbard *et al.* 1997), a status NHST holds to this day. While the technical calculation of statistical procedures may be correct, the underlying problem is that NHST results are almost universally misinterpreted in the social-science literature. Along with this fundamental misinterpretation, there has grown a supporting mythology about statistical significance, and the combined effect creates an “amazing persistence of a probabilistic misconception” regarding NHST (Falk & Greenbaum 1995). Collectively, NHST and its emergent mythology have created a statistically and scientifically unsound basis for the evaluation and publication of b-school research. I refer to this as the Generally Accepted Soft Social Science Publishing Process, or GASSSPP. It is as generally accepted as is the GAAP in US accounting; it also incorporates the general research procedures common to “soft” psychology (Meehl 1978).

What cannot be debated is that as use of NHST has grown, so has the misinterpretation of statistical significance. Unfortunately, this fact is often a revelation for many, and because it is a fundamental issue in the GASSSPP, must be explained more fully. As Carver (1978) put it:

What is the probability of obtaining a dead person (label this part D) given that the person was hanged (label this part H); this is, in symbol form, what is $P(D|H)$? Obviously, it will be very high, perhaps 0.97 or higher. Now, let us reverse the question. What is the probability that a person has been hanged (H), given that the person is dead (D); that is, what is $P(H|D)$? This time the probability will undoubtedly be very low, perhaps 0.01 or lower. No one would be likely to make the mistake of substituting the first estimate (0.97) for the second (0.01); that is, to accept 0.97 as the probability that a person has been hanged given that the person is dead. Even though this seems to be an unlikely mistake, it is exactly the kind of mistake that is made with interpretations of statistical significance testing—by analogy, calculated estimates of $P(D|H)$ are interpreted as if they were estimates of $P(H|D)$, when they clearly are not the same.

Statistical significance tells only $P(\text{Data}|\text{Hypothesis})$ —it never tells us $P(\text{Hypothesis}|\text{Data})$, which obviously has a completely different base rate, but is nearly universally interpreted as $P(H|D)$ in our journals. This confusion of base rates is extremely important, as Bakan (1966: 425) related in a brief example:

Some years ago, the author had occasion to run a number of tests of significance on a battery of tests collected on about 60,000 subjects from

Kmetz

all over the United States. Every test came out significant. Dividing the cards by such arbitrary criteria as east versus west of the Mississippi River, Maine versus the rest of the country, North versus South, etc., all produced significant differences in means. In some instances, the differences in the sample means were quite small, but nonetheless, the p values were all very low.

In a case like this, confusing $P(H|D)$ for $P(D|H)$ would seem unlikely because we know that any differences on a battery of standard test means have to be attributable to random variation. But suppose we are doing a study where we are doing nothing more than arbitrarily parsing data, and have no *a priori* reason to disregard such differences? By incorrectly interpreting p as $P(\text{Hypothesis}|\text{Data})$ and using only the p level to decide what our data tell us, we have just reached an incorrect conclusion based on sampling error, which in fact probably accounts for most NHST results in the GASSSP journals (Hunter, Schmidt, & Jackson 1982).

This is bad enough, but a mythology about p levels has grown within the GASSSP, such that it is falsely believed to reveal insights into statistical outcomes that it simply does not and cannot. A brief summary of this mythology is:

- p tells us the odds that our rejection of the null hypothesis is due to chance; it only tells us the likelihood of the data being obtained by chance under the null hypothesis
- statistical significance establishes existence of a statistical effect; in fact, significance and effect are independent of each other for any sample of reasonable size, and large samples assure statistical significance
- $p < .05$ proves we have support for an hypothesis; in fact, p alone is never proof of anything, and the .05 level is a convention lacking any scientific basis whatsoever
- $p < .05$ is a “significant” outcome, $p < .01$ is “very significant,” and $p < .001$ is “highly significant;” in fact, there is no scale of outcome strength as a function of the p level, and these all-too-common statements are always incorrect
- p is the appropriate metric for those interested in theory development, and effect sizes matter only when practical application is the issue; in fact, p alone is never the appropriate metric to evaluate outcomes, and it establishes neither practical nor theoretical importance
- the p level indicates the likelihood that an outcome would not replicate if the study were repeated; in fact, p provides absolutely no information about replicability
- the p level predicts the number of statistical outcomes that would be significant by chance; this would be true only if one is certain that $P(\text{type II error}) = 0$, and it never is
- a null hypothesis is a scientific hypothesis; it is not—the null hypothesis is an artifact used to construct an empirical question (and never used in real science), where a scientific hypothesis is usually a tentative explanation of a

Kmetz

- phenomenon based on limited evidence; the two are completely unrelated
- rejecting a null hypothesis means the alternative is correct; it does not—if a specific alternative is true, that must be demonstrated independently
- reliability can be substituted for validity; they are not the same, of course, but in GASSSPP journals they usually are treated as if they are—literally, consistency is considered the same as accuracy, even if it means simply repeating the same mistake, and one study (Scandura & Williams, 2000) concluded that the validity of measures in several top journals has declined in recent years, not become stronger.

Within the GASSSPP, NHST is the root process that perpetuates the mythology of statistical significance, and this mythology has been forcefully discredited in the research-methods literature for decades (Abelson 1995; Abrahamson 2005; Aiken *et al.* 1990; Armstrong 2007a, 2007b; Bakan 1966; Becker 1987; Berkson 1938, 1942; Bolles & Messick 1958; Carver 1978; Cohen 1962, 1990, 1994; Dawes *et al.* 1993; Edwards 2008; Faulkner *et al.* 2008; Fidler *et al.* 2004; Hays 1963; Hubbard & Armstrong 2006; Hubbard & Bayari 2003; Hunter 1997; Hunter, Schmidt & Jackson 1982; Hunter & Schmidt 2004; Ioannidis 2005; Jones 1955; Kalinowski & Fidler 2010; Kish 1959; Kmetz 1992, 2007; Loftus 1996; McCloskey 1992; McCloskey & Ziliak 1996; Meehl 1967, 1978, 1986, 1990; Moonesinghe *et al.* 2007; Murray 1993; Nix & Barnette 1998; Nunnally 1960; Oakes 1986; Rosnow & Rosenthal 1988, 1989; Rozeboom 1960; Schmidt 1992, 1996; Shrout 1997; Sterne 2001; Thompson 1998; Tukey 1960; Tyler 1931; Wang 2010; Ziliak & McCloskey 2008, among others—this list is not exhaustive). Nevertheless, any examination of our GASSSPP journals demonstrates the dominance of statistical significance as the evidential metric.

Many researchers attempt to dismiss the obvious flaws of the GASSSPP as mere quibbles over methods which ultimately make no difference in terms of scientific progress. This is simply not true. Consider the findings from Jackson & Dutton's (1988) study of how people process information cues. Table 2 shows the data for the second part of their study, where they evaluated responses to cues under four information conditions—threat, opportunity, neutral, or ambiguous. Based on whether paired comparisons indicated by the superscript numbers in Table 1 were statistically significant, where all but the comparison for hypothesis 4 were, they concluded:

In Study 2, the hypotheses we tested were developed from the simple assumption that the presence of issue characteristics that were consistent with threat (or opportunity) would strengthen threat (or opportunity) inferences, while the presence of issue characteristics that were discrepant with threat (or opportunity) would weaken such inferences. The specific predictions in hypotheses 1 through 4 were straightforward extensions of this assumption. Had all hypotheses been supported, we could have concluded that managers follow simple logical rules of information processing to discern threats and opportunities. The results suggest, however, that threat and opportunity inferences

Kmetz

cannot be accurately predicted from such a simple model of information processing. Instead, they indicate that managers are more sensitive to information that suggest the presence of a threat than they are to information that suggests the presence of an opportunity.... (p. 384).

TABLE 2
Jackson and Dutton's (1988) Results for Perceived Threat and Opportunity in Four Information Conditions (Study 2)

Information condition	Opportunity ratings		Threat ratings	
	Mean	SD	Mean*	SD
Neutral	3.8	1.1	3.7 ^{1,3,4}	1.2
Ambiguous	3.5	1.3	4.1 ^{2,3}	1.0
Opportunity-distinctive	4.1	1.2	3.6 ⁴	1.3
Threat-distinctive	3.1	1.2	4.3 ^{1,2}	0.9

n = 83 (of 400)

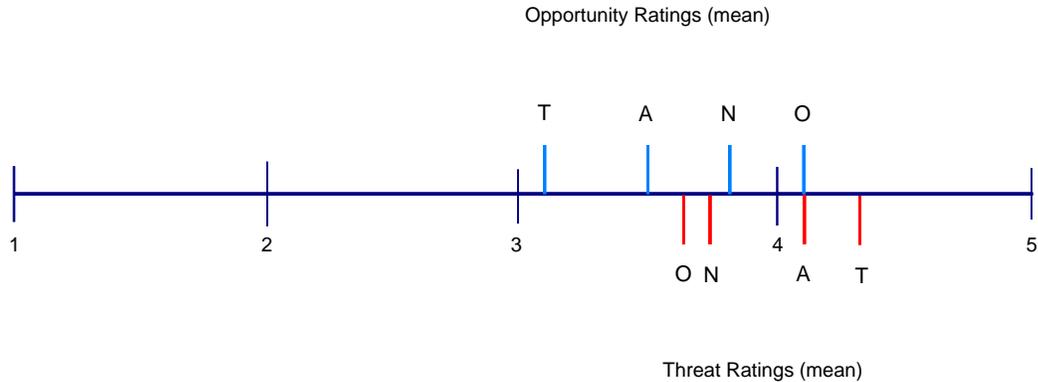
* Superscript numbers refer to planned comparisons in the authors' discussion.

However, an examination of the effect sizes in their table supports an entirely different conclusion. Figure 1 shows a linear plot of the means in Table 2, to scale.

As Figure 1 shows, the decision-frame conditions affect the mean values for the four types of cues only slightly; the threat and opportunity means are the extremes in both frames; opportunity-specific and threat-specific ratings are highest under their respective frames; and the order of means exhibits mirror symmetry between the frame conditions. Figure 1 clearly portrays a simple and highly consistent information-processing model. However, these effect sizes were never directly evaluated in the paper, and now that this study has been accepted through peer review and is part of our GASSSPP literature in a highly prestigious journal, it is a "fact" that managers use different rules to process differently-framed cues.

Kmetz

FIGURE 1
Effect-size Plot of Jackson & Dutton's (1988) Study 2 on Threat and Opportunity Perceptions.*



Legend: Letters indicate nature of information condition embedded in experimental stimulus booklet given to subjects. Means are located to scale, with opportunity-specific decision frame above line, and threat-specific frame below.

T = Threat-distinctive
O = Opportunity-distinctive
A = Ambiguous
N = Neutral

* Source: Kmetz, 2011, in review

What might once have been characterized as only widespread misunderstanding of the meaning of p values has grown into outright substitution of statistical significance for statistical effect, where effect should always be the primary criterion to determine what an outcome means. Because significance and effect are independent of each other, this substitution means that we include misleading findings in our published research, exacerbated by the GASSPP mythology. In terms of Figure 2, we select only the left column of outcomes, and thus automatically sweep in disinformation from cell 3. At the same time, we ignore potentially interesting outcomes in cell 2. This results in a 50 percent error rate in the body of published research, which is effectively a coin toss (Edwards 2008; Edwards & Berry 2010; Hunter 1997; Meehl 1967). In a study of medical literature, Ioannidis (2005) concluded that at least half of the published research is wrong. Meehl (1990) concluded that most findings in the social sciences were “discovery” of the low level of relatedness between all things, which he referred to as the “crud factor,” and Lykken (1968) earlier labeled “ambient correlation noise.”

Kmetz

FIGURE 2
The Independence of Statistical Significance and Statistical Effect

		Statistical Significance	
		Significant	Not Significant
Statistical Effect	Effect	(1) Both significance and effect	(2) Effect without significance
	No Effect	(3) Significance without effect	(4) Neither significance nor effect

For example, GASSSPP journals (including the “top” journals) are filled with what I refer to as “revelation by regression,” where researchers seem to literally believe that a significant coefficient or multiple R is evidence of an underlying relationship that has been “revealed” through regression analysis. I am personally appalled at the number of presentations I have seen in the past few years where the researcher dutifully outlines the research questions, develops null hypotheses, explains data collection and analysis, and summarizes the support (or lack thereof) for each hypothesis solely on the p level obtained; when I ask about the effect sizes associated with these p levels, it has been increasingly common to be told that the researcher doesn’t know them offhand, but has them somewhere in a table! Worst of all, multiple regression is typically being used for exploratory studies where there is no justification for labeling variables “dependent” or “independent” except in the mind of the researcher.

Unfortunately, GASSSPP research has collectively degraded into a body of junk science, where meaningful progress in pursuing a question or topic of interest is undone by the GASSSPP mythology. Studies are not replicated, a cornerstone of real science, owing to the false belief that a published significant outcome automatically predicts the result of such a replication. One cannot compare studies, cumulate studies, or estimate parameters from such a fragmented body of work; even in cases like the reified concept of “absorptive capacity,” where Lane *et al.* (2006) found 289 studies published between 1991 and 2002, mostly focused on R&D environments, they concluded (p. 858) that “the cross-citations between the papers in this body of literature show little evidence of an accumulated body of knowledge.” This may seem surprising in light of the large number of studies on the subject, but the “coin toss” nature of GASSSPP outcomes means that the signal-to-noise ratio in our research is typically dominated by noise.

Kmetz

In closing this discussion, I want to be absolutely clear that I am not questioning the technical statistical skills of researchers, nor suggesting that their calculations are incorrect. The issue here is the *interpretation* of statistical outcomes, and this goes well beyond the mathematics of crunching numbers. In the interpretation of outcomes and the understanding of scientific procedures in our research, an enormous body of erroneous beliefs has grown through incorrect textbooks, mimicry of others' practices, and creative but incorrect "interpretive innovation" over the years. We place the ultimate in electronic and technologically sophisticated aids to navigation in our boats, but we steer by watching the wake behind us.

4. What Can We Do?

Having been engaged in attempts to bring this issue to the attention of US researchers for nearly 20 years (Kmetz 1992, 1998, 2002, 2007, 2011), and having observed the same lack of attention befall others having much more recognition in the field than I (Starbuck 2005, 2006; Barnett 2007), I am forced to conclude that for multiple reasons, the GASSSPP disciplines simply disregard clear evidence of major technical and scientific shortcomings. Some of this, I must conclude, is defensive resistance on the part of those who feel threatened that acknowledging the GASSSPP mythology would undermine reputations they have worked very hard to build. Peters & Ceci (1982) resubmitted 12 papers to the same psychology journals that had published them 18 to 32 months earlier, but changed the names of the authors and their institutional affiliations; the result was rejection of eight of the 12 papers on resubmission. The reaction to this experiment was two years of anger, retribution, and professional attack (Ceci & Peters 1982), rather than a call to investigate obvious flaws in the review process. The APA Task Force report (Wilkinson 1999) on NHST supposedly created momentum toward improved publication practices, but has resulted in little change (Finch *et al.* 2001; Sohn 2000). But I suspect a large part of this is attributable to two persistent influences: (1) many statistics texts, having confused the Fischer-Neyman & Pearson debate on NHST years ago, simply continue to publish incorrect interpretations of p and α (Carver 1978; Huberty 1993; Hubbard & Armstrong 2006); and (2) many researchers, having gone through the "castor oil" experience of studying research methods (Edwards 2008), never return to the methodology literature to further question what they think they know. Given the pressure to publish, they emulate others and do what editors and reviewers tell them to do (see point (1)), and so the GASSSPP lives on.

Whatever the causes, there are a number of things that those who want to raise the quality standards of our research can do, in terms of technical changes to our methods, authorship changes to our papers, and institutional changes to the research process. A list of these suggestions, the most important part of this paper, follows.

4.1 Technical Changes

Despite the long list of interpretational issues the GASSSP mythology raises, correction of the most problematic of these is actually quite simple. It requires only that researchers: (1) focus their assessment of outcomes on effect sizes; (2) use and report confidence intervals as the principal tool to assess the degree of uncertainty associated with outcomes; and (3) relegate p levels to the limited but important status they deserve in research, as an *a priori* value of acceptable risk of Type 1 error before data are collected. Scholars do not need to relearn statistical methods or software, but frequently will need to learn more about how to interpret the results. The marked references at the end of this paper provide a “quick start” list to accelerate this process.

There are many opportunities to improve the methods used in scholarly social-science research, many of which are more technical than the simple steps above. Ten such changes have recently been suggested by Edwards (2008), who also recommends confidence intervals and alternatives to NHST, and his paper is recommended for those who wish to consider more technical and specific changes to research practices.

4.2 Authorship Changes

Push (Back). Assuming that many contributors, editors, and reviewers in the GASSSP journals are not aware of the flaws in the GASSSP, a beginning step that authors can take is to prepare papers so that results are discussed in terms of effect sizes and confidence intervals. The confidence interval is favored by nearly all statisticians and methodologists who recommend metrics other than p for proper interpretation of results, and while other measures may be used, this one is familiar to anyone who has had basic instruction in statistics. Sadly, it appears to me that many researchers do not recognize that when they discuss correlation coefficients, difference scores, and many other metrics, these are the effect sizes that should be evaluated with a confidence interval; the fixation on p causes these important metrics to be overlooked.

Since many in the research review process must be assumed to be unfamiliar with the methodology literature, it falls on scholars to provide sufficient citations to this literature to make the case that p levels alone are never adequate to assess results. To facilitate this process, I have placed an extensive bibliography of references on my University of Delaware weblog, “Management Junk Science” (<http://sites.udel.edu/mjs/>). In addition to establishing a forum where authors can seek and find assistance in dealing with these problems, several sections of the blog provide prepared bibliographies that support proper interpretation of research outcomes. These can be used to “push” the case for correct interpretation in initial submissions.

An unfortunate element of the GASSSP is that the “peer review” process is actually not peer review, but rather a process that places a new author in the role of being a

Kmetz

supplicant before superiors who have published before (Starbuck 2003). Despite evidence of bias against null outcomes (Atkinson *et al.* 1982; Ceci & Peters 1984; Greenwald 1975) and growing recognition that peer review is really a process of reinforcing professional norms (Bedeian 2003, 2004) which does not assure quality in complex multivariate problems like social and biological sciences (Wager & Jefferson 2001; Jefferson *et al.* 2002a, b), it is considered the *sine qua non* of scientific quality. Thus, authors need strong support when their methods are questioned by reviewers, and these references can be used to “push back” when a reviewer or an editor incorrectly asserts that authors erred in not relying on p levels to interpret results. While diplomacy is always needed, it is possible to contest the judgments of reviewers, and in the area of interpretation of outcomes, it is quite likely this need will arise often.

Challenge. Closely related to pushing back is the need to challenge published, and often erroneous, reported results. This is a practice that is fundamental to real science, but is notable for its absence in the public work of the GASSSPP, and the social sciences in general. Since it is well established that the “voting” methods used in literature reviews to summarize previous work on a question (where previous work is often problematic in itself) yield incorrect and misleading results (Hunter & Schmidt 2004; Cohen 1990, 1994; Meehl 1967, 1990), authors are entirely capable of challenging many of the “methodological or statistical weaknesses” that are frequently cited as the reason for rejection of papers. The simple fact is that most reviewers do not know the deficiencies that characterize current research practices.

For example, colleagues in marketing inform me that it is increasingly common to require “replications” of research in the submission of marketing papers. When I inquired into this, I found that these “replications” are actually repetitions of the same study procedure by the same authors. I reject the characterization of this type of work as a “replication”—in my view, true replication implies independent work by another author or group. Repetition does not meet this standard. If a dedicated astronomer like Sir Percival Lowell can “see” canals on Mars because he “knows they are there” from the earlier work of Giovanni Schiaparelli, I find it difficult to accept the objectivity or detachment of a single author in a field as subjective as marketing.

Use the internet to full advantage. While the Internet has done much to ease and improve the research process, it has been an underutilized resource in the struggle to overcome the GASSSPP. Several things can be done with and through the Internet in this regard. First, scholars who want to see real science in the study of organizations need to form a global network. A group perhaps called the True Science Hypothesis Network (TSHN, or NHST backwards) could be organized to promote contact, support, and visibility for those who want to promote better science in b-school research. Given the near hegemony of the GASSSPP in the US, another reason to create a TSHN is to form a critical mass of scholars devoted to better science, and many of these will be outside the US and hence less geographically and organizationally concentrated. It is likely that a non-US university will be the host of such a network, given that past

Kmetz

experience shows little support for change away from the GASSSPP.

Second, existing organizations like the Social Science Research Network (SSRN) can be used to make information available quickly and efficiently, and without having to go through the gatekeeping process of peer review if the author desires. In that connection, I have agreed with Dr. Michael Jenkins, Chairman of SSRN, to establish a section of the Management Research Network tentatively entitled “Rosenthal’s File Drawer,” where papers can be published in a manner that addresses deficiencies of the GASSSPP. Specifically, there has long been need for a place to publish papers that resulted in a null outcome—sound studies that did not find what was expected, and generally are rejected for publication for that reason, and end up in the “file drawer.” A bias against negative findings is a problem in all of the sciences, and SSRN provides a repository where such studies can be made public.

In addition, replication studies can also be published through the SSRN “file drawer.” These may be from many sources, but one of the underutilized resources we have available to reexamine studies are our graduate students. It is not uncommon to find methodology courses where doctoral students are assigned studies to replicate as a training tool (and many of these studies do not replicate!). These can and should be made available to other researchers, and can be a disproportionately valuable form of information precisely because they are replications, regardless of the outcomes obtained.

Sponsor conferences requiring real science. Conferences and research forums are a hallmark of academic research, and our disciplines have seen remarkable global growth in the number of such conferences, the present one being an excellent example. An unexploited opportunity for advancement of real management science would be to organize a conference around such practices. Several key characteristics would be to: (1) solicit keynote addresses from scholars who champion and represent the correct approach to science and statistical interpretation; (2) invite papers that conform only to the standards of real science or revisit key issues with appropriate methods and reappraisals, such as meta-analyses; and (3) actively publish and promote such standards for all research and research institutions electronically as recommended above. A preliminary step would be to prepare a short “procedure and style guide” which makes it clear that standard GASSSPP research is ineligible for consideration, as well as stipulating what is; this would be announced with the call for papers. Planning will also require advance contact with the scholars and statisticians who would make the major addresses and chair several of the paper sessions.

4.3 Institutional Changes

There are many things that our universities and research organizations can do to lessen the grip of the GASSSPP on international and b-school research. Many of these are long-term changes that will require cooperative effort, and so I will summarize a few of these here, focusing on those things that we might influence as individuals.

Kmetz

The academic reward system is a fundamental obstacle to effecting change in US b-schools, and in most Tier 1 universities the journal-article “bean count” is the principal driver of academic success or failure, thus giving those with academic careers little choice but to publish in the GASSSPP journals. I am afraid that I must agree with Bennis & O’Toole (2005) that most US b-schools have “lost their way.” For non-US b-schools to avoid the same fate, they will need to be more creative in structuring these reward systems. Clearly, promotion and tenure documents that place higher value on truly scientific research than on GASSSPP productivity will be needed. Ironically, some of the potential for creation of reward systems that measure achievements against multiple objectives exists within our own HR literature, but we have not exploited that to our own academic purposes.

A second institutional change for many b-schools is to develop closer ties with the business community and management practitioners, similar to the professional relationships found in medicine, engineering, law, and other professional disciplines. Indeed, Khurana (2007) has concluded that development of management as a profession, one of the implied aims of the Ford and Carnegie Foundations, has yet to be fulfilled. In the US, many GASSSPP scholars are actually taking the position that joint or relevant research between academics and practitioners is not necessary (Bartunek 2007). I consider the latter to be evidence of how far b-schools have lost their way. Contact with the professional community can be a significant source of feedback and testable ideas, as has been demonstrated by the Marketing Science Institute in the US. No comparable organization exists within the broader realm of b-school research.

5. Conclusion

International business research is fortunate in many ways—this is a time in history when its potential to have a major positive impact on the development of the global economy has never been better, as is true of the global recognition of the role that business education can play in this development. The growth of business schools around the world unquestionably demonstrates that conviction. Yet the nature of that research is affected by many variables specific to the universities where it is done; among these the oldest and arguably most influential are the US b-schools, in which the dominant research model is the GASSSPP model. This model has ironically been the major obstacle to the achievement of the scientific objectives that are the heart of research work, and have been a major reason for the fact that the empirical research produced by US b-schools is of nearly no value whatsoever, for either the development of strong theories or improved practices. Fifty years of potential progress have been lost, and the community of international scholars must not fall into the same trap through emulation of the US model.

In my wanderings around 30-odd countries over the course of my career, I have encountered many international scholars and administrators who are concerned about the poor quality of international business research, and b-school research in general,

Kmetz

but do not know what to do about the problem. Unfortunately, I no longer believe that the US organizations steeped in the ways of the GASSSPP are willing to seriously confront these issues, and so it falls on international scholars to take those steps necessary to set b-school research on the path it should have followed decades ago. I believe that one of the most important things we must do to counter the GASSSPP is to support each other; in that connection, I humbly request that whenever someone who might benefit from the arguments in this paper comes to mind, please pass the reference to it in the online conference proceedings on to that individual or group. All of us know other scholars and journal editors; most of us know academic administrators who would like to know more about this issue; and many of us know administrators and colleagues in funding and donor organizations who might find it useful to evaluate proposals and results. Please feel free to pass this along—none of us can afford to lose more years.

References

- Anonymous 2005, '...and statistics', *The Economist*, **376**, 72
- Anonymous 2004, 'Economics focus: Signifying nothing?', *The Economist*, **370**, 76
- Anonymous 2004, 'Panning for gold in New Orleans', *The Economist*, **372**, 57-58.
- Anonymous 2007, 'Practically irrelevant?', *The Economist online*, **384**,
- Anonymous 2007, 'Signs of the times', *The Economist*, **382**, 94
- AACSB International 2007, *Impact of Research Draft*, St. Louis, Missouri: AACSB International.
- Abelson, RP 1995, *Statistics as principled argument*, 1st edn. Hillsdale, N.J.: Lawrence Erlbaum.
- Abrahamson, E. (2008) Torturing data until it speaks: Significance testing and the file drawer problem in organizational research. Paper presented during the symposium on Rejecting the Null Hypothesis, Academy Of Management National Meeting, Miami, Florida.
- Aiken, LS, West, SG, Sechrest, L, Reno, RR, Roediger III, HL, Scarr, S, Kazdin, AE and Sherman, SJ 1990, 'Graduate training in statistics, methodology, and measurement in psychology: a survey of PhD programs in North America', *American Psychologist*, **45** (6):721-734.
- Armstrong, JS 2007, 'Significance tests harm progress in forecasting', *International Journal of Forecasting*, **23**, 321-327.
- Armstrong, JS 2007, 'Statistical significance tests are unnecessary even when properly done and properly interpreted: Reply to commentaries', *International Journal of Forecasting*, **23**, 335-336.
- Atkinson, DR, Furlong, MJ and Wampold, BE 1982, 'Statistical significance, reviewer evaluations, and the scientific process: is there a (statistically) significant relationship?', *Journal of Counseling Psychology*, **29**, 189-194.
- Bakan, D 1966, 'The test of significance in psychological research', *Psychological Bulletin*, **66**, 423-437.

Kmetz

- Barnett, ML 2007, '(Un)learning and (mis)education through the eyes of Bill Starbuck: An interview with Pandora's playmate', *Academy of Management Learning & Education*, **6**, 114-127.
- Bartunek, JM 2007, 'Academic-practitioner collaboration need not require joint or relevant research: Toward a relational scholarship of integration', *Academy of Management Journal*, **50**, 1323-1333.
- Becker, BJ 1987, 'Applying tests of combined significance in meta-analysis', *Psychological Bulletin*, **102**, 164-171.
- Bedeian, AG 2003, 'The manuscript review process: the proper role of authors, referees, and editors', *Journal of Management Inquiry*, **12**, 331-338.
- Bedeian, AG 2004, 'Peer review and the social construction of knowledge in the management discipline', *Academy of Management Learning & Education*, **3**, 198-216.
- Bedeian, AG, Taylor, SG and Miller, AN 2010, 'Management science on the credibility bubble: Cardinal sins and various misdemeanors', *Academy of Management Learning & Education*, **9**, 715-725.
- Behrman, JN and Levin, RI 1984, 'Are business schools doing their job?', *Harvard Business Review*, **62**, 140-147.
- Bennis, WG and O'Toole, J 2005, 'How business schools lost their way', *Harvard Business Review*, 1-9.
- Berkson, J 1938, 'Some difficulties of interpretation encountered in the application of the chi-square test', *Journal of the American Statistical Association*, **33**, 526-542.
- Berkson, J 1942, 'Tests of significance considered as evidence', *Journal of the American Statistical Association*, **37**, 325-335.
- Beyer, JM and Trice, HM 1982, 'The utilization process: a conceptual framework and synthesis of empirical findings', *Administrative Science Quarterly*, **27**, 591-622.
- Bolles, R and Messick, S 1958, 'Statistical utility in experimental inference', *Psychological Reports*, **4** 223-227.
- Buckley, MR, Ferris, GR, Bernardin, HJ and Harvey, MG 1998, 'The disconnect between the science and the practice of management', *Business Horizons*, **41** (2):31-38. 0007-6813.
- Byrne, JA 1997, 'Commentary: Management theory—or fad of the month?', *Business Week*,
- Byrne, JA 1990, 'Is research in the ivory tower, 'fuzzy, irrelevant, pretentious'?', *Business Week*, 62-66.
- Carver, RP 1978, 'The case against statistical significance testing', *Harvard Education Review*, 378-399.
- Ceci, SJ and Peters, D 1982, 'Peer review: a study of reliability', *Change*, **14**, 44-48.
- Cohen, DJ 2007, 'The very separate worlds of academic and practitioner periodicals in human resource management: Reasons for the divide and concrete solutions for bridging the gap', *Academy of Management Journal*, **50**, 1013-1019.
- Cohen, J 1994, 'The earth is round ($p < .05$)', *American Psychologist*, **49**, 997-1003.

Kmetz

- Cohen, J 1962, 'The statistical power of abnormal-social psychological research: a review', *Journal of Abnormal and Social Psychology*, **65**, 145-153.
- Cohen, J 1990, 'Things I have learned (so far)', *American Psychologist*, **45**, 1304-1312.
- Crainer, S 1998, *Key management ideas: Thinkers that changed the management world*, 3rd edn. London: Financial Times Prentice Hall (Pearson Education).
- Cummings, TG 2007, '2006 Presidential Address: Quest for engaged Academy', *Academy of Management Review*, **32**, 355-360.
- David, RJ and Strang, D 2006, 'When fashion is fleeting: Transitory collective beliefs and the dynamics of TQM consulting', *Academy of Management Journal*, **49**, 215-234.
- Dawes, RM, Mirels, HL, Gold, E and Donahue, E 1993, 'Equating inverse probabilities in implicit personality judgments', *Psychological Science*, **4**, 396-400.
- Deadrick, DL and Gibson, PA 2007, 'An examination of the research-practice gap in HR: Comparing topics of interest to HR academics and HR professionals', *Human Resource Management Review*, **17**, 131-139.
- Duncan, WJ 1974, 'Management theory and the practice of management', *Business Horizons*, **17**, 48-52.
- Duncan, WJ 1974, 'Transferring management theory to practice', *Academy of Management Journal*, **17**, 724-738.
- Dunnette, MD 1966, 'Fads, fashions, and folderol in psychology', *American Psychologist*, **21**, 343-352.
- Dunnette, MD and Brown, ZM 1968, 'Behavioral science research and the conduct of business', *Academy of Management Journal*, **11**, 177-188.
- Edwards, JR 2008, 'To prosper, organizational psychology should ...overcome methodological barriers to progress', *Journal of Organizational Behavior*, **29**, 469-491.
- Edwards, JR and Berry, JW 2010, 'The presence of something or the absence of nothing: Increasing theoretical precision in management research', *Organizational Research Methods*, **13**, 668-689.
- Falk, R and Greenbaum, CW 1995, 'The amazing persistence of a probabilistic misconception', *Theory & Psychology*, **5**, 75-98.
- Faulkner, C, Fidler, F and Cumming, G 2008, 'The value of RCT evidence depends on the quality of statistical analysis', *Behaviour Research and Therapy*, **46**, 270-281.
- Ferguson, CJ 2009, 'An effect size primer: a guide for clinicians and researchers', *Professional Psychology: Research and Practice*, **40**, 532-539.
- Fidler, F, Thomason, N, Cumming, G, Finch, S and Leeman, J 2004, 'Editors can lead researchers to confidence intervals, but can't make them think', *Psychological Science*, **15**, 119-26.
- Finch, S, Cumming, G and Thomason, N 2001, 'Colloquium on Effect Sizes: the Roles of Editors, Textbook Authors, and the Publication Manual: Reporting of Statistical Inference in the *Journal of Applied Psychology*: Little Evidence of Reform', *Educational and Psychological Measurement*, **61**, 181-205.

Kmetz

- George, CS, Jr. 1968, *The history of management thought*, 1st edn. Englewood Cliffs, N.J.: Prentice-Hall.
- Gopinath, C and Hoffman, RC 1995, 'The relevance of strategy research: Practitioner and academic viewpoints', *Journal of Management Studies*, **32**, 575-594.
- Gordon, RA and Howell, JE 1959, *Higher education for business*, 1st edn. New York: Columbia University Press.
- Greatbatch, D and Clark, T 2005, *Management speak: Why we listen to what management gurus tell us*, 1st edn. London: Routledge.
- Greenwald, AG 1975, 'Consequences of prejudice against the null hypothesis', *Psychological Bulletin*, **82**, 1-20.
- Guest, DE 2007, 'Don't shoot the messenger: a wake-up call for academics', *Academy of Management Journal*, **50**, 1020-1026.
- Hambrick, DC 1994, '1993 Presidential address: What if the Academy actually mattered?', *Academy of Management Review*, **19**, 11-16.
- Hambrick, DC 2007, 'The field of management's devotion to theory: too much of a good thing?', *Academy of Management Journal*, **50**, 1346-1352.
- Hays, WL 1963, *Statistics*, 1st edn. New York: Holt, Rinehart and Winston.
- Hoffman, RC and Gopinath, C 1994, 'The importance of international business to the strategic agenda of U.S. CEOs', *Journal of International Business Studies*, **25**, 625-638.
- Hubbard, R and Armstrong, JS 2006, 'Why we don't really know what "statistical significance" means: a major educational failure', *Journal of Marketing Education*, **28**, 114-120.
- Hubbard, R and Bayari, MJ 2003, 'Confusion over measures of evidence (p 's) versus errors (α 's) in classical statistical testing (with comments)', *The American Statistician*, **57**, 171-182.
- Hubbard, R, Parsa, RA and Luthy, MR 1997, 'The spread of statistical significance testing in psychology: the case of the *Journal of Applied Psychology*', *Theory & Psychology*, **7**, 545-554.
- Hubbard, R and Ryan, PA 2000, 'The historical growth of statistical significance testing in psychology—and its future prospects', *Educational and Psychological Measurement*, **60**, 661-81.
- Huberty, CJ 1993, 'Historical origins of statistical testing practices: the treatment of Fisher versus Neyman-Pearson views in textbooks', *Journal of Experimental Education, Statistical Significance Testing in Contemporary Practice*, **61**, 317-333.
- Huczynski, A 2006, *Management gurus*, 1st edn. London: Routledge.
- Hunter, JE 1997, 'Needed: A ban on the significance test', *Psychological Science*, **8**, 3-7.
- Hunter, JE and Schmidt, FL 2004, *Meta-analysis: Correcting error and bias in research findings*, 2nd edn. Thousand Oaks, Ca.: Sage.
- Hunter, JE, Schmidt, FL and Jackson, GB 1982, *Meta-analysis: cumulating research findings across studies*, 1st edn. Beverly Hills, Ca.: Sage.

Kmetz

- Ioannidis, JPA 2005, 'Why most published research findings are false', *PLoS Med*, **2**, e124
- Jackson, SE and Dutton, JE 1988, 'Discerning threats and opportunities', *Administrative Science Quarterly*, **33**, 370-387.
- Jefferson, T, Alderson, P, Wager, E and Davidoff, F 2002, 'Effects of editorial peer review: A systematic review', *JAMA, The Journal of the American Medical Association*, **287**, 2784-2786.
- Jefferson, T, Wager, E and Davidoff, F 2002, 'Measuring the quality of editorial peer review', *JAMA, The Journal of the American Medical Association*, **287**, 2786-2790.
- Jones, LV 1955, 'Statistics and research design', *Annual Review of Psychology*, **6**, 405-430.
- Kalinowski, P and Fidler, F 2010, 'Interpreting significance: the differences between statistical significance, effect size, and practical importance', *Newborn & Infant Nursing Reviews*, **10**, 51-54.
- Khurana, R 2007, *From higher aims to hired hands: the social transformation of American business schools and the unfulfilled promise of management as a profession*, 1st edn. Princeton, NJ: Princeton University Press.
- Kilmann, RH, Thomas, KW, Slevin, DP, Nath, R and Jerrell, SL 1994, *Producing useful knowledge for organizations*, San Francisco: Jossey-Bass.
- Kish, L 1959, 'Some statistical problems in research design', *American Sociological Review*, **24**, 328-338.
- Kmetz, JL 1992, '*Proposals to improve the science of organization*', Paper presented to the Research Methods Division of the national meeting of the Academy of Management, Las Vegas, NV.
- Kmetz, JL 1998, *The information processing theory of organization: Managing technology accession in complex systems*, 1st edn. Aldershot, England: Ashgate.
- Kmetz, JL 2002, *The skeptic's handbook: Consumer guidelines and a critical assessment of business and management research*, 1st edn. Social Science Research Network.
- Kmetz, JL 2007, '*Science and the study of management: an opportunity to set the global standard for valid social science research*', Second International Conference on Corporate Governance and Corporate Social Responsibility, State University Higher School of Economics, Moscow, Russia.
- Kmetz, JL 2011, 'What "food chain?" The disregard of academic research in best-selling business books', *article in second review*.
- Lane, PJ, Koka, BR and Pathak, S 2006, 'The reification of absorptive capacity: a critical review and rejuvenation of the construct', *Academy of Management Review*, **31**, 833-863.
- Loftus, GR 1996, 'Psychology will be a much better science when we change the way we analyze data', *Current Directions in Psychological Science*, **5**, 161-171.
- Lykken, DT 1968, 'Statistical significance in psychological research', *Psychological Bulletin*, **70** (3):151-159.

Kmetz

- Lyles, MA 1990, 'A research agenda for strategic management in the 1990s', *Journal of Management Studies*, **27**, 363-375.
- McCloskey, DN 1992, 'The bankruptcy of statistical significance', *Eastern Economic Journal*, **18**, 359-361.
- McCloskey, DN and Ziliak, ST 1996, 'The standard error of regressions', *Journal of Economic Literature*, **XXXIV**, 97-114.
- McKelvey, B 2006, 'Response—Van de Ven and Johnson's "engaged scholarship": nice try, but...', *Academy of Management Review*, **31**, 822-829.
- Meehl, PE 1967, 'Theory-testing in psychology and physics: a methodological paradox', *Philosophy of Science*, **34**, 103-115.
- Meehl, PE 1978, 'Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology', *Journal of Consulting and Clinical Psychology*, **46**, 806-834.
- Meehl, PE 1986, What social scientists don't understand. In: Fiske, D.W. and Shweder, R.A., (Eds.) *Metatheory in social science pluralisms and subjectivities*, Chicago: University of Chicago Press.
- Meehl, PE 1990, 'Why summaries of research on psychological theories are often uninterpretable', *Psychological Reports*, 66 (Monograph Supplement 1-Vol. 66), 195-244.
- Miller, DW Aug 6, 1999 'The black hole of education research', *The Chronicle of Higher Education*,
- Mol, MJ and Birkinshaw, J 2007, *Giant steps in management: Innovations that change the way we work*, 1st edn. London: Financial Times Prentice Hall (Pearson Education).
- Moonesinghe, R, Khoury, MJ and Janssens, ACJW 2007, 'Most published research findings are false—but a little replication goes a long way', *PLoS Med*, **4**, e28.
- Murray, ML 1993, 'On the significance of significance (the significance of statistical information)', *CPCU Journal*, **46**, 67-69.
- Nix, TW and Barnette, JJ 1998, 'The data analysis dilemma: Ban or abandon. A review of null hypothesis significance testing', *Research in the Schools*, **5**, 3-14.
- Nunnally, JC 1960, 'The place of statistics in psychology', *Educational and Psychological Measurement*, **20**, 61-650.
- Oakes, MW 1986, *Statistical inference: a commentary for the social and behavioural sciences*, 1st edn. New York: Wiley.
- Oviatt, BN and Miller, WD 1989, 'Irrelevance, intransigence, and business professors', *Academy of Management Executive*, **III**, 304-312.
- Pearce, JL 2004, 'Presidential address: What do we know and how do we really know it?', *Academy of Management Review*, **29**, 175-179.
- Peters, D and Ceci, SJ 1982, 'Peer-review practices of psychological journals: the fate of published articles, submitted again', *The Behavioral and Brain Sciences*, **5**, 187-195.
- Pfeffer, J and Fong, CT 2002, 'The end of business schools? Less success than meets the eye', *Academy of Management Learning & Education*, **1**, 78-95.

Kmetz

- Pierson, FC 1959, *The education of American businessmen; a study of university-college programs in business administration*, 1st edn. New York: McGraw-Hill.
- Podsakoff, PM, MacKenzie, SB, Bachrach, DG and Podsakoff, NP 2005, 'The influence of management journals in the 1980s and 1990s', *Strategic Management Journal*, **26**, 473-488.
- Podsakoff, PM, MacKenzie, SB, Podsakoff, NP and Bachrach, DG 2008, 'Scholarly influence in the field of management: a bibliometric analysis of the determinants of university and author impact in the management literature in the past quarter century', *Journal of Management*, **34**, 641-720.
- Porter, LW and McKibbin, LE 1988, *Management education and development: Drift or thrust into the 21st century?*, 1st edn. New York: McGraw-Hill.
- Rosenthal, R 1979, 'The "file drawer problem" and tolerance for null results', *Psychological Bulletin*, **86**, 638-641.
- Rosnow, RL and Rosenthal, R 1988, 'Focused tests of significance and effect size estimation in counseling psychology', *Journal of Counseling Psychology*, **38**, 203-208.
- Rosnow, RL and Rosenthal, R 1989, 'Statistical procedures and the justification of knowledge in psychological science', *American Psychologist*, **44**, 1276-1284.
- Rousseau, DM 2006, 'Presidential address: Is there such a thing as "evidence-based management"?', *Academy of Management Review*, **31**, 256-269.
- Rousseau, DM and McCarthy, S 2007, 'Educating managers from an evidence-based perspective', *Academy of Management Learning & Education*, **6**, 84-101.
- Rozeboom, WW 1960, 'The fallacy of the null-hypothesis significance test', *Psychological Bulletin*, **57**, 416-428.
- Rynes, SL, Giluk, TL and Brown, KG 2007, 'The very separate worlds of academic and practitioner periodicals in human resource management: Implications for evidence-based management', *Academy of Management Journal*, **50**, 987-1008.
- Safón, V 2007, 'Factors that influence recruiters' choice of B-schools and their MBA graduates: Evidence and implications for B-schools', *Academy of Management Learning & Education*, **6**, 217-233.
- Scandura, TA and Williams, EA 2000, 'Research methodology in management: Current practices, trends, and implications for future research', *Academy of Management Journal*, **43**, 1248-1264.
- Schmidt, FL 1996, 'Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers', *Psychological Methods*, **1**, 115-129.
- Schmidt, FL 1992, 'What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology', *American Psychologist*, **47**, 1173-1181.
- Shrout, PE 1997, 'Should significance tests be banned? Introduction to a special section exploring the pros and cons', *Psychological Science*, **8**, 1-2.
- Skapinker, M Jan 7, 2008, 'Why business ignores the business schools', *Financial Times online*.

Kmetz

- Sohn, D 2000, 'Significance testing and the science', *American Psychologist*, **55** (8):964-965.
- Starbuck, WH 2006, *The production of knowledge: the challenge of social science research*, 1st edn. New York: Oxford.
- Starbuck, WH 2003, 'Turning lemons into lemonade: Where is the value in peer reviews?', *Journal of Management Inquiry*, **12**, 344-351.
- Sterne, JAC 2001, 'Sifting the evidence—what's wrong with significance tests? ', *British Medical Journal*, **27**, 115-129.
- Tahai, A and Meyer, MJ 1999, 'A revealed preference study of management journals' direct influences', *Strategic Management Journal*, **20**, 279-296.
- Thompson, B 1998, 'Statistical significance and effect size reporting: Portrait of a possible future', *Research in the Schools*, **5**, 33-38.
- Trank, CQ and Rynes, SL 2003, 'Who moved our cheese? Reclaiming professionalism in business education', *Academy of Management Learning & Education*, **2** (2):189-205.
- Tukey, JW 1960, 'Conclusions vs. decisions', *Technometrics*, **26** (4):
- Tyler, RW 1931, 'What is statistical significance?', *Educational Research Bulletin*, **10**, 115-142.
- Van de Ven, AH 2000, 'Presidential address: Our roles in the changing field of management education and research', *Academy of Management News*, **31**, 1-2.
- Van de Ven, AH and Johnson, PE 2006, 'Knowledge for theory and practice', *Academy of Management Review*, **31**, 802-821.
- Wager, E and Jefferson, T 2001, 'Shortcomings of peer review in biomedical journals', *Learned Publishing*, **14**, 257-263.
- Wang, LL 2010, 'Retrospective statistical power: Fallacies and recommendations', *Newborn & Infant Nursing Reviews*, **10**, 55-59.
- Webber, RA, Ference, TP, Fox, WM, Miles, RE and Porter, LW 1971, Behavioral science and management: Why the troubled marriage?, Atchison, T.J. and Ghorpade, J.V., (Eds.) *Academy of Management Proceedings Annual Meeting*, 377-395.
- Wharton History, <http://www.wharton.upenn.edu/about/wharton-history.cfm>. 2010.
- Wilkinson, L and the Task Force on Statistical Inference 1999, 'Statistical methods in psychology journals: Guidelines and explanations', *American Psychologist*, **54** (8):594-604.
- Ziliak, ST and McCloskey, DN 2008, *The cult of statistical significance: how the standard error costs us jobs, justice, and lives*, 1st edn. University of Michigan Press.

Appendix*

Books Selected for Reference Analysis in Table 1

Books are ordered following the code sequence in Table 1, not alphabetically. Author academic affiliations, if any, are noted with the reference.

- 8H Covey, S.R.
2004 The 8th habit: from effectiveness to greatness. New York: Free Press.
Academic affiliation: None (formerly Brigham Young)
Theme: Leadership, Management
- 101 Albright, M. and C. Carr.
1997 101 biggest mistakes managers make and how to avoid them. Paramus, NJ: Prentice-Hall.
Academic affiliation: None
Theme: Management
- AL George, B.
2003 Authentic leadership: Rediscovering the secrets to creating lasting value. San Francisco, CA: Jossey-Bass.
Academic affiliation: None
Theme: Leadership, Performance
- BDB Carter, C.B. and J.W. Lorsch
2004 Back to the drawing board: designing corporate boards for a complex world. Boston, MA: Harvard Business School Publishing.
Academic affiliation: CC, None; JL, Harvard
Theme: Governance, Organization
- BOS Kim, W.C. and R. Mauborgne
2005 Blue ocean strategy: how to create uncontested market space and make competition irrelevant. Boston, MA: Harvard Business School Publishing.
Academic affiliation: both INSEAD
Theme: Marketing, Strategy
- CFF+ Hamel, G. and C.K. Prahalad
1994 Competing for the future. Boston: Harvard Business School Press.
Academic affiliation: GH, London BS; CP, Michigan
Theme: Strategy, Performance
- CR Bossidy, L. and R. Charan
2004 Confronting reality: Doing what matters to get things right. New York: Crown Business.
Academic affiliation: None

Kmetz

Theme: Globalization, Strategy

- CSH Micklethwait, J. and A. Wooldridge
2005 The company: a short history of a revolutionary idea. New York: Modern Library.
Academic affiliation: None
Theme: Organization, History
- CST McAfee, R.P.
2002 Competitive solutions: the strategist's toolkit. Princeton, NJ: Princeton University Press.
Academic affiliation: Cal Tech
Theme: Marketing, Strategy
- CWP Abrahamson, E.
2004 Change without pain: how managers can overcome initiative overload, organizational chaos, and employee burnout. Boston, MA: Harvard Business School Publishing.
Academic affiliation: Columbia Univ.
Theme: Management, Organization
- DC Moore, J.F.
1996 The death of competition: Leadership and atrategy in the age of business ecosystems. New York: HarperCollins.
Academic affiliation: None
Theme: Leadership, Strategy
- EEQ Cooper, R.K.
1996 Executive EQ: Emotional intelligence in business. New York: Berkley Publishing Group.
Academic affiliation: None
Theme: Leadership, Management
- FP+ Micklethwait, J. and A. Wooldridge
2000 A future perfect: the challenge and hidden promise of globalization. New York: Crown.
Academic affiliation: None
Theme: Globalization
- FRK Levitt , S.D. and S.J. Dubner
2005 Freakonomics - a rogue economist explores the hidden side of everything. New York: HarperCollins.
Academic affiliation: SL, Chicago; SD, None
Theme: Economics
- G2G Collins, J.

Kmetz

- 2001 Good to great: Why some companies make the leap... and others don't. New York: HarperCollins.
Academic affiliation: None (formerly Stanford lecturer)
Theme: Leadership, Strategy
- GBL Csikszentmihalyi, M.
2003 Good business: Leadership, flow, and the making of meaning. New York: Penguin.
Academic affiliation: Claremont
Theme: Leadership, Performance
- GN Salacuse, J.W.
2003 The global negotiator: Making, managing, and mending deals around the world in the twenty-first century. New York: Palgrave Macmillan.
Academic affiliation: Tufts
Theme: Negotiation, Globalization
- HOW Carlson, T.A.
2005 The how of WOW: a guide to giving a speech that will positively blow 'em away. New York: AMACOM.
Academic affiliation: None
Theme: Communication
- IS Christensen, C.M. and M.E. Raynor
2003 The innovator's solution: Creating and sustaining successful growth. Boston, MA: Harvard Business School Publishing.
Academic affiliation: CC, Harvard; MR, None
Theme: Innovation, Strategy
- M6S Breyfogle, F.W.III., J.M. Cupello, and B. Meadows 2001
Managing Six Sigma: A Practical Guide to Understanding, Assessing, and Implementing the Strategy That Yields Bottom-Line Success. New York: John Wiley.
Academic affiliation: None
Theme: Quality, Strategy
- MFO Roberts, J.
2004 The modern firm: Organizational design for performance and growth. New York: Oxford.
Academic affiliation: Stanford
Theme: Organization, Performance
- PZ Slywotzky, A.J., D.J. Morrison, and B. Andelman
2001 The profit zone: how strategic business design will lead to tomorrow's profits. New York: Three Rivers Press.

Kmetz

Academic affiliation: None
Theme: Strategy, Performance

- RC MacAvoy, P. and I. Millstein
2003 The recurrent crisis in corporate governance. New York: Palgrave Macmillan.
Academic affiliation: PM, Chicago; IM, None
Theme: Governance, Finance
- SFO Kaplan, R.S. and D.P. Norton
2001 The strategy-focused organization: how balanced scorecard companies thrive in the new business environment. Boston, MA: Harvard Business School Publishing.
Academic affiliation: both Harvard
Theme: Strategy, Performance
- TC Luttwak, E.N. and N. Weidenfeld
1999 Turbo-capitalism: Winners and losers in the global economy. New York: HarperCollins.
Academic affiliation: None
Theme: Globalization, Strategy
- TM Martin, C.
2005 Tough management: the 7 winning ways to make tough decisions easier, deliver the numbers, and grow the business in good times and bad. New York: McGraw-Hill.
Academic affiliation: None
Theme: Decision making, Performance
- WD Micklethwait, J. and A. Wooldridge
1996 The witch doctors: Making sense of the management gurus. New York: Times Books.
Academic affiliation: None
Theme: Management, Consulting
- WMI Magretta, J.
2002 What management is: how it works and why it's everyone's business. New York: Free Press.
Academic affiliation: None
Theme: Management, Performance
- WNG Welch, J. and S. Welch
2005 Winning. New York: HarperCollins.
Academic affiliation: None
Theme: Management, Leadership

Kmetz

WRW+ Joyce, W., N. Nohria, and B. Roberson
2003 What really works: The 4+2 formula for sustained business success. New York:
Harper Business.
Academic affiliation: WJ, Dartmouth; Nohria, Harvard; BR, None
Theme: Management, Organization

+ Books added to the sample by the author.

* Source: Kmetz, John L., What “food chain?” The disregard of academic research in best-selling business books.” Paper in review, 2011.