Does Experience Improve Acquisition Performance? – It's Complicated, and That is When it Helps Most

Magdalena Langosch
University of Greenwich
Faculty of Business
Department of International Business & Economics
Queen Mary Building, Room 162
Old Royal Naval College, Park Row
London SE10 9LS, United Kingdom
e-mail: m.langosch@greenwich.ac.uk

Justin Tumlinson
University of Exeter
Science, Innovation, Technology, and Entrepreneurship
73 Streatham Court
Exeter EX4 4PU, United Kingdom

Institute for Strategy, Technology and Organization
Ludwig-Maximilian-Universität München
80539 Munich, Germany
e-mail: j.tumlinson@exeter.ac.uk

Acknowledgements: The authors would like to thank Wilfred Dolfsma, Mehmet Ugur, Thomas Triebs, Tobias Kretschmer, Gerhard Schnyder, Anna Grossman, Stefano Ghinoi, Michael Borenstein, Christopher Tucci and Paul Ingram for their input.

Abstract:
Across social science the observation that experience in X increases performance in X is broadly established. The empirical literature on quantifying the effect of acquisition experience on the performance of future acquisitions is an anomaly—only half the studies published in top management journals report such a positive association. We meta-analyze this literature. Our study contributes three primary discoveries: (1) We robustly establish the positive relationship between acquirer experience and performance after accounting for the statistical quality of each study. Just as this result is non-obvious to academic observers of this conversation, it is also apparently non-obvious to investors—the effect size from studies using stock market reaction as a proxy for performance is indistinguishable from zero in all specifications and significantly less than from those using accounting-based measures. (2) The positive association of experience to performance strengthens in study settings is characterized by a moderator previously explored among industrial workers: Complexity. In particular, experience is more positively associated with performance in cross-border and multi-industry settings as well as those where the performance metrics reflect information more available to insiders than outsiders. (3) We document the considerable discord in this literature and highlight its probable sources and remedies.

Keywords: Merger & Acquisitions, Experience, Performance, Complexity, Meta-Analysis

INTRODUCTION

According to the Institute for Mergers, Acquisitions and Alliances the value of global Merger and Acquisition (M&A) deals topped $4 Trillion in 2018.[1] With so much value at stake, knowing whether one can get good at acquisitions or not matters. How much experience influences M&A success has occupied the attention of hundreds of academics for decades.

On the surface it seems obvious—more experienced acquirers should pick better target firms and integrate them into their existing operations more smoothly, because, with each successive deal, acquirers learn of potential pitfalls to be avoided and opportunities to seized upon (Comier & Hagman, 1987). This learning should be embedded in the firm's routines, structures, and its employees, standing ready to deploy in the next deal (Argote & Miron-Spektor, 2011). Hence it seems natural that M&A deals involving more experienced acquirers should perform better—and serial acquirers should be especially successful. This is the intuitive logic of *Organizational Learning* (Levitt & March, 1988).

Empirical confirmation has not been universal, though. *Figure 1* plots the partial (i.e., fully controlled) correlations between experience and M&A performance from 89 studies published in top management journals.[2] These correlations are so evenly distributed about zero that the casual reader could scarcely determine the sign of the median effect, much less identify a consensus magnitude (Barkema & Schijven, 2008; Cartwright & Schoenberg, 2006; Zollo & Singh, 2004). Many empirical studies do show the natural positive relationship between acquisition performance and experience (Ahammad, Tarba, Liu, & Glaister, 2016; Barkema, Bell, & Pennings, 1996; Hebert, Very, & Beamish,

[1] https://imaa-institute.org/mergers-and-acquisitions-statistics/
[2] Construction of the sample and measures used in *Figure 1* will be discussed in detail later.

2005), but as just as many report a negative (Parola et al., 2015; Uhlenbruck et al., 2006) or insignificant effect (Zollo, 2009; Zollo & Singh, 2004).

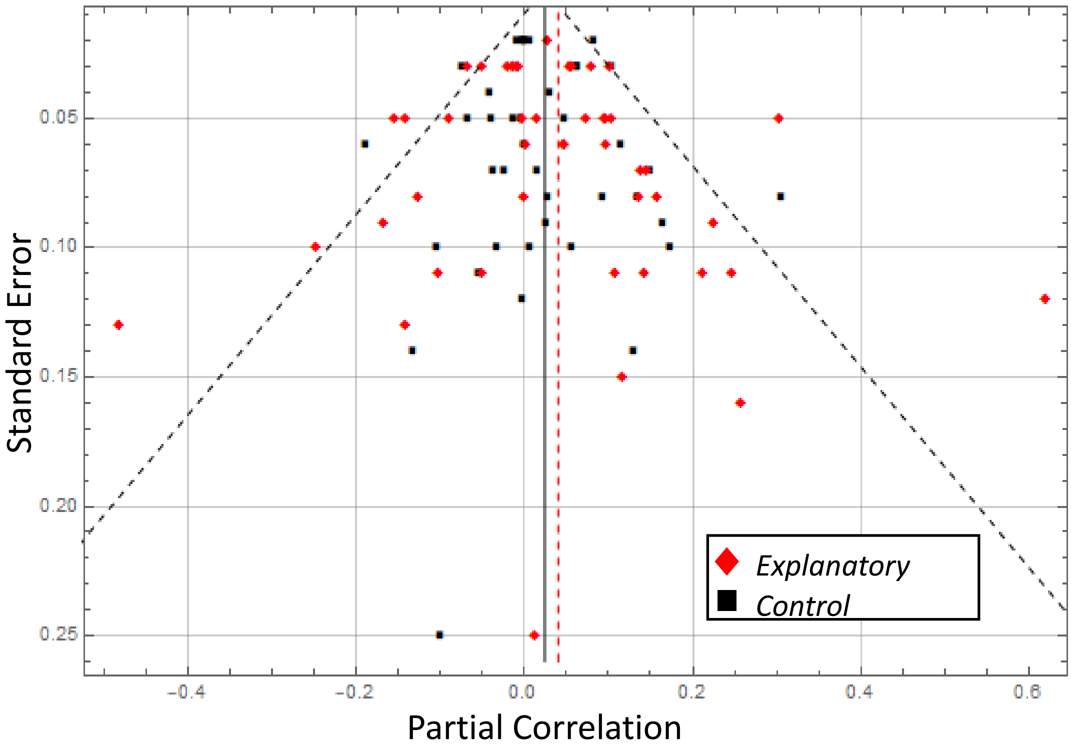


Figure 1 *scatterplots partial correlations between experience and performance versus their associated standard errors for primary studies in our meta-sample (i.e., Figure 1 is a funnel plot). Red diamonds denote effect sizes where experience is an explanatory variable. The red dashed line at* $\bar{r}_{EV} = 0.040$ *depicts the (unweighted) average of this group. Black squares denote effect sizes where experience is a control. The black solid line at* $\bar{r} = 0.024$ *depicts the (unweighted) average of the entire meta-sample. Dashed diagonal lines plot the lower and upper bounds of the 99% confidence interval around* $\bar{r}$ *given by the standard error on the vertical axis.*

Null and negative results have prompted more nuanced theories to explain them. These generally derive from *Transfer Theory's* core insight: for experience to be helpful, it must be applicable (Haleblian & Finkelstein, 1999). If a firm tries to apply learnings generated in previous deals to one that is insufficiently similar, then that experience may prove useless, or worse, if previous experience engenders overconfidence, details of the new transaction may be overlooked and go unattended to, leading to a performance *reduction*. Haleblian & Finkelstein (1999) relay the following example: Cocksure from their successful integration of Miller Brewing Company in 1969, Philip Morris felt, "7-Up was reasonably similar to Miller Beers," and bought 7-Up in 1979. Philip Morris tried the same, previously successful, integration and market strategies again, only to fail and divest 7-Up within five years after racking up $25 million in losses.

Again, at first blush the *Transfer Theory* explanation for observed negative effects of experience on M&A performance resonates: most readers, like Philip Morris, have been led astray *on occasion*, in business or life, by misapplying previous experience to a fundamentally different situation. Yet, there is a major difference between explaining individual *ex post* errors in judgement and claiming that *on average*, gaining more experience *systematically* leads to worse outcomes across a population, even if that statistical statement is scoped to specific settings. Among other things, it means there are contexts where the usual learning processes consistently do not hold—agents cannot even generally learn that experience is not only not helpful in these contexts (i.e., leading to a zero effect), but it consistently induces them to do the wrong thing (i.e., leading to a negative effect). What are these peculiar contexts?

So, in this paper, we use meta-analyses over this very diverse literature to synthesize what we know about the role of experience on M&A performance. In particular, we seek answers to the following: (1) Can we generalize a homogeneous and significant effect (either positive or negative) of experience on M&A performance in the existing literature? (2) Is the relationship between experience and performance influenced by measurement and context differences in the studies? (3) Since the direction of reported effects are so mixed, can we identify contexts where the most extreme implications of *Transfer Theory* hold—causing experience to reduce acquisition performance?

THEORY

Phenomenon

Firms acquire other firms for synergies, an attempt to create an entity greater than the sum of its parts (Krishnan & Park, 2002). Were there no synergies, the price paid would be less than post-merger value and result in losses, due to the inevitable costs of integration. The literature categorizes these synergies in two ways: (1) Operating synergies, which result from resource combinations, like costs savings through economies of scale and revenue growth from new product offerings from access to new markets or clients and purchase of externally generated innovations (Barney, 1991; Hitt et al., 1996).

(2) Financial synergies, which arise from combining the firms' financial structures, tax savings and obtaining additional profits from well-managed undervalued target firms (Rabier, 2017). Of course, the presence of potential synergies is not the same as realizing them.

Experience's potential impact on acquisition performance is manifold. Before intentions are even announced, experienced acquirers know when to buy and when not to. They have better access to and know-how to utilize outside resources, financials, and legal assistance to close the deal. Many argue that just knowing what the key integration success factors are, requires experience (Bruton et al., 1994). Previous acquisitions build routines for successfully integrating new targets (Haleblian & Finkelstein, 1999) . Experience still matters long after the deal closes—management that has been through previous mergers are better equipped to juggle diverse product portfolios across varied demographic and geographic markets, and capitalize on complementarities (Hayward, 2002; Kim & Finkelstein, 2009).

Need for a Meta-Analysis

Figure 1 raises a paradox. It scatterplots the partial (i.e., fully controlled) correlations between experience and acquisition performance from 89 studies published in top management journals from 1980 to 2019. As we noted above, without the help of the vertical lines denoting the average effect sizes, any pattern would be hard to perceive. However, that experience in *X* produces better performance in *X* is arguably the most well-established fact in social science (see e.g.: Argote & Epple, 1990). Why would acquisitions be such an outlier? Resolving this puzzle requires first quantifying it; something that has, to our knowledge, not been done.

Hence, there is a need to systematically synthesize this large body of evidence on experience's effect on M&A performance. *Figure 1* reveals that much of the evidence comes from small sample studies, which may yield noisier estimates. Indeed, among studies where experience is a variable of interest, only one is based on more than 2,500 observations.[3] An advantage of a statistical, rather than narrative, review of independent primary studies is that meta-analyses accounts for variations in sample sizes and the statistical significance of individual study findings, as well as the variance across them.

Aguinis et al. (2011a) emphasize two capabilities of meta-analyses: (1) estimating an overall direction and strength of the studied relationship of the variables, and (2) investigating across-study variance of the individual effect sizes to derive potential moderators that explain such dispersion. *Figure 1* depicts considerable dispersion in primary study effect sizes, visually exposing the literature's disagreement about the relationship between experience and performance. By grouping primary studies according to shared features, we can identify contexts or metrics where the effects are stronger or weaker and more or less agreement exists in the literature—indicating where future efforts should be focused.

Meta-analyses of M&A research have focused mostly on strategic and financial complementarities (King et al., 2004) and the role of culture (Stahl & Voigt, 2008) but have not yet studied the overall effects of experience separately as a variable of interest. King et al. (2004) in their meta-analysis of M&A performance touch upon prior acquisitions experience as a potential moderator – previous acquisition experience facilitates routines for improving future integration and deal performance. Their estimated effect size is insignificant. However, as only seven studies were included, their results remain inconclusive. By focusing solely on experience, we aim to complement the established findings in the M&A field and to test the direct effects of the concept on post-deal performance.

By now so many studies have measured the relationship between experience and M&A performance that it is often relegated to a control variable status or omitted entirely but understanding its potential value as a control is critical for assessing the effect of other M&A performance drivers. To illustrate, suppose that a future M&A setting offered no data on acquirers' prior deals. Then one may worry,

[3] Recall that under *conditional mean independence*, OLS generates unbiased estimates for the effects of explanatory variables but estimates for controls *may* be biased. Hence, we check whether this distinction drives the lack of consensus.

since experience is likely to influence performance, that leaving it out of the regressions will introduce omitted variable bias. This worry is genuine, but does this bias the regressor of interest upwards or downwards? The answer depends on the correlation of the variable of interest to experience and the direction of experience's effect on performance. Furthermore, the answer to that question determines whether the coefficient of interest is conservatively or liberally estimated. In the former case, the study might be published, with the acknowledgement that the estimate of the focal effect is a lower bound, while in the latter case, a suitable control for experience must be found; otherwise, the study is of little value. Hence, a meta-analysis to synthesize the empirics of what we know about the role of experience on M&A performance is timely, even as attention progresses to other acquirer attributes, which may be correlated to its experience. However, a meta-analysis cannot synthesize the accumulated theory. So, we summarize the most relevant concepts in the next subsection.

Theories of Experience in M&A Research

Organizational Learning Theory describes how organizations transform experience into knowledge for continuous improvement. It was originally developed to understand manufacturing processes, a setting where more experience reliably leads to better outcomes (Zollo & Singh, 2004). The strategy literature gradually applied the logic to other management arenas, like strategic decision making and integrating acquisitions (Meschi & Metais, 2013). *Organizational Learning Theory* has been further refined by the concept of the *Learning Curve.* It describes (1) the rate at which improvements occur as a function of level of experience, and (2) the timing between when an insight generating experience occurs until when it pays dividends in higher performance. Following the logic of Bayesian updating that each successive piece of information adds less to the total understanding, the *Learning Curve* implies that more experience leads to better performance, but the improvement rate declines with more experience. Formally, performance is increasing and strictly concave in experience, and is typically modelled as linear-log or quadratically.[4] There is more dispute in the timing dimension: Ingram and Baum (1997) propose that the salience of experience decays, implying that recent experiences matter most; however, others argue that organizations need time to turn experience into acquisition relevant competencies, potentially leading to inverted-U relationships between the age of experience and impact on performance (Liu & Zou, 2008; Zollo & Singh, 2004). The application of *Organizational Learning Theory* to M&As has been critiqued, though—they are more heterogenous than manufacturing lines and such a montone positive relationship may not always hold (Hayward, 2002).

In contrast, *Transfer Theory* argues that in order for experience to improve future performance, past transactions must resemble future ones, thereby making experience transferable (Comier & Hagman, 1987). If this criterion does not hold, then experience may be of limited, even no, value (Gick & Holyoak, 1987; Ingram, 2002). In fact, *Transfer Theory* allows that when an agent falsely believes that future scenarios resemble past ones, and this causes the agent to blindly apply past 'learnings' to future situations rather than exploring the new situation, the performance implications of experience can be negative (Finkelstein & Haleblian, 2002). Relatedly, Basuil & Datta (2015) show that generic experience measures do not influence performance, but specific types of experience are positively related. Haleblian & Finkelstein (1999) explain the U-shaped relationship they observe—the best performing acquirers either have no experience at all or a lot, and the ability to decide when generalization of previous experience is appropriate for the focal deal—using *Behavioral Learning Theory*. They link these findings back to *Transfer Theory*—naiveté reduces generalization errors.

Experiential Learning emphasizes, in subtle differentiation from *Transfer Theory*, that diversity of previous experience, rather than similarity to current situations, is what matters—heterogenous experience provide a broader pallet of solutions from which to draw in future challenges (Zollo, 2009).

[4] The term "curvilinear" is often used, but it is imprecise, encompassing any non-linear functional form.

The empirical evidence that more diverse prior acquisitions leads to better future M&A outcome is mixed (Galavotti, 2019; Meschi & Metais, 2006)

In the following subsection we describe the moderators we use to detect the influence of the above theoretical forces on experience's role in M&A performance in broad cross-section of studies, and whether detected the associations are influence by the metric used.

Moderators

The primary effect considered in a meta-analysis can vary in magnitude and/or heterogeneity depending on the *context* of, or the *measures* of the independent or dependent variable used in individual primary studies (Aguinis et al., 2011b). These context and measurement differences are known as moderators. They can be intuitively interpreted much like interaction effects, also called moderators in regression analysis. However, in meta-analysis, the unit of observation is the primary study. By comparing meta-analyses of primary study subgroups sharing common moderator values, we can determine in what settings the association between experience and acquisition performance is strongest, where its direction might change, or what metrics yield the most consistent effects. Our choice of moderators is driven by the prevailing theories of experience above, as well as what is possible, that is to say, a large enough subsample of studies must share a particular attribute to permit a meaningful meta-analysis. Because the unit of analysis is a primary study, any moderation must be over differences between study-level metrics and contexts, not deal-level attributes. So, for example, while deal size may play an interesting role in moderating the effect of experience at the deal level, moderating over deal size in a meta-analysis would require that some group of primary studies examining experience only included large deals and another group included only small ones, something that to our knowledge does not exist.

Measurement Moderators

Market- vs. Accounting-based performance. The M&A literature generally either measures performance as shareholder value creation or synergy realization. The former is measured in stock price movement and the latter from accounting metrics (Stahl & Voigt, 2008).

Market-based performance is most commonly measured using an event study (Zollo & Meier, 2008). These estimate cumulative abnormal returns (CARs) as the difference between the actual stock return, around a window (typically 0-30 days) of the acquisition announcement and the return that would be expected without an acquisition announcement, conditional on the broader market performance (Brown & Warner, 1980). Recognizing that M&As' effects may take longer to accrue to shareholders (Oler et al., 2008), some consider somewhat longer event windows and Buy-and-Hold Abnormal Returns (BHARs) (Basuil & Datta, 2015).

Acquisitions are meant to increase revenues, reduce costs, or create synergies in terms of new products, knowledge, and technologies (Chatterjee, 1986; Devos et al., 2009; Houston et al., 2001; Krishnan et al., 2007). Synergy realization is typically measured using accounting-based metrics, such as Return on Assets (ROA), Return on Equity (ROE), or innovation-based metrics, such as the creation of new products, knowledge, technologies in terms of patents, which Birkinshaw et al. (2000) call a "synonym of synergy realization". These metrics are generally assumed to capture long-term performance.

There are two distinct reasons to moderate over the way performance is measured. First, certain metrics may capture the notion of performance more precisely or accurately than others. Our *ex-ante* expectation was that the greater consistency with which market-based measures were applied and the ready availability of stock-price data facilitating large sample studies would lead to greater homogeneity in market-based estimates of experience's effect on performance. As we will see, though, those expectations were wrong.

Second, not all performance measures contain the same information available to the same audience at the same time. Overall, market-based performance metric aggregate public information available about expected near- and long-term performance around the time of the acquisition announcement. They are an *ex-ante* measure. On the other hand, *ex-post* accounting performance measures capture what actually happened (Zollo & Meier, 2008).[5] Were experience to matter more in studies using accounting-based than market-based metrics, one might speculate that the market both does not generally recognize the value of experience and that experience matters particularly for working through deal intricacies that are invisible or inscrutable to the public.

Linear vs. Logarithmic experience. Most primary studies in our sample measure experience as the number of acquisitions within a specified time window prior to the focal deal. A minority (about 10 per cent), though, measure experience as the natural logarithm of that count. The former metric weights every additional deal as equally important, while the latter weights each additional deal proportional to the percentage of new experience that it adds. In other words, logging the number of deals treats each additional deal as less important. As we measure effect sizes as (partial) correlations, a larger coefficient on the subsample using logged experience would support the *Learning Curve* view.

Context Moderators

Experience Recency. When accounting experience, the length of the time window prior to the focal deal matters. Some studies count only experience within the two years prior to the focal deal, while others count any prior recorded deal as material, resulting in more than 18 years of experience accumulation in some cases (e.g., Shi & Prescott, 2012). The typical experience counted in the former studies is more recent (from the perspective of the focal deal) than the typical experience counted in the latter. Hence, we use the length of the time window as a measure of experience recency—shorter

[5] Note that we do not mean that market metrics capture short term synergies and accounting ones longer term value. The efficient market hypothesis implies that prices include available information about expected future synergies.

windows proxy for more recent experience. *Transfer Theory* suggests subsamples using shorter windows will produce stronger effect sizes.

Domestic vs. Cross-Border deals. The literature acknowledges that cultural and legal differences could drag on the performance of cross-border deals. Firms that cross international borders by acquiring a target firm abroad face an additional layer of complexity. Some scholars explicitly avoid such confounding factors by limiting their sample to domestic transactions only (Cording et al., 2008). Cross-border studies proxy for greater environmental complexity.

Industry Relatedness. Whether the acquirer and its target are in similar or the same industry impacts both, acquisition behavior and performance (Haleblian & Finkelstein, 1999). If both firms are operating in similar industries, or at least have an overlap in their value chain, identifying synergies and integration are more straightforward (Basuil & Datta, 2015). We identify three distinct sample settings in decreasing levels of industry relatedness: (1) single industry – scholars explicitly examine deals within a single industry, mostly to rule out industry level drivers of performance (Kim & Finkelstein, 2009); (2) two industries – authors chose two related industries that tend to acquire from each other (e.g., mining and manufacturing) (Kroll et al., 1997); and (3) multiple industries – scholars open the sample to deals across all industries (e.g., Cuypers et al., 2017). Like our cross-border moderator, study level industry relatedness captures a dimension of average deal complexity.

METHODS

We conducted a series of CMAs (comprehensive meta-analyses) following Borenstein, Hedges, Higgins & Rothstein (2011). The method requires data to be drawn from (1) statistically equivalent and (2) conceptually comparable primary studies (Lipsey & Wilson, 2001). All included studies report the measured effect between our variable of interest and dependent variable either as a correlation or regression coefficient, both of which can be transformed into equivalent partial correlations, satisfying the first criterion. We include only studies that measure the direct effect of *prior acquisition experience* on *post-deal financial performance*, satisfying the second.

Identifying and Coding Studies

The search for primary studies proceeded in several steps. First, we included all studies (1) published between 1980-2019 in top-100 journals as ranked by Scimago 2014 impact factors in the subject areas "Business, Management and Accounting" and "Economics, Econometrics and Finance," (2) having *merger(s)*, *acquisition(s)*, *takeover(s)*, or *M&A* in the title, and *experience*, *performance*, or *post-performance* in the abstract, (3) using market- or accounting-based measures (e.g., CARs or ROAs) for post-acquisition performance, and (4) reporting correlations or regression coefficients (with *t*-statistics or standard errors) of acquirer experience to post-announcement acquirer performance.[6] This yielded 84 studies. 'Snowballing' from their bibliographies yielded eight more relevant papers.[7] Each of these 92 included studies explicitly investigate acquisitions rather than mergers (e.g., merger of equals). Because we aggregate the individual samples from primary studies, treating them as independent draws from a meta-sample, each underlying data source can only be included once (Borenstein et al., 2011; Stahl & Voigt, 2008). Where several papers used the same data, we retained only the most recently published, leaving 89 studies in the final sample.[8]

Combined, these studies create a meta-sample of 83,132 M&A deals with study-level sample sizes ranging from 24 (Bednarczyk et al., 2010) to 9,419 (Agyei-Boapeah, 2019) and averaging 934.07 deals. Each primary study was coded for: (1) performance (our dependent variable), (2) experience (our variable of interest), (3) whether experience was used as a (i) variable of interest or (ii) control, (4) the effect size (zero-order correlation or regression coefficient), and (5) sample size.

[6] We added *Administrative Science Quarterly*. Although listed in sociology, it is also a top management outlet.
[7] We are grateful to an anonymous referee for this suggestion to expand our sample in this way. Qualitatively similar results using only the 84 independently drawn studies are available on request.
[8] Ahammad *et al.*, (2016) and Ahammad & Glaister (2011) share a dataset. Cording et al. (2014) and Cording et al. (2008) share a dataset. Puranam et al. (2009) and Puranam et al., (2006) share a dataset.

We further coded each study according to how our primary studies measured performance:

(i) Market-based performance: Cumulated Abnormal Returns (CARs) are commonly used to isolate the immediate market reaction, but broader windows are also used. Hence in some models we distinguish between short event study windows up to 90 days versus longer ones.

(ii) Accounting-based performance: We consider accounting-based ratios (e.g., ROAs, ROEs), management self-assessments based on accounting data, and innovation-based (patent count) measures of synergy realization. Except in *models 9.1.a-9.1.c* these metrics are treated together.

Our sample of studies measured experience in various ways. Most counted the number of acquisitions completed prior to the focal deal as the acquiring firm's 'acquisition experience' (Lin, 2012; Uhlenbruck et al., 2006; Zollo & Singh, 2004). The windows for accounting prior experience range from two to thirty years. Eight papers use the natural logarithm of prior deals anticipating a better model fit due to the *Learning Curve* (Barkema & Schijven, 2008; Ellis et al., 2011) and another eight a simple dummy to indicate whether the firm acquired before at all (Finkelstein & Haleblian, 2002). Others measure more nuanced concepts of experience combining management teams' self-assessment with financial ratios (Ahammad & Glaister, 2011; Slangen, 2006). We thus coded four measures of experience: (i) number of prior deals, (ii) binary (1 if acquired before; 0 otherwise), (iii) the natural logarithm, 'ln,' of deal count, and (iv) use of a self-assessment questionnaire.

We also coded each study for the following context moderators: (6) (i) cross-border sample vs. (ii) domestic sample, (7) whether the study occurred in (i) a single industry, (ii) two related industries or (iii) encompassed more industries, and (8) the time window for which experience counted as relevant. Regarding this final moderator, primary studies admit experience occurring varying amounts of time backwards from the focal deal as relevant. We divide studies into three categories: those using (i) short windows of less than five years, (ii) medium windows of less than ten years, and (iii) longer windows. Table 1 lists the primary studies included in our sample together with several of the coding above.

Table 1: Sample of Primary Studies

Reference	Performance	Experience	Cross-Border	Ind. Setting	N	Pearson Correlation	Partial Correlation	Reference	Performance	Experience	Cross-Border	Ind. Setting	N	Pearson Correlation	Partial Correlation
Agyei-Boapeah 2018	A,R	#		M	9419	0.13	0.08	Kim & Finkelstein 2009	M,L	#		S	2204	-0.07	-0.02
Ahammad & Glaister 2011	A,Q	Q	×	M	65	0.33	0.13	King *et al.* 2008	M,S	#		M	133	0.08	0.06
Arena & Dewally 2017	M,S	#		M	3627	-	0	Kroll *et al.* 1997	M,S	#	×	M	209	0.05	0.03
Barkema & Schijven 2008	A,R	0/1		M	523	0.14	0.31	Laamanen & Keil 2008	M,L	0/1		M	541	-	-0.09
Basuil & Datta 2015	M,L	#	×	M	431	0.05	0.02	Lee & Kim 2014	O	Q		M	607	0.45	-0.04
Basuil & Datta 2017	M,L	#	×	S	222	0.03	-0.03	Lin 2012	M,L	#		M	154	0.20	0.17
Bauer *et al.* 2016	A,Q	ln	×	M	528	0.22	0.07	Ma *et al.* 2016	A,R	#	×	M	272	0.11	0.15
Bauer *et al.* 2018	A,Q	#		M	101	0.15	0.14	McDonald *et al.* 2008	M,L	#		S	1916	-0.09	-0.01
Bednarczyk et al. 2010	M,S	#		S	24	-	0.02	Meschi & Metais 2006	M,S	#		M	291	-0.02	0
Benson & Ziedonis 2009	M,L	#		M	242	0.06	-0.02	Nowiński 2017	M,S	ln		M	104	-	-0.1
Bruton *et al.* 1994	A,R	#		S	51	0.19	0.12	Orsi *et al.* 2015	O	#		M	152	0.47	0.23
Buckley *et al.* 2014	A,R	#	×	M	570	0.03	0	Papadakis 2005	A,Q	ln		M	72	-0.16	-0.14
Campbell *et al.* 2016	A,R	#	×	M	2403	0.01	-0.01	Parola *et al.* 2015	M,S	Q	×	M	310	0.12	0.10
Cefis *et al.* 2019	A,R	#		M	1736	0.1	-0.07	Popli *et al.* 2017	M,L	ln	×	M	292	0.12	0.12
Chang & Tsai 2013	M,L	#		R	4293	-	0	Porrini 2004	A,R	#		M	437	-0.1	0.05
Chao 2018	A,R	#		S	2223	0.06	-0.01	Puranam & Srikanth 2007	O	#	×	M	97	0.15	-0.05
Chemmanur *et al.* 2019	M,S	#		S	1293	-	0.06	Puranam *et al.* 2009	O	0/1	×	R	207	0.3	0.09
Cho & Arthurs 2018	A,R	#		M	270	0.06	0.02	Rabier 2017	M,L	#		S	1222	-0.03	-0.07
Colombo *et al.* 2007	A,Q	0/1	×	M	67	0.63	0.62	Ragozzino & Reuer 2010	A,R	#		S	445	0.03	0
Cording *et al.* 2014	M,L	#		S	129	-0.10	-0.12	Ragozzino 2006	M,S	#		M	409	-0.16	0
Cuypers *et al.* 2017	M,S	#		S	1241	0.13	0.11	Ransbotham & Mitra 2010	M,S	0/1		R	140	0.68	-0.03
Dicova & Sahib 2013	M,S	0/1	×	M	1223	-0.08	-0.05	Reus *et al.* 2016	A,Q	#	×	M	99	0.15	0.21
Ellis *et al.* 2009	A,Q	#	×	R	67	-0.2	-0.13	Reus & Lamont 2009	M,L	#	×	S	118	0.23	0.18
Ellis *et al.* 2011	A,R	ln		M	107	-0.38	0.11	Saboo *et al.* 2017	M,S	#	×	M	319	0.06	0.05
Fang *et al.* 2015	A,R	0/1	×	M	1096	-	0.03	Sears & Hoetker 2014	M,S	#		M	97	0.14	0
Field & Mkrtchyan 2017	M,S	#		S	1766	-	0.06	Shen *et al.* 2014	M,S	Q		M	2948	0.03	0.01
Finkelstein & Haleblian 2002	M,L	#		S	192	-0.16	-0.12	Shi & Prescott 2012	M,L	#		M	421	-0.15	-0.01
Fowler & Schmidt 1989	M,L	#		R	42	0.28	0.26	Slangen 2006	A,Q	ln		M	102	0.17	-0.05
Francis *et al.* 2014	M,L	#	×	S	317	-	0	Slangen & Hennart 2008	A,Q	#	×	M	191	0.31	0.16
Galavotti 2019	A,R	#		M	469	0.09	0.1	Stettner & Lavie 2014	A,R	ln	×	M	435	0.24	-0.06
Goranova *et al.* 2010	M,L	#		M	1131	0.13	0.11	Trichterborn *et al.* 2016	A,Q	#		M	205	0.23	0.16
Gubbi & Elango 2016	M,S	0/1	×	S	589	-0.05	0.1	Tseng & Chou 2011	M,S	#		M	117	-0.26	-0.24
Haleblian & Finkelstein 1999	M,L	#		M	449	-0.14	-0.14	Uhlenbruck *et al.* 2006	M,S	#		S	363	-0.07	-0.19
Haleblian *et al.* 2006	M,S	#		M	6714	0.06	0.03	Vaara *et al.* 2014	A,Q	#	×	M	92	0.26	0.25
Hayward 2002	M,L	#		S	214	0.15	0.14	Vasilaki 2011	A,Q	#	×	M	109	0.23	0.01
He *et al.* 2019	A,R	#	×	S	8725	-	0	Vasilaki 2012	A,Q	#	×	M	139	0.13	0.03
Hebert *et al.* 2005	A,R	#		R	216	0.04	0	Walters *et al.* 2008	M,S	#		S	342	0.05	0.05
Heimericks *et al.* 2012	A,Q	#	×	M	30	-0.05	-0.1	Wright *et al.* 2002	M,S	#	×	S	163	0.19	0.14
Huang *et al.* 2017	M,L	#		R	2115	-0.06	0	Yang 2015	M,S	#		S	1358	0.08	0.06
Humphery-Jenner *et al.* 2019	M,S	#		R	1955	-0.07	0.07	Zaheer *et al.* 2010	M,S	#		S	503	-0.07	-0.17
Humphery-Jenner *et al.* 2017	M,S	#	×	M	4023	-	0	Zhu & Qian 2015	A,R	#	×	M	1191	0.07	0.06
Hutzschenreuter *et al.* 2014	M,S	#		M	65	-0.19	-0.48	Zollo 2009	M,S	#		S	167	0.31	0.14
Jo, Park & Kang 2016	O	ln		M	212	0.24	0.15	Zollo & Reuer 2010	A,R	#	×	M	150	0.03	-0.17
Kedia & Reddy 2016	M,L	ln		M	1120	0.09	0.08	Zollo & Singh 2004	A,R	#	×	M	577	0.03	-0.15
Kim & Davis 2019	A,R	#		M	417	-0.01	0.11								

Notes: Performance $\in \{\boldsymbol{A}ccounting - \boldsymbol{R}atio, \boldsymbol{A}ccounting - \boldsymbol{Q}uestionaire, \boldsymbol{M}arket - \boldsymbol{L}ong\ term, \boldsymbol{M}arket - \boldsymbol{S}hort\ term\}$. *Experience* $\in \{\#Deals, \mathbf{ln}\ \#Deals, \boldsymbol{B}inary, \boldsymbol{Q}uestionaire\}$,. *Industry Setting* $\in \{\boldsymbol{S}ingle, \boldsymbol{R}elated, \boldsymbol{M}ultiple\}$, *Greyed = Variable of Interest.*

Meta-Analytical Procedure

We compute the meta-analytic mean correlations between acquisition experience and deal performance using a random-effects model (Borenstein et al., 2011; Hunter & Schmidt, 2004). Most primary studies in our meta-sample report the relationship between experience and performance as regression coefficients. Unfortunately, the units used to measure these variables, and hence the units associated with the reported regression coefficients, vary from study to study, rendering them incomparable without transformation. Instead, we work with correlations for their consistent, unitless interpretation. Pearson (zero-order or bivariate) correlations are common in meta-analyses of management studies (Duran et al., 2019; Geyskens et al., 2009). However, partial correlations, like regression coefficients, control for correlations beyond the focal pair to mitigate omitted variable bias (OVB) (Stanley & Doucouliagos, 2012). For this reason, partial correlations are now being adopted as effect sizes in some management literature and have long been the standard in economics (Doucouliagos & Ulubasoglu, 2006; Doucouliagos & Ulubas, 2008; Efendic et al., 2011), where identification is paramount (Carney et al., 2011). By using partial correlations, we can control for the effects of omitted variables as well as the original studies did; however, to the extent that experience correlates to other variables not included in a primary study that directly affect performance, the partial correlation for that study remains subject to OVB. We follow Duran et al. (2019) and perform our meta-analyses using both partial and Pearson correlations. Partial correlations have the additional advantage that they can be equivalently derived from a table of Pearson correlations and from individual regression coefficients. Since not all studies in our meta-sample report Pearson correlations, sample sizes for meta-analyses using partial correlations are somewhat larger. Since the results are qualitatively similar, we relegate those from Pearson correlations to the Appendix (see Table A.2, page 48).

We use two equivalent methods to compute partial correlation coefficients. Where primary studies report regression coefficients, we compute partial correlations ($r_{xy,z}$) between experience (x) and performance (y), given the entire set of variables used in the primary study as controls (z) from the reported t-statistics and degrees of freedom as

$$r_{xy,z} = \frac{t_{xy}}{\sqrt{t_{xy}^2 + DF}},$$

where $r_{xy,z}$ denotes the partial correlation between x and y, given a set of control variables z, t_{xy} denotes the t-statistic for the x^{th} independent term in the linear model, y denotes the dependent variable and DF the degrees of freedom in the primary study's analysis (Stanley & Doucouliagos, 2012).[9] In each case, we use the t-statistic corresponding to the coefficient on experience in the authors' preferred regression, typically the most controlled model. In a few cases, t-statistics had to be manually computed from reported standard errors. For studies not reporting regression coefficients, we derived partial correlations from the reported zero-order correlation matrix Ω reported as follows:

$$r_{xy,z} = -\frac{p_{xy}}{\sqrt{p_{xx}^2 p_{yy}^2}},$$

where $r_{xy,z}$ denotes the partial correlation between the independent variable x and the dependent variable y, controlling for the full set of variables in Ω except x and y, and p_{xy} is the xy^{th} element of the precision matrix $P = \left(p_{xy}\right) = \Omega^{-1}$. The two methods generate identical partial correlations.

[9] For example, in a typical regression analysis with an intercept, the degrees of freedom are given by the number of observations N minus the total number of regressors k and the dependent variable ($N - k - 1$).

Homogeneity & Moderator Analysis

Heterogeneity in meta-analysis refers to the dispersion of the true effect sizes rather than random errors in outcomes between studies. Confidence intervals, typically 95 percent, convey the precision with which the mean effect sizes are estimated—they give the range in which the estimates could be expected to vary between new hypothetical samples (Borenstein et al., 2011; Lipsey & Wilson, 2001).

We compute Cochran's Q and I-squared statistics to test the significance of heterogeneity of the sample variance of our mean effect size estimates (Borenstein et al., 2011; Lipsey & Wilson, 2001). The former tests the null hypothesis that all studies share the same effect size. When Q exceeds the $(1-\alpha)$-quantile of the chi-squared distribution with $k-1$ degrees of freedom, where α is the desired significance level (set to 0.05) and k is the number of studies, then the null is rejected, and we can conclude that the studies do not share a common effect size. The I-squared statistics provides the proportion of variance across the studies resulting from heterogeneity rather than chance (Higgins et al., 2003). A high I-squared indicates that the variance of included effect sizes exceeds the expected level due to chance alone, suggesting other variables *moderate* the relationship between experience and M&A performance. Higher values of both indicate thus greater heterogeneity, but the latter is more intuitively interpreted. Higgins et al. (2003) suggest the following benchmarks for I-squared interpretations: up to 25per cent as low, up to 50 per cent moderate, and over 75 per cent as high. They suggest that if 75 per cent is exceeded, the observed variance is largely due to differences in the primary studies (rather than measurement error), indicating that an appropriate subgrouping of the primary studies may reduce the variance, thereby pointing to specific conditions which may moderate the effect of experience on M&A performance. It is also generally more difficult to derive significant results from more heterogeneous studies, because heterogeneous samples naturally tend to yield larger confidence intervals than homogeneous ones, and these tend to be less informative, because heterogeneity suggests that the effect depends on conditions, which do not uniformly hold.

In meta-analyses, moderators are examined via subsample analysis rather than through interaction effects, as is common in linear regression methods. By subsequently re-analyzing sub-groups of studies (using the same protocol and specifications than before) one can increase the homogeneity of effect sizes, and thereby determine whether any significant results are driven by a particular subgroup of studies—say those using a certain measure for a dependent variable or variable of interest. To compare the mean effect sizes of groups and test whether those are different as a result of the membership in a specific group, we apply the Q-test based on the analysis of variance. This test reports the fraction of variance between the groups in the analysis and the grand mean of all combined effect sizes ($Q_{Between}$) (Borenstein et al., 2011). In this context, the variable over which subgroups are divided is called the moderator (DeCoster, 2004). We defined our subgroups in previous sections.

RESULTS

Table 2: Results of meta-analyses (base and moderators)

Moderator		Sample	*k*	*N*	Point Est.	CI	Cochran's Q	I²	Q_{Bet}
None		(1) Full Sample	89	83,132	**0.03 (0.00)**	(0.01, 0.05)	385.95 (0.00)	77.20	
Performance Measurement		(2.a) Market-based	49	48,258	0.01 (0.20)	(-0.01, 0.03)	155.21 (0.00)	**69.07**	5.22 (0.02)
		(2.b) Accounting-based	40	34,874	**0.06 (0.00)**	(0.02, 0.08)	224.18 (0.00)	82.60	
Experience Measurement		(3.a) Deal Count	69	71,000	**0.02 (0.02)**	(0.00,0.04)	290.50(0.00)	76.59	2.45 (0.48)
		(3.b) Dummy	8	5,257	0.02 (0.54)	(-0.04, 0.07)	12.98 (0.07)	**46.10**	
		(3.c) ln(Deal Count)	8	5,568	0.05 (0.29)	(-0.04,0.14)	63.02 (0.00)	88.89	
		(3.d) Self-Assessment	4	737	**0.13 (0.08)**	(-0.01, 0.27)	10.33 (0.02)	**70.94**	
Perf. Measurement (Deal Count Only)		(4.a) Market-based	39	40,276	0.01 (0.31)	(-0.01, 0.03)	141.52 (0.00)	**73.14**	2.25 (0.02)
		(4.b) Accounting-based	30	30,724	**0.04 (0.01)**	(0.01, 0.08)	143.55 (0.00)	79.79	
International		(5.a) Domestic	56	38,692	**0.02 (0.07)**	(0.00, 0.04)	211.85 (0.00)	**74.03**	0.94 (0.22)
		(5.b) Cross-border	33	44,228	**0.04 (0.00)**	(0.01, 0.06)	173.40 (0.00)	81.54	
Industries		(6.a) Single	23	24,292	0.00 (0.95)	(-0.03, 0.03)	73.51 (0.00)	**70.07**	6.45 (0.03)
		(6.b) Related	7	3,420	-0.02 (0.54)	(-0.11, 0.06)	19.06 (0.00)	**68.07**	
		(6.c) Multiple	59	55,420	**0.04 (0.00)**	(0.02, 0.06)	283.21 (0.00)	79.52	
Recency of Experience		(7.a) < 5 years	21	12,250	**0.05 (0.00)**	(0.02, 0.08)	44.27 (0.00)	**54.82**	5.02 (0.17)
		(7.b) < 10 years	53	49,906	**0.02 (0.04)**	(0.00, 0.04)	205.92 (0.00)	**74.74**	
		(7.c) < 18 years	13	20,976	0.02 (0.56)	(-0.04, 0.06)	113.39 (0.00)	89.41	
Market Reaction (Market Based)	Event Windows	(8.1.a) ≤ 90 days	28	30,493	0.02 (0.14)	(-0.01, 0.04)	85.72 (0.00)	**68.77**	0.43 (0.51)
		(8.1.b) > 90 days	21	17,765	0.01 (0.73)	(-0.02, 0.03)	64.05 (0.00)	**68.50**	
	International	(8.2.a) Domestic	35	32,696	0.01 (0.49)	(-0.02, 0.03)	130.47 (0.00)	**46.94**	0.22 (0.64)
		(8.2.b) Cross-border	14	15,562	0.02 (0.18)	(-0.01, 0.04)	24.50 (0.02)	**73.94**	
	Industries	(8.3.a) Single	14	14,741	0.01 (0.96)	(-0.02, 0.03)	31.52 (0.00)	**58.75**	2.60 (0.27)
		(8.3.b) Related	4	7,454	-0.11 (0.28)	(-0.3, 0.1)	15.32 (0.00)	80.42	
		(8.3.c) Multiple	31	26,063	**0.02 (0.07)**	(-0.01, 0.04)	104.61 (0.00)	**71.32**	
	Recency of Experience	(8.4.a) < 5 years	8	8,942	0.04 (0.10)	(-0.01, 0.09)	25.31 (0.00)	**72.34**	5.25 (0.15)
		(8.4.b) < 10 years	32	31,234	0.02 (0.15)	(-0.01, 0.04)	92.14 (0.00)	**66.49**	
		(8.4.c) < 18 years	8	8,082	-0.03 (0.25)	(-0.01, 0.09)	32.14 (0.00)	78.22	
Synergy	Performance Measurement	(9.1.a) Ratios	21	31,732	0.03 (0.13)	(-0.01, 0.06)	145. 09 (0.00)	86.21	2.57 (0.01)
		(9.1.b) Self-Assessed	14	1,867	**0.13 (0.01)**	(0.04, 0.23)	50.53 (0.00)	**74.27**	
		(9.1.c) Others	5	1,275	0.07 (0.17)	(-0.03, 0.18)	12.93 (0.01)	**69.07**	

	International	(9.2.a) Domestic	21	6,208	**0.05 (0.06)**	(0.00, 0.11)	79.60 (0.00)	**74.84**	3.32 (0.00)
		(9.2.b) Cross-border	19	28,666	**0.06 (0.01)**	(0.02, 0.11)	144.55 (0.00)	87.54	
	Industries	(9.3.a) Single	9	3,838	0.01 (0.87)	(-0.07, 0.08)	41.81 (0.00)	80.86	7.05 (0.03)
		(9.3.b) Related	3	2,880	0.00 (0.89)	(-0.04, 0.04)	2.22 (0.32)	**9.92**	
		(9.3.c) Multiple	28	28,156	**0.08 (0.00)**	(0.04, 0.11)	171.65 (0.00)	84.27	
	Recency of Experience	(9.4.a) < 5 years	13	3,308	**0.07 (0.00)**	(0.02, 0.11)	14.77 (0.25)	**18.80**	2.33 (0.51)
		(9.4.b) < 10 years	21	18,672	**0.04 (0.07)**	(-0.01, 0.08)	109.22 (0.00)	81.69	
		(9.4.c) < 18 years	5	12,864	**0.10 (0.04)**	(0.01, 0.19)	50.81 (0.00)	92.13	

Notes: k = number of included studies; N =total number of deals across included studies; point estimate= weighted mean effect size (p-values in parentheses; bold typeface indicates significance at less than 10%); CI= 95% Confidence Interval of estimates, Q = value of chi-square distributed homogeneity statistics (p-values in parentheses); I^2 = proportion of the observed variance reflecting differences in true effect sizes rather than sampling error ($I^2 < 75$ are bolded).

Table 2 reports all models of our meta-analysis including moderators. *Model 1* estimates a modest, positive and significant association between experience (*i.e.,* using all metrics) and performance (*i.e.,* both *market-* and *accounting-based* metrics) with $r_{xy,z} = 0.03, p = 0.00$. However, our heterogeneity metrics ($Q = 385.95$ and $I^2 = 77.2$) indicate that the effect is not uniform across our sample of studies and that other factors moderate this overall correlation.

Measurement-based Moderators

Models 2.a and 2.b moderate our analysis by the performance metric used in the primary studies. The magnitude and significance of experience's association doubles in the subsample using *accounting-based* performance measures ($r_{xy,z} = 0.06, p = 0.00$), while in studies using *market-based* measures it is insignificant. Based on the variance between the groups ($Q_{Bet} = 5.22, p = 0.02$), we reject the null hypothesis of identical effect sizes in both subsamples and conclude that the groups are different from each other.

Next, we moderated by the experience metric used. The effect size of the largest subgroup, *deal count* (model 3.a), falls ($r_{xy,z} = 0.02$) relative to the baseline model, and remains significant ($p = 0.02$) but heterogeneous ($Q = 290.50$, $I^2 = 76.59$). The effect among studies measuring experience *binarily* (model 3.b) was insignificant. Those using a common *self-assessment* survey (model 3.d) yielded an effect size significant at conventional levels and larger than the baseline model ($r_{xy,z} = 0.13, p = 0.08$) and was moderately homogenous ($Q =$

$10.33, I^2 = 70.94$). While the coefficient for model 3.c ($r_{xy,z} = 0.05, p = 0.29$) exceeds the coefficient for model 3.a, suggestive of the *Learning Curve*, it is neither statistically significant, nor homogenous ($Q = 63.02,\ I^2 = 88.89$). We cannot reject the null hypothesis that the effect in all four subgroups is the same ($Q_{Bet} = 2.45, p = 0.48$). Still, experience positively relates to deal performance in all subsamples over experience metrics (models 3.a-3.d). Given that, except for *deal count*, the number of studies using any single measurement for experience is small, the lack of significance is unsurprising. Finally, by restricting the subsample of studies to those using *accounting-based* performance measures and the most widely utilized experience metric, namely *deal count*, (model 4.b), a significant positive association between experience and performance of 0.04 ($p = 0.01$) could be identified, however heterogeneity remains high ($Q = 10.33, I^2 = 79.79$). Again, we reject that the size of this effect is the same in studies using market-based performance metrics ($Q_{Bet} = 2.45, p = 0.02$).

Hence, we conclude that experience positively relates to *accounting-based* measures of M&A performance. This effect size of 0.04 (model 4.b) to 0.06 (model 2.b) is both significantly different from zero and from the effect size on *market-based* performance. The effect is also positive in the subsample analysis over experience measures (models 3.a-d), though not universally significant. The general heterogeneity associated with subgroups generating significant effects suggests that the literature has not employed an experience measure that consistently captures the same underlying quantity in enough studies to deliver homogenous results.

Since *market-based* performance metrics typically capture (short-term) investor returns, while *accounting-based* metrics are generally taken as proxies for (longer-term) synergy realization, one may ask, "Is the difference between the relationship of experience and the two distinct ways to measure performance meaningful?" There are at least two possibilities: (1) Investors fail to recognize the true value of acquirers' M&A experience in future deals. If so, it suggests that the activities where experience is most useful to are too subtle for the market to perceive or too complex for it to unravel, but that *ex post* accounting-based measures are sensitive enough to capture gains from such complex mechanisms. (2) Alternatively, there are factors observed by investors but unobserved by the empiricists of our primary studies, which offset gains in accounting measures of performance and *correlate to experience*. For example, if experience improves return on assets (ROA), but experienced buyers systematically pay for more acquisitions, then experienced acquirers' deal announcements will not trigger a valuation increase. Meta-analysis cannot disentangle these interpretations. Hence, we recommend future primary studies target these specific possibilities raised by our meta-analysis.

Context-moderators

Models 5.a and 5.b respectively examine whether experience matters more in *domestic* or *cross-border* settings. The subsample of domestic deals exhibits moderate homogeneity ($Q = 211.85, I^2 = 74.03$) and although both subsamples yield positive and statistically significant effect sizes ($r_{xy,z} = 0.02, p = 0.07$, $r_{xy,z} = 0.04, p = 0.00$ respectively), we cannot reject the null hypothesis that they are identical.

Models 6.a-c divide the sample into studies including acquisitions within a *single industry*, *two related industries* or across *many industries*, respectively. Only the final subsample yields a significant effect; it is again positive ($r_{xy,z} = 0.04, p = 0.00$), though heterogeneity remains high ($Q = 283.21, I^2 = 79.52$). We reject the null of common effect sizes across subgroups ($Q_{Bet} = 6.45, p = 0.03$).

Authors generally restrict primary samples to one or two related industries as a method of controlling for unobserved (industry-level) heterogeneity in deals (Kim & Finkelstein, 2009).

Our purpose here is different. We posit that cross-industry acquisitions may be more complex than deals in the same industry. Since these are more prevalent, indeed can only occur, in primary samples covering M&As in multiple industries, we use our subgroups as a proxy for complexity in the dimension of industrial distance in acquisitions. To the extent that cross-industry transactions are more complex, this lends support to our view that experience is more positively related to performance in complex settings.

Models 7.a – 7.c moderate our analysis by the *recency of experience* considered in the primary studies. All studies included in model 7.a capture only M&A experience *less than 5 years* before the focal deal, those in model 7.b consider experience *up to 10 years* old, while studies included in model 7.c do not restrict experience based on when it accrued. The estimated association is positive in all subgroups and decreases from $r_{xy,z} = 0.05$ ($p = 0.00$) in the sample using only experience less than five years old to $r_{xy,z} = 0.02$ ($p = 0.04$) in the sample allowing experience less than 10 years old. The effect in studies allowing even longer experience windows is insignificant. Perhaps intuitively, the heterogeneity of variance increases as larger experience windows are allowed ($Q = 44.27, I^2 = 54.82; Q = 205.92, I^2 = 74.74; Q = 113.39, I^2 = 89.41$). These findings support *Transfer Theory*'s view that only more "relevant" experience improves performance but given that we cannot reject the null that the underlying effect sizes are identical ($Q_{Bet} = 5.02, p = 0.17$), this evidence is weak. Further, as all effects are positive, even in studies with longer windows, we find no evidence for *Transfer Theory*'s most extreme predictions that experience can be harmful.

Joint Measurement & Context Moderators

Our moderation over metrics reveals what types of measurement yield consistent results for the association between experience and M&A performance and what types do not. This informs us about the robustness of the literature's answers so far, and the fact that experience relates more positively to accounting-based performance measures than market-based ones raise some consequential questions for future research; however, revelations for managers are limited. On the other hand, our context moderators might yield new management insights, but although the effects in several subgroups are significant, the between-group heterogeneity remains too high to determine whether effects differ between the "treated" (e.g., *cross-border*) and "untreated" (e.g., *domestic*) subsample. So, to find more homogeneous subgroups yielding more significantly distinguishable effects over dimensions with business impact, we moderate over performance measurement and context simultaneously. In particular, we divide our sample of primary studies into those using *market-based* (models 8.1.a. to 8.4.c) versus *accounting-based measures* (models 9.1.a to 9.4.c), and then further divide these groupings by finer metrics measures and by our context moderators.

Overall, these tighter subgroup analyses reconfirm the above results with a higher degree of homogeneity. Models 8.1.a and 8.1.b divide the sample of studies using *market-based* performance metrics into those examining abnormal market reactions over windows less than or equal to 90 days versus longer windows—although the homogeneity of the samples increases ($Q = 85.72, I^2 = 68.77; Q = 64.05, I^2 = 68.50$), neither produces effects significantly different from zero. This general pattern is seen in all the moderation analyses in models 8.1.a – 8.4.c, excepting *multiple industries* (model 8.3.c; $r_{xy,z} = 0.02, p = 0.07$). This corroborates the view that stock market reactions to deal announcements largely ignore M&A experience.

Meanwhile, models 9.1.a – 9.4.c subdivide studies using *accounting-based* metrics into finer resolution yielding more meaningful results. Models 9.1.a – 9.1.c subdivide these studies into those using *accounting ratios*, *self-assessment,* and *others* to measure performance—the effect size of all is positive but only on *self-assessment* is it significant ($r_{xy,z} = 0.13, p = 0.01$). Homogeneity increases in the latter two groupings ($Q = 50.53, I^2 = 74.27; Q = 12.93, I^2 = 69.07$). Taken together this reaffirms our findings that the association between experience and synergy realizations in M&As is positive. Turning to our context moderators, experience relates positively to accounting performance in both *cross-border* deals ($r_{xy,z} = 0.06, p = 0.01$) and *domestic deals* ($r_{xy,z} = 0.05, p = 0.06$), but now we reject the null that this effect is the same ($Q_{Bet} = 3.32, p = 0.00$). Again, the strength of the positive association increases as the industrial breadth of included M&As increases (models 9.3.a-c), though only in samples including deals across many industries is it significantly different than zero ($r_{xy,z} = 0.08, p = 0.00$), and the null of common effect sizes is rejected ($Q_{Bet} = 7.05, p = 0.03$).

Here *Transfer Theory*'s prediction (models 7.a – c) that the effect of experience declines (even to the point of negativity) as experience becomes less recent (models 9.4.a-c) is absent: While studies considering only experience *less than five years* yield a precisely estimated association of 0.07 ($p = 0.00$) from a very homogeneous group of studies ($Q = 14.77, I^2 = 18.80$) and those considering *any experience* (< 18 years) an even larger estimate ($r_{xy,z} = 0.10, p = 0.04$), the group admitting experience *up to ten years* old yielded an imprecise estimate smaller than either ($r_{xy,z} = 0.04, p = 0.07$). In the end, we cannot reject the null hypothesis of identical true effect sizes ($Q_{Bet} = 2.33, p = 0.51$). Overall, evidence for *Transfer Theory* across our meta-sample is weak.

We ran several robustness checks. First, we tested the sensitivity of our results to outliers. Stanley & Doucouliagos (2012) distinguish between "outliers" and "leverage points". They recommend funnel plots to identify both. The former are (implausibly) large effect sizes estimated with low precision. Two small sample studies of this type stand out (the horizontally most extreme red diamonds) in our funnel-plot (*Figure 1*): Hutzschenreuter et al., (2014) with effect size -0.48, and Colombo et al., (2007) with effect size of 0.62. These are also the two studies picked up by so-called "three-sigma (3σ)" and "1.5×interquartile-range (IQR) rules" used across empirical sciences to identify outliers.[10] Standard meta-analysis weights effect sizes according to their precision, and hence, even without special treatment, these "outliers" barely influence the results. Leverage points though, defined by as extreme effect sizes having high precision, can influence results strongly.[11] Neither *Figure 1*, 3σ- nor 1.5×IQR rules identify any leverage points. Hence, unsurprisingly dropping these two studies does not change our results meaningfully, but these results can be found in *Table A.1* of the Appendix.

Second, we reran the above analysis using Pearson correlations, rather than partial ones (see (Carney et al., 2011) and (Duran et al., 2019) for examples of meta-analysis using both). These results, reported in *Table A.2* of the Appendix, qualitatively resemble those using partial correlations. In particular, the positive association between experience and performance is consistent, if not stronger and more robust. Again, this is truer when accounting metrics capture performance. The effect sizes between domestic and cross-border settings as well as across industry settings cannot be statistically distinguished though. Evidence for *Transfer Theory*'s predictions never materializes. Since the meta-sample of studies reporting Pearson correlations is somewhat smaller, the variations in magnitude and significance are to be expected, even without considering the controls that partial correlations bring.

Complexity

Synthesizing the findings of our moderator analyses—in particular, that the *positive* association between experience and M&A performance independently strengthens in: (1) cross-border

[10] See for example, https://en.wikipedia.org/wiki/68%E2%80%9395%E2%80%9399.7_rule and https://en.wikipedia.org/wiki/Interquartile_range#Outliers.

[11] Unless leverage points represent an error, they are informative and should be retained (Stanley & Doucouliagos, 2012).

deals (model 5.b), (2) multiple industry settings (model 6.c), and (3) when performance is measured with accounting (model 2.b) rather than market-based metrics—collectively suggests that experience may be more useful in more complex acquisition environments.

How do these factors relate to complexity? Because the pre-acquisition systems, procedures, supply chains and so on of the respective partners likely aligned with their domiciles, cross-border M&A integrations tend to be between more different entities and hence be more complex on average. Others have recognized the relationship between complexity and cross-border transactions—Cording *et al.* (2008) explicitly limit their study to domestic transactions to minimize complexity. The same logic applies to acquisitions across industrial boundaries—technological and market differences could well be greater between firms in different industries (Haleblian & Finkelstein, 1999). Likewise, many authors (e.g., Kim & Finkelstein, 2009) limit their samples to specific industries, like banking, to rule out industry level factors—a source of complexity both for the empiricist and manager alike.

The relationship between accounting metrics and complexity is subtler. The primary researcher's choice to use accounting-based performance measures does not, in and of itself, indicate that the studied settings are more complex. Rather, as we noted previously, when accounting-based performance metrics are sensitive to the effects of experience, but market ones are not, then apparently, experience is particularly valuable for navigating obstacles that are invisible or inscrutable to public investors. To the extent that complexity induces such opacity, the difference between experience's effect, as captured by accounting metrics but not market ones, positively relates to the underlying complexity of the deal. Hence, the relationship between performance metric and deal-level complexity is indirect. Further, there may be other factors besides complexity that drive a measurable *ex-post* relationship between performance (through accounting metrics) and experience where an *ex ante* one cannot be perceived (through market metrics). Therefore, we include an indicator for whether accounting metrics were used in the primary study as an additional proxy for complexity in the analysis below, fully cognizant of its more speculative nature. We hope that this *post-hoc* sense-making of the study-level moderators available in our meta-sample stimulate rigorous primary studies focused on more precisely measured deal-level complexity.

How then, might complexity increase the strength of experience's positive effect on M&A performance? First, notice that such a moderation effect does not contradict the natural intuitions that complexity reduces performance or learning complex things or learning in complex environments is harder. On the contrary, it is this difficulty that likely makes experience increase performance more. Intuitions from *Learning Curve Theory* help. When encountering a series of simple acquisitions, mastery is close, even on the first one, and non-acquisition experiences, which would not be recorded by a researcher as experience at all, get managers to the curve's flatter tail—there is simply less weakness for previous experience to improve upon. On the other hand, when a series of acquisitions are complex, then managers start low on the *Learning Curve*, where difficulty is great and performance is not, but this is precisely where the derivative of the *Learning Curve* is steepest, and novel experience (not just everyday business history) gained in complex deals early in the series helps most.

We are not the first to consider the impacts of complexity on learning (Bohlen & Barany, 1976). In most cases, complexity is simply introduced as a control when studying the performance effects of various types of experience (see e.g., (Huckman et al., 2009), among others) or a descriptor to distinguish the entire study setting (e.g., Ackerman, 1992). Its moderating effect on performance is little studied. Nembhard (2000) provides a notable exception in the industrial engineering literature.[12] He observed that experienced textile line workers learned more quickly relative to inexperienced ones, and the difference became more pronounced the more complex the new task. Given that other observations in *Organizational Learning Theory*, like the *Learning Curve*, have successfully translated from such low-level manufacturing tasks to higher order managerial ones, the positive moderating effect of complexity may also. To our knowledge, though, the moderating effect of complexity on performance of managerial-level tasks has not been previously investigated.

Although meta-analysis' ability to test the effect of novel independent variables not included in the primary studies is fundamentally limited, we can dig deeper. To do so, we create a set of composite moderators of increasing setting complexity. We define $Complexity_1$ as a dummy variable set to unity for any primary study that is *either* set in a cross-border *or* multiple-industry setting *or both.* The $Complexity_2$ indicator is set to unity for any study that is *both* cross-border *and* multiple-industry. Finally, $Complexity_3$ studies are cross-border *and* multiple-industry *and* use accounting-based performance metrics. Hence, $Complexity_{i+1} \subset Complexity_i$. We assume that an environment that is complex in more dimensions is more complex. So, complexity increases in i.

These complexity metrics are admittedly coarse. First, like all independent variables in meta-analysis, they are study-level since we do not observe any individual deal-level attributes. Second, just because a primary study allowed for individual deals to be more complex in some dimension, we may not know how many deals satisfied the complexity criterion. Hence, we can learn the direction of complexity's effect but not its magnitude. To illustrate, suppose that

[12] McDaniel et al. (1988) examine the relationship between individual job experience and individual job performance. They perform a subsample analysis by the cognitive complexity of the job. They estimate a slightly higher effect for low complexity jobs, but the difference from the effect in high complexity jobs is not statistically significant.

we wished to know the effect of oranges versus apples (symbolizing complex and simple deals respectively) on X. Further suppose that we cannot count the number of apples versus oranges in each box (a primary study), but we know that some boxes contain only apples, while others have both. If we observe that "mixed" boxes' affect X differently from "apples only" boxes, then we know oranges matter and even in which direction but cannot say anything sensible about the *magnitude* of an orange's effect. Our subsample of *multiple-industry* studies suffers this complication, since some indicated primary studies allow cross-industry deals to be included alongside within-industry ones. In any case, the fact that our more complex subsamples are contaminated by simple (within industry) acquisitions that are unobservable at the study-level, will attenuate any detected effect of complexity in our subsample analysis. It empirically works against us finding a statistically significant effect. Hence, our reported effect sizes should be interpreted as a lower bound. In principle, our *cross-border* subsample could have the same problem, but it does not, because our *cross-border* subsample of primary studies all excluded domestic only deals.

Table 3: Complexity

Moderator	Sample	*k*	*N*	Point Est.	CI	Cochran's Q	I^2	Q_{Bet}
$Complexity_1$ (Cross-border | Multiple)	(10.a) Complex	67	7,0553	**0.04 (0.00)**	(0.02, 0.06)	283.51 (0.00)	76.60	7.56 (0.00)
	(10.b) Non-complex	22	12,579	-0.02 (0.30)	(-0.06, 0.02)	89.97 (0.00)	76.72	
$Complexity_2$ (Cross-border & Multiple)	(11.a) Complex	24	24,027	**0.07 (0.00)**	(0.03, 0.10)	140.24 (0.00)	83.59	2.99 (0.00)
	(11.b) Non-complex	65	59,105	**0.02 (0.08)**	(0.00, 0.04)	242.67 (0.00)	**73.62**	
$Complexity_3$ (X-border & Mult. & Acc.)	(12.a) Complex	15	12,050	**0.14 (0.00)**	(0.06, 0.22)	120.07 (0.00)	**72.39**	9.78 (0.00)
	(12.b) Non-complex	74	71,082	0.01 (0.13)	(0.00, 0.03)	264.45 (0.00)	88.34	

Notes: k = number of included studies; N =total number of deals across included studies; point estimate= weighted mean effect size (p-values in parentheses; bold typeface indicates significance at less than 10%); CI = 95% Confidence Interval of estimates, Q = value of chi-square distributed homogeneity statistics (p-values in parentheses); I^2 = proportion of the observed variance reflecting differences in true effect sizes rather than sampling error.

Cognizant of these limitations, Table 3 presents the results of layered subsample analyses over our *complexity* moderators. Models 10.a and 10.b show that studies classified as $Complexity_1$

have a significantly ($Q_{Bet} = 7.56, p = 0.00$) larger effect size ($r_{xy,z} = 0.04, p = 0.00$) than those without the designation ($r_{xy,z} = -0.02, p = 0.30$). In models 11.a and 11.b, for $Complexity_2$ studies the effect size increases ($r_{xy,z} = 0.07, p = 0.00$) and statistically differs ($Q_{Bet} = 2.99, p = 0.00$) from those without the designation ($r_{xy,z} = 0.02, p = 0.08$). Finally, in models 12.a and 12.b, for $Complexity_3$ studies, which adds our indirect complexity measure, an indicator for the use of accounting performance metrics, the effect size doubles ($r_{xy,z} = 0.14, p = 0.00$) and again significantly differs ($Q_{Bet} = 9.78, p = 0.00$) from those not in the subgroup ($r_{xy,z} = 0.01, p = 0.13$). Even accounting for 2 per cent ($r^2_{xy,z} = (0.14)^2 = 2\%$) of the variation in M&A performance seems economically quite large when one considers the monetary value at stake and how little of M&A performance the literature explains. Two caveats are worth recognizing: (1) This effect applies only in the most "complex" subsample we could build, a group whose size we cannot estimate in the population. (2) Although partial correlations correct for biases that could occur between performance drivers and factors correlated to experience, it may be that some primary studies in our subsample could have done better in adding controls. Nevertheless, the increase in effect size with complexity seems meaningful. However, since the "complex" subgroups are subsets of one another, a comparison of $Q_{Between}$ statistics for them would not be; instead, we simply note that the individual effect sizes are precisely estimated and monotonically increasing in the degree of complexity. Running our complexity moderators on the sample removing outliers and using Pearson correlations revealed the same basic pattern—as complexity increases, so does its positive moderating effect. On balance, we find support for the view that complexity positively moderates the association of experience on performance and recommend deal-level examination as the next step.

DISCUSSION

Our study makes three contributions: (1) We establish the positive relationship between acquirer experience and M&A performance. While intuitive, this result is non-obvious in that it cannot be perceived by a non-meta-analytical review of the literature—the number of studies reporting contrary results equals the confirmatory, but the former studies are weaker. The result is also apparently non-obvious to investors, as the effect size from studies using stock market reaction as a proxy for performance is indistinguishable from zero in all specifications and significantly less than from those using accounting-based measures. The consistency of the positive association in our summary effects across many subsamples suggests that reported individual negative effects in the literature are anomalous rather than reflecting the strongest implications of *Transfer Theory*. (2) We find that *Complexity* positively moderates experience's association with M&A performance. Until now, the positive moderation of complexity on experiences' performance effect has only been observed in low-level manufacturing tasks. There is still more work to do, though, to understand the mechanism. (3) Despite confirming the association between experience and M&A performance as positive, we formally report the considerable discord in this literature. This dissension in the study of M&A matters, precisely because the fact that "experience in X improves performance in X" is confirmed with virtual unanimity in most other settings, suggesting that learning in M&A differs from the standard. We highlight probable reasons why measuring learning in M&A (and other business partnerships) generates so much disagreement and how it may be remedied. We expand on these contributions below.

Positive Relationship

Our unmoderated analysis revealed a positive correlation between experience and acquisition performance ($r_{xy,z} = 0.03, p = 0.00$) as *Organization Learning Theory* would suggest. In some specifications, the effect size grows (see e.g., model 9.4.c.: $r_{xy,z} = 0.10, p = 0.04$) but never becomes large by traditional management literature standards; however, one should bear in mind that our effect sizes are fully controlled. Still, a marginal correlation of even just 0.03 in explaining what happens to a $4 trillion a year investment is economically significant, especially when finding dominant explanations for M&A performance has proved so elusive. However, this substantial effect, even though statistically significant, needs to be interpreted with caution—it is simply an average effect across many studies, and there may not be even a single setting where this effect size can be observed.

Our analysis offers two ways to look at the consistency or robustness of the above positive relationship. The first is technical: by examining the heterogeneity of the effect sizes in the overall sample of primary studies and in various subsamples, as measured by Cochran's Q or I^2 metrics. From this vantage, the field's conclusions remain murky—*both* significant and homogenously estimated effects arose in just ten of 44 specifications (models 3.d, 5.a, 7.a, 8.3.c, 7.b, 9.1.b, 9.2.a, 9.4.a, 11.b and 12.a). Without a strong theoretical connection between these, we attribute the higher homogeneity in these specifications to chance. We could not identify a context where the literature provides a unified answer.

The second way to look for robustness is holistically. In total, we sliced the 89 primary studies 44 different ways. Every single significant estimate of the association of experience on performance, 22 in all, is positive, though the subsamples generating those estimates range from extremely heterogenous to very homogenous. Although nearly half of the individual primary studies (41) report non-positive effects our findings strongly suggest that the analyzed relationship between performance and acquisition experience is, in fact, positive. Support for *Organizational Learning* is broad.

These findings suggest that settings where such forces can completely reverse the natural positive direction of experience's effect on performance in a population (rather than small sample) are rare or non-existent—evidence *against* these strongest interpretations of *Transfer Theory*. This is not to say we can reject those milder implications of *Transfer Theory* are at work in some studies, but we also find no statistically robust evidence that more recent experience improves performance.

Our sample of studies using logarithmic experience metrics is too small to find conclusive evidence for the *Learning Curve*, but suggestively, the coefficient nearly doubles relative to the linear model (see model 3.c: $r_{xy,z} = 0.05, p = 0.29$).

Complexity Positively Moderates

The second key finding is that this positive relationship between experience and acquisition performance strengthens in primary studies set in complex contexts. From empirically strongest to weakest, this is evidenced by larger effect sizes in each of our "natural" subsamples characterized by more complexity vis-à-vis less: accounting- vs. market-based performance metrics, multiple vs. related vs. single industry contexts, and cross-border vs. domestic settings. We also create subsamples over "derived" complexity variables: the first includes all studies set in *either* cross-border *or* multiple industry settings *or* both, the second, all studies that are *both* cross-border *and* multiple industry, and the third, all studies that are cross-border *and* multiple industry *and* utilize accounting-based performance measures. Hence, the three subsamples increase in complexity, and likewise so does the estimated relationship between experience and performance in each.

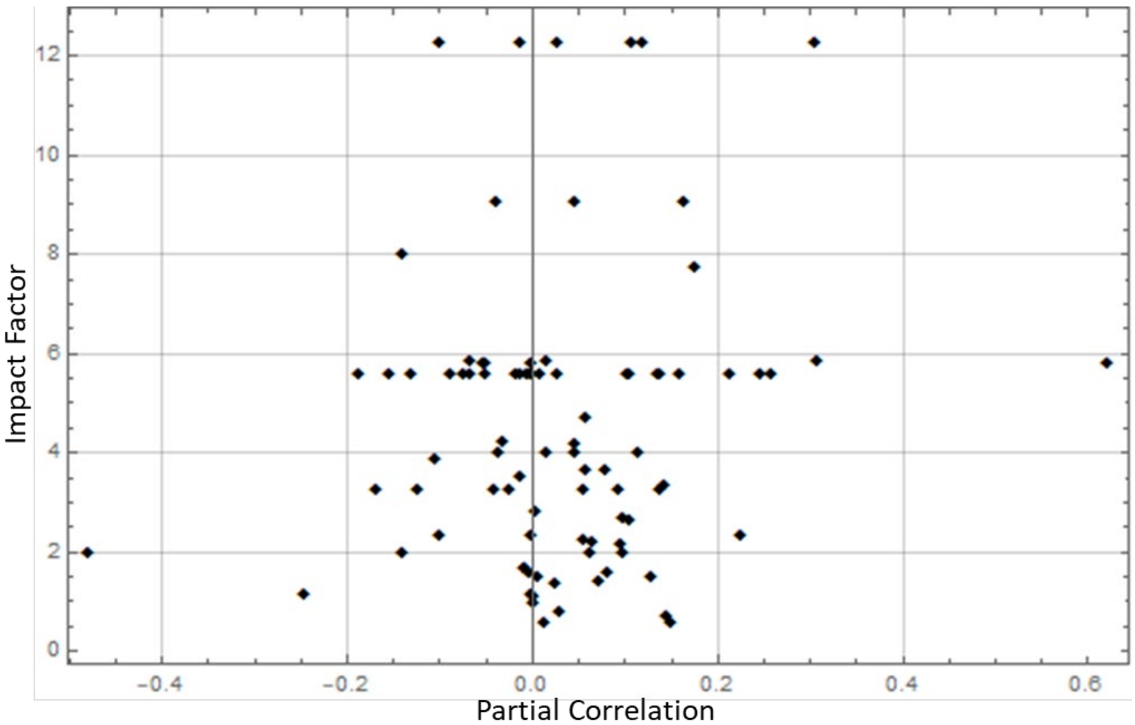
Although our meta-analytic results suggest that experience improves performance more in complex environments, they cannot substitute for an empirical analysis of complexity at the deal level. Future primary studies should seek to understand how the type and diversity of previous experience interacts with the complexity and novelty of the focal setting.

Potential Causes of Discord

A final analytical contribution of our study is to make plain how much empirical disagreement exists about the relationship of experience to acquisition performance. While our meta-analysis can help make sense of conflicts in the literature that arise due to random variation in samples used in individual studies, meta-analysis cannot settle all debates. At a high level, since the methods of meta-analysis (and the methods of the included studies themselves) are observational rather than (quasi-) experimental, we cannot conclude that the statistically significant relationships that we observe are causal. Here, we highlight three specific sources of estimation bias: (1) Although meta-analysis cannot correct it, it can help diagnose whether a literature suffers from publication bias. (2) Self-selection- and (3) omitted variable-bias could be especially problematic in studying the relationship between experience and M&A performance and may systematically drive the discord seen in the literature.

Publication Bias

Publication bias occurs when papers or results are selected to confirm expected or desired outcomes. Editorial teams may be predisposed to accept papers confirming conventional views, researchers themselves my decline to submit contradictory result for publication, and statistically less significant results are less likely to be published (Card & Krueger, 1995). Funnel plots often reveal publication bias (Stanley & Doucouliagos, 2012). Since the individual effect sizes plotted in Figure 1 distribute symmetrically around the mean at all precision levels publication bias favoring a conventional belief seems unlikely. Since the distribution of points in the dimension of standard error is reasonably uniform, we also do not suspect bias driven by



so-called ‘p-hacking’ (Nuzzo, 2014). However, because of the way we collected our sample, we checked for an additional potential source of publication bias.

Impact Factor. Our sample construction biased it toward articles published in top journals. This provided needed quality assurance on the meta-data used for our analysis, but these journals reputedly put a special premium on novel, and even counter-intuitive, findings. While such a selection criterion pushes the frontiers of knowledge quickly it might exaggerate heterogeneity in our sample of studies. Although our sample of studies does not include articles published in the long tail of journals with impact factors (IF) below 0.57, by examining the pattern as IF increases above that level, we can infer whether IF is likely to influence either effect size or heterogeneity. Figure 2 graphically shows this is unlikely to be the case—there is no significant relationship between either the effect size or variance of it and the IF of the journal publishing the results. Although not depicted here we also found no relationship between publication date and effect size or heterogeneity.

Figure 2: Scatterplot of effect-size and 5-year impact factor, ending in 2019, of the journals in which each of the primary studies in our meta-sample were published.

File Drawer Problem. A general worry in meta-analyses and potential source of heterogeneity when true effect sizes are small, is that only primary studies with findings statistically significant enough to merit publication are included, while those yielding less significant results are simply relegated to the "file-drawer." We tested for potential publication bias by calculating Rosenthal's (1979) Fail-safe N. Using his notation, the combined critical value Z_c for the $k = 89$ studies in our sample is given by, $Z_c = \Phi^{-1}\left[\frac{0.0014}{2}\right] = -3.19$, where Φ^{-1} denotes the inverse cumulative distribution function of a standard normal random variable and 0.0014 is the *p-value* of the combined effect size (Table 2, model 1). Rosenthal's original formulation can be transformed into the following:

$$N_{fs} = k\left\{\left(\frac{Z_c}{\hat{Z}}\right)^2 - 1\right\} = 147$$

where $\hat{Z} = \Phi^{-1}\left[\frac{0.05}{2}\right] = -1.96$ is the critical value for a significance level of 5 per cent. So, it would take 147 addition primary studies with null results to reduce the significance of the overall positive association of experience to performance so much that it would no longer be significant at a 5 per cent level. This calculation implies that it would take a "file-drawer" problem one and half times the size of our collected sample to push our effect estimate out of traditional significance levels. Furthermore, *Figure 1* clearly shows many published effect sizes near zero, which casts doubt on whether the "file drawer" has been a disproportionate resting place for studies yielding small effects in this literature.

Selection on Match

However, a more fundamental selection problem does potentially plague studies of M&A performance—self-selection bias over experience. It stems from the fact that acquirer and target are not randomly paired in an M&A transaction. On the contrary, a deal arises precisely because the parties anticipate positive synergies. That selection can bias the estimated effect of any attribute on final performance is widely known (and mostly ignored). What is less understood is that in the examination of partnership performance, the direction of the bias is very likely to be toward zero and proportional to the strength of the performance driver's true effect. To see why, suppose that two potential acquirers stand to enjoy the same long-term synergies Y from successfully integrating any subsidiary, but they differ in the costs of doing so. In particular, if the total acquisition cost decreases in some attribute X (like experience), then the acquirer with more X can profitably afford to buy firms with lower Y, which biases the positive estimates effect of X on Y, when measured on a per acquisition basis (as M&A performance always is), downward. Hence, this unaccounted selection effect attenuates the uncorrected estimate of performance and could explain why it has been so difficult to identify a consistently positive effect of experience on acquisition performance—indeed it may explain why the discord in the literatures studying drivers of partnership performance of all types is generally so high.

OVB

M&A experience correlates to many other acquirer attributes, like firm age, revenues, industry, human capital, and labor flexibility, to name just a few, that likely also directly drive the performance of any acquisition. Our final potential source of discord in the literature that meta-analysis cannot resolve—OVB—can occur whenever an unobserved (to the empiricist) factor drives the dependent variable and correlates to the variable of interest. When this happens, the effect of the unobserved (literally: 'omitted' from the regression) variable is attributed to the correlated variable of interest, potentially biasing its estimated effect. Suppose, for example, that acquirer CEO intelligence improves M&A performance, but CEO intelligence also positively correlates to the acquirer's M&A experience. Typically, the empiricist cannot observe executive IQs. If the empiricist does not control for it (say, with the league table ranking of the CEO's alma mater), then estimates of experience's effect on performance will be upwardly biased. Now suppose, instead, that the omitted variable is the current number of targets that the acquirer is currently trying to integrate. *Ceteris paribus* it is plausible that attempting to integrate more firms simultaneously adversely affects the performance of any single deal. Yet, the number of simultaneous integrations positively correlates to experience. Omitting the number of simultaneous integrations will downwardly bias the estimated effect of experience, perhaps even making it negative. These are just two of many potential examples. Given the plethora of potential correlates of experience that could drive performance, OVB could be significant.

Control vs. Variable of Interest. Neither meta-analysis nor large samples can resolve OVB. Primary studies typically add controls for the omitted variables to ensure that *conditional mean independence* is satisfied for the variable of interest. To the extent that included primary studies employ these, our computed partial correlations account for them too, but where OVB affects the primary estimates, these biases will aggregate in meta-analyses. Limiting our sample to studies published in top-100 outlets should ameliorate the problem, because the standards for empirical rigor should be higher. However, within this set, one might anticipate that OVB affects a particular sub-group of our studies more, namely studies where experience is a control rather than variable of interest. This is because it is not necessary that *conditional mean independence* hold for control variables—these variables are included so that estimated coefficients for the variables of interest are unbiased, not so that the estimated coefficients on the controls can be interpreted themselves, as these may be biased.

Figure 1 scatterplots the effect- vs. sample-size of the 89 primary studies in our analyses. The red diamond points denote the 48 primary studies treating experience as a variable of interest, while the black square points denote the 41 studies treating it merely as a control. Although a difference between the two groups is not visually obvious, the subsample analysis presented in Table 4, shows there is one. The positive effect size in the 'Explanatory' subsample ($r_{xy,z} = 0.04, p = 0.02$) is twice as large as in the 'Control' one ($r_{xy,z} = 0.02, p = 0.03$) and they differ significantly from one another ($Q_{Bet} = 3.14, p = 0.00$). The estimate for the 'Explanatory' subsample is heterogeneous ($Q = 244.78, I^2 = 80.80$). Furthermore, a manual review of the papers in the 'Explanatory' subsample reveals that the controls that are added in these studies are not well-argued and demonstrating the robustness of the estimates of the variable(s) of interest to variation in controls is uncommon. Overall, these three factors suggest that OVB may be problematic in the study of experience's effect on M&A performance.

Table 4: Experience as Explanatory vs. Control Variable

Moderator	Sample	k	N	Point Est.	CI	Cochran's Q	I^2	Q_{Bet}
Purpose of 'Experience' in Primary Study	(13.a) Explanatory	48	34,690	**0.04 (0.02)**	(0.01, 0.06)	244.78 (0.00)	80.80	3.14 (0.00)
	(13.b) Control	41	48,442	**0.02 (0.03)**	(0.00, 0.04)	140.22 (0.00)	**71.47**	

Notes: k = number of included studies; N =total number of deals across included studies; point estimate= weighted mean effect size (p-values in parentheses; bold typeface indicates significance at less than 10%); CI= 95% Confidence Interval of estimates, Q = value of chi-square distributed homogeneity statistics (p-values in parentheses); I^2 *= proportion of the observed variance reflecting differences in true effect sizes rather than sampling error.*

Recommendations

Meta-analyses summarize literatures, and in so doing, may expose systematic problems or uncover subtle patterns in them that may have hitherto been unidentified. In the end, though, because the unit of analysis is a primary research study rather than an economic source of variation, they are ill-suited to test novel theories operating at the level of economic agents. Hence, we conclude with emergent directions that future primary studies should probe to resolve the issues of Selection-over-Match and OVB that likely drive discord in the literature about the role of acquirer experience on M&A performance, and to expose the exact mechanisms behind our novel observation that the positive association between experience and performance strengthens in complex environments.

The standard way to address selection issues is to apply (Heckman, 1979) selection estimation methods, which correct for biases introduced because the selected sample (here, of observed acquisitions) is non-random with respect to the variable of interest (here, experience). In neighboring literatures, this approach has been used to isolate the performance effects of agent attributes on match performance, when the match itself is driven by the attributes of interest (Hegde & Tumlinson, 2014). However, the technique has not apparently been applied in the

M&A literature.[13] This represents an opportunity to better answer to the focal question of this paper. Nevertheless, we would caution that Heckman's approach depends on distributional assumptions, which may not hold in the M&A performance setting and requires an exclusion restriction—inclusion of some variable in the first stage that drives match but has no other impact on performance in the second stage. These instrument-like variables are non-trivial to identify, and the drivers of M&A matches are virtually unstudied.

Hence, we recommend a more tractable starting point: focus on the drivers of matches—the underlying selection, which has received scant attention. Under the assumption that firms are rational—they choose matches optimally—then measuring the effect of a posited driver, like experience, on match probability, will provide a more reliable, proportional indicator of the driver's role in performance than trying to measure performance directly (see e.g., (Langosch & Tumlinson, 2020). Furthermore, a more thorough understanding of the selection process is the first step in correcting for it.

As with selection on match, the state-of-the art solutions to other forms of OVB, like instrumental variables and (quasi-)experimental methods, tend to be difficult to implement in the M&A setting. Nevertheless, a straightforward, if mundane, solution to most kinds of OVB, which will plague almost any study of experience's effect, is almost completely overlooked in our sample: just two included acquirer fixed effects—(Puranam et al., 2006) and (Puranam et al., 2009).[14] Including acquirer fixed effects controls for all unobserved, time-invariant factors about an acquiring firm and its management (see e.g., the IQ example above), which both correlate to the level of experience it has and drives its corporate performance. Identification

[13] Several studies in our sample employed Heckman (1979) selection estimation methods to account for various sources of potential selection. However, none considered the selection that arises in matching due anticipated effects of experience.
[14] Twenty-one of 89 studies mention fixed effects, but these were generally over time or industry, rather than firm, and thus are not particularly helpful in controlling for unobserved firm-level heterogeneity.

will then be over the *change in a given firm's increase in experience*, which, is in fact, the level at which all theories of experience operate, anyway.

In closing, although meta-analysis led to the initial discovery that the Complexity of the study context positively moderates the relationship of acquisition experience to performance, it can shed little light on the mechanism. To do so requires quantifying Complexity at the deal rather than study level. Furthermore, because the included primary studies measure only how much experience acquirers have but do not consider how varied the experiences are, we cannot connect the concept of Experiential Learning to Complexity. Future primary studies are needed. These require the panel regression of serial acquirers—with fixed firm-level effects—and time-varying measures of experience breadth as well as depth. The payoff to such future research is to offer managers a quantifiable way to amortize the experience gained on the present acquisition over the future stream of anticipated M&As.

APPENDIX
(for online publication)

Table A.1: Main & Complexity Results Omitting Outliers (According to 3σ & 1.5 IQR Rules)[15]

Panel 1

Moderator	Sample	*k*	*N*	Point Est.	CI	Cochran's Q	I²	Q_{Bet}
None	(1) Full Sample	87	83,000	**0.03 (0.00)**	(0.01,0.04)	335.66 (0.00)	**74.38**	
Performance Measurement	(2.a) Market-based	48	48,193	**0.02 (0.09)**	(0.00, 0.03)	137.31 (0.00)	**65.77**	3.07
	(2.b) Accounting-based	39	34,807	**0.05 (0.00)**	(0.02, 0.08)	193.15 (0.00)	80.33	(0.08)
Experience Measurement	(3.a) Deal Count	67	70,868	**0.02 (0.02)**	(0.00, 0.04)	240.19 (0.00)	**72.52**	
	(3.b) Dummy	8	5,827	0.02 (0.55)	(-0.04. 0.07)	12.98 (0.07)	**46.10**	2.52
	(3.c) ln(Deal Count)	8	5,568	0.05 (0.29)	(-0.04, 0.13)	63.01 (0.00)	88.89	(0.47)
	(3.d) Self-Assessment	4	737	**0.13 (0.08)**	(-0.01, 0.27)	10.32 (0.02)	**70.94**	
Perf. Measurement (Deal Count Only)	(4.a) Market-based	38	40,211	0.02 (0.15)	(0.00, 0.04)	123.66 (0.00)	**70.08**	0.81
	(4.b) Accounting-based	29	30,657	**0.03 (0.03)**	(0.00, 0.06)	112.42 (0.00)	75.10	(0.34)
International	(5.a) Domestic	55	38,839	**0.02 (0.03)**	(0.00, 0.05)	193.65 (0.00)	**72.11**	0.14
	(5.b) Cross-border	32	44,161	**0.03 (0.01)**	(0.00, 0.05)	141.69 (0.00)	78.11	(0.70)
Industries	(6.a) Single	23	24,292	0.01 (0.95)	(-0.03, 0.03)	73.51 (0.00)	**70.07**	
	(6.b) Related	6	3,355	-0.01 (0.96)	(-0.03, 0.03)	2.53. (0.77)	**0.00**	6.04 (0.05)
	(6.c) Multiple	58	55,353	**0.04 (0.00)**	(0.02, 0.06)	251.87 (0.00)	77.37	
Recency of Experience	(7.a) < 5 years	21	12,250	**0.05 (0.00)**	(0.02, 0.08)	44.27 (0.00)	**54.83**	
	(7.b) < 10 years	52	49,839	**0.02 (0.09)**	(0.00, 0.04)	172.83 (0.00)	**70.49**	5.69 (0.13)
	(7.c) < 18 years	12	20,911	0.03 (0.18)	(-0.01, 0.08)	93.41 (0.00)	88.23	
Market Reaction (Market Based): Event Windows	(8.1.a) ≤ 90 days	27	30,428	**0.02 (0.03)**	(0.00, 0.05)	67.27 (0.00)	**61.35**	0.92
	(8.1.b) > 90 days	21	17,765	0.01 (0.72)	(-0.03, 0.04)	64.05 (0.00)	**68.77**	(0.34)
Market Reaction (Market Based): International	(8.2.a) Domestic	34	32,631	0.01 (0.26)	(-0.01, 0.03)	112.46 (0.00)	**70.66**	0.05
	(8.2.b) Cross-border	14	15,562	0.02 (0.18)	(0.00, 0.04)	24.50 (0.03)	**46.94**	(0.83)
Market Reaction (Market Based): Industries	(8.3.a) Single	14	20,454	0.00 (0.96)	(-0.02, 0.03)	31.52 (0.00)	**58.75**	
	(8.3.b) Related	3	475	0.00 (0.96)	(-0.08, 0.09)	0.31 (0.85)	**0.00**	1.26 (0.53)
	(8.3.c) Multiple	31	27,264	**0.02 (0.07)**	(0.00, 0.05)	104.61 (0.00)	**71.32**	
Market Reaction (Market Based): Recency of Experience	(8.4.a) < 5 years	8	8,942	**0.04 (0.09)**	(0.00, 0.09)	25.31 (0.00)	**72.34**	
	(8.4.b) < 10 years	32	31,234	0.02 (0.15)	(0.00, 0.04)	92.52 (0.00)	**66.50**	3.74 (0.29)
	(8.4.c) < 18 years	7	7,913	0.00 (0.69)	(-0.5, 0.03)	15.37 (0.02)	**60.95**	
Synergy Realization (Accounting Based): Performance Measurement	(9.1.a) Ratios	21	31,732	**0.03 (0.04)**	(0.00, 0.06)	145.09 (0.00)	86.22	
	(9.1.b) Self-Assessed	13	1,800	**0.10 (0.02)**	(0.02, 0.16)	26.11 (0.01)	**54.05**	2.57 (0.27)
	(9.1.c) Others	5	1,275	**0.07 (0.10)**	(-0.03, 0.18)	12.93 (0.01)	**69.07**	
Synergy Realization (Accounting Based): International	(9.2.a) Domestic	21	6,208	**0.05 (0.06)**	(0.00, 0.11)	79.60 (0.00)	**74.87**	0.02
	(9.2.b) Cross-border	18	28,599	**0.04 (0.03)**	(0.00, 0.08)	113.47 (0.00)	85.02	(0.79)
Synergy Realization (Accounting Based): Industries	(9.3.a) Single	9	3,838	0.01 (0.87)	(-0.07, 0.08)	41.81 (0.00)	80.60	
	(9.3.b) Related	3	2,880	0.00 (0.89)	(-0.04, 0.05)	2.22 (0.33)	**10.00**	5.11 (0.07)
	(9.3.c) Multiple	27	28,089	**0.07 (0.00)**	(0.03, 0.10)	141.28 (0.00)	81.60	
Synergy Realization (Accounting Based): Recency of Experience	(9.4.a) < 5 years	13	3,308	**0.07 (0.00)**	(0.03, 0.11)	14.79 (0.25)	**18.80**	
	(9.4.b) < 10 years	20	18,605	0.02 (0.27)	(-0.02, 0.06)	74.96 (0.00)	**74.65**	4.74 (0.19)
	(9.4.c) < 18 years	5	12,864	**0.01 (0.04)**	(0.00, 0.19)	50.82 (0.00)	92.13	

Panel 2

Moderator	Sample	*k*	*N*	Point Est.	CI	Cochran's Q	I²	Q_{Bet}
$Complexity_1$ (Cross-border | Multiple)	(10.a) Complex	66	56,584	**0.04 (0.00)**	(0.02, 0.05)	252.04 (0.00)	**74.21**	5.28
	(10.b) Non-complex	21	26,416	-0.01 (0.54)	(-0.05, 0.03)	73.42 (0.00)	**72.76**	(0.02)
$Complexity_2$ (Cross-border & Multiple)	(11.a) Complex	23	18,312	**0.05 (0.01)**	(0.02, 0.08)	107.37 (0.00)	79.51	2.45
	(11.b) Non-complex	64	64,688	**0.02 (0.04)**	(0.01, 0.04)	224.02 (0.00)	**71.88**	(0.12)
$Complexity_3$ (X-border & Mult. & Acc.)	(12.a) Complex	14	7,628	**0.11 (0.00)**	(0.04, 0.18)	88.90 (0.00)	85.34	6.68
	(12.b) Non-complex	73	75,372	**0.02 (0.07)**	(0.00, 0.03)	246.17 (0.00)	**70.75**	(0.01)

Notes: k = number of included studies; N =total number of deals across included studies; point estimate= weighted mean effect size (p-values in parentheses; bold typeface indicates significance at less than 10%); CI= 95% Confidence Interval of estimates, Q = value of chi-square distributed homogeneity statistics (p-values in parentheses); I^2 *= proportion of the observed variance reflecting differences in true effect sizes rather than sampling error (*$I^2 < 75$ *are bolded)*

[15] Omitting Hutzschenreuter et al., 2014; and Colombo, Gonca & Gnan, 2007

Table A.2: Main & Complexity Results using Pearson Correlations

Panel 1

Moderator		Sample	k	N	Point Est.	CI	Cochran's Q	I²	Q_{Bet}
None		(1) Full Sample	78	57,323	**0.06 (0.00)**	(0.04, 0.06)	**650.89 (0.00)**	88.17	
Performance Measurement		(2.a) Market-based	40	32,270	0.01 (0.37)	(-0.02, 0.04)	274.13 (0.00)	85.77	12.25
		(2.b) Accounting-based	38	25,053	**0.11 (0.00)**	(0.06, 0.14)	274.13 (0.00)	85.16	(0.00)
Experience Measurement		(3.a) Deal Count	61	49,612	**0.07 (0.00)**	(0.03, 0.09)	546.29 (0.00)	89.01	
		(3.b) Dummy	5	1,406	0.03 (0.52)	(-0.06, -0.1)	9.34 (0.05)	**57.18**	7.54
		(3.c) ln(Deal Count)	8	5,568	-0.03 (0.55)	(-0.12, 0.06)	68.19 (0.00)	89.73	(0.05)
		(3.d) Self-Assessment	4	737	**0.17 (0.00)**	(0.05, 0.27)	6.55 (0.08)	**54.21**	
Perf. Measurement (Deal Count Only)		(4.a) Market-based	33	28,709	0.02 (0.29)	(0.02, 0.05)	252.50 (0.00)	87.32	12.29
		(4.b) Accounting-based	28	20,903	**0.12 (0.00)**	(0.08, 0.16)	180.08 (0.00)	85.00	(0.00)
International		(5.a) Domestic	52	31,011	**0.06 (0.00)**	(0.02, 0.09)	468.01 (0.00)	89.10	0.14
		(5.b) Cross-border	26	26,312	**0.05 (0.01)**	(0.01, 0.08)	165.66 (0.00)	84.81	(0.71)
Industries		(6.a) Single	23	22,999	**0.08 (0.01)**	(0.02, 0.13)	340.19 (0.00)	93.53	
		(6.b) Related	6	3,396	0.06 (0.19)	(-0.03, 0.15)	20.43 (0.00)	75.52	0.86 (0.65)
		(6.c) Multiple	49	30,928	**0.05 (0.00)**	(0.02, 0.08)	283.39 (0.00)	83.06	
Recency of Experience		(7.a) < 5 years	19	7,910	**0.05 (0.06)**	(0.00, 0.11)	89.87 (0.00)	79.97	
		(7.b) < 10 years	45	28,541	**0.08 (0.00)**	(0.04, 0.12)	417.13 (0.00)	89.45	4.31 (0.29)
		(7.c) < 18 years	13	20,872	0.01 (0.77)	(-0.05, 0.06)	130.19 (0.00)	90.78	
Market Reaction (Market Based)	Event Windows	(8.1.a) ≤ 90 days	22	19,656	0.03 (0.19)	(-0.02, 0.08)	175.62 (0.00)	88.04	1.49
		(8.1.b) > 90 days	18	12,614	-0.01 (0.71)	(-0.05, 0.03)	72.84 (0.00)	76.66	(0.22)
	International	(8.2.a) Domestic	31	24,803	0.02 (0.40)	(-0.02, 0.06)	238.69 (0.00)	87.43	0.07
		(8.2.b) Cross-border	9	7,467	0.01 (0.76)	(-0.04, 0.06)	34.02 (0.00)	76.48	(0.78)
	Industries	(8.3.a) Single	14	19,161	0.03 (0.36)	(-0.03, 0.09)	178.71 (0.00)	92.73	
		(8.3.b) Related	3	516	0.01 (0.86)	(-0.11, 0.13)	3.33 (0.19)	**39.97**	0.27 (0.87)
		(8.3.c) Multiple	23	12,593	0.01 (0.56)	(-0.03, 0.05)	90.91 (0.00)	75.80	
	Recency of Experience	(8.4.a) < 5 years	6	4,602	0.02 (0.67)	(-0.07; 0.10)	29.78 (0.00)	83.21	
		(8.4.b) < 10 years	26	19,690	0.03 (0.18)	(-0.01, 0.07)	195.18 (0.00)	87.19	2.48 (0.28)
		(8.4.c) < 18 years	8	7,978	-0.03 (0.33)	(-0.07, 0.11)	38.42 (0.00)	81.78	
Synergy Realization (Accounting Based)	Performance Measurement	(9.1.a) Ratios	19	21,911	**0.05 (0.02)**	(0.01, 0.09)	111.30 (0.00)	83.82	
		(9.1.b) Self-Assessed	14	1,867	**0.17 (0.00)**	(0.12, 0.22)	16.34 (0.23)	**20.47**	17.25 (0.00)
		(9.1.c) Others	5	1,275	**0.26 (0.00)**	(0.10, 0.39)	28.73 (0.00)	86.08	
	International	(9.2.a) Domestic	21	6,208	**0.13 (0.00)**	(0.05, 0.20)	171.92 (0.00)	88.36	1.17
		(9.2.b) Cross-border	17	18,845	**0.08 (0.00)**	(0.04, 0.12)	75.17 (0.00)	78.71	(0.27)
	Industries	(9.3.a) Single	9	3,838	**0.14 (0.02)**	(0.02, 0.26)	107.37 (0.00)	92.54	
		(9.3.b) Related	3	2,880	0.12 (0.16)	(-0.05, 0.27)	17.03 (0.00)	88.26	0.80 (0.67)
		(9.3.c) Multiple	26	18,335	**0.08 (0.00)**	(0.06, 0.13)	105.50 (0.00)	76.30	
	Recency of Experience	(9.4.a) < 5 years	13	3,308	**0.07 (0.07)**	(-0.01, 0.14)	44.86 (0.00)	**73.25**	
		(9.4.b) < 10 years	19	8,851	**0.14 (0.00)**	(0.07, 0.21)	163.88 (0.00)	89.01	3.24 (0.35)
		(9.4.c) < 18 years	5	12,864	**0.07 (0.06)**	(0.00, 0.14)	32.57 (0.00)	87.72	

Panel 2

Moderator	Sample	k	N	Point Est.	CI	Cochran's Q	I²	Q_{Bet}
$Complexity_1$ (Cross-border | Multiple)	(10.a) Complex	57	46,510	**0.05 (0.00)**	(0.02, 0.07)	344.52 (0.00)	83.74	0.78
	(10.b) Non-complex	21	10,813	**0.09 (0.03)**	(0.01, 0.17)	304.93 (0.00)	**93.44**	(0.38)
$Complexity_2$ (Cross-border & Multiple)	(11.a) Complex	19	7,231	**0.08 (0.01)**	(0.02, 0.13)	87.29 (0.00)	79.37	0.73
	(11.b) Non-complex	59	50,092	**0.05 (0.00)**	(0.02, 0.08)	559.12 (0.00)	89.62	(0.39)
$Complexity_3$ (X-border & Mult. & Acc.)	(12.a) Complex	14	3,325	**0.13 (0.00)**	(0.08, 0.17)	17.88 (0.16)	**27.29**	9.46
	(12.b) Non-complex	64	53,998	**0.04 (0.00)**	(0.01, 0.08)	613.03 (0.00)	89.72	(0.00)

Notes: k = number of included studies; N =total number of deals across included studies; point estimate= weighted mean effect size (p-values in parentheses; bold typeface indicates significance at less than 10%); CI= 95% Confidence Interval of estimates, Q = value of chi-square distributed homogeneity statistics (p-values in parentheses); I^2 *= proportion of the observed variance reflecting differences in true effect sizes rather than sampling error (*$I^2 < 75$ *are bolded).*

REFERENCES

Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology*, *77*(5), 598–614.

Aguinis, H., Pierce, C. A., Bosco, F. A., Dalton, D. R., & Dalton, C. M. (2011a). Debunking myths and urban legends about meta-analysis. *Organizational Research Methods*, *14*(2), 306–331.

Aguinis, H., Pierce, C. A., Bosco, F. A., Dalton, D. R., & Dalton, C. M. (2011b). Debunking myths and urban legends about meta-analysis. *Organizational Research Methods*, *14*(2), 306–331.

Agyei-Boapeah, H. (2019). Foreign acquisitions and firm performance: The moderating role of prior foreign experience. *Global Finance Journal*, *42*.

Ahammad, M. F., Tarba, S. Y., Liu, Y., & Glaister, K. W. (2016). Knowledge transfer and cross-border acquisition performance: The impact of cultural distance and employee retention. *International Business Review*, *25*(1), 66–75.

Ahammad, M., & Glaister, G. (2011). Postacquisition Management and Performance of Cross-Border Acquisitions. *International Studies of Management and Organization*, *41*(3), 59–75.

Argote, L., & Epple, D. (1990). Learning Curves in Manufacturing. *Science*, *247*, 920–924.

Argote, L., & Miron-Spektor, E. (2011). Organizational Learning: From Experience to Knowledge. *Organization Science*, *22*(5), 1123–1137.

Barkema, H. G., Bell, J. H. J., & Pennings, J. M. (1996). Foreign Entry, Cultural Barriers, and Learning. *Strategic Management Journal*, *17*(2), 151–166.

Barkema, H. G., & Schijven, M. (2008). Toward Unlocking The Full Potential Of Acquisitions: The Role Of Organizational Restructuring. *Academy of Management Journal*, *51*(4), 696–722.

Barney, J. (1991). Firm Resources and Sustained Competetive Advantage. *Journal of Management*, *17*(1), 99–120.

Basuil, D. A., & Datta, D. K. (2015). Effects of Industry- and Region-Specific Acquisition Experience on Value Creation in Cross-Border Acquisitions: The Moderating Role of Cultural Similarity. *Journal of Management Studies*, *52*(6), 766–795.

Bednarczyk, T. P., Schiereck, D., & Walter, H. N. (2010). Cross-border acquisitions and shareholder wealth: Evidence from the energy and industry in Central and Eastern Europe. *Journal of East European Management Studies*, *15*(2), 106–127.

Birkinshaw, J., Bresman, H., & Hakanson, L. (2000). Managing the post-acquisition integration process: How the human integration and task integration processes interact to foster value creation. *Journal of Management Studies*, *37*(3), 395–425.

Bohlen, G. A., & Barany, J. W. (1976). A learning curve prediction model for operators performing industrial bench assembly operations. *International Journal of Production Research*, *14*(2), 295–303.

Borenstein, M., Hedges, L. v., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to Meta-Analysis*. John Wiley & Sons.

Brown, S. J., & Warner, J. B. (1980). Measuring Security Price Performance. *Journal of Financial Economics*, *8*, 205–258.

Bruton, G. D., Oviatt, B. M., & White, M. A. (1994). Performance of Acquisitions of Distressed Firms. *Academy of Management Journal*, *37*(4), 972–989.

Card, D., & Krueger, A. B. (1995). Time-Series Minimum-Wage Studies : A Meta-analysis. *The American Economic Review*, *85*(2), 238–243.

Carney, M., Gedajlovic, E. R., Heugens, P. P. M. A. R., The, S., Journal, M., & June, N. (2011). Business Group Affiliation, Performance, Context, and Strategy: a Meta-Analysis. *Academy of Management Journal*, *54*(3).

Cartwright, S., & Schoenberg, R. (2006). 30 Years of Mergers and Acquisitions Research : Recent Advances and Future Opportunities. *British Journal of Management*, *17*(1), 1–5.

Chatterjee, S. (1986). Types of synergy and economic value: The impact of acquisitions on merging and rival firms. *Strategic Management Journal*, *7*(2), 119–139.

Colombo, G., Conca, V., Buongiorno, M., & Gnan, L. (2007). Integrating Cross-Border Acquisitions: A Process-oriented Approach. *Long Range Planning*, *40*(2), 202–222.

Comier, S. M., & Hagman, J. D. (1987). *Transfer of learning: Contemporary research and application.* Academic Press.

Cording, M., Christman, P., & King, D. (2008). Reducing Causal Ambiguity in Acquisition Integration: Intermediate Goals as Mediators of Integration Decisions and Acquisition Performance. *Academy of Management Journal*, *51*(4), 744–767.

Cording, M., Harrison, J. S., Hoskisson, R., & Jonsen, K. (2014). Walking the Talk : a Multistakeholder. *Academy of Management Perspectives*, *28*(1), 38–56.

Cuypers, I. R. P., Cuypers, Y., & Martin, X. (2017). When the target may know better: Effects of experience and information asymmetries on value from mergers and acquisitions. *Strategic Management Journal*, *38*(3), 609–625.

DeCoster, J. (2004). *Meta-Analysis Notes*.

Devos, E., Kadapakkam, P. R., & Krishnamurthy, S. (2009). How do mergers create value? A comparison of taxes, market power, and efficiency improvements as explanations for synergies. *Review of Financial Studies*, *22*(3), 1179–1211.

Doucouliagos, C., & Ulubasoglu, M. A. (2006). Economic freedom and economic growth: Does specification make a difference? *European Journal of Political Economy*, *22*(1), 60–81.

Doucouliagos, H., & Ulubasoglu, M. A. (2008). Democracy and Economic Growth: A Meta-Analysis. *American Journal of Political Science*, *52*(1), 61–83.

Duran, P., van Essen, M., Heugens, P. P. M. A. R., Kostova, T., & Peng, M. W. (2019). The impact of institutions on the competitive advantage of publicly listed family firms in emerging markets. *Global Strategy Journal*, *9*(2), 243–274.

Efendic, A., Pugh, G., & Adnett, N. (2011). Institutions and economic performance: A meta-regression analysis. *European Journal of Political Economy*, *27*(3), 586–599.

Ellis, K., Reus, T., Lamont, B., & Ranft, A. (2011). Transfer Effects in Large Acquisitions: How Size-Specific Experience Matters. In *Academy of Management Journal* (Vol. 54, Issue 6, pp. 1261–1276).

Finkelstein, S., & Haleblian, J. (2002). Understanding acquisition performance: The role of transfer effects. *Organization Science*, *13*(1), 36–47.

Galavotti, I. (2019). Firm-level recent profitability and acquisition performance: exploring competing theoretical perspectives. *Eurasian Business Review*.

Geyskens, I., Krishnan, R., Steenkamp, J. B. E. M., & Cunha, P. v. (2009). A review and evaluation of meta-analysis practices in management research. *Journal of Management*, *35*(2), 393–419.

Gick, M. L., & Holyoak, K. J. (1987). The cognitive basis of knowledge transfer. In *Transfer of learning: Contemporary research and applications* (pp. 9–46).

Haleblian, J., & Finkelstein, S. (1999). The influence of organizational acquisition experience on acquisition performance: A behavioral learning perspective. *Administrative Science Quarterly*, *44*(1), 29–56.

Hayward, M. L. A. (2002). When Do Firms Learn from Their Acquisition Experience ? Evidence from 1990-1995. *Strategic Management Journal*, *23*(1), 21–39.

Hebert, L., Very, P., & Beamish, P. W. (2005). Expatriation as a Bridge Over Troubled Water: A Knowledge-Based Perspective Applied to Cross-Border Acquisitions. *Organization Studies*, *26*(10), 1455–1476.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *The Econometrica*, *47*(1), 153–161.

Hegde, D., & Tumlinson, J. (2014). Does Social Proximity Enhance Business Partnerships? Theory and Evidence from Ethnicity's Role in U.S. Venture Capital. *Management Science*, *60*(9), 2355–2380.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*(7414), 557–560.

Houston, J. F., James, C. M., & Ryngaert, M. D. (2001). Where do merger games come from? Bank mergers from the perspective of insiders and outsiders. *Journal of Financial Economics*, *60*(2), 285–331.
Huckman, R. S., Staats, B. R., & Upton, D. M. (2009). Team familiarity, role experience, and performance: Evidence from Indian software services. *Management Science*, *55*(1), 85–100.
Hunter, J. E., & Schmidt, F. L. (2004). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings* (second). Sage Publication.
Hutzschenreuter, T., Kleindienst, I., & Schmitt, M. (2014). How mindfulness and acquisition experience affect acquisition performance. *Management Decision*, *52*(6), 1116–1147.
Ingram, P. (2002). Interorganizational learning. In J. A. Baum (Ed.), *The Blackwell companion to organizations*. Blackwell.
Ingram, P., & Baum, J. A. C. (1997). Opportunity and Constraint: Organizations' Learning From the Operating and Competitive Experience of Industries. *Strategic Management Journal*, *18*(S1), 75–98.
Kim, J. Y., & Finkelstein, S. (2009). The effects of strategic and market complementarity on acquisition performance: Evidence from the U.S. commercial banking industry, 1989-2001. *Strategic Management Journal*, *30*(6), 617–646.
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*(2), 187–200.
Krishnan, H. A., Hitt, M. A., & Park, D. (2007). Acquisition premiums, subsequent workforce reductions and post-acquisition performance. *Journal of Management Studies*, *44*(5), 709–732.
Krishnan, H. A., & Park, D. (2002). The impact of work force reduction on subsequent performance in major mergers and acquisitions: An exploratory study. *Journal of Business Research*, *55*(4), 285–292.
Kroll, M., Toombs, L., & Leavell, H. (1997). Form Of Control: A Critical Determinant Of Acquisition Performance And CEO Rewards. *Strategic Management Journal*, *18*(2), 85–96.
Langosch, M., & Tumlinson, J. (2020). Whom to Acquire – Culture, Country, or Kilometers in Target Selection. *Academy of Management Proceedings*, *2020*(1).
Levitt, B., & March, J. M. (1988). Organizational Learning. *Annual Review of Sociology*, *14*, 319–340.
Lin, L. (2012). Organizational Structure and Acculturation in Acquisitions : Perspectives of Congruence Theory and Task Interdependence. *Journal of Management*, *40*(7), 1831–1856.
Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE Publications, Inc.
Liu, X., & Zou, H. (2008). The impact of greenfield FDI and mergers and acquisitions on innovation in Chinese high-tech industries. *Journal of World Business*, *43*(3), 352–364.
McDaniel, M. A., Schmidt, F. L., & Hunter, J. E. (1988). Job Experience Correlates of Job Performance. *Journal of Applied Psychology*, *73*(2), 327–330.
Meschi, P. X., & Metais, E. (2006). International acquisition performance and experience: A resource-based view. Evidence from French acquisitions in the United States (1988-2004). *Journal of International Management*, *12*(4), 430–448.
Meschi, P. X., & Metais, E. (2013). Do Firms Forget about Their Past Acquisitions? Evidence from French Acquisitions in the United States (1988-2006). *Journal of Management*, *39*(2), 469–495.
Nembhard, D. A. (2000). The effects of task complexity and experience on learning and forgetting: A field study. *Human Factors*, *42*(2), 272–286.
Nuzzo, R. (2014). Statistical Errors : P Values. *Nature*, *506*(13 February), 150–152.
Oler, D. K., Harrison, J. S., & Allen, M. R. (2008). The danger of misinterpreting short-window event study findings in strategic management research: an empirical illustration using horizontal acquisitions. *Strategic Organization*, *6*(2), 151–184.
Parola, H. R., Ellis, K. M., & Golden, P. (2015). Performance effects of top management team gender diversity during the merger and acquisition process. *Management Decision*, *53*(1), 57–74.

Puranam, P., Singh, H., & Chaudhuri, S. (2009). Integrating Acquired Capabilities : When Structural Integration Is (Un)necessary. *Organization Science*, *20*(2), 313–328.
Puranam, P., Singh, H., & Zollo, M. (2006). Organizing for Innovation: Managing the Coordination-Autonomy Dilemma in Technology Acquisitions. *The Academy of Management Journal*, *49*(2), 263–280.
Rabier, M. R. (2017). Acquisition Motives and the Distribution of Acquisition Performance. *Strategic Management Journal*, *38*(13), 2666–2681.
Shi, W., & Prescott, J. E. (2012). Rhythm and entrainment of acquisition and alliance initiatives and firm performance: A temporal perspective. *Organization Studies*, *33*(10), 1281–1310.
Slangen, A. H. L. (2006). National cultural distance and initial foreign acquisition performance: The moderating effect of integration. *Journal of World Business*, *41*(2), 161–170.
Stahl, G., & Voigt, A. (2008). Do Cultural Differences Matter in Mergers and Acquisitions? A Tentative Model and Examination. *Organization Science*, *19*(1), 160–176.
Stanley, T. D., & Doucouliagos, H. (2012). *Meta-regression analysis in Economics and Business*. Routledge.
Uhlenbruck, K., Hitt, M. A., & Semadeni, M. (2006). Market value effects of acquisitions involving internet firms: A resource-based analysis. *Strategic Management Journal*, *27*(10), 899–913.
Zollo, M. (2009). Superstitious Learning with Rare Strategic Decisions : Theory and Evidence from Corporate Acquisitions Superstitious Learning with Rare Strategic Decisions : Theory and Evidence from Corporate Acquisitions. *Organization Science*, *20*(December 2015).
Zollo, M., & Meier, D. (2008). What Is M & A Performance ? *Academy of Management Perspective*, *22*(3), 55–78.
Zollo, M., & Singh, H. (2004). Deliberate learning in corporate acquisitions: Post-acquisition strategies and integration capability in U.S. bank mergers. *Strategic Management Journal*, *25*(13), 1233–1256.

Short Biographies:

Magdalena Langosch

Magdalena Langosch is Assistant Professor in International Business at the University of Greenwich Business School (London, UK). She studies behavior and performance of International Business Partnerships, mainly focusing on M&As and the effect of regional culture and experience.

Justin Tumlinson

Justin Tumlinson is Associate Professor of Business Analytics and Director of the Research Methods Centre at the University of Exeter Business School. He studies cooperative ties between economic actors, from families to firms to nations, when they form, when they do not and how these linkages matter to society.