CONTINGENCY HYPOTHESES IN STRATEGIC MANAGEMENT RESEARCH: USE, DISUSE, OR MISUSE?

Brian K. Boyd
W.P. Carey School of Business
Arizona State University
Tempe, AZ 85287
briankboyd@asu.edu

Katalin Takacs Haynes
Alfred Lerner College of Business and Economics
University of Delaware
Newark, DE 19716
Katalin.takacs@gmail.com

Michael A. Hitt
Mays Business School
Texas A&M University
College Station, TX 77843-4221
mhitt@cgsb.tamu.edu

Donald D. Bergh
Daniels College of Business
University of Denver
Denver, CO 80208
dbergh@du.edu

David J. Ketchen, Jr.
College of Business
Auburn University
Auburn, AL 36849-5241
ketchda@auburn.edu

Journal of Management 2012 Annual Review, forthcoming

Keywords: contingency theory, strategic management, research methods

Running Head: Contingency Hypotheses

CONTINGENCY HYPOTHESES IN STRATEGIC MANAGEMENT RESEARCH: USE, DISUSE, OR MISUSE?

ABSTRACT

The answer to many strategic management research questions is often summarized as 'It depends.' Faced with marginal results of many main effect hypothesis tests of one variable on another variable, strategy researchers began developing contingency hypotheses that explored more nuanced relationships involving multiple variables. Herein, we examine the development of contingency thinking in strategic management via a review of all empirical papers published in *Strategic Management Journal* from its inception in 1980 through 2009. Using Venkatraman's (1989) framework, we identify all contingency studies within this sample. Our analysis reveals that, while contingency hypotheses are becoming more common, there is less diversity in the way the effects are tested. Additionally, while the framing of contingency hypotheses has become more sophisticated over time, there remain many opportunities for methodological improvements. Based on this content analysis, we offer both theoretical and methodological guidelines for future strategic management studies.

"The world changed from having the determinism of a clock to having the contingency of a pinball machine."

Physicist Heinz R. Pagels, 1982

Distinguishing between effective and ineffective management practices is a primary focus for much of management research. In the early part of the 20th century, Frederick Taylor and other management pioneers sought to identify the 'one best way' of solving managerial problems. From this point of view, organizations were best understood in terms of direct, linear relationships between variables often represented by main or direct effects. For example, dozens of studies have examined potential direct effects between aspects of corporate governance such as CEO duality (i.e., whether the CEO also chairs the board of directors) and the balance of inside and outside directors on the one hand and firm performance on the other (Finkelstein &Hambrick, 1996; Finkelstein, Hambrick & Cannella, 2008). Dalton, Daily, Ellstrand and Johnson's (1998) synthesis of these studies suggested that very little evidence exists in favor of simple direct effects and this was confirmed in a recent review of the research (Dalton, Hitt, Certo& Dalton, 2008). This is perhaps not surprising: Direct effects are often very important, but they seem incapable of fully capturing the complexity of organizations. Similar findings, or non-findings, can be found in meta-analyses and quantitative reviews of a number of core research streams in the field of strategic management: For example, the direct effect of strategic planning on performance is marginal at best (Boyd, 1991; Miller & Cardinal, 1994), acquisitions have little direct effect on performance (King, Dalton, Daily & Covin, 2004), and tests of transaction cost economics assumptions are evenly split between supportive and non-supportive results (David & Han, 2004).

In our opening epigraph, the citation of Pagels refers to changes within the field of physics but his analogy applies equally well to shifts within the management field that began in the 1950s and 1960s in reaction to the limitations of previous theorizing. In contrast to the deterministic approach of the 'one best way' perspective, a contingency perspective on organizations suggests that, similar to

the complex movement of pinballs, most relationships between two variables are influenced by other variables. Burns and Stalker (1961), for example, theorized that the manner in which an organization achieves fit with its environment depends on the nature of the environment. If environmental change is slow and fairly predictable, organizations benefit from creating mechanistic structures that maximize efficiency. If change is rapid and disjointed, organic structures that sacrifice efficiency but allow flexibility create the most value. Following other landmark conceptual efforts such as books by Lawrence and Lorsch (1967) and Thompson(1967), contingency thinking became a highly viable perspective on management by the early 1970s. Additional scholars began to apply contingency thinking to new areas, such as Galbraith's (1973) work on organization design.

The 1970s also gave rise to strategic management as an organized subdiscipline within the management field. Strategic management coalesced around the question of why some organizations succeed while others fail (e.g., Child, 1972; Miles & Snow, 1978) and this question remains the cornerstone of strategic management research(Nag, Hambrick& Chen, 2007). Indeed, the first empirical article published in *SMJ* (Jauch, Osborn & Glueck, 1980) applied contingency methods in an attempt to answer this very question. Across the ensuing decades, contingency-based studies have made important contributions to strategic management (Peteraf& Reed, 2007;Zott&Amit, 2008). Consider the null findings for direct effects regarding governance mentioned in the preceding paragraph. Following the findings by Dalton, Daily, Ellstrand and Johnson (1998), subsequent work has been characterized by more nuanced theorizing that has fueled the identification of contingency effects. For example, Combs, Ketchen, Perryman and Donahue (2007) drew on agency and power circulation theories to show that CEO power moderates the relationship between board composition and performance. Similarly, Westphal (1998) developed a mediated model involving board factors, CEO behavior, and subsequent firm performance. Findings such as these have helped refine theories of corporate governance and offered empirically-grounded normative recommendations for

managers. This offers one example among many in which contingency thinking has provided the foundation for advancing knowledge about important relationships.

An important clarification is the distinction between contingency hypotheses and contingency theory. Schoonhoven (1981: 350) characterized contingency not as an actual theory, but rather "an orienting strategy or metatheory, suggesting ways in which a phenomenon ought to be conceptualized or an approach to the phenomenon ought to be explained." Contingency as a theory came under criticism in the 1980s, for a variety of methodological and theoretical issues. In recent years, however, there has been a resurgence with strategy scholars citing contingency theory as the basis for their hypotheses – a finding demonstrated in our subsequent content analysis.[1]

One of the main critiques of contingency has been the ambiguity surrounding the analysis of contingency hypotheses. While "it depends" may be viewed as the foundation of a general research question, there are many ways to frame specific research questions related to it. Consequently, there is much uncertainty regarding how contingency hypotheses are examined within strategic management research. Contingency takes several different specific forms and an array of analytical techniques are available to test contingency hypotheses. Given this complexity, significant potential exists for mismatches between how hypotheses are framed and the analytic tools used to test them (Hitt, Boyd & Li, 2004). This dilemma and its possible negative implications for research have been discussed for several decades. For example, Galbraith and Nathanson (1979) commented on how a lack of precision within contingency studies in strategic management could inhibit knowledge creation, and their thoughts were echoed and extended by Venkatraman (1989).

To date, however, no systematic examination of contingency hypotheses and how they are tested within strategic management studies exist in the literature. One implication is that researchers and managers are unaware of the validity and robustness of the insights derived from contingency

[1]Peteraf and Reed (2007) offer an excellent history of contingency as a theoretical perspective.

studies. To help fill this gap in knowledge, we content analyzed 1,173 empirical strategic management articles published between 1980 and 2009 to determine the extent to which contingency perspectives have been used, disused, and misused. Based on our findings, we offer a series of suggestions for improving future contingency-based studies. Our suggestions encompass both theory and methods, and we integrate these two aspects by proposing a number of contingency research opportunities on several strategic management topics.

CONTENT ANALYSIS OF CONTINGENCY STUDIES

There is a growing research stream that focuses on methodological issues in strategic management research. The intent of this stream is to identify flaws or weaknesses associated with current practices, and subsequently improve the methodological rigor of future studies. We used reviews of this stream (Bergh, Perry & Hanke, 2006; Boyd, Gove & Hitt, 2005a; Hitt et al, 2004) to benchmark design attributes of the current project. Because our goal was to assess trends in the use of contingency models, we sampled a broad time horizon, using a 30-year window of strategic management scholarship. Because we also expected contingency models to be used in only a subset of studies, we sampled all available articles during this time horizon.

We limited our scope to papers published in *Strategic Management Journal*(*SMJ*), versus a broader journal pool. Therefore, the sampling frame for our content analysis includes all articles published in *SMJ* from its inception in 1980 through the end of 2009. We focused on *SMJ* articles because they unequivocally represent strategic management research. Indeed, Nag, Hambrick, and Chen (2007: 938) suggest that the publication of an article in *SMJ* is seen by strategy researchers as "providing *prima facie* evidence that it was an SM article." This approach avoids the complication of deciding which papers published in general management outlets (e.g., *AMJ*,*ASQ*, *JOM*, or *JMS*) are in the strategy domain. Many other methodological reviews have used this single journal approach

(e.g., Bergh, Hanke, Balkundi, Brown & Chen, 2004; Bergh & Holbein, 1997; Ferguson & Ketchen, 1999; Stone-Romero, Weaver & Glenar, 1995).

Further, we collected a smaller, second sample to assess the generalizability of our results. Specifically, we analyzed articles in *AMJ* that were published in 1980, 1990, and 2000. After reading each article in these three volumes, we identified a subset of articles (a) whose content represented the domain of strategic management, and (b) which used at least one of the contingency tools in Venkatraman's (1989) taxonomy. We then compared these articles to the *SMJ* contingency papers in corresponding years. This comparison revealed very strong similarities between articles in the two journals. For example, there was no statistically significant difference in the sample size across journals in any of the three comparison periods. The mean number of contingencies used per article was also not significantly different. Furthermore, the relative use of different contingency tools was stable across the two samples: Interaction was the predominant tool in both journals, followed by subgroup analysis. For a more formal comparison, results of a Spearman test based on rank orders of the different tools indicated that the two samples correlated at 0.79. These findings enhanced our confidence that *SMJ* articles could capably represent strategic management research.

More broadly, several multi-journal comparisons support the generalizability of our findings. Mone, Mueller and Mauland (1996), for example, reported that the power levels of small, medium, and large effect sizes did not differ significantly between *SMJ* and *AMJ*. In a similar vein, Boyd, Gove and Hitt (2005a) found no significant differences in measurement practices between *SMJ* and strategy articles published in *AMJ, ASQ,* and *Management Science*. Moreover, the methodological practices documented in other journals, including the *Academy of Management Journal, Personnel Psychology* and the *Journal of Applied Psychology* are similar to those we observed in the *SMJ* (Aguinis, Beaty, Boik,& Pierce,2005; Aguinis & Gottfredson, 2010; Aguinis, Petersen, & Pierce, 1999). Thus, focusing on the *SMJ* provides a type of indepth longitudinal case study opportunity to

examine the development of theories and research practices within the field of strategy while simultaneously observing rules and procedures that appear to represent universal practices across a range of literatures.

Starting with an initial pool of 1,715 articles that were published between 1980 and 2009, we initially identified all quantitative studies for further analysis. We excluded conceptual papers, meta-analyses, simulations, and purely qualitative articles, which left 1,173 articles for additional evaluation. Next, two expert raters reviewed these 1,173 empirical articles to determine which used contingency tools as described in Venkatraman's (1989) framework. Coders discussed each article until they came to a consensus regarding its classification. For articles that included a contingency analysis, coders also collected data regarding sample size, nature of the data, article keywords, theories used to frame hypotheses, and topic areas. Additional data were also collected based on the specific contingency models used in the individual articles, and are described in more detail in subsequent sections.

Our unit of analysis is the individual article. When an article reported multiple tests of a particular tool, we coded initially at the level of the individual test, and then aggregated our data at the article level. A number of articles used multiple types of contingencies – for example, a set of mediation hypotheses might be compared across multiple subgroups. Other studies combined both interaction and subgroup analyses to test hypotheses. In these cases of multiple types of contingencies, we included an article in the content analysis for each respective tool. Finally, we report our analysis using composite time windows versus individual years. Specifically, we compare trends in tool use across the 1980s, 1990s, and 2000s. We initially developed comparisons using a number of smaller time windows (e.g., 2, 3, and 5 year breaks), but the 10-year windows provide a better sense of overall trends over time.

Our content analysis is organized as follows: First, we provide information regarding basic characteristics of the sample, including frequency of different types of articles, sample size and type of data used over time. Next, we examine trends in the framing of these studies, including topic areas and theoretical perspectives. Third, we describe overall trends in the types of contingency tools used, including the nature of these tests as well as the result of hypothesis tests. Finally, we review the specific characteristics of studies vis-à-vis the specific contingency tools used (e.g., mediation or interaction).

Descriptive characteristics

Overall article trends. The field of strategic management has grown substantially over the last three decades, a trend underscored by the growth of *SMJ* over time. For example, the initial volume of *SMJ* in 1980 consisted of four issues and 386 pages. In comparison, the 2009 volume of *SMJ* consisted of 12 issues, and 1,348 pages. The number of articles published annually in *SMJ* has grown from less than 30 in 1980 to upwards of 70 in the late 2000s. *SMJ* publishes predominantly empirical articles, and the prevalence of contingency-based studies has risen steadily since 1980, both in terms of the overall number of such studies, and as a percent of all empirical articles. In the 1980s, less than one in three empirical articles had a contingency element. In contrast, by the 2000s over half of all empirical articles had a contingency component. We identified a total of 568 contingency-based studies, which we describe in greater detail in the subsequent sections.

Data sources. Overall, approximately two-thirds of contingency studies relied on archival sources, while one quarter of studies used surveys. The remaining articles (11.8 percent) used other approaches, such as laboratory studies. The type of data used became more homogeneous with the passage of time: In the 1980s, 42 percent of studies relied on archival data. That percentage grew to 57.6 percent in the 1990s, and 66.2 percent in the 2000s. The proportion of studies relying on survey data has declined from 34 percent in the 1980s to 24 percent in the 2000s. Studies using other data

types declined from one quarter of all articles in the 1980s to less than ten percent of articles in the 2000s.

Sample size. The sample size used in contingency studies varied widely: While the average sample size was 4,523[2], it ranged from less than a dozen to more than one million observations. Additionally, as shown in Figure 1, sample size statistics were also very different for studies relying on survey versus archival data. While it appears from the Figure that sample sizes steadily increased over time, this trend is largely not statistically significant. To perform t-tests on the average sample size, we ran pairwise comparisons of studies published in the 1980s, 1990s, and 2000s, and separately for studies relying on archival versus survey data. For survey-based studies, the gain in sample size from the 1980s to subsequent decades was statistically significant at p=.07 and .06 levels, respectively, while there was no statistically significant difference between the 1990s and 2000s. For studies using archival data, the difference between the 1980s and 2000s was marginally statistically significant (p=.054), while other comparisons were also marginally statistically significant at p=.10. Overall, then, there is relatively little gain in average sample size among contingency studies during the period studied. Sample size has implications for statistical power, as discussed further in the context of specific contingency tools.

==========
Insert Figure 1 about here
==========

Framing of studies

Topic areas. In order to categorize the topic areas of contingency studies, we developed a taxonomy of fourteen categories. We considered multiple approaches to develop the classification scheme, including a review of all article keywords used in the sample pool, histories of the field, and listings used to classify conference and journal submissions. After careful consideration, we adopted

[2] The median sample size was 246.

the approach used by a recent content analysis of industrial and organizational psychology articles (Cascio & Aguinis, 2008), that synthesized topic chapters used in psychology textbooks. Consequently, we reviewed a number of prominent strategic management textbooks, and also discussed the taxonomy with an author of multiple strategic management textbooks. Based on this review, we developed a list of main topics, each encompassing a set of subgroup topics. These topics were identical to those identified in historical reviews of the strategic management literature (see Bowman, Singh & Thomas, 2002). We then made additional refinements after pilot testing randomly selected years of articles. Our final taxonomy consists of 14 categories, each with 4-7 subcategories. Main categories and representative subcategories are shown in Table 1, along with the proportion of studies referencing each topic over time. Because some articles contain multiple topic areas, the percentages sum to greater than one.

==========

Insert Table 1about here

==========

Overall, the most prominent topic areas were competitiveness and returns, external and internal environments, and organizational structure and controls. Strategic intent and mission and strategic entrepreneurship were the least emphasized areas. However, as shown in Table 1, emphasis on these different topics changed substantially over time. During the 1980s, articles in *SMJ* were likely to focus on one of three primary areas: Competitiveness, the external environment, and the firm's internal environment. During the 1990s, interest in these areas diminished, but they still represented dominant topics. A major shift in research during the 1990s was for contingency studies to emphasize cooperative strategies, such as joint ventures, alliances, and outsourcing arrangements. Strategic intent and mission statements also experienced greater attention in the 1990s, as did corporate acquisitions.

In the 2000s, firm environments and competitiveness continued their decline in emphasis, but still remained prominent areas: For example, roughly twenty percent of contingency studies in the 2000s included either the internal or external environment as a primary topic. Similarly, more than a third of studies in the 2000s addressed competitiveness. Given their prominence over three decades, then, these three topic areas might be considered to constitute the core domain of strategic management contingency studies. Another major shift in studies during the 2000s was increased scholarly attention to knowledge and learning topics: roughly 17 percent of studies in our sample included an emphasis in this area. Focus on cooperative strategies also grew substantially during this time period, as did on international topics.

Theoretical perspectives used. We also classified articles based on their theoretical framing. One challenge in completing this stage of coding was the ambiguity in what constitutes a "theory". For example, some articles are quite explicit in the particular theory used as the basis for hypotheses. In contrast, other articles draw on a particular research stream or area, such as networks, or knowledge. Thus, in order to classify the perspective of each article, our two coders entered up to three theoretical labels per article. Additionally, coders assigned labels to some articles based on their evaluation of the article, even if this designation was not explicit in the text of the article. For example, an article published in the early 1980s on top manager characteristics did not use the label "upper echelons", because the label had not been coined at that time.

Table 2 lists the top ten perspectives used to develop hypotheses for each of the three decades of our sample. As shown in the table, there is both stability and change in the perspectives used to frame contingency hypotheses. In the 1980s, Industrial Organization (IO) economics was the most frequently used perspective, followed by contingency. The Structure-Conduct-Performance model

and organizational theory[3] were the next most common perspectives, with the remaining perspectives experiencing only limited use. In the 1990s, IO economics retained its top spot, but was followed closely by both the resource-based view (RBV) and agency theory. Contingency theory and TCE were the fourth and fifth most common frameworks used in this decade. Again, the remaining theoretical perspectives on the top ten list were used sparingly. For the 2000s, the resource-based view became the clear perspective of choice for contingency studies, appearing more than twice as often as any other framework. Several new categories appeared in the 2000s: Networks and knowledge were ranked as the third and fourth most commonly used frameworks, and social capital and signaling theory appeared in the ninth and tenth ranks. Transaction cost economics remained in the fifth slot, and was followed closely by contingency. The greatest change was for IO economics, which dropped from first, in the 1980s and 1990s, to the eighth position in the 2000s.

==========
Insert Table 2and Figure 2 about here
==========

Trends in the use of contingency tools

We used Venkatraman's typology of approaches to examine "fit" between three or more variables (1989). We reviewed the hypothesis and methods sections of each of the 1,173 empirical articles published during our time window to identify the use of contingency tools. Thus, our content analysis included possible contingencies for both hypothesized and control relationships. Figure 2 shows the distribution of *SMJ* articles in aggregate over time. As depicted in the figure, the overall distribution of articles is roughly split into thirds between conceptual papers, papers using contingency tools, and empirical papers without contingency tools. Interaction is clearly the most commonly used contingency tool, with twice as many interaction effects as subgroup analyses; both of which are used to test for moderation. Mediation was the next most frequently used tool, although

[3] While 'organizational theory' seems a broad label – e.g., it could include resource dependence, population ecology, or environmental determinism, for example – it was the term used to describe the theoretical basis of the studies in question.

it was clearly less prevalent: For example, for each test of mediation, there were more than six tests of moderation. Each of the remaining tools was only rarely used.

The distribution of contingency tool use also varied substantially over time. In the 1980s, any of the six categories of tools was found in only a small percentage of articles – e.g., interaction effects were reported in roughly five percent of articles published in the early 1980s, with smaller numbers of subgroup analyses or mediation tests. In the 1990s, the use of interaction, subgroups, and mediation became more widespread, while use of any of the other tools largely remained below the five percent mark. In the 2000s, the disparity in tool use grew substantially. While both subgroups and mediation continued to be more prevalent, the gap between interaction and other tools became progressively larger.

Application of individual tools

Interaction moderation. Moderation exists if a predictor variable has a differential effect on an outcome variable depending on the level of another variable (Schoonhoven, 1981; Venkatraman, 1989). The moderator can be a categorical variable (e.g., different country environments) or a continuous variable (e.g., the degree of relatedness). Strategic management researchers frequently test moderation using either interactions of predictor variables (e.g., cross-product multiplicative terms) that are added to a regression model or subgroups, which involves splitting the sample into two or more groups based on the level of the contextual variable. In this latter approach, t-tests of the correlation coefficients are commonly used to indicate differences between two groups and chi-square tests are used to indicate differences among several subgroups (Arnold, 1982, 1984; Zedeck, 1971). Interaction moderation is defined as form moderation, as the regression slope coefficient of the X – Y relationship varies depending on Z. In contrast, subgroup moderation is characterized in terms of the strength (Venkatraman, 1989) or degree (Arnold, 1982) of the X – Y relationship.

Interaction is relevant for contingency hypotheses regarding the form of the relationship between variables (Arnold, 1982; Stone-Romero & Anderson, 1994). Overall, we identified 328 articles that used interactions. In Table 3, we identify key statistics and trends regarding the use of interactions. Most articles used multiple interaction terms. Overall, we found tests of 849 interaction effects in our article pool. Among articles using interactions, the mean number of such terms grew steadily over time – from 1.9 in the 1980s, to 2.2 in the 1990s, and 2.8 in the 2000s. The maximum number of interactions in a given article also increased substantially during this period.

The type of analytic tool used varied widely across our sample time window. Regression analysis was consistently the primary tool in all three time windows, although it was partially replaced by logit and probit models in the 2000s. ANOVA was the second most common tool in the 1980s, and its use rapidly declined in subsequent decades. Structural equation modeling was only rarely used to test interaction effects.

As moderation can be tested via either interactions or subgroups, clarity in the type of hypothesized effect can be helpful in understanding theoretical predictions. As shown in Table 3, there is a growing trend to hypothesize interactions as affecting the strength of the relationship– e.g., this approach was used by one third of interaction studies in the 1980s, and by nearly eighty percent of interaction studies in the 2000s. The remainder of interaction studies simply referred to a moderation effect, or used other phrasing. However, as noted previously, interaction is more appropriately characterized as form versus strength moderation.

Overall, interaction was used predominantly for hypothesis tests – three quarter of interactions were theoretical predictions in the 1980s and 1990s, and this proportion grew to approximately 90 percent in the 2000s. The remainder of the interaction effects used were either as control variables, or unspecified tests such as post hoc analyses.

In terms of results, approximately sixty percent of tests reported either strong or moderate effects. Mixed results – i.e., either non-significant or effects contrary to prediction – were found in roughly one third of articles, and showed a slight upward trend over time. Null findings were considerably more rare in the 2000s – 2.2. percent of studies –compared to prior periods.

Despite the frequency of mixed and null effects, studies rarely addressed the issue of statistical power. For all three decades of our sample, more than eight in ten articles did not explicitly address statistical power. Measurement reliability can also affect power via attenuation (Boyd, Gove & Hitt, 2005a), so we examined coverage of reliability in our sample pool as well. In the 1990s and 2000s, more than 70 percent of papers either did not report reliability statistics, or used single indicators to measure variables.

Next, we examined some methodological issues which may arise when using interaction terms. In the 1980s and 1990s, approximately three quarters of articles did not explicitly address the question of mean centering. The 2000s saw a major shift, with 40 percent of articles acknowledging this issue. However, while this topic was addressed far more often, only a small proportion of articles in the 2000s (4.3 percent) reported actual use of mean centering. Multicollinearity is also a potential concern when using product terms. As shown in the Table, the relative proportion of papers which mention this topic changed dramatically across our sample window: In the 1980s, sixty percent of studies made no mention of multicollinearity concerns. In contrast, by the 2000s approximately 62 percent of articles discussed this issue.

Finally, we examined whether studies included graphs of their results, as visual representation facilitates the interpretation of findings. As reported in the Table, the use of graphs depicting the interaction effects has been inconsistent over time: For the different periods, between 8.6 and 33.3 percent of articles included graphic models of their findings.

Subgroup moderation. Overall, we identified 119 articles which used subgroups. Strategy scholars have sometimes used subgroup comparisons for describing and testing the strength of the relationship between variables within contingency models and hypotheses. For example, Prescott (1986) reported that firm environments moderated the strength of the relationship between strategy variables and subsequent performance. In Table 4, we identify key trends regarding the use of this approach.

==========
Insert Table 4about here
==========

First, we examined the analytic tool used to compare subgroups. In the 1980s, there was an even split between regression models and simple correlation coefficients. As shown in the table, the use of simple correlations declined dramatically – down to 10.2 percent in the 1990s and less than two percent in the 2000s. There has been a corresponding increase in the use of other tools, including logit and probit models, which accounted for one-fifth of subgroup studies in the 2000s. For this most recent period, one in ten studies compared structural equation models across subsamples.

Recognizing the imprecision associated with the presentation of early contingency studies, Venkatraman (1989) emphasized the need to clearly distinguish between interaction and subgroup moderation. As shown in the table, substantial progress has been made in this regard. For both the 1980s and 1990s, virtually none of the subgroup studies acknowledged the distinction between subgroup and interaction approaches to moderation. By the 2000s, however, approximately 20 percent of subgroup studies explicitly addressed this distinction, with the bulk of these studies also conducting supplementary analyses using interaction terms as an alternative. The framing of hypotheses also changed substantially during the time window of our sample. In the 1980s and 1990s, many studies did not reference the nature of the moderating effect. In contrast, by the 2000s,

two-thirds of subgroups studies framed their hypotheses using some variant of the strength of the relationship between variables.

More than eighty percent of studies used subgroup analyses to test specific hypotheses. A small number of studies – about two percent – used subgroups as a type of control, in order to address differences between subsets of their sample. Between fifteen and twenty percent of these studies used subgroup models for post hoc analysis – either to provide greater detail regarding some aspect of a hypothesis test, or, more frequently, to provide normative suggestions based on a comparison of high and low performers.

Next, we examined the outcome of model tests. As with other contingency tools in our review, and in similar content analyses (e.g., Miller, et al., 2007), the vast majority of studies reported either moderate or strong support for hypothesis tests. Still, nearly one in five studies reported mixed or null results, in aggregate.

We also examined issues associated with statistical power. Subgroup moderation has been criticized for low statistical power (Stone-Romero & Anderson, 1994). Consequently, large sample sizes are recommended – between 100 and 200 observations per group (Ping, 1996). As with other contingency tools in our sample, studies rarely address statistical power explicitly, even in the presence of mixed or null results. Similarly, few studies report reliability statistics, which can be useful for readers when trying to assess weak or inconsistent findings.

We also examined the attributes of the subgroups, as they also have several implications. The mean number of groups remained relatively stable over our thirty year window, and ranged from 2-3 groups in a typical study. The maximum number of groups ranged between 5 and 7, and showed no substantial trends over time. The grouping variable was predominantly categorical, such as two different industry groups, or two nations or regions. While the different groups are proxies for some underlying phenomenon – stable versus turbulent environments, for instance, the phenomenon was

not measured formally. By the 2000s, the remainder of the groups were based on either ordinal or continuous variables. Studies typically used median splits, with a small number of studies deleting values on either side of the median.

Finally, we examined whether studies reported statistical significance tests when comparing results across subgroups. As shown in the table, 4 out of 5 subgroup studies published in the 2000s did not conduct any formal test to determine if differing findings across groups were substantive. In these cases, authors referred to differences in explained variance across groups, or differences in the magnitude, directionality, or the statistical significance of individual coefficients. Of concern is that the proportion of subgroup studies reporting statistical significance declined over time.

Mediation tests the existence of a variable that intervenes between a predictor and an outcome variable, specifying the existence of an indirect effect.[4] The mediator is anchored by a criterion variable (Venkatraman, 1989) and is assumed to account for a significant amount of the variance of the outcome variable when included in the model. Path analysis is appropriate to test mediation. Researchers must also address the issue of partial or full mediation. Full mediation reflects cases in which the direct relationship between the original predictor and the criterion variable disappears in the presence of the mediator (i.e., the coefficient of the predictor is zero in the presence of the mediator). If the coefficient of the predictor is other than zero, partial mediation exists.

Overall, we identified 76 articles, or 4.4 percent of studies which used mediation. This statistic is comparable to the norm for similar journals: For a comparable time period (1981 – 2005), *AMJ* had 6.2 percent of articles that used mediation, versus 5.1 percent in *ASQ*, and 2.8 percent in *Personnel Psychology* (Wood, Goodman, Beckman & Cook, 2008). Thus although mediation was

[4]Because mediation focuses on linear relationships, mediation is not viewed by all scholars as a form of contingency. Semantics aside, including mediation in our study was vital because it was discussed as a form of fit by Venkatraman (1989). We appreciate an anonymous reviewer who insightfully noted the need to offer this clarification.

the third most commonly used contingency tool, it was still only infrequently applied. In Table 5, we identify trends in key aspects of the methodology of these studies.

==========
Insert Table 5about here
==========

The first attribute that we examined was the analytic tool used to test for mediation. In broad terms, the general categories of tools include regression, simultaneous equations, and analysis of correlations.[5] Based on the results of our content analysis, we added a fourth category which we termed faux mediation. As shown in the Table, structural equation modeling was the primary tool of choice, used in more than two thirds of all mediation studies in *SMJ*. This distribution represents a marked difference from other management journals, where regression analysis is the tool of choice for mediation: Wood and colleagues (2008) reported more than two regression mediation studies for every SEM study in a pool of five management journals. Additionally, a number of studies use what we label as faux mediation. In these articles, the authors develop a conceptual model that includes paths between predictor, intermediary, and outcome variables. However, the different paths are tested independently, such as in separate OLS regressions. In some cases, this approach is used because an intermediate variable is binary (e.g., a choice or outcome), so authors use regression for one portion of their hypotheses, and logit or probit for another portion.

The next aspect we examined was the nature of the contingency. A simple mediation exists when a model presents a single mediator intervening between and predictor and outcome variable. A complex mediation exists when there are multiple mediated paths between variables. As shown in the table, roughly 90 percent of mediation articles in any given time period used complex mediation. The proportion of complex mediation studies is higher for *SMJ* than other management journals

[5]Given our broader focus on contingency methods, we chose not to address the execution of Baron and Kenny's (1986) 'four-step' method for framing contingency studies. Miller and colleagues (2007) reported that the Baron-Kenny approach was used widely but inconsistently in strategic management studies. Our data provides comparable results.

(Wood, et al, 2008), and possibly explains the prevalence of structural equation modeling in this series of studies; SEM is more efficient and effective in analyzing concurrent relationships among a set of multiple predictors and dependent variables than regression analysis.

Third, we examined the temporal sequencing of variables. As shown in the table, a purely cross-sectional design is by far the dominant approach. While SEM advocates argue that this tool can extract causality from cross-sectional data, there are also caveats regarding generalization of findings under these conditions (Mathieu, DeShon & Bergh, 2008). Encouragingly, the use of cross-sectional designs is both substantially lower than in other management journals (Wood, et al., 2008), and also decreased over time (Bergh et al., 2004; Bergh & Holbein, 1997. By the 2000s, approximately one-third of mediation studies had either partial or full temporal gaps between variables.

Fourth, we examined the outcome of model tests. As reported in the table, the overwhelming majority of studies reported either moderate or strong support for mediation tests, a proportion similar to other comparisons of mediation studies (Miller, Triana, Reutzel & Certo, 2007).

Next, we examined issues associated with statistical power. Power is a concern with mediation studies, particularly with regression approaches (MacKinnon, Lockwood, Hoffman, West & Sheets, 2002), and power issues may be exacerbated in strategic management research, as this field often has studies with small effect sizes (Hitt et al, 2004). Also, weak reliability of measures result in attenuation, further diminishing the observed effect sizes (Bergh & Fairbank, 2002; Boyd et al, 2005a). Consequently, we examined how studies addressed the dual issues of statistical power and reliability. In terms of power, our results are straightforward: None of the papers we reviewed explicitly addressed power in either the methods or results sections of their articles. In other words, we cannot draw conclusions whether non-significant paths are an accurate representation of the paths between variables, or simply an artifact of insufficient power. On a more positive note, however,

mediation studies were increasingly likely over time to provide partial or complete data regarding the reliability of their variables.

Next, we examined whether authors considered alternative models. Particularly in the case of complex mediation there are many other possible configurations among the variables (Boyd, Bergh & Ketchen, 2010). Approximately one-third of studies reviewed tested at least one alternative model, while a handful of studies discussed but did not test alternatives.

Our final coding items addressed the presentation of results. As shown in the table, over 90 percent of studies report correlations. These matrices are useful for subsequent meta-analyses, and also enable readers to replicate findings via structural equation models or other tools. However, about 15 percent of studies report only partial correlation matrices – e.g., possibly only showing correlations among exogenous variables, or, in the case of multi-indicator models, reporting only composite correlations between latent constructs. Finally, roughly three quarters of studies provide a visual representation of their mediation models. However, it is more common to report a conceptual model versus final results.

Gestalt. Venkatraman (1989: 433) defined a gestalt as the "degree of internal coherence among a set of theoretical attributes." As opposed to a linear association between predictors and outcomes, in a gestalt model a number of variables forms a holistic pattern, archetypes (Miller & Friesen, 1977), or configurations. Venkatraman (1989) makes an important distinction between gestalts and taxonomies – such as Miles and Snow's Prospectors, Defenders, and Analyzers. The latter, Venkatraman argues, does not sufficiently address the issue of internal congruence. Adding to the complexity of this contingency, gestalts are sometimes used in the same vein as several other labels, including archetypes, strategic groups, and configurations. In Venkatraman's framework, a number of criteria are used to validate a gestalt, including formal statistical tests, use of holdout samples, and a theoretical rationale for the ensuing gestalt. The two most important analytical issues

regarding gestalts are descriptive validity and predictive validity. The most common methodologies for identifying gestalts are cluster analysis and q-factor analysis. Given the scarcity of gestalts meeting Venkatraman's criteria, we conducted a content analysis on the broader pool of configuration studies in *SMJ*. Both gestalts and configurations use overlapping issues and have a number of common issues. Key characteristics of the configuration studies are shown in Table 6. As shown in the table, much of the basis for different taxonomies has historically been inductive, relying on exploratory analysis of variables. In contrast, over time the individual dimensions of taxonomies have been based increasingly on deductive logic– i.e., either through connection to a specific theory, or a review of research connected to a particular industry being studied. Configurations based on expert opinion (cognitive) or combinations of approaches were rare. Studies largely relied on a single clustering algorithm, with the proportion of articles triangulating through multiple methods has increased over time. Additionally, sample size is historically a concern with configuration studies (Short, Payne & Ketchen, 2008). There was no statistically significant improvement in the mean sample size of configuration studies over the three decade window of our review.

In a prior review of configuration studies, Ketchen and Shook (1996) reported that researchers rarely used techniques to validate their models such as split half reliabilities (based on sample subsets) or the use of hold-out samples. The latter approach was viewed by Venkatraman (1989) as essential to gestalt development. As shown in the table, our results indicate that virtually no progress has been made since Ketchen and Shook's assessment of these two techniques – neither approach is used with regularity, nor are there clear temporal trends in the use of them. However, on a more positive note, the majority of studies report criterion validity tests, with their results largely consistent with prediction.

Profile deviation. Conceptually, profile deviation is divergence from a specific set of attributes. Adherence to the specified profile has a positive effect on performance, the criterion

variable, while divergence from it is expected to produce a negative effect. According to Venkatraman (1989) three analytical steps are included in this approach: developing an ideal profile, assigning appropriate weights to each dimension, and creating a baseline model to assess power and ensure that results are not due to random variation.

Overall, we identified nine articles which used profile deviation as a methodology. Profile deviation studies are spread over the last two decades, and no temporal trends appear to exist in the design of these studies, based on our reading of the articles. Consequently, rather than assessing design features over time, as we did with moderation and mediation, we instead offer a general comparison of the attributes of these studies.

A starting point is the development of specific settings or contexts to be studied. Profile deviation asserts that a given context will have its own distinctive issues. Thomas, Litschert and Ranawswamy (1991), for example, proposed that firms pursuing Prospector and Defender strategies each had unique needs vis-à-vis the characteristics of its top managers, and that the alignment of strategy and human capital would shape subsequent performance. Alternately, the context might be a firm's entrepreneurial style (e.g., Naman & Slevin, 1993) or market orientation (Dobni & Luffman, 2003). In practice, the creation of the specific contexts begins with a theoretical specification, followed by empirical analyses such as discriminant or cluster analysis. A given study might examine a single context, or multiple ones.

Once the contexts have been identified, an ideal profile must be developed. A common approach is to identify relevant predictors of firm performance for each of the respective contexts, using higher performers as the metric of "ideal" – e.g., studying the top 10 percent of performers in a given context (e.g., Thomas, et al, 1991; Venkatraman& Prescott, 1990). A second method is to use the dominant approach employed in a specific context (Nath & Sudharshan, 1994). Zajac, Kraatz and

Bresser (2000) took yet another approach in their study of change, in which they modeled expected levels of change, and measured firm departures from the predicted change level.

Venkatraman (1989) made several recommendations regarding the construction of ideal profiles which have largely been ignored by subsequent analyses. First, he strongly advocated weighting the individual contributions to the ideal profile; Venkatraman and Prescott (1990), for instance, used regression beta weights to combine the profile. In contrast, Naman and Slevin (1993) offered a strong rationale as to why unweighted profiles are the desired approach. None of the studies we reviewed empirically examined the role of weighting and how this might affect the outcome of hypothesis tests. Additionally, the distinction between weighted and unweighted ideal profiles was typically not mentioned explicitly in our article pool. A second issue raised by Venkatraman was the use of a separate sample for developing the profile than for hypothesis testing. Venkatraman and Prescott (1990) used the two sample approach, which allows comparison of results and further supports the generalizability of findings. However, none of the other studies we examined utilized multiple samples.

Using the ideal profile, departures from the ideal are then connected to the dependent variable. A variety of approaches have been used to date, including correlations (e.g., Thomas et al, 1991) and regression analysis (e.g., Naman & Slevin, 1993). Some articles embellished the basic regression model in pursuit of more robust findings. Venkatraman and Prescott (1990), for instance, developed a baseline model, which would be conceptually analogous to a null model in structural equation modeling, thereby providing some insight into the relative predictive power of the hypothesized model. Separately, Zajac and colleagues (2000) argued that departures from the ideal can have differential effects – e.g., in the context of change, too much change might have different implications for firm performance than too little change. Consequently, they incorporated a dummy variable to capture this differential effect.

Covariation is internal consistency among a theoretically linked set of variables. For example, Mintzberg's concept of megastrategy (1978), or other conceptualizations of strategy as a pattern of decisions illustrate the concept of covariation. Namely, in order to pursue a successful strategy, a firm may need to concurrently pay attention to several resource allocation decisions and make adjustments to each one in a patterned fashion. Analytically, confirmatory or exploratory factor analyses are often used to test covariation. Additional tests may include a comparison of the coefficients of determination, the calculation of the coefficient and the statistical significance of the second order factor loadings (Venkatraman, 1989).

This tool was also rarely used, with seven articles applying this methodology. All of these studies were published in the 2000s, however making it more frequently used in the last decade than tools such as profile deviation. Venkatraman (1989: 436) commented that "This perspective requires a much greater precision in the patterns of logical consistency among the factors and the explication of the underlying logical link among the attributes. Conceptually, covariation is related to gestalt, as fit is seen as a consistent pattern across a number of discrete areas. Hult and Ketchen (2004), for instance, used resource-based theory to examine how four capabilities combine to create a unique advantage for a firm. The individual elements were seen as necessary, but insufficient, conditions for success. Drawing on the RBV notion of bundling, the authors concluded "We do not suggest that market orientation, entrepreneurship, innovation, and organizational learning constitute unique resources independently but rather that they *collectively* contribute to the creation of a new resource (2001: 900; emphasis in original)." Similarly, Tippins and Sohi (2003: 748) identified three attributes that contribute to a firm's information technology competency. They noted: "…while independent, all three components are required to be present in order to achieve IT competency." Procedurally, covariation hypotheses are analyzed using a second order factor model. The degree of

fit among dimensions is assessed via a comparison of the first versus second order factor structures, and the significance levels of the loadings of the first order factors on the higher order dimension.

Higher order factor models require at least three first order factors to meet minimum identification thresholds, with multiple measures of each factor (Rindskopf & Rose, 1988). The papers that we reviewed used an average of four to five first order factors in their models, and two to seven indicators per factor. The studies in our sample reported differing practices in the assessment of their models: While all of the studies reported factor loadings for the first and second order factors, just over half of these papers made any formal comparison of the overall factor structure. Of these papers, Tanriverdi and Venkatraman (2005) used the 'T' statistic, while the remainder utilized more traditional SEM summary fit measures.

We also observed a number of concerns and potential threats to the findings of these studies. Given the combination of multiple factors with multiple indicators, the studies we analyzed typically utilized a large number of individual variables. In most of these studies, there were less than ten observations per indicator, which is less than ideal (Aguinis& Harden, 2009). Additionally, in nearly half of the studies, there were no control variables included in the models. Third, only one of the articles we reviewed reported a full correlation matrix, limiting the ability of other researchers to replicate findings. The other studies reported either partial correlation matrices, or matrices aggregated at the composite level. Fourth, many of these studies did not take full advantage of their higher order factor mode, vis-à-vis structural modeling software. Two studies, for instance, used the factor models to create composite measures, while a third validated the higher order model, but used only first order factors for hypothesis tests.

However, the covariation studies we analyzed were also more likely than other approaches to use a concurrent combination of contingency effects. Kale and Singh (2007), for instance, used their second order factor as a moderator, while Tippins and Sohi (2003) used mediation in conjunction

with two higher order factors. Additionally, Koka and Prescott (2002) compared the effects of their social capital model using subgroups based on firm nationality.

Matching is a model of fit that is not anchored to a criterion variable and is defined by Venkatraman (1989: 430) as "a measure of fit between two variables…developed independently of any performance anchor." Chandler (1962) illustrates matching by proposing that the multidivisional structure is needed to support and implement a diversification strategy, whereas field units are required for geographical expansion (Venkatraman, 1989). The fit measure is an interval-level variable, developed using one of three approaches: deviation scores, residual analysis, or a three way interaction in an ANOVA. Deviation score analysis, or analysis of difference scores can be methodologically problematic, due to potentially decreased reliability, the possibility of spurious associations, and concerns about discriminant validity (Bergh & Fairbank, 2002; Edwards, 1994a, 1994b).

In our content analysis, we encountered a number of articles which used the label "matching" in either the theory development or actual hypotheses. However, in the overwhelming majority of these papers, matching was operationalized as some form of moderation, typically as an interaction term in a multiple regression. Powell (1992) provides an example of the deviation score approach, while Habib and Victor (1991) is an example of the three-way interaction. Given the limited applications of this tool, we refer readers to Venkatraman (1989) for a more detailed discussion of the methodological issues associated with this approach.

RECOMMENDATIONS FOR FUTURE STUDIES

As demonstrated by our content analysis, strategic management research has a rich history of contingency-based studies and hypotheses. While the field has clearly advanced over the last three decades, there are also multiple opportunities for further refinement. In this section, we identify a number of avenues for improving the quality of future contingency studies. We begin with a high

level focus, offering recommendations that would enhance both the theoretical framing and methodology of contingency research and articles. We then integrate and illustrate these suggestions by proposing a series of contingency research opportunities on strategic management topics.

Avenues for future research: Theory and methodology

Framing more nuanced hypotheses. Doing quality research begins with a valuable research question in which the answer will make a contribution to theory. This often means that the research provides an extension of an existing theory or a refinement of it. In fact, Edwards (2010) suggests that we need more research that provides theoretical refinement identifying the boundaries (and limitations) of existing theories, increasing the precision of theoretical predictions and/or comparing the ability of a particular theory to explain phenomena against competing theories. Our analysis of the research using a contingency framework/approach showed that much of the empirical work integrates the contingency approach with a specific theory (e.g., agency, resource-based view, transaction cost economics, etc.). Thus, a contingency framework was used to examine the conditions under which the particular theory applies and the way it operates (or does not). This approach then examines the boundaries and limitations of the theory and helps to refine our understanding of it. Therefore, using a contingency approach helps to provide more theoretical precision. For example, Sirmon, Hitt and Ireland (2007) developed a model to explain how firms manage resources to achieve a competitive advantage. The model responded to criticisms of the resource based view (RBV) and to the need for refining our understanding of it. The model included the contingency of environmental uncertainty thereby better defining the boundaries of the RBV. In this way, the model and later empirical tests of it (e.g., Sirmon & Hitt, 2009) provided more theoretical precision.

The use of multi-theoretic hypotheses can help to respond to the needs expressed by Edwards (2010). There are several approaches that may be pursued when framing multi-theoretic hypotheses. One method is to consider hypothesis testing as a tournament, “testing which theory or model is best

suited to explain a specific phenomenon, or the conditions under which a specific theory provides more powerful predictions (Boyd, Haynes &Zona, forthcoming)." For example, Ocasio (1994) found that theoretical predictions of CEO tenure had differing utility depending on the stage of a CEO's career. An alternate approach to multi-theory hypotheses posits a synergistic role, such as through interaction terms (e.g., Combs & Ketchen, 1999).

Another opportunity for framing more nuanced hypotheses is to examine the influence of multiple contingencies. While we do not advocate a universal "more is better" position regarding contingency tools, we identified a number of studies that combined multiple tools in order to provide unique insights for theory and practice. Simonin (1999), for instance, used structural equation modeling to develop a fully mediated model of the determinants and consequences of ambiguity in the setting of joint venture knowledge transfers. In addition to finding support for the mediated model, the author found that the path coefficients varied when considering the moderating role of learning capacity or collaboration know-how. As a result, this paper offered substantial advancements in understanding how to develop more effective knowledge management practices. In a related area, Kotabe, Martin and Domoto (2003) developed a set of interaction hypotheses regarding the performance of supplier relationships for US and Japanese automakers. Using a Chow test, the authors determined that their hypothesized relationships differed significantly enough across the two regional samples that they should be analyzed separately. The authors observed that the contrasting findings across countries had important implications: "…adopting a straight 'Japanese' model with extensive technological cooperation may be difficult and potentially wasteful for some US firms (2003: 312)." In both of these examples, the use of an initial contingency tool provided new insights into different phenomena, while the use of a second tool provided additional insight.

Combined mediation and moderation would provide benefits from greater use. The bundling of the two methods concurrently is not new – e.g., Baron and Kenny (1986) briefly mention this in

their seminal article. However, there has been limited use of this combination in strategic management contingency studies to date. Research on strategic consensus is an example where multiple contingencies have helped to advance theory. Consensus studies have applied environmental uncertainty as a moderator of the consensus – performance relationship, with limited success (Kellermans, Walter, Lechner & Floyd, 2005). A forthcoming paper (Gonzalez-Benito, et al., in press) applied both mediation and moderation, and concluded that moderation alone was insufficient to capture the nuance of the consensus – performance relationship. One impediment to the use of these two tools concurrently was the dearth of detailed information regarding the procedure for their use. Both Edwards and Lambert (2007) and Mueller, Judd and Yzerbyt (2005) offer more detailed guidance for combining these tools.

One problem in the use of contingency tools, especially when used in combination, is the lack of knowledge regarding their application. This deficit points to the need for doctoral programs to reassess the methods portion of their curriculum. A longitudinal review of studies published in *SMJ* found that linear models were by far the dominant tool used, and that the proportion of regression studies increased over time. Additionally, tools that might be useful for contingency studies, such as SEM or cluster analyses, were less common over time (Shook, Ketchen, Cycyota& Crockett, 2003). These authors also surveyed recent strategy Ph.D. graduates, and found that junior faculty were well trained in traditional techniques, but less so in other methodological tools. Faculty also tended to specialize in a few techniques, instead of a broader portfolio of tools. More disconcerting, perhaps, is the fact that faculty did not improve their methods skills in the years following graduation. Based on these findings, Shook and colleagues concluded:

> While not every doctoral student needs to master every technique, our results highlight that doctoral training may not be keeping pace with data analytic trends and future research needs. The potential implications are dramatic. To the extent that a strategy scholar lacks sufficient methods training, his/her ability to contribute to the body of knowledge is inhibited. Most scholars may limit themselves to research opportunities that fit their narrow skill set (2003: 1236).

This conclusion resonates strongly with the results of our content analysis. Over our thirty-year window, interactions are clearly becoming the *de facto* mode to operationalize contingency hypotheses. The use of other approaches or tools has remained relatively stable over our observation window. In practical terms, this means there is considerably less variety in the framing of contingency hypotheses over time. Therefore, doctoral programs in strategic management should increase the examination of other contingency methodologies in their curriculum. Shook et al. (2003) also offer further helpful suggestions for enhancing the methods skills of strategic management faculty. We examine them in the following paragraphs.

Design studies with sufficient power. Statistical power is a critical concern, for three reasons, yet is rarely addressed. First, studies of strategic management topics have been found on multiple occasions to have insufficient power (Ferguson & Ketchen, 1999;Mazen, Hemmasi & Lewis, 1987; Mone, et al., 1996). The weak trend in larger sample sizes suggests that this concern has not been ameliorated in recent years. Second, constructs in strategic management studies are often broad in nature, yet typically measured with single indicators (Boyd, et al., 2005a). The ensuing attenuation exacerbates power issues, thereby limiting the ability of empirical tests to advance theory (Boyd, Gove & Hitt, 2005b). Third, statistical power is a particular issue with many of the contingency tools examined in our analysis. Many of the studies that we reviewed reported weak or inconsistent results, and articles typically did not address power issues as a potential explanation for these findings. Two recommendations are easily implemented to address this problem: First, consider statistical power when designing a contingency study, and set a sample size target that would rule out power as an explanation for weak or null findings. This issue is critical, because, as shown in our content analysis, contingency-based strategic management studies have reported only limited gains in average sample size since the 1980s. Second, power analyses should

be reported in the manuscript in the event of unsupported hypotheses. Such actions might prevent inaccurate interpretation of the results.

Continue advances with mediation. Mediation represents a potentially valuable application of contingency tools in strategic management. On one hand, mediation is rarely used – e.g., our analysis suggests that less than five percent of studies in *SMJ* used mediation; similar to its use in other journals. However, despite this fact, the studies using this tool are also fairly sophisticated. Most management studies that involve mediation also use regression analysis (Wood, et al., 2008), driven by the widespread adoption of Baron and Kenny's (1986) 'four step' approach. However, there are limitations with the 'four step,' and structural equation modeling is the preferred alternative (LeBreton, Wu & Bing, 2009). Thus, it is encouraging that approximately 70 percent of all mediation studies in *SMJ* have used some type of structural model. Additionally, a number of these studies have included moderation as well, although there is no clear trend of using a combination of moderation and mediation. Overall, then, strategic management researchers have conducted relatively advanced tests of mediation hypotheses.

Still, there are many opportunities to further refine mediation hypotheses. Strategic management studies still rely largely on purely cross-sectional data, and only small improvements have been made in this area over time. And, although studies often report moderate or mixed results, power analyses have not been addressed as possible explanations for such findings. Improvements in both reliability and the use of alternate models are reflected in our analysis, but both could be used with greater frequency.

Pay greater attention to the type of moderator effects. The approaches used for analyzing moderation also warrant further consideration, as concerns about the merits and drawbacks of the various testing methods were voiced in the early 1970s (Zedeck, 1971). In particular, the subgroup approach has drawn the most concerns. Although subgroups allow for the identification and

description of differences in coefficients (e.g., correlation, regression) relative to different levels of a moderator, their use also has potential problems. First, the process for creating subgroups can produce a loss of information (Cohen, 1988), and the creation of a categorical variable to replace a continuous one increases the likelihood of spurious statistical associations (Maxwell & Delaney, 1993). Second, subgroup comparisons tend to have lower statistical power than moderated multiple regression analysis, cannot be considered relative to alternative explanations (group comparative tests such as t-tests do not include control variables), and the tests do not capture the slope differences for various subgroups, thereby inhibiting the ability to assess differential predictions (Aguinis & Stone-Romero, 1997; Stone-Romero & Anderson, 1994). By contrast, regression analyses provide the opportunity to test continuous and categorical variables, simultaneously control for other potential predictors, test main effects and interaction terms, provide specific estimates of hypothesized relationships, and offer specifications of slope sign and statistical significance. Further, regression analyses can provide estimators for higher-order moderators (e.g., three-way interactions) (Aiken & West, 1991), associations that are likely difficult to understand within subgroups. Thus, for testing moderators, subgroups may allow descriptive and supplemental insights, but modeling interaction terms within regression equations generally offers more robust statistical advantages.

Subgroup moderation relates to the strength of an effect, while interaction refers to the form of an effect (Arnold, 1982). While the two approaches to moderation are interrelated, they do not always yield comparable results (Gerdin & Greve, 2004). Goll and Rasheed (1997), for example, found that the moderating effect of munificence and dynamism on the decision rationality – performance relationship differed depending on the type of moderation used. Subsequent to the publication of Venkatraman's (1989) article, only a small number of the moderation studies we reviewed discuss the distinction between subgroup and interaction moderation, while a growing number of articles actually report supplementary analysis in this regard. Additionally, Venkatraman

(1989) framed the choice between these two approaches as having a theoretical component as well. In a seminal article, Prescott (1986: 362) found that organizational environment moderated the strength but not the form of the strategy – performance connection. He concluded: "Contingency theory should focus on identifying meaningful subenvironments and on examining strategy – performance relationships within and across these subenvironments." Similar findings on subgroup moderation can be found in other examinations of organizational environments (e.g, Boyd, 1995; Henderson, Miller and Hambrick, 2006). More recent studies have attempted to address some of the concerns associated with earlier approaches to subgroups (i.e., comparison of correlations across groups) with more powerful alternatives, such as multi-sample analyses in structural equation modeling (Simonin, 1999).Kor and Misangyi (2008) and Geletkanycz and Boyd (2011) offer examples of framing strength hypotheses, and applications in regression and SEM, respectively.

Limited dependent variable and interaction. A rapidly growing proportion of interaction terms are used with limited dependent variables. Interpretations of moderation tests vary based upon whether the dependent variable (y)is continuous or of limited range (e.g., binary, discrete) (Hoetker, 2007). In contrast to model fit and coefficient estimators in ordinary least squares regression (OLS), limited dependent variable models (LDV) are nonlinear, which has important methodological implications (Wiersema& Bowen, 2009). LDV models, which are oftentimes represented in logit and probit regression analyses, have different approaches with respect to model fit and individual variable coefficient meaning. Indeed, unlike in OLS, Bowen and Wiersema (2004) note that the "size and magnitude of a coefficient estimate in a LDV model is *almost never* (italics in original) an accurate guide to the direction and magnitude of the underlying relationship between the dependent and independent variables" (page 88).

More specifically, LDV model tests involve computing a Likelihood Ratio (LR), which compares the maximized value of the log-likelihood function of an unrestricted model (all variables)

to the maximized value of a restricted log-likelihood model (just the constant term). The LR test is the difference between these two models (multiplied by -2) and its significance test statistic *p* is a Chi-square distribution with the degrees of freedom equal to the number of variables. The measure of fit represented by the maximum likelihood has meaning by comparison across models, with smaller values (closer to zero), indicative of improved fit.[6] The testing of individual coefficients is also different in LDV models. In LDV, the ratio of a coefficient to its standard error is a normal *z*-value (rather than a *t*-statistic) and the value (direction) refers to the change in the probability of the dependent variable when it is equal to 1. Further, because the relationship between the dependent and independent variables is nonlinear, the size of the coefficient has different meaning than in OLS models. Therefore, the researcher needs to compute the marginal effect for each independent and moderating variable (the two common approaches are to calculate the marginal effect at the mean and the odds effect). Mathematically, the interaction effect x_j and x_k variables become the cross-partial derivative with respect to each (Hoetker, 2007: 336).

Testing and interpreting moderation in LDV models suggest that the influence of a moderator (the cross-product of two predictor variables) on the relationship between an explanatory and dependent variable is "rarely indicated by the sign and significance of the estimated coefficient on the interaction term in the model. Instead, [it] is itself a marginal effect…[as] the equation for the moderator will be nonlinear, its value will depend on the values taken by all model variables…and it will not equal the coefficient on the model's interaction term" (Wiersema & Bowen, 2009: 685). Instead, a moderator variable is tested by examining the sign and statistical significance of the values of the variable's marginal effect on the relationship between the independent and dependent variable. It is important to note that the sign of the interaction coefficient may not indicate the direction of the interaction effect; instead, the entire interaction effect must be calculated at a given value, which can

[6] Other measures of model fit exist, such as pseudo R^2. See Hoetker (2007) for a discussion.

indicate whether it is positive, null or negative for some observations and not others (Hoetker, 2007). Further, there can be variation in the significance levels of the interaction effects, even if the overall test statistic is non-significant. Therefore, graphing the relationships can be informative.

Overall, the process for testing moderation in LDV models requires first determining whether the coefficient for the interaction variable is statistically significant and then computing the marginal effect. The relationships can be graphed according to specific values of the moderator (e.g., low, medium, high), minimum and maximum z-statistic values are then reported, and an indication of whether the relationship changes form over this range can be made (Hoetker, 2007; Wiersema & Bowen, 2009).

Potential research opportunities

Management scholars' increased attention to uncommon contingency modeling, such as gestalt, matching, co-variation and profile deviation provide an opportunity to explore new research questions, or to find better, more refined answers to old ones. For example, the notion that the institutional context influences firm outcomes (North, 1990; Scott, 2005) is generally accepted among management scholars. The institutional context, described in the economic tradition (North, 1990) as composed of formal and informal institutions, and in the sociological tradition as a built on normative, cultural-cognitive and regulative pillars (Scott, 2005), consists of multiple, dynamic elements that simultaneously influence firm-level outcomes. Theories of the institutional context, such as theories of culture (e.g. House et al., 2004; Hofstede, 1980; Trompenaars& Hampden-Turner, 1997; Hall, 1966) or the political environment (e.g. Henisz, 2004) generally describe a complex system with internally coherent components (Venkatraman, 1989), and "frequently recurring clusters of attributes or gestalts" (Miller, 1981: 5). Yet, strategic management researchers tend to decompose the institutional context and model the impact of its specific elements on firm-level outcomes in isolation. An alternative approach would be to conceptualize the institutional context as a gestalt and

examine its elements "as a set of relationships which are in a temporary state of balance" (Miller & Friesen, 1977: 264; Venkatraman, 1989: 432). A step in this direction has been taken by Holmes, Miller, Hitt and Salmador (2011) who modeled the institutional environment using four factors composed of multiple items representing regulatory, economic and political institutions.

A related issue is that institutions change over time (North, 1990), creating a dynamic and evolving context in which organizations function. Yet, in strategic management, the elements of the institutional context are most often modeled as static, rather than dynamic. Contingency modeling using profile deviation might be a more fitting modeling approach in this case. In profile deviation, an ideal profile might be created that consists of differentially weighted elements (Venkatraman, 1989: Drazin & Van de Ven, 1985). Such a profile might describe the ideal context for corporate governance, for combating corruption, or for the development of a particular industry. Actual institutional contexts might be contrasted with the ideal, in order to reveal the factors that might prevent desired outcomes.

Aside from institutions, the industry context might also be re-examined in light of less common contingency models. In the strategic management literature, industry is most often included as a control variable. Few studies compare a number (2-4) of industries across a number of dimensions. The impact of industry is often described using a single numerical score or category (e.g. managerial discretion; slow, standard or fast-cycle markets). Yet, because the industry environment is complex and dynamic, profile deviation or co-variation might provide a richer description (Venkatraman, 1989); instead of assigning a single score to describe industry, these approaches allow researchers to capture multiple facets of the contingency variable. Using such an approach might reveal that industry effects on managerial discretion are not uniform. For example, industry effects on managerial discretion might differ across industries. To wit, corporate governance norms in an industry might allow high levels of managerial discretion. However, other influences unique to this

industry might affect strategic decision-making in a different way. Using profile deviation or co-variation allows for multiple variables to be considered concurrently, providing a more refined and precise description of the variable in question. These are but a few opportunities to expand the theoretical and methodological modeling of context in new ways, relying on some of the less commonly used contingency methods. There are other theories and/or research streams that also offer opportunities for application of contingency approaches, to include signaling theory and strategic groups.

Signaling theory. Signaling theory is becoming popular in management research and generally involves signalers, signals, receivers, feedback and a signaling environment (Connelly, Certo, Ireland and Ruetzel, 2011). This perspective offers rich opportunities for contingency hypotheses, as signals as along with feedback can be modeled as moderators of the relationship between the signaler and receivers. Specifically, signaling theory posits that the relationship between the signaler and the receiver could vary based on the signal's cost, reliability, fit, honesty, and, effectiveness. For example, in their study of stock market reactions to CEO certification, Zhang and Wiersema(2009) posit that the CEO's background (shareholdings, directorships, tenure, financial restatements)moderates investors' reactions to certifications. In this instance, the signaler is the firm and its CEO's certification, the receiver is the stock market, and the signal is the CEO's background, with the hypothesis that the receiver's reaction to the CEO's certification will be higher or lower depending on the CEO's background attributes.

Methodologically speaking, applications of signaling theory could model the receiver's reaction as the dependent variable, and the independent variable(s) the signal (measured in terms of its cost, quality, and fit). In this case, the signaler serves as the contingency variable. In a regression model, the aforementioned multiplicative interaction analyses could apply, so the form and/or strength of the interaction term would be used to test the signaling theory hypothesis. Another

approach would be to use subgroups, with the sample split on the basis of the signal, with comparisons across the receiver's reaction to determine whether the theoretical predictions held. From our assessment of signaling theory tests reported in Connelly et al (2011), abnormal stock returns are a common receiver reaction and regression analysis is used frequently for testing the hypothesized signaling effect.

In addition, signaling theory can be used for making more complicated contingency predictions than the foregoing discussion implies. Signaling theory also involves an underlying Nash Equilibrium solution, wherein signalers and receivers can pursue different maximization solutions. For example, if a signal's quality cannot be ascertained before or after its consumption by a receiver, a low quality signaler has an incentive to portray itself as better than it may be, and to gain from the difference in cost savings that it did not have to incur to be a high quality signaler. In this "pooling equilibrium" solution, the above regression analytical approach could be applied. However, if the receiver has the ability to verify the quality of the signal, or if the signal is prohibitively costly for the low quality signaler, the signalers have different alternatives, each of which may provide optimizing outcomes.

Spence (1974) demonstrated this equilibrium format using his labor selection problem, in which uneducated employees are paid less for not having an education, but optimize their opportunity costs in not attaining one, while educated employees are paid more for getting theirs, as the opportunity cost of getting the education is higher for them than foregoing it. Education serves as a signal that represents multiple optimal cost and benefit alternatives for the employees and the employer. In this case, the signal serves as the basis for what is known as a "signaling equilibrium," and an additional component is needed to test this extra condition. Using a regression-analytic model, one methodological alternative is to structure the dependent variable as the receiver's performance after the signal's consumption, the main effect is the receiver's reaction to the signal, and the

contingency (moderator) is the interaction of the signaler and signal. The interaction of the signal and the receiver's reaction to the signal also could be modeled.

Strategic groups. The notion of "fit as gestalt" (Venkatraman, 1989) has played a unique role throughout the evolution of strategic management research. Following its emergence in the 1970s, strategic groups research quickly became a prominent stream of gestalt-driven inquiry. Serious questions were later posed about the merit of the strategic groups concept (Barney & Hoskisson, 1990), the caliber of theory development surrounding the concept (McGee & Thomas, 1986), and the rigor of the methods – such as cluster analysis – used to identify strategic groups (Ketchen & Shook, 1996). By the end of the 1990s, the controversy surrounding how fit as gestalt had been examined by researchers appears to have exceeded the prevalence of this form of contingency within the literature.

In recent years, however, inquiry that views fit as gestalt has been advanced via important theoretical and empirical contributions. For example, Fiss (2007) presented a set-theoretic approach to deriving groups that avoids many of the problems inherent to traditional clustering methods. Short, Ketchen, Palmer, and Hult (2007) leveraged systems theory to empirically establish that strategic groups explain performance variance among firms even after accounting for firm and industry effects. More generally, it appears that something of a renaissance is underway for research on fit as gestalt (cf. Short, Payne & Ketchen, 2008). We believe this is no coincidence. As the theory and methods surrounding fit as gestalt have advanced in recent years, the explanatory value of this form of contingency has increased significantly.

Looking to the future, it appears likely that knowledge generation will be maximized to the extent that researchers integrate fit as gestalt with other forms of contingency. Gestalt-based inquiry is effect inidentifying sets of firms and explaining differences across these sets, but it leaves unanswered many important questions about the workings *inside* sets of organizations. Examining fit

as moderation, for example, within the context of strategic groups could unveil the sources of competitive advantage that are enjoyed by the members of a particular strategic group (Rouse & Daellenbach, 1999), while applying fit as profile deviation might uncover the extent to which closely following the core features of a strategic group either enhances or detracts from performance. Work that examines the internal attributes of strategic groups has long been represented in the literature (e.g., Lawless, Bergh & Wilsted, 1989), but this type of inquiry has remained relatively rare. Thus, the simultaneous use of fit as gestalt and other forms of contingency appears to be a promising approach for future inquiry.

CONCLUSION

The development of contingency hypotheses is central to strategic management. Based on a content analysis of three decades of research, we find that there have been major changes in the use of contingency hypotheses during this time. There are several key areas for improvement that should be addressed in future contingency studies. First, we observed a dramatic decline in the diversity of tools used to frame contingency hypotheses. It is ironic that strategic management having rejected the Taylorist 'one best way', mindset, is at risk of interactions becoming accepted as the 'best way' for any type of contingency question. Yet there are a number of alternatives that should be considered depending on the specific research question and the data available. Second, authors need to think in greater depth regarding the mechanism for the contingencies they are hypothesizing, such as greater clarification regarding the type of moderation they expect. Third, there is often little consistency in the way each of these contingency tools is implemented. Finally, the sample size has remained relatively stable in these studies over time. Combined with single indicator measurement schemes, the frequency of mixed findings, and a lack of power analysis, there are major threats to theoretical or normative recommendations based on unsupported hypotheses.

There are, however, many strengths of the research that we reviewed as well. First, contingency hypotheses have become increasingly prevalent over the last three decades, reflecting greater sophistication in the application of different theories. Second, these studies encompass a broad range of topic areas and theoretical perspectives, indicating that more nuanced hypotheses have become more common in the breadth of the strategic management domain. Third, while many tools are only rarely used, there are a number of exemplars available to guide the development of new studies. Fourth, many of the design elements that we reviewed – such as the reliance on SEM versus regression for mediation hypotheses – indicate that strategic management contingency designs are often more advanced than research in other management subdisciplines.

On a final note, we encourage strategic management researchers to search broadly when seeking methodological insights for new studies. Historically, strategy researchers have been adept at introducing new theories and tools into the field – e.g., consider the widespread adoption of agency, transaction cost economics and resource based perspectives, or the application of network analysis to topics as diverse as corporate boards and knowledge transfer. Research in other fields, such as marketing and economics, can provide useful examples regarding the design and execution of different contingency models. Closer to the home discipline, the Research Methods Division of the Academy of Management is another resource for insights on study design. Lastly, there are also a number of methods-oriented discussion boards (e.g., RMNET, SEMNet) which can also facilitate the design of contingency studies.

REFERENCES

Aguinis, H, & Gottfredson, R. K. 2010. Best-practice recommendations for estimating interaction effects using moderated multiple regression. *Journal of Organizational Behavior*, 31, 776-786.

Aguinis, H. & Stone-Romero, E. 1997. Methodological artifacts in moderated multiple regression and their effects on statistical power. *Journal of Applied Psychology*, 82: 192-206.

Aguinis, H., & Hardin, E.E. 2009. Sample size rules of thumb: Evaluating three common practices. In, Lance, C.E., & Vandenberg, R.J., Eds., *Statistical and Methodological Myths and Urban Legends.* New York: Routledge. Pp. 267-286.

Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. 2005. Effect size and power in assessing moderating effects of categorical variables using multiple regression: A 30-year review. *Journal of Applied Psychology*, 90, 94-107.

Aguinis, H., Petersen, S. A., & Pierce, C. A. 1999. Appraisal of the homogeneity of error variance assumption and alternatives to multiple regression for estimating moderating effects of categorical variables. *Organizational Research Methods*, 2, 315-339.

Aiken, L. S., & West, S. G. 1991.*Multiple Regression: Testing and Interpreting Interactions.* Newbury Park, CA: Sage.

Arnold, H. 1982. Moderator variables: A clarification on conceptual, analytic and psychometric issues. *Organizational Behavior and Human Performance*, 29: 143-174.

Arnold, H.J. 1984. Testing moderator variable hypotheses: A reply to Stone and Hollenbeck. *Organizational Behavior and Human Performance*, 34: 214-224.

Barney, J. B., &Hoskisson, B. E. 1990. Strategic groups: Untested assertions and research proposals. *Managerialand Decision Economics*, 11: 187-198.

Baron, R.M., & Kenny, D.A. 1986. The moderator-mediator distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51: 1173-1182.

Bergh, D. D. & Holbein, G. F. 1997. Assessment and redirection of longitudinal analysis: Demonstration with a study of the diversification and divestment relationship. *Strategic Management Journal*, 18: 557-571.

Bergh, D.D., & Fairbank, J.F. 2002.Measuring and testing change in strategic management research. *Strategic Management Journal*, 23: 359-366.

Bergh, D.D., Hanke, R., Balkundi, P, Brown, M., & Chen, X. 2004. An assessment of research designs in strategic management research: The frequency of threats to internal validity. In D. Ketchen & D. D. Bergh (eds.) *Research Methodology in Strategy and Management*, 1: 347-363.

Bergh, D.D., Perry, J., & Hanke, R. 2006. Some predictors of SMJ article impact. *Strategic Management Journal*, 27: 81-100.

Bowen, H.P., &Wiersema, M.F. 2004. Modeling limited dependent variables: Guidelines for researchers of strategic management. In D. Ketchen, Jr. & D. Bergh (eds.), *Research Methodology in Strategy and Management*, Volume 1, JAI/Elsevier, Oxford, UK, 87-134.

Bowman, E.H. Singh, H., & Thomas, H. 2002. The domain of strategic management: History and evolution. In A. Pettigrew, H. Thomas & R. Whittington (eds.), *Handbook of Strategy and Management*, 1: 31-54.

Boyd, B. K., Gove, S. &Hitt, M.A. 2005a. Construct measurement in strategic management research: Illusion or reality? *Strategic Management Journal*, 26: 239-257.

Boyd, B.K. 1991. Strategic planning and financial performance: A meta-analytic review. *Journal of Management Studies*, 28: 353-374.

Boyd, B.K. 1995. CEO duality and firm performance: A contingency model. *Strategic Management Journal*, 16: 301-312.

Boyd, B.K., Bergh, D.D., & Ketchen, D. 2010. Reconsidering the reputation – performance relationship: A resource-based view. *Journal of Management*, 36: 588-609.

Boyd, B.K., Gove, S., & Hitt, M.A. 2005b. Consequences of construct measurement problems in strategic management research: The case of Amihud and Lev. *Strategic Management Journal*, 26: 367-375.

Boyd, B.K., Haynes, H., & Zona, F. (forthcoming). Dimensions of CEO – board relations. *Journal of Management Studies*, in press.

Burns, T.& Stalker, G.1961. *The Management of Innovation*, London: Tavistock Press.

Cascio, W.F., &Aguinis, H. 2008. Research in industrial and organizational psychology from 1963 to 2007: Changes, choices, and trends. *Journal of Applied Psychology*, 93: 1062-1081.

Chandler, A.D. 1962.*Strategy and Structure: Chapters in the History of the American Industrial Enterprise*. Cambridge, MA: MIT Press.

Child, J. 1972. Organizational structure, environment and performance: The role of strategic choice. *Sociology*, 6: 1-22.

Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*. Second Edition. Hillsdale, NJ: Erlbaum.

Combs, J. & Ketchen, D. 1999. Explaining interfirm cooperation and performance: Toward a reconciliation of predictions from the resource-based view and organizational economics. *Strategic Management Journal*, 20: 867-888.

Combs, J. G., Ketchen, D. J., Perryman, A. A. & Donahue, M. S. 2007. The moderating effect of CEO power on the board composition–firm performance relationship.*Journal of Management Studies*, 44: 1299-1323.

Connelly, B.L., Certo, S.T., Ireland, R.D., &Reutzel, C.R. 2011. Signaling theory: A review and assessment. *Journal of Management*, 37: 39-51.

Dalton, D. R., Daily, C. M. ,Ellstrand, A. E. and Johnson, J. L. 1998. Meta-analytic reviews of board composition, leadership structure and financial performance. *Strategic Management Journal*, 19: 269 - 90.

Dalton, D. R., Hitt, M. A., Certo, S. T., & Dalton, C. M. 2008.The fundamental agency problem and its mitigation: Independence, equity, and the market for corporate control. In J. P. Walsh & A. P. Brief (Eds.), *Annals of the Academy of Management* 1: 1-64. New York: Lawrence Erlbaum.

David, R.J., & Han, S.K. 2004. A systematic assessment of the empirical support for transaction cost economics. *Strategic Management Journal*, 25: 39-58.

Dobni, C.B., & Luffman, G. 2003. Determining the scope and impact of market orientation profiles on strategy implementation and performance. *Strategic Management Journal* 24: 577–585.

Drazin, R. & Van de Ven, A. 1985. Alternative forms of fit in contingency theory. *Administrative Science Quarterly*, 30: 514-539.

Edwards, J.R. 1994a. Regression analysis as an alternative to difference scores. *Journal of Management*, 20: 683-690.

Edwards, J.R. 1994b. The study of congruence in organizational behavior research: Critique and a proposed alternative. *Organizational Behavior and Human Decision Processes*, 58: 51-100.

Edwards, J.R. 2010. Reconsidering theoretical progress in organization and management research. *Organizational Research Methods*, 13: 615-619.

Edwards, J.R., & Lambert, L.S. 2007. Methods for integrating moderation and mediation: A general analytic framework for using moderated path analysis. *Psychological Methods, 12: 1-22.*

Ferguson, T. D. & Ketchen, D. J. 1999. Organizational configurations and performance: The role of statistical power in extant research. *Strategic Management Journal*, 20: 385-395.

Finkelstein, S., & Hambrick, D. 1996.*Strategic Leadership: Top Executives and their effects on Organizations*. St. Paul: West.

Finkelstein, S., Hambrick, D., & Cannella, A. Jr. 2008. Strategic Leadership: Theory and Research on Executives, Top Management Teams, and Boards. Oxford Press.

Fiss, P.C. 2007. A set-theoretic approach to organizational configurations. *Academy of Management Journal*, 32: 1180-1198.

Galbraith, J. R. 1973.*Designing Complex Organizations*. Reading, MA: Addison-Wesley.

Galbraith, J. R., &Nathanson, D.A. 1978. *Strategy Implementation: The Role of Structure and Process*. St. Paul, MN: West Publishing.

Geletkanycz, M.A., & Boyd, B.K. 2011.CEO outside directorships and firm performance:A reconciliation of competing perspectives. *Academy of Management Journal*, 55: 335-352.

Gerdin, J., & Greve, J. 2004. Forms of contingency fit in management accounting research – a critical review. *Accounting, Organizations and Society* 29: 303-326.

Goll, I., & Rasheed, A.M.M. 1997. Rational decision-making and firm performance: The moderating role of environment. *Strategic Management Journal,* 18: 583-591.

Gonzalez-Benito, J., Aguinis, H.A., Boyd, B.K., & Suarez- Gonzalez, I. (forthcoming). Coming to consensus on strategic consensus: A mediated moderation model of consensus and performance. *Journal of Management*, in press

Habib, M.M., & Victor, B. 1991. Strategy, structure, and performance of U.S. manufacturing and service MNCs: A comparative analysis. *Strategic Management Journal*, 12: 589-606.

Hall, E. T. 1966.*The Hidden Dimension*. Garden City, NY: Doubleday.

Hampden-Turner, C. & Trompenaars, A. 1997. *Riding the waves of culture: Understanding diversity in global business.* New York: McGraw-Hill.

Henderson, A.D., Miller, D., & Hambrick, D.C. 2006. How quickly do CEOs become obsolete? Industry dynamism, CEO tenure, and company performance. *Strategic Management Journal*, 27: 447-460.

Henisz, W. J. & Macher, J. T. 2004. Firm- and county-level tradeoffs and contingencies in the evaluation of foreign investment: The semi-conductor industry, 1994-2002. *Organization Science*, 15: 537-554.

Hitt, M.A, Boyd, B. P. & Li, D. 2004. The state of strategic management research: A vision of the future. In D. Ketchen & D. D. Bergh (eds.) *Research Methodology in Strategy and Management* 1: 1-31.

Hoetker, G. 2007. The use of Logit and Probit models in strategic management research: Critical issues. *Strategic Management Journal*, 28: 331-343.

Hofstede, G. 1980.*Culture's consequences: International differences in work-related values.* Newbury Park, CA: Sage.

Holmes, R. M., Miller, T. Hitt, M.A. &Salmador, M. P. 2011. The interrelationships among informal institutions, formal institutions, and foreign direct investment. *Journal of Management,* in press.

House, R. J., Hanges, P. W., Javidan, M., Dorfman, P. & Gupta, V. (Eds.) 2004.*Culture, Leadership and Organizations: The GLOBE Study of 62 Societies.* Thousand Oaks: SAGE.

Hult, G.T., & Ketchen, D.J. 2001. Does market orientation matter?: A test of the relationship between positional advantage and performance. *Strategic Management Journal*, 22: 899-906.

Jauch, L.R., Osborn, R.N., & Glueck, W.F.1980. Short term financial success in large business organizations: The environment-strategy connection. *Strategic Management Journal*, 1: 49-63.

Kale, P., & Singh, H. Building firm capabilities through learning: The role of the alliance learning process in alliance capability and firm-level alliance success. *Strategic Management Journal*, 28: 981-1000.

Kellermanns, F.W., Walter, J., Lechner, C., & Floyd, S.W. 2005. The lack of consensus about strategic consensus: Advancing theory and research. *Journal of Management*,31: 719-737.

Ketchen, D. J.& Shook, C. L. 1996. The application of cluster analysis in strategic management research: An analysis and critique. *Strategic Management Journal*, 17: 441-458.

King, D.R., Dalton, D.R., Daily, C.M., & Covin, J.G. 2004. Meta-analyses of post-acquisition performance: Indications of unidentified moderators. *Strategic Management Journal*, 25: 187-200.

Koka, B.R., & Prescott, J.E. 2002. Strategic alliances as social capital: A multidimensional view. *Strategic Management Journal*, 23: 795-816.

Kor, Y.Y., & Misangyi, V.F. 2008.Outside directors' industry-specific experience and firms' liability of newness. *Strategic Management Journal*, 29: 1345-1355.

Kotabe, M., Martin, X., & Domoto, H. 2003. Gaining from vertical partnerships: Knowledge transfer, relationship duration, and supplier performance improvement in the U.S. and Japanese auto industries. *Strategic Management Journal*, 24: 293-316.

Lawless, M.W., Bergh, D.D.& Wilsted, W.D. 1989.Performance variations among strategic group members: An examination of individual firm capability. *Journal of Management*, 15: 649-661.

Lawrence, P. R., and Lorsch, J. W. 1967. *Organizations and Environment*. Boston: Harvard University Graduate School of Business Administration.

LeBreton, J.M., Wu, J., & Bing, M.N. 2009.The truth(s) on testing for mediation in the social and organizational sciences. In, Lance, C.E., & Vandenberg, R.J., Eds., *Statistical and Methodological Myths and Urban Legends.* New York: Routledge. Pp. 107-141.

MacKinnon, D.P., Lockwood, C.M., Hoffman, J.M., West, S.G., & Sheets, V. 2002.A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7: 83-104.

MacMillan, I. C. 1991.The emerging forum for business policy scholars. *Strategic Management Journal,* 12: 161-165.

Mathieu, J.E., DeShon, R.P., & Bergh, D.D. 2008. Mediational inferences in organizational research: Then, now, and beyond. *Organizational Research Methods*, 11: 203-223.

Maxwell, S.E., & Delaney, H.D. 1993. Bivariate median splits and spurious statistical significance. *Psychological Bulletin*, 113: 181-190.

Mazen, A.M.M., Hemmasi, M., & Lewis, M.F. 1987.Assessment of statistical power in contemporary strategy research. *Strategic Management Journal*, 8: 403-410.

McGee, J., & Thomas, H. 1986. Strategic groups: Theory, research, and taxonomy. *Strategic Management Journal*, *7*: 141-160.

Miles, R. E., & Snow, C. C. 1978. *Organizational Strategy, Structure, and Process*. New York: McGraw-Hill.

Miller, C. C. & Cardinal, L. B. 1994. Strategic planning and firm performance: A synthesis of more than two decades of research. *Academy of Management Journal*, 37: 1649–1665.

Miller, D. & Friesen, P. H. 1977. Strategy-making in context: Ten empirical archetypes. *Journal of Management Studies*. 14: 253- 280.

Miller, D. 1981. Toward a new contingency perspective: The search for new organizational gestalts. Journal of Management Studies. 18: 1-26.

Miller, T.L., del Carmen Triana, M., Reutzel, C.R., &Certo, S.T. 2007. Mediation in strategic management research: Conceptual beginnings, current application, and future recommendations. In D. Ketchen & D. D. Bergh (eds.) *Research Methodology in Strategy and Management*, 4: 295-318.

Mintzberg, H. 1978. Patterns in strategy formation. *Management Science*, 24: 934-948.

Mone, M. A., Mueller, G. C., & Mauland, W. 1996.The perceptions and usage of statistical power in applied psychology and management research. *Personnel Psychology*, 49: 103-120.

Mueller, D., Judd, C.M., & Yzerbyt, V,Y. 2005. When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology*, 89: 852-863.

Nag, R., Hambrick, D.C. & Chen, M.J. 2007. What is strategic management, really? Inductive derivation of a consensus definition of the field. *Strategic Management Journal* 28: 935-955.

Naman, J., & Slevin, D. 1993. Entrepreneurship and the concept of fit: A model and empirical tests. *Strategic Management Journal*, 14: 137-153.

Nath., D., & Sudharsha**n,** D. 1994. Measuring strategy coherence through patterns of strategic choices. *Strategic Management Journal*, 15: 43-61

North, D. 1990.*Institutions, institutional change and economic performance*. Cambridge, UK: Cambridge University Press.

Ocasio, W. 1994. Political dynamics and the circulation of power: CEO succession in U.S. industrial corporations, 1960-1990. *Administrative Science Quarterly*, 39: 285 – 312.

Peteraf, M., & Reed, R. 2007.Managerial discretion and internal alignment under regulatory constraints and change. *Strategic Management Journal*, 28: 1089-1112.

Ping, R.A. 1996. Estimating latent variable interactions and quadratics: The state of the art. *Journal of Management*, 22: 163-183.

Powell, T.C. 1992. Organizational alignment as competitive advantage. *Strategic Management Journal*, 13: 119-134.

Prescott, J.E. 1986. Environments as moderators of the relationship between strategy and performance..*Academy of Management Journal*, 29: 329-346.

Rindskopf, D., & Rose, T. 1988. Some theory and applications of confirmatory second-order factor analysis. *Multivariate Behavioral Research*, 23: 51-67.

Rouse, M.J., & Daellenbach, U.S. 1999. Rethinking research methods for the resource-based perspective: Isolating sources of sustainable competitive advantage. *Strategic Management Journal*, 20: 487-494.

Schoonhoven, C. B. 1981. Problems with contingency theory: Testing assumptions hidden within the language of contingency theory. *Administrative Science Quarterly*, 26: 349-377.

Scott, R. 2008.Institutions and organizations: Ideas and Interests 3rd ed. Thousand Oaks: Sage.

Shook, C.L., Ketchen, D.J., Cycyota, C.S., & Crockett, D. 2003. Data analytic trends and training in strategic management. *Strategic Management Journal*, 24: 1231-1237.

Short, J. C., Ketchen, D. J., Palmer, T. B., & Hult, T. 2007.Firm, strategic group, and industry influences on performance. *Strategic Management Journal*, 28: 147-167.

Short, J., Payne, G.T., & Ketchen, D. 2008.Research on organizational configurations: Past accomplishments and future challenges. *Journal of Management*, 34: 1053-1079.

Simonin, B.L. 1999. Ambiguity and the process of knowledge transfer in strategic alliances. *Strategic Management Journal*, 20: 595-623.

Sirmon, D. G., & Hitt, M. A**.** 2009.Contingencies within dynamic managerial capabilities: Interdependent effects of resource investment and deployment on firm performance**.** *Strategic Management Journal*, 30: 1375-1394.

Sirmon, D.G., Hitt, M.A., & Ireland, R.D. 2007. Managing firm resources in dynamic environments to create value: looking inside the black box. *Academy of Management Review,* 32: 273–292.

Spence, M. 1974.*Market Signaling: Informational Transfer in Hiring and Related Processes*. Cambridge, MA:Harvard University Press.

Stone-Romero, E.F., & Anderson, L.E. 1994. Relative power of moderated multiple regression and the comparison of subgroup correlation coefficients for detecting moderating effects. *Journal of Applied Psychology*, 79: 354-359.

Stone-Romero, E.F., Weaver, A.E., & Glenar, J.L. 1995.Trends in research design and data analytic strategies in organizational research. *Journal of Management*, 21: 141-157.

Tanriverdi, H., &Venkatraman, N. 2005. Knowledge relatedness and the performance of multibusiness firms. *Strategic Management Journal*, 26: 97-119.

Thomas, A. S., Litschert, R. J., & Ramaswamy, K. 1991. The performance impact of strategy-manager coalignment: An empirical examination. *Strategic Management Journal*, 12: 509-522.

Thompson, J.D. 1967. *Organizations in Action*. New York: McGraw-Hill.

Tippins, M.J., & Sohi, R.S. 2003. IT competency and firm performance: Is organizational learning a missing link? *Strategic Management Journal*, 24: 745-761.

Venkatraman, N. 1989. The concept of fit in strategy research: Toward verbal and statistical correspondence. *Academy of Management Review*, 14: 423-444.

Venkatraman, N. and Prescott, J.E. 1990. Environment-strategy coalignment: An empirical test of its performance implications. *Strategic Management Journal*, 11: 1-23.

Westphal, J. D. 1998. Board games: How CEOs adapt to increases in structural board independence from management. *Administrative Science Quarterly*, 43: 511-37.

Wiersema, M.F., & Bowen, H.P. 2009. The use of limited dependent variable techniques in strategy research: Issues and methods. *Strategic Management Journal*, 30: 679-692.

Wood, R.E., Goodman, J.S., Beckmann, N., & Cook, A. 2008. Mediation testing in management research: A review and proposal. *Organizational Research Methods*, 11: 270-295.

Zajac E.J., Kraatz M.S., Bresser R.K.F. 2000. Modeling the dynamics of strategic fit: A normative approach to strategic change. *Strategic Management Journal*, 21:429–53.

Zedeck, S. 1971. Problems with the use of "moderator" variables. *Psychological Bulletin*, 76: 295-310.

Zott, C., & Amit, R. 2008. The fit between product market strategy and business model: Implications for firm performance. *Strategic Management Journal*, 29: 1-26.

Table 1

TOPIC AREA CLASSIFICATIONAND TRENDS

Topic Area	Area	Emphasis Over Time			
		1980s	1990s	2000s	Overall
External environment	• Industry structure • Environmental scanning • General and industry environments	58.0%	42.4%	22.9%	32.1%
Internal environment	• Resources • Core competencies	36.0%	32.8%	19.4%	25.0%
Strategic intent and mission	• Formalization of planning processes • Mission and vision statements	8.0%	12.4%	6.5%	8.5%
Business level strategy	• Differentiation, cost leadership • Focused and broad scope	22.0%	16.4%	5.9%	10.6%
Competitive rivalry and dynamics	• First mover, second mover, late mover • Mutual forbearance • Multi-market competition	20.0%	13.0%	10.9%	12.3%
Corporate level strategy	• Product diversification • Relatedness	16.0%	18.1%	7.1%	11.3%
Acquisitions and restructuring	• Mergers and acquisitions • Takeover	6.0%	13.6%	11.2%	11.5%
International strategy	• National advantage • Entry modes • Global integration vs. responsiveness	8.0%	13.0%	19.7%	16.6%
Cooperative strategy	• Alliances • Joint ventures • Outsourcing and supply chain management	0.0%	14.1%	21.2%	17.1%
Organizational structure and controls	• Organizational forms • Compensation • Board composition	22.0%	26.6%	20.3%	22.4%
Strategic leadership	• CEO tenure • TMT demography • Succession	16.0%	19.8%	11.2%	14.3%
Strategic entrepreneurship	• New product development • Survival of IPO firms	10.0%	5.1%	7.1%	6.7%
Competitiveness and returns	• Above average returns • Risk	74.0%	58.8%	20.9%	37.4%
Knowledge and learning	• Absorptive capacity • Knowledge transfer	2.0%	1.7%	17.1%	10.9%

Table 2

MOST COMMON THEORETICAL PERSPECTIVES OVER TIME

Ranking	**1980s**	**1990s**	**2000s**
1	IO economics	IO economics	Resource-based view
2	Contingency	Resource-based view	Agency
3	Structure Conduct Performance Model	Agency	Knowledge
4	Organizational theory	Contingency	Networks
5	Upper echelons	Transaction cost economics	Transaction cost economics
6	Agency	Upper echelons	Contingency
7	Multinationals	Strategic choice	Upper echelons
8	Grand strategy	Population ecology	IO economics
9	Strategic choice	Resource dependence	Social capital
10	Behavioral theory	Diversification	Signaling

Table 3

KEY CHARACTERISTICS OF INTERACTION STUDIES

		1980s	1990s	2000s
Interactions per Article	Mean	1.9	2.2	2.8
	Max	7	12	21
Analytic Tool	Regression	53.3	78.0	68.4
	SEM	13.3	1.2	3.0
	Logit/probit	0	9.7	17.3
	ANOVA	33.3	7.3	2.2
	Other	0	3.6	9.1
Rationale for Subgroups vs. Interactions	Not discussed	66.7	64.6	66.2
	Discussed, not analyzed	6.7	4.9	7.8
	Supplementary analyses	26.7	30.5	26.0
Framing	Moderating	53.3	6.1	7.4
	Strength	33.3	65.9	79.2
	Other	13.3	28.0	13.4
Hypothesis	Hypothesized	73.3	74.4	89.2
	Control	0	9.8	3.0
	Unspecified	26.7	15.9	7.8
Results	Strong	33.4	18.3	19.1
	Moderate	26.7	26.8	42.6
	Mixed	26.7	35.3	36.1
	Null	13.3	19.5	2.2
Power	Not addressed	80	86.6	81.4
	Addressed	20	14.4	18.7
Reliability	Not reported, or single-item measures	46.7	74.4	71.0
	Partial reported	13.3	7.3	12.1
	All reported	40.0	18.3	16.9
Mean Centered	Not mentioned	73.3	78.0	58.0
	Mentioned	0	22.0	37.7
	Used	26.7	0	4.3
Multicolinearity	Not addressed	60	59.8	37.7
	Addressed	40	40.2	62.3
Graph	No	86.7	91.4	66.7
	Yes	13.3	8.6	33.3

All numbers expressed as percentages

Table 4

KEY CHARACTERISTICS OF SUBGROUP STUDIES

		1980s	1990s	2000s
Analytic Tool	Regression	50	55.1	60
	SEM	0	8.2	10
	Logit/probit	0	14.3	21.7
	Correlation	50	12.2	1.7
	Other	0	10.2	6.7
Rationale for Subgroups vs. Interactions	Not discussed	100	97.6	80.4
	Discussed, not analyzed	0	0	5.4
	Supplementary analyses	0	2.4	14.3
Framing	Moderating	11.1	0	30.8
	Strength	22.2	28.6	67.3
	Other	66.7	71.4	1.9
Hypothesis	Hypothesized	80	85.7	83.1
	Control	0	2	1.7
	Unspecified	20	12.2	15.3
Results	Strong	50	57.4	61.4
	Moderate	20	25.5	19.3
	Mixed	30	17	15.8
	Null	0	0	3.5
Power	Not addressed	100	98	98.3
	Addressed	0	2	1.7
Reliability	Not reported, or single-item measures	100	85.7	88.3
	Partial reported	0	8.2	6.7
	All reported	0	6.1	5.0
Grouping Variable	Nominal	60	49	58.3
	Ordinal	0	20.4	16.7
	Continuous	40	30.6	18.3
Groups	Mean	3.2	2.4	2.7
	Minimum	2	2	2
	Maximum	7	5	6
Significance Tests	No	70	71.4	81.7
	Yes	30	28.6	18.3

All numbers expressed as percentages

Table 5

KEY CHARACTERISTICS OF MEDIATION STUDIES

		1980s	1990s	2000s
Analytic Tool	Regression	25	16.5	21.7
	SEM	75	76.9	63
	Correlation	0	3.8	2.2
	Faux mediation	0	7.7	13
Mediation Type	Simple	0	3.8	13
	Complex	100	96.2	87
Temporal Sequencing	Cross-sectional	75	61.5	60.9
	Mixed	25	3.8	15.2
	All different	0	26.9	17.4
	Unclear	0	7.7	6.5
Results	Strong	75	53.8	56.5
	Moderate	25	26.9	28.3
	Mixed	0	15.4	13
	Null	0	3.8	2.2
Power	Not addressed	100	100	100
	Addressed	0	0	0
Reliability	Not reported, or single-item measures	100	69.2	45.7
	Partial reported	0	11.5	15.2
	All reported	0	19.2	39.1
Alternate Models	Not addressed	100	61.5	73.9
	Mentioned, not reported	0	7.7	4.3
	Results reported	0	30.8	21.7
Moderators	No	75	96.2	73.9
	Yes	25	3.8	26.1
Correlation Matrix	None	75	23.1	8.7
	Complete	25	61.5	76.1
	Partial	0	15.4	15.2
Diagram	None	0	7.7	28.3
	Conceptual	100	38.5	45.7
	Results	0	53.8	26.1

All numbers expressed as percentages

Table 6

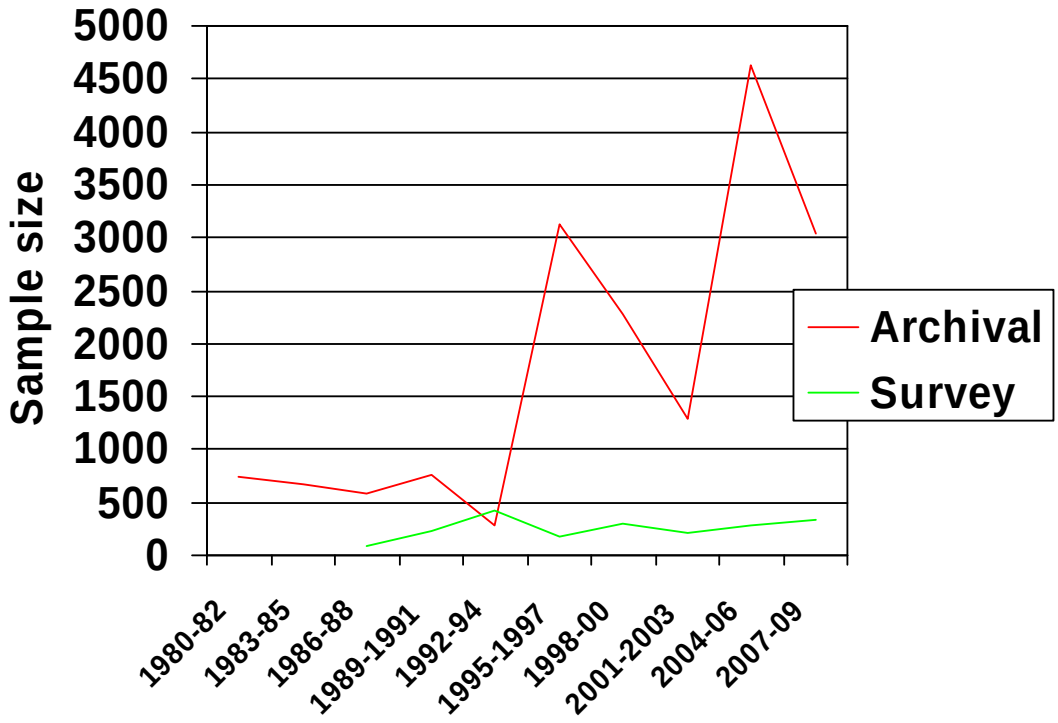
KEY CHARACTERISTICS OF CONFIGURATION STUDIES

		1980s	**1990s**	**2000s**
Clustering Basis	Inductive	63.6	30.8	33.3
	Deductive	27.3	53.8	55.6
	Cognitive	9.1	15.4	0
	Combination	0	0	11.1
Clustering Method	Single	81.8	65.4	66.7
	Multiple	9.1	11.5	33.3
	Unspecified	9.1	23.1	0
Standardization	No	100	95.8	87.5
	Yes	0	4.2	12.5
Sample Size		519.45	194.07	300.78
Split Half Reliability	None	90.9	96	100
	Clusters tested on subsets	9.1	4	0
Hold-out sample	None	100	92	100
	Cluster replicated in separate sample	0	8	0
Criterion Test Used	None	9.1	42.3	11.1
	Yes	90.9	57.7	88.9
Criterion Test Results	Strong	60	60	87.5
	Moderate	20	40	12.5
	Mixed	0	0	0
	Null	20	0	0

All numbers expressed as percentages

FIGURE 1

SAMPLE SIZE CHARACTERISTICS



Results of t-tests that compared samples size across decades are as follows:

Archival Studies			Survey Studies		
	1980s	1990s		1980s	1990s
1990s	1.67 (.10)		1990s	1.81 (.07)	
2000s	1.94 (.054)	1.65 (.10)	2000s	1.89 (.064)	0.58 (ns)

FIGURE 2

DISTRIBUTION OF SMJ APPROACHES OVER TIME