From Failures to Innovation – How Organizations Learn from Failures to Innovate.

Abstract

Kiran Awate
The Ohio State University
Fisher College of Business, Columbus IH 43210
awate.3@buckeyemail.osu.edy

Key Words: Technology and Innovation Management, Innovation Process, organizational Learning, Adaptive learning from Failures.

Preliminary Draft
January, 2016

INTRODUCTION

LITERATURE REVIEW

Role of failures in Innovation

Experimentation and failures play pivotal role in innovation (Sitkin, 1992; March 1991). Failures not only provide crucial insights about causal relationships when complete understanding of the underlying novel mechanism is unavailable but also create options to choose from for further experimentation. Organizations learn from failures and alter their search based on the feedback (Weick, 1979; Sitkin, 1992; Argyris & Schön; 1978; Austin, Devin & Sullivan, 2012; Khanna Guler & Nerkar, 2015).

Two streams of literature have examined issues related to innovation and learning from failures. First stream of literature has examined whether and how diverse knowledge elements can lead to new discoveries (Utterback, 1994; Galunic & Rodan, 1998; Hargadon and Douglas, 2001; Rivkin, 2000; Fleming 2001; Ahuja & Katila 2004). This literature argues that innovation occurs through combination and recombination of existing elements into novel artifacts. While predictions from this literature provide insights into the factors that may lead to innovation, this research does not provide systematic explanations about processes that precede innovation. Specifically, this research remains silent on how organizations learn from prior recombination attempts that failed.

Second stream of literature – learning from failures – rests its logic on organizational learning to investigate whether and how organizations learn from failures (Baum & Dahlin, 2007; Haunschild & Sullivan, 2002; Ramanujam & Goodman, 2003). This stream draws some of its key insights from experiential learning literature (Argote, Beckman, & Epple, 1990; Lieberman, 1987) to argue that organizations learn from failures to improve operational efficiency. While this literature elaborates key mechanisms of learning from failures, the studies are limited to examining role of learning in improving operational efficiency or productivity. Thus, role of learning from failures in the context of experimentation and innovation remains unexplored.

In aggregation, both sets of literatures fail to explore role of learning from failures in navigating path from experimentation to novel discoveries. In the next section, I review these two literatures in greater depth to elaborate this argument.

Knowledge Recombination

A stream of literature argue that most innovations derive either from combining technologies in a novel manner (Fleming, 2001; Nelson and Winter, 1982; Schumpeter, 1934) or from reconfiguring existing combinations so that they can be put to new uses and applications (Henderson and Clark, 1990; Yayavaram and Ahuja, 2008). Scholars have characterized organizations' ability to recombine its knowledge for innovation as combinative capabilities (Kogut & Zander, 1992), reconfiguring competence (Galunic & Rodan, 1998), or recombinant capabilities (Carnabuci and Operti, 2013). Broadly, this work suggests that either a combination of new knowledge elements or new relationship between previously combined knowledge components can lead to novel discoveries.

In their seminal study of the photolithographic equipment industry, Henderson and Clark (1990) showed that established firms often lose their technological leadership due to their inability to link well-understood technologies in new and more fruitful ways. This was one of the first studies that highlighted how linkages between knowledge components can have significant competitive implications for organizations. Similarly, Yayavaram and Ahuja (2008) showed that semiconductor firms generate more impactful ideas when they recombine technological components from otherwise decoupled knowledge domains. Focusing on distinction between capabilities that create new components and capabilities that reuse components, Carnabuci and Operti (2013) examined factors that favor or hinder each type of capability and innovation. Other empirical studies corroborate finding that innovation comes either from combining technological components in a novel manner, or through reconfiguring existing combinations. (Katila, 2002; Katila and Ahuja, 2002; Karim & Kaul 2015; Ahuja & Lampert; 2001)

Basing on idea of recombination as the source of novelty, a sub-stream within knowledge recombination has focused on understanding sources of novel knowledge components that lead to innovation. Focusing on sources within organizations, Fleming (2001) demonstrated that familiar components are better understood and therefore result in more valuable innovations. Examining age of knowledge, Katila (2002) found that older intra-industry knowledge hurts innovation but older extra-industry knowledge actually helps innovation. Similarly, Galunic and Rodan (1998) proposed that knowledge characteristics such as tacitness, context specificity and knowledge dispersion moderates knowledge recombination. Organization design (Karim & Kaul 2015), role of founders' imprinting (Hsu & Lim 2014), prior experience (Ghosh et al., 2014) are some of the other sources examined by researchers.

Taking locus of attention outside of organization, a set of scholars have examined external sources such as suppliers, alliance partners, and network ties that help organizations tap into novel knowledge elements. For example, Hargadon & Sutton (1997) illustrated how organizations act as technology brokers among their customers to generate new designs and products. Taking similar view, Ahuja (2000) and Carnabuci and Operti (2013) demonstrated how organizations' network positions affect innovation output. Rosenkopf and Nerkar (2001) showed importance of overcoming technological and organizational boundaries for ideation.

While this literature has brought to bear robust prediction about role of knowledge components and linkages in discovering new ideas, this work lack clarity in two ways. First, barring a few exceptions (Fleming 2001; Ahuja and Katila 2001), this work either assumes or ignores role of experimentation in navigating path from knowledge combination to innovation. It almost seems that this work treats experimentation as a black box and only factor that matters in the innovation process is input of knowledge components or linkages. As highlighted in the section above, experimentation plays crucial role in discovering new ideas and this lack of clarity restricts our understanding of the phenomenon.

Second, given that there could be very large number of possible combinations of knowledge elements and organizations have limited resources, it is unclear how organizations prioritize combination process that increases probability of success. Further, this work remains silent on the role of experience from prior recombination that assist organizations in prioritizing search. Specifically, it is unclear whether and how prior failed attempts of recombinations guide future recombination decisions. To conclude, the literature on knowledge recombination offers little insights about what valuable lessons organizations draw from failed experiments and how that helps them in discovering new innovations.

Learning from failure

One of the most influential and robust finding from the organizational learning literature is that organizations improve their performance as they accumulate experience in that field. Specifically, learning-curve literature documents how productivity improves with increasing experience (e.g. Argote 1999; Dutton and Thomas 1984; Yelle 1979). This overarching assertion has been tested in a variety of contexts such as manufacturing plant productivity (Argote, Beckman, and Epple, 1990), medical procedure completion time (Pisano et al. 2001) and continuous-process product quality (Lieberman 1984).

Recently, a stream of literature that branched out of experiential learning work has begun to explore the role of prior success and failures experience. This literature investigates whether and how organizations learn from prior success and failures to improve performance (Baum & Ingram, 1998; Haunschild & Sullivan, 2002; Miner, Kim, Holzinger, & Haunschild, 1999). These events provide important feedback that helps organizations evaluate their course of action and improve their performance (March & Simon 1958; March & Olsen, 1976).

This literature has brought to bear important insights about how organizations learn from their successes and failures to improve outcomes. This literature, however, makes certain assumptions that need to be reexamined in the context of innovation. First, learning required for discovering a novel solution is different from leaning required to improve productivity. Second, organizations' ability to learn from failures in the context of innovation may differ significantly than when organizations engage in reducing errors to improve productivity. In the next section, I examine each of these assumptions in depth to argue why the predictions from the extant literature cannot be extended to the context of innovation. Thus, this void has limited our understanding about how organizations learn from their failures and success to innovate.

Learning from failures to improve operational performance.

The literature on learning from failures builds on the idea that organizations evaluate outcomes of their actions against an expected value – aspiration level - to determine future course of action. Organizations alter their course of action when performance either falls short of or exceed expected value (Cyert & March, 1963; March & Olsen, 1976). Findings from this literature resonate with a broader rationale of experiential learning which asserts that organizational performance improves with experience acquired intentionally or unintentionally (Argote, Beckman & Epple, 1990; Huber 1991). While the studies on learning from failure have offered interesting insights about a peculiar kind of learning, these studies focus mainly examine how leaning helps organizations enhance operational performance.

For example, Baum and Dahlin (2007) studied the effects of prior operating experience and accident costs on future accident costs among large U.S. railroad carriers. They argued that experiential and vicarious learning helped U.S. railroad carriers reduce accident costs thereby improving operational performance. They noted:

> *..[w]e expect a cumulative benefit from a railroad's own and other railroads' accidents. A railroad that has accumulated greater accident experience has had the opportunity to learn how to avoid similar accidents through training or adoption of innovative safety systems, reducing its own current accident-related costs. Since accidents represent operational failures, accumulated experience with successful operations should also contribute to a reduction in accident costs. ……In this study, therefore, we explicitly link organizations' performance feedback with experiential learning……*

Similarly, Haunschild and Rhee (2004) drew on experience curve research in the context of automotive product recalls and found that prior production experience reduces subsequent automotive products recall rates. While they examined role of volition in learning, broadly their results confirmed the important role of experience in improving productivity (Argote 1999). They summarized:

> *[W]e find a basic learning effect in that prior production experience reduces subsequent recall rates, but only when the learning target is reducing voluntary recalls, not involuntary recalls.*

Similar measures have been used by other researchers to understand role of failures in improving operational performance (Hayward, 2002; Haunschild & Rhee, 2004; Henderson & Stern, 2004). In a recent study, Desai (2015) used a panel of hospitals performing CABG surgery within hospital in the state of California to investigate whether and how distribution of failures within an organization affect learning from failures. To measure performance, he captured productivity improvement by measuring average risk-adjusted mortality rate across CABG surgeries. Desai (2015) mentioned:

> *To measure each hospital's failure rate, the study's dependent variable tracks the average risk-adjusted mortality rate across CABG surgeries performed at the hospital during each two-year period, with higher values reflecting a worse patient survival rate. It is important to ensure that this measure reflects hospital quality (i.e., patient deaths due to poor surgeon performance) rather than patient deaths due to preexisting conditions or other patient-specific surgical complications.*

While organizations learn effectively from their own experience, organizational learning has also emphasized role of vicarious learning that occurs through observation of experience of other organizations (Beckman & Haunschild, 2002; Denrell, 2003). Taking this reasoning further researchers have argued that organizations learn from others' failures (Chuang & Baum, 2003; Kim, 2000; Miner, Kim, Holzinger & Haunschild, 1999). For example Kim Miner (2007) examined how vicarious learning helped banks reduce either liquidation and or merger with other bank. Madsen and Desai (2010) tested their hypotheses about vicarious learning in the context to of global orbital launch industry. This sub-stream of literature, similar to learning from owns failures, has largely focused on gains derived in terms of improving operational efficiency.

Review of this literature suggests that failures play important role in improving operational efficiency in three ways. First, failures expose faulty routines that may have led to failures. Second, failures prompt organizations to take corrective actions in terms of either replacing faulty routines or creating additional routines, such as standardization (Levinthal & Rerup 2006), to avoid failures in future. Finally, these studies highlight how organizations learn from failure to bring their performance either back to or exceed industry-mean level (March 1991; McGrath 2001). Taken together, this stream of research informs that organizations are less tolerant towards failures and that sole purpose of learning from failures is to avoid failures in future.

To the contrary, failures play different role in the case of innovation, Failures are deemed valuable as they provide important information about the experiment undertaken. Given the novel nature of experimentation, failures may point to un-explored or less-known causal mechanisms and may prompt organizations to create novel routines. Organizations are most likely to undertake innovation to exceed industry-mean performance or seek competitive advantage and not to seek industry-mean performance. This suggests that organizations have higher tolerance towards errors and failure induced leaning that may lead to new set of experimentation.

Learning from High Magnitude Failures

A sub-stream of research focusses on magnitude of failures in the context of space shuttles, aircraft carriers, nuclear power generation plants, and air traffic control systems (Starbuck and Farjoun 2005, Perrow 1984, Roberts and Rousseau 1989, Rochlin 1993, Starbuck and Milliken 1988, Vaughan 1996). Unlike failures that have minor impact, large magnitude failures are very costly and threaten organizational survival. Failures in the context of innovation share this attribute in the sense that experiments may absorb crucial resources but fail to deliver thereby threatening organizations' survival. Hence it would be worthwhile to examine if this literature offers any insights that may be extended to the context of innovation.

While building on similar principles of learning from failures, these studies examines how catastrophic failures expose flaws in the operational processes and force organizations to make significant operational changes. Researchers argue that high magnitude errors provide important learning opportunity for organizations (Baum & Dahlin, 2007). Therefore these failures force to organizations to draw lessons to make necessary operational changes. Lampel, Shamsie and Shapira (2009) noted:

> *When rare events have a major impact on the organization, there is clear motivation to draw lessons and make the necessary operational and cognitive adjustments, whereas when rare events have a minor impact learning is kept to a minimum, and the experience is seldom transformed into implementable lessons.*

Further, drawing on cognitive heuristics reasoning some researchers argue that mindfulness of organizations is critical within the context of high reliability systems (Weick & Roberts, 1993; Weick & Sutcliffe 2001) [1]. For example, Levinthal and Rerup (2006) argue that mindfulness provides the potential for attending to more subtle cues and feedback that emerge from ongoing operations as a basis for effective adaptation in such circumstances. Broadly, this research assert that organizations develop cognitive templates that facilitate anticipation and

[1] Langer (1989) specifies the concept of mindfulness as a state of active awareness characterized by the continual creation and refinement of categories, an openness to new information, and a willingness to view contexts from multiple perspectives.

sensemaking of unknown that might help organizations avoid catastrophic events (Feldman 2000; Rerup 2009; Christianson et al.2009).

While high magnitude failures help organizations improve performance, organizations typically strive to avoid failures that threaten organizations' survival Organization build early warning systems that give cues about potential high magnitude failures so that such failures can be avoided. Any deviation of a routine activity from planned state prompts course correction. Learning in such settings occurs via local search, refinement and selection and reuse of existing routines.

To the contrary, most often, novel discoveries come about when organizations take deviations from well-established paths. Processes that lead to the most original innovations cannot result from available predictive logic. Not surprisingly, many scholars have argued that innovators benefit from escaping established routines (e.g. Austin, Levin & Sullivan 2012). Learning ensues a process of deliberate variations in experimentations (Baum, Li & Usher, 2000).

Given that high magnitude failures inhibit innovation and triggers different learning mechanisms, insights from this sub-stream of literature provides little guidance on how organizations learn from their failures during the process of innovation.

Learning from failure where causes of failures are clear and interpretable

At the core of the literature on learning from failures lies the idea of reinforcement learning (Levitt & March 1988; Argote 1999; March & Simon 1958; March & Olsen, 1976) in which organizations evaluate outcomes of their actions on the basis of aspiration levels Outcome classified as exceeding the aspiration is evaluated as a success and is reinforced in subsequent periods, while outcome not meeting the aspiration is judged as failure that triggers search for a new way of doing business In other words, these adaptive processes distinguish between learning in response to success and failure.

While this logic has brought to bear robust predictions about how organizations improve their performance, it assumes that causes of successes and failures are unambiguous or can easily be identified and interpreted. For example, Haunschild and Sullivan (2002) investigated variations in learning by examining whether organizations learn more from errors with heterogeneous causes or homogenous causes. They analyzed causes of airline accident and incident rate and classified failures based on whether the cause of an accident could be attributed to multiple and diverse reasons or a single reason. They observed:

> *The results of this study contribute to our understanding of variation in how organizations learn from their prior error experience. It appears that, in general, experiences with heterogeneous causes produces more learning than experience with homogeneous causes.*

Clearly, this attribution of errors to specific causes was possible because the errors resulted not from novel experiments but from routine activities that airlines were doing for years. While Haunschild and Sullivan (2002) argued that heterogeneous causes are beneficial in some context, Yang, Li and Delios (2015) posited a counter argument that heterogeneity in causes of errors can inhibit vicarious learning. They demonstrated how size, location and ownership structure of prior entrants in China moderated relationship between new entrants' learning from prior entrants' failure experience. They noted:

> *Our conceptual arguments and empirical findings demonstrate that learning from the failure experiences of prior entrants increases a new entrant's survival chances when*

entering China. Further, we find that the value of this learning is less effective when there is a greater level of heterogeneity in the causes of these failures.

Similarly, Madsen and Desai (2010) narrate how failures in NASA's *Columbia* and *Atlantis* prompted it to address shuttle safety issues that stemmed from foam losses. Using global orbital launches as a context to test their hypotheses, Madsen and Desai (2010) assert that failure not only indicates gap in organizational knowledge but also, but also provides clear information about which search activities may be most productive.

Failure motivates organization members to correct problems, challenge old assumptions, and innovate (Sitkin, 1992). .....Furthermore, failure indicates not only the existence of a gap in organizational knowledge, but in many cases also provides a clear indication of where that gap may be (Turner, 1978; Wildavsky, 1988). Therefore, failure not only increases organization members' willingness to search for new knowledge, but also provides a roadmap showing where search activities may be most productive (Levinthal & March, 1981).

While these studies explicitly examined the causes of failures, a few others have imputed the causes of failures by observing managers' actions post-failures. For example, Chuang and Baum (2003) examined Ontario nursing home chains' naming strategies. They demonstrated how organizations changes chains' names from commonly-named to locally-named and vice-versa in the event of failures. This implies that organizations were able to identify the cause of the failure viz., chains' naming strategy.

In aggregation, these studies examine learning effects of failures in contexts where it is possible to identify and interpret causes of failures unambiguously. In the case of innovation, however, there are often important ambiguities in the relationship between aspirations, actions, and outcomes (March and Olsen 1976; Van de Ven & Polley, 1992). Ambiguity may simply be the result of stochastic effects that obscure the link between actions and outcomes. Alternatively, ambiguity may stem from the temporal and spatial distance between actions and observed outcomes, in that feedback may be delayed or get confounded with other factors. The full consequence of actions taken today may not be felt until some future time or may have reverberations in other parts of the organization. Ambiguity may also result from outcomes that are neither clear failures (e.g., near failure) nor clear successes (e.g., near successes) (Rerup 2006a). Thus contexts that encounter ambiguous outcomes such as innovation makes feedback from failures difficult to interpret and pose challenges to subsequent learning from failures.

Behavioral effects of failures

Using behavioral perspective to analyses organizations' response to failures, a line of research suggests that organizations learn from their experience and adjust their responses accordingly (Cyert & March 1963, Levitt and March 1988). According to this behavioral theory of organizational learning, poor performance prompts decision makers to search for solutions and take risks to address shortfalls (Greve, 1998). This line of enquiry focusses on managers' behaviors and their effect organizations' decision making in the wake of failures.

For example, some studies building on behavioral theory suggest that managers in organizations that are threatened by failure may focus on a survival level and become more conservative, since risks taken to repair performance shortfalls could instead threaten survival of the organization (March & Shapira 1987, 1992; Miller & Chen 2004). Building on this logic, researchers have turned attention to factors such as organizations experience, legitimacy etc, that moderate managers' decision making (Audia & Greve, 2006; Desai 2008 Kaplan & Eggers 2009).

To the contrary, some scholars argue that managers in organizations facing actual or expected performance short falls tend to take greater risks, motivated in part by decision makers' beliefs that the resulting variation in performance may help to exceed performance targets in the future (Bromiley 1991; Greve 2003; Palmer and Wiseman 1999). This baseline hypothesis has been supported in prior empirical studies (e.g., Gooding et al. 1996; Greve 1998; Miller and Chen 2004). Mixed empirical results on whether failures prompt risk-seeking or risk avoidance behavior indicates that managers' responses to failures may be contingent on situational factors (Desai, 2008). Lack of consensus notwithstanding, it is reasonable to state that failures are important drivers of behavioral changes that either inhibit or foster risk-seeking behaviors.

The idea put forth by these researchers is appealing given that its helps understand decision making process in organizations. It is unclear, however, how managers take these decisions? While current failure or success may prompt managers to alter the course of actions, what role do prior experience of failures or successes play in that decision making. For example, if the current success was preceded by a string of failures of high magnitude then managers' decision may still get influenced by failures and demonstrate conservative behavior. In effect, we don't know how learning from prior failures shape current behaviors.

Recent work on learning from failure in the context of Innovation

In a very recent past, a few scholars have begun to investigate role of learning from failures in the context of innovation. Study by Khanna et al. (2015) investigates how characteristics of failures affect quality and focus of R&D output in pharmaceutical companies. They empirically demonstrate that timing, importance and scale of failures affect subsequent research output. In building their arguments, they embed an assumption that early and small failures provide important causal mechanism that improve odds of innovation. This assumption is needs to be examined given that early failures may not provide complete understanding of the underlying causal mechanisms.

Further, on empirical front, their use patents to measure innovation is questionable. Patents are good measure of new knowledge creation within organizations. They are, however, as argued by many scholars, imperfect measure of innovation. Patents do not measure the economic value of these technologies (Acs, Anselin & Varga, 2002; Hall et al., 2001). Schumpeter (1939) defined innovation as the commercial application or adoption of an invention. He argued that, "the making of the invention and the carrying out of the corresponding innovation are, economically and sociologically, two entirely different things" (p. 85) (also see Kapoor & Klueter, 2015; Lewin et al. 1987). Griliches (1979) and Pakes and Griliches (1980) argue that patents are a flawed measure of innovative output particularly since not all new innovations are patented and since patents differ greatly in their economic impact. Second it is unclear how organizations learn from failures to come up with innovation. While this study takes a step towards understanding role of failures in the context of innovation,

Similarly, Austin et al. (2014) examine how accidents provide valuable information in innovation and how digital systems moderate this relationship. By identifying specific factors such as characteristics of technologies, they predict that digital systems support innovation processes by capturing valuable uncertainty in the process. While this study brings failures to the discussion of innovation, it does not provide mechanisms about how capturing uncertainty through digital systems aid organizational learning. The information captured after accidents can offer important information to organizations. It remains unclear, however, how organizations use that information to learn and avoid accidents in future. In other words, by focusing on role information gathering in the aftermath of accidents, this study assumes role of organizational learning in improving odds of innovation.

To conclude, while we see attempts being made by scholars to address this gap, we still lack clarity about how organizations learn from failures in the context of innovation.

Summary of Literature review.

Experimentation and failures are germane to innovation. Organizations learn important causal mechanisms from failures that aid them in their search of new ideas. Extant literature has studied role of experimentation with knowledge components and their linkages to predict factors that foster or inhibit innovation. Further, focusing on failures and experiential learning, literature on learning from failure have examined how organizations learn from prior failure to improve operational performance and productivity. While these two literatures have examined the question – how firms learn from failures to innovate – in parts, collectively lack understanding of this question. While one literature ignores role of organizational learning, other assumes that mechanisms of learning from failures are equally applicable to productivity improvement and innovation.

We predict that role of learning from failure in the context of innovation is peculiar for three reasons. First, given the presence of high causal ambiguity. Organizations that learn more about causal ambiguity from failures are more prepared to direct their search that increases odds of innovation.

THEORY & HYPOTHESIS

Innovation is a highly ambiguous process in which organizations strive to discover new knowledge for implementation into products and/or services. The path tends to be nonlinear, dynamic, and unpredictable where cause-effect relationships are unknown. Further, unknown interactions among knowledge entities involved in the process exacerbates the problem of causal ambiguity. In short, Innovation a causally ambiguous and unpredictable process by which outcomes emerge.

Organizations' *a priori* lack clear understanding about how inputs link with desired novel outcomes. This *ex ante* ambiguity restricts organizations ability to establish a predictable model that provides desired outcomes. Causal ambiguity could also stem from *ex post* ambiguity in the sense that not all actions may result in a discrete outcome; results from some actions may be unclear. Some organizational outcomes are neither successes nor failures. For example, organizations experience moderate success or partial failures. This limits organizations' ability to interpret the outcomes. Further, causal ambiguity in the process gets amplified as the novelty of the desired outcome increase. In sum, organizations lack clear understating of the steps that usher them to innovation.

As a consequence, a central problem that organizations encounter in the process of innovation is determining how to reduce ex ante and ex post causal ambiguity. Researchers seek to study not only whether one variable affects another but also how such a causal relationship arises. Discovering causal mechanism that links inputs to outcomes becomes a key to reducing causal ambiguity for two reasons. First, deeper understanding of causal mechanisms helps organizations gain insights about knowledge entities, their relevant properties, interactions among these entities (both spatially and temporally) and factors that could prevent or modify the outcome.

Second, deciphering multiple causal mechanisms can be very effective when they form a hierarchy such that higher level of mechanisms presupposes existence of lower level of mechanisms thereby describing entire system of simultaneous relationships among the entities. In this way deeper understanding of causal mechanisms play a crucial role in shedding light into the black-box of the innovation process.

Organizations engage in adaptive learning or feedback-loop learning to acquire new knowledge about causal relationships. To elaborate, organizations undertake a course of action based on their prior knowledge; there is some outcome of that action. Organizations evaluate response from the action with reference to a desired outcome and decide on subsequent course of action. Evaluation of the outcome or learning from feedback helps organizations update their knowledge about the causal mechanisms.

While outcomes can broadly be dichotomized as success or failures, failures play a vital role in discovering causal mechanisms in at least three ways. First, failures provide immediate opportunity to examine causal relationships between inputs and outputs. This resembles hypothesis-testing in which organizations first constructs tentative rules about the relationships between actions and outcomes and then alter these rules based on feedback. Second, failures may help narrow down the options that have higher potential of achieving desired outcomes and thus help avoiding long random walks that plague innovation process. Finally, failures when combined with organizational memory of past experience enriches organizations' understanding of the underlying science that assist them in their pursuit of novel discoveries.

Failure Timing

Scholars argue that early failures are important than late failures since early failures provide timely feedback about a project that saves valuable resources (Khanna et al 2015). Early feedback in the R&D process allows firms to manage available resources and limit allocation of resources to unproductive arenas. Some other scholars argue that delayed feedback reduces efficacy of feedback as causal linkage between actions and outcomes may get confounded by noise (Skinner, 1954; Denrell, Fang, & Levinthal, 2004).

Given the stochastic nature of the innovation process, however, we argue that late failures play a crucial role in the innovation process for three reasons. First, while early feedback is crucial from efficient use of resources perspective, it may not shed light on all underlying causal mechanisms. Stochastic nature of the process suggests that causal mechanisms can unfold at any time during the search process for innovation. It is likely that some causal mechanisms unfold only when certain time is elapsed or a few actions are undertaken. In this sense, feedback from late failures can offer important insights into the process. This suggests that failures in the late stage can also provide important insights that early-stage failures cannot provide.

Second, in the case of R&D, the logic of time-compression diseconomies implies that maintaining a given rate of R&D spending over a particular time interval produces a larger increment to the stock of R&D know-how than maintaining twice this rate of R&D spending over half the time interval (Dierickx & Cool; 1989). Thus time is an important factor in accumulating capabilities or gaining understanding of the causal mechanisms (Levinthal & March 1993). In other words, there is an upper limit to learning rate which implies that the process of innovation can't be rushed (Sosa, 2011; Knott, Bryce & Posen 2003). Late failures defy the limitations brought about by time compression diseconomies in two ways. First, late failures provide organizations adequate time to build in-depth understanding of the underlying science. Organizations are not bound by the learning-rate constraints. Second, late failures provide organizations to accumulate absorptive capacity on one period that will positively affect efficient accumulation in the next period (Cohen & Levinthal 1990).

Finally, late failures tend be of large magnitude and hence have higher impact on organizations' strategies. Failures of large magnitudes provides an opportunity for transformative learning (Christianson et al. 2009; Lampel Shamsie Shapira 2009) Failures of substantial magnitude are recorded with more attention and more detailed lessons are drawn and recorded (Weick and Sutcliffe 2006; Rerup; 2009).

Therefore late failures are key to deciphering causal mechanism and subsequent innovation. This leads to our first hypothesis:

H1: An organization's experience with late failures is positively related to its innovation output in the subsequent period.

Failures in new projects

Innovation can result from either combining novel technologies (Fleming, 2001; Nelson and Winter, 1982; Schumpeter, 1934) or from reconfiguring existing technologies so that they can be put to novel uses (Henderson and Clark, 1990). While failures are intrinsic part of both processes, learning from these failures are not symmetrical. Failures from combining novel technologies provide greater opportunities to learn about the causal mechanisms underlying science than failures from reconfiguring existing technologies. Two reasons explain this asymmetric learning.

First, failures that occur while combining novel technologies often force organizations to depart from existing routines and protocols as the existing routines offer limited guidance for solving novel problems. Failures in novel areas mean that standardized approaches are not less effective. A need to deal with nonstandard and unexplored failures forces organizations and the scientists involved to analyze the underlying structure of a problem in more depth, thereby enhancing their understanding of the issue and thereby updating their knowledge about the underlying causal mechanisms. Thus failures in the unexplored areas provide distinctive new variations of knowledge that are necessary to understand causal mechanisms (March, 1991). To the contrary, failures that occur while reconfiguring existing technologies result in adopting previously explored approaches that were successful and may not result in new insights (Stan & Vermeulen, 2013).

Second, failures from unexplored causal mechanisms stimulate learning by greater coordination. Given that such cases necessitate a departure from standard procedures, organizational members depart from established norms of coordination. New forms of coordination provides a fresh perspective to in understanding the causal mechanisms. On the other hand, failures that involve known knowledge components, organizations may rely on established norms and procedures of coordination which may not provide new insights about the underlying science. Thus we propose the following hypothesis.

H2: An organization's experience with failures while combining novel technologies is positively related to its innovation output in the subsequent period.

Failure Concentration

The literature on organizational learning has noted the advantages of variance in learning that comes thru creation of novel knowledge and routines (March 1991; Levinthal & March 1993l;

McGrath 2001). Literature on knowledge recombination also underscores the importance of variety in knowledge (Fleming, 2001; Henderson & Clark, 1990). Specifically in the context of innovation, failures that originate from heterogeneous knowledge domains offer richer insights than failures that originate from similar knowledge domains for three reasons.

First, failures from heterogeneous domains have potential to provide a variety of insights than failures from homogenous domains fail to offer. Learning from failures from one group might solve problems of another. When existing solutions are connected with the problems across the boundaries, existing ideas often appear new and creative as they change form, combining with other knowledge entities (Hargadon Sutton 1997). Thomas Edison's laboratory is a classic example of this where Edison and his colleagues used their knowledge of electromagnetic power from the telegraph industry the lighting, telephone, phonograph, railway, and mining industries (Hughes, 1989; Millard, 1990).

This is particularly appealing in the case of innovation because it involves multiple interdependent processes filled with causal ambiguity at various stages. Absent a base of cause and effect understanding, a variety of failures from heterogeneous knowledge domains offer important opportunity to gain understanding of the causal mechanisms. On the other hand, errors from homogenous domains fail to generate variance in learning and offer novel insights. Thus, when learning from one domain is applied to another, an innovation in the form of a new synthesis, new patterns or new configuration emerges (Usher 1929; Petrovski, 1992: 44).

Second, the variety of information obtained due to failures heterogeneous domains leads to better causal inferences because an organization can draw on a sample with a wider set of action-outcome possibilities. Variance means that some examples are different from others, so a firm can see how different actions may result in different outcomes. The experimental framework is grounded in varying conditions and in using different treatments to assess causality (Cook and Campbell, 1979). Stronger causal inferences can be made by examining both situations in which the action in question happens and situations in which it does not (Mill, 1893). In other words, good experimental methodology suggests that varying experiences will assist firms in making correct inferences about underlying causality.

Third, failures from heterogeneous domains shift loci of learning from one unit to multiple units. This collaboration enhances organizational learning and subsequent innovation as sources of innovation often reside at interstices between multiple units within organizations, universities and research laboratories (Powell, 1990; Hamel, 1991; Dodgson, 1993). The literature on team performance echoes this finding that multi-team tasks sets diversity in motion that tends to stir constructive conflict around the issues at hand, which leads people to deliberate about the appropriate action (Jehn, Northcraft & Neale, 1999). This deliberation, in turn, improves group performance, especially on complex tasks. These assertions lead to our next hypothesis:

H3: An organization's experience with failures from heterogeneous domains is positively related to its innovation output in the subsequent period.

Generalist versus Specialist

While failures provide important insights by deciphering causal mechanisms, organizations vary in their ability to learn from these failures. Among others, ability to identify, accumulate and utilize knowledge from internal and external sources differentiates organizations in their pursuit of learning from failures, specifically, in the context of innovation (Cohen & Levinthal 1990).

Generalist organizations that offer a broader range of products differ in their abilities to learn from failures than specialist organizations for two reasons.

First, because generalist organizations deal with a wider variety of issues than specialist organizations do, they have access to diverse knowledge within organizations. Collaboration among technical teams between units from multiple domains enhances organizational learning (Hamel, 1991; Dodgson, 1993; Lahiri & Narayan; 2013). Further, accumulating broader knowledge in one period influences organizations' absorptive capacity thereby by assisting them to absorb more knowledge in the subsequent period (Cohen & Levinthal 1990). Second, generalist organizations with higher absorptive capacity are better positioned to utilize knowledge from external sources such as alliances. With higher absorptive capacity, organizations can form alliances with partners that are technologically more distant thereby deriving more novel insights (Ahuja & Katila??). This assertion has been supported by empirical research which shows that biotechnology firms with diverse alliances are more successful (Baum, Calabrese & Silverman 2000; Powell, Kogut & Smith-Doerr, 1996). This suggests that generalists organizations that are exposed to more heterogeneous contexts benefit from exposure to diverse ideas and experiences that allow for thinking "out of the box" than specialist organizations do (Levinthal & March, 1993; Levitt & March, 1988).

The heterogeneity in access to diverse knowledge among generalists and specialist organizations determines how they learn about causal mechanisms when encountered with failures. Generalist organizations with prior exposure to diverse knowledge are better positioned to extract superior insights from the failures than specialist organizations. Access to wider internal and external knowledge resources help generalists make novel use of their knowledge to decipher causal associations between action and consequences. This logic suggests that keeping number of failures constant, generalist organizations will extract superior insights that in turn help them innovate. In other words, generalist organizations, with prior experience of diverse knowledge will need lesser number of failures than specialist organizations to learn about the underlying causal mechanisms. This leads to our final hypothesis:

H4: Generalist organization (compared to specialist organizations) are less likely to experience enhanced benefits of increasing number of failures on innovation output.

METHODS

Industry Context

To test the hypotheses, we use the context of pharmaceutical industry - specifically, pharmaceutical firms engaged in the discovery, development, and commercialization of drugs. This industry is well-suited for this research for three reasons. First, continuous innovation is one of the pharmaceutical industry’s central characteristics. Innovations from the industry have high impact on quality and duration of human life. Second, the industry offers rich settings to examine issues concerning organizational learning. Innovations are a result of continuous learning that converge multiple technologies and know-how, including, for example, those associated with molecular biology, immunology, genetics, combinatorial and bioinformatics, and thus draws on many interdisciplinary skills ((Sorensen and Stuart, 2000; Russell, 1999).

Finally, failures are the most defining characteristics of the industry. The odds of creating a marketable drug are minuscule: only 1 in every 5,000–10,000 potential compounds investigated by the US-based pharmaceutical companies is granted FDA approval (Petrova, 2014). More than 90 percent of compounds that enter Phase I trials are destined to fall out of the development pipeline (Mckinsey, 2010). This provides an opportunity to examine failures and understand how organizations learn from failures to innovate.

Drug Discovery Process

New product development in the pharmaceutical industry is regulated and thus, proceeds along a series of well-defined steps illustrated in Figure 1. Drug development starts with an investigation of the chemical and biological properties of a compound in the lab (basic research). Scientists search for a lead compound (an organic or other drug molecule) that may act on the target to alter the disease course, for example by inhibiting or stimulating the functions of the target biomolecule. Preclinical trials precede clinical trials on humans where drugs are tested *in vivo* and *in vitro* on animals.

Finally, three stages of clinical trials or trials in human subjects (phase I, II, and III) are undertaken. In Phase 1 trials the candidate drug test conducted with about 20–100 healthy volunteers to discover if the drug is safe in humans. In Phase 2 trials scientists evaluate the candidate drug's effectiveness in about 100–500 patients suffering from disease under investigation. The main objective of this step is to examine if the drug is working by the expected action mechanism and whether it improves the condition in question. In Phase 3 trials scientists study the drug candidate in a large number of patients (about 1,000–5,000) to generate statistically significant data about safety, efficacy, side effects, and determine the ultimate tradeoffs between benefits and risks.

A drug can fail at any stage of the process. Our study is designed around the failures that can occur at any stage - from development to phase 3.

Sample

Data on innovation was gathered from US Food and Drug Administration website. FDA lists drugs that are approved every month. These drugs include totally new drugs based on New Chemical Entities (NCE) or significant enhancement to the existing drug. Information on failures was obtained using Pharmaprojects. It is a comprehensive database tracking new drug development in the pharmaceutical industry and has been used in a number of prior studies (e.g., Kapoor & Klueter, 2015; Adegbesan & Higgins, 2011; Hess & Rothaermel, 2011; Sosa, 2014).

This data has been supplemented with accounting data obtained from COMPUSTAT via the Wharton Research Data Services' (WRDS), patent data from USPTO, alliance data from Recombinant Capital[2] (Reap). Data regarding projects and company's year of establishment is obtained from Pharmaprojects.

We started with firms that have failures at various stages from 1991 to 2014. For each firm-focal year, characteristics of failures for the last seven years were examined. In the pharmaceutical industry the R&D process, including both drug discovery and development, is lengthy, averaging seven to eight years from concept to market (Buzzell 1983, Pharmaceutical Council Panel 1983, U.S. Industry Outlook 1986). Additionally, US FDA approval takes 12 to 18 months for approval. Hence it may take 9- to 10-year period for a drug to come out. Hence a period starting 9 years and ending 2 years before a drug approval is considered for a examining failures. It is not necessary that a firm discovers a new drug every year. For all such records. new drug counts are listed as zeros.

This study focusses on drug discovery in the S market only. US FDA report (2012) claims that 75% of the innovative drugs were approved first by the FDA for the US market. Some

---

[2] ReCap, a proprietary database tracking the life science industry, is one of the most comprehensive publicly available industry data sources (Kapoor & Klueter, 2015)

estimates indicate that 64 % of all research on new drugs approved in the last 10 years was done in the USA. This makes US market an appropriate setting for this analysis.

Dependent Variable

Dependent variable for this research is firms' innovation output measured as new drug application (NDAs) that are approved by US Food and Drug Administration (US FDA) from 1991 to 2015. Once approved by US FDA, firms can market and sell these drugs in the US market. FDA uses classification system to define several major categories of new drugs, which fall on a continuum from very to slightly innovative[3]. While FDA approves drugs that are based on small (chemical based) and large molecules (biotechnology based), only small molecules are considered for this analysis.

Independent Variables

Failure Timing: Drug discovery process broadly consists of five major steps – discovery, preclinical, Phase 1, Phase 2 and Phase 3[4]. Discovery process identifies a drug candidate that holds a promises for treating one or more disease indications. Preclinical stage tests the drug candidate on animals. Phase 1 through 3 – also termed as clinical trials – test the drug on human subjects. Phase 1 tests the drug on a small set of healthy human subjects to establish safety. Phase 2 and 3 test the drug on small and large groups of healthy volunteers and patients respectively to demonstrate safety and efficacy. This is a sequential process and a drug candidate can fail at any stage. To measure failure timing, we capture the stage at which the drug fails. A drug can fail at

[3] The FDA classifies all new drugs on two dimensions: by the type of chemical entity (NDA Chemical Type) and by their therapeutic potential (Review Classification). See appendix for more details.

[4] Refer appendix for more details on the drug discovery process.

five different stages mentioned above and we count failures at each stage for the past seven years before a new drug is approved by FDA.

Failure in new areas:

Failure Concentration: To measure failure concentration, we classify errors based on the disease indications they intended to treat. These indication are classified in fifteen different therapeutic classes. Failure concentration is calculated using Herfindahl index

$$\mathrm{FC} = \sum (P_i)^2$$

Control Variables

We included control variables to rule out alternative explanations that predict innovation success. All control variables, were measured at the focal year.

Firm Size: We controlled for firms' size as large firms' often have superior financial and human resource endowments that enable them to nurture superior R&D capabilities. Large firms may also have greater market power or positional advantages compared to their smaller rivals that affect their incentives and abilities to adapt to changing environments (e.g., Delacroix and Swaminathan, 1991; Baum and Oliver, 1991;Baum and Mezias, 1992).We measured the size by taking firms' revenues and transforming it to log scale for each year.

R&D Unit: We controlled for the size of firms' R&D units, as large R&D units are likely to have higher output. We calculated the size of the R&D unit by counting the number patents filed with US Patents and Trademark Office.

Firm Tenure: We controlled for firms' tenure as older firms may have higher older firms benefit from accumulated experience (e.g., Amburgey, Kelly, and Barnett, 1993; Carroll and Delacroix, 1982). It was measured as the number of years that had passed since the firm was first founded

Number of Alliances: Prior research shows that firms source innovation externally (Ahuja, 2000; Ahuja & Katila, 2001; Sampson, 2007). To account for this, we controlled for the number of alliances for each firm in our sample for each year.

Number of Projects: Bigger firms often start greater number of projects and encounter larger number of failures compared to smaller firms. This provides larger firms to learn more from their failures. This may confound our propositions regarding learning from failures. Therefore, we controlled for the R&D projects that firms start each year.

Nationality: Firms from developed countries have advantage over firms from developing countries. Developed countries often have strong intellectual property rights protection that encourage innovation. Firms from these countries also have better access to talent pool and resources. We controlled for these confounding effects by including country location of firms headquarters in the estimation model.

Estimation Method

The following models test the relationships between failures and innovation in a panel dataset for 102 firms for a period of 25 years:

$$NDC_{it} = \beta_1 T_{it-7} + \beta_2 NCE_{it-7} + \beta_3 TC_{it-7} + \beta_4 G_{it} + \beta_5 C_{it} + u_{it}$$

In the above equation NDC_{it} is novel drug count, measured as the number of new drugs approved by US FDA for a firm i in year t. Tit-7 is a vector of timing of the failures during past 7 years. NCE_{it-7} represents proportion of failures from projects that used new chemical entity or new compound vis-à-vis existing compound. TC_{it-7} represents concentration of failures across 14 therapeutic classes. G_{it} represents whether the firm is a generalist or a specialist. C_{it} is the matrix of control variables, and includes revenue, number of alliance, number of patents filed with US PTO, number of projects started by the firm, firm age and country of head quarter location.

Several considerations need to be addressed while estimating the model outlined above. First, new drugs invented is a count variable which suggests that either Poisson or negative binomial estimation models are appropriate for estimation. Second, the data shows significant evidence of overdispersion of the dependent variable[5]. This make poisson model unsuitable as when the assumption of equidispersion is violated, standard errors may be underestimated and the results may be unreliable (Cameron and Trivedi, 2009). Negative binomial regression model more appropriate in such settings as it can be used for over-dispersed count data or when the conditional variance exceeds the conditional mean (Hausman, Hall, and Griliches, 1984). It can be considered as a generalization of Poisson regression since it has the same mean structure as Poisson regression and it has an extra parameter to model the over-dispersion.

Finally, we need to account for the effects of unobserved firm-specific effects. To determine the choice between a random- and a fixed-effects model, I employed a Hausman (1978) specification test on the baseline model. The Hausman test was strongly significant (p = 0.01), indicating that a fixed-effects model is the appropriate model for this analysis[67]. The negative binomial model for panel data is estimated using the XTNBREG command in STATA (13.1).

---

[5] See appendix for goodness-of-fit test.

[6] See appendix for more details about Hausman test.

[7] Consistent with Wooldridge (2002), the fixed effects model did not include the time-invariant variables such as country dummies that captures firms' head quarter locations. As such, we tested for the country effects in the random effects model. Country effects were found to be insignificant. There was no other material change in other effects – as such, effects were consistent between the fixed and random effects model.

RESULTS

Table 1 shows the descriptive statistics and correlations among the all variables used in the analysis. Since revenue, patent count, and number of projects started are skewed, we transformed these variables using logarithmic function to correct for non-normality of the distribution (Chaganti & Damanpour, 1991; Ruef & Patterson, 2009). Some of the independent variables are significantly correlated. To ensure that estimation is not affected due to multicollinearity, we estimated variance inflation factors (VIF) for all variables. VIFs for all variables are less than 10 with an average value of 2.98 for models that testing hypotheses 1 through 3 and 1.85 for variables testing hypothesis 4. The factors are within the acceptable range hence no further action is required.

Table 2 displays the results of the models prediction innovation output. Model 1 displays baseline model with only control variables. Model 2 through model 6 test effects of individual independent variable in conjunction with control variables. Model 8 tests hypotheses 1 through 3 will all independent variables and control variable together. Similarly, model 7 tests hypothesis 4 with independent and control variables.

Hypothesis 1 predicted that late failures should positively affect innovation output. This was based on the assertion that late failures, like early failures, can offer important insights about the causal mechanisms. The result partially supports this hypothesis. Failures in the preclinical phase positively affect innovation output as the coefficient ($\beta = 0.005$) is statistically significant ($p < 0.05$). While these failures cannot be classified completely as late failures, however, these failures occur after early stage (discovery stage). Ideally, errors in the Phase 2 and Phase 3 are late stage failures. Coefficients of these are statistically not significant. Negative coefficients on Phase 1 and Phase 3 suggest that these failures hinder innovation. Intuitively, costs of failures in these stages are prohibitive. Any such failure occurrence, however beneficial for learning, may hamper funding for subsequent R&D projects that in turn may influence innovation output. Thus we conclude that hypothesis 1 finds partial support in this analysis.

Hypothesis 2 predicted that experience with failures while combining novel technologies is positively associated with innovation output. Learning that occurs from such failures generate more novel insights than failures that occur while building on existing knowledge that enhance understanding of the causal mechanisms. Model 3 and model 8 in Table shows results for this hypothesis. While model 3 shows partial support for this hypothesis ($\beta = 0.380$, $p < 0.1$), model 8 shows support for this hypothesis ($\beta = 0.528$, $p < 0.05$). Overall, these results support our assertion that insights gained from failures combining novel technologies are superior and enhance odds of innovation output.

Model 4 and 8 display results for hypothesis 3 which predicted that an organization's experience with failures from heterogeneous domains is positively related to its innovation output. Both model show strong support for the hypothesis (β = 1.173 and 1.216; $p < 0.01$). The empirical results support the logic that an organization's experience with heterogeneous information help it think 'out of the box'. The finding also echoes the findings of the literature on knowledge recombination which assert that existing solutions when applied to a novel context can yield innovative ideas.

Finally, we hypothesized that generalist organizations, compared to specialist organizations are less likely to experience enhanced benefits of increasing number of failures on innovation output. This assertion rests on the logic that generalist organizations, with prior experience of diverse knowledge, will need lesser number of failures than specialist organizations to learn about the underlying causal mechanisms. Model 6 and 7 shows results for this hypothesis. In the full model, while total failures have positive impact on the innovation output ($\beta = 0.004$; p

< 0.01), the interaction term with generalist form of organizations is negative and statistically significant (β = -0.003; $p < 0.01$). The negative interaction term supports our claim that positive effect of increasing number of failures is dampened for generalist organizations as they are able to derive superior extracts from smaller number of errors.

Robustness Checks

To test the stability of our results we conducted multiple robustness checks. First, despite our extensive data collection effort, data on financial characteristics is missing for some firms in the sample. Organizations with missing financial data tend to be slightly younger, less experienced, and hold fewer patents than the industry average. Further, inclusion of firm fixed effects in our estimation models means that organizations drop out of any regression where there is no variation in the dependent variable. To ensure that regression results are not materially affected by the missing data or exclusion of organizations from the analysis, we reran all models without firm fixed effects, to recover dropped observations. While keeping the overall estimation method same (Negative Binomial method) we used random effects. We also included country dummies in this model to capture as much firm level heterogeneity as possible. The results from this estimation model are documented in the table 4 of appendix. This exercise produced essentially identical coefficient estimates for all our main variables of interest, indicating that the results reported in the table 2 are not dependent on particular sample restrictions.

Second, while overdispersion of the dependent variable suggests that negative binomial is most suitable model for estimation, we also ran Poisson fixed effects model to test the robustness of the estimates. The results are documented in Table 4 of the appendix. Except for the weak support for the second hypothesis, results from Poisson fixed effects model are in line with the main model.

Third, a stream of research on organizational learning argues that learning depreciates over time (Darr, Argote & Epple; 1995). While learning persists through time, knowledge gained and stored in organizational memories suffers from decay. This has important implication for our analysis as failures that occurred a few years before may have lesser impact than failure that occurred in the current year. In other words, simple aggregation of failures for the past 7 years may ignore knowledge depreciation phenomenon. To examine if the results are biased due to this, we ran a separate model applying a decaying factor of 0.9 for previous years in a geometric progression. Apart from the main model (negative binomial with fixed effects), the effect of knowledge depreciation is also tested in the Negative Binomial random effects model and Poisson fixed effects model. The results are showed in the table 3 of the appendix. While we observe a slight variation in the strength of the coefficients of variables of interests, overall direction and significance of the variables are in the expected direction. This shows that our results are robust to the effects of knowledge depreciation.

Finally, while we examine failures for past 7 years based on the previous research and feedback received from the scientist, we also conducted estimation for 8, 6, and 5 years period to assess the stability of the results. Apart from the main model (negative binomial with fixed effects), the effect of different year periods is also tested in the Negative Binomial random effects and Poisson fixed effects model. The results are largely stable and in the same direction. The results are documented in the appendix section. (Table 3)

DISCUSSION AND CONCLUSION

Despite prolific research on organizational learning, our understanding of role of organizational learning as a source of innovation is incomplete. We take a step towards addressing this gap. In this research, we develop and test hypotheses that shed light on the causal association between experiential learning and innovation. Our work contributes to theory and practice in the following ways.

First, we contribute to the literature on innovation. Despite multiple calls on understanding sources of innovation (Lahiri & Narayan 2013), role of learning from experience as a source of innovation has received meagre attention. In most cases, innovation is treated as a black box and more attention is given towards either factors of input or the environment. Such treatment is inappropriate as it undermines organization's persistent efforts at learning causal mechanisms that drive the process of innovation. This research sheds light on the black box by highlighting the role of experiential learning in the process of innovation.

Second, we examine how learning from failures affects organizations innovation performance. Failures experience is an important form of experiential learning as failures provide rich information about organizations' practices and routines (Haunschild & Sullivan, 2002; Desai 2015). While extant literature on learning from failure has examined how organizations learn from prior failure to improve operational performance and productivity, our work extends past research in this area by highlighting the role of failures in the context of innovation. Innovation is a causally ambiguous process and this research highlights the role of failures that play a central role in reducing the causal ambiguity thereby increasing the odds of innovation.

Third, our results, unlike previous research, indicate that not all failures are equal. Failures that occur when organizations combine novel technologies are more valuable than failures that occur while building on existing knowledge. Further, failures emanating from heterogeneous domains act as a catalyst in reducing causal ambiguity and offering rich insights into the innovation process. In sum, our work highlight the specific characteristics of failures that enrich organizational learning from failures in the context of innovation.

Finally, our contribution lies at the intersection of two literatures – knowledge recombination and learning from failures. On the one hand, the literature on knowledge recombination has studied role of experimentation with knowledge components and their linkages to predict factors that foster or inhibit innovation. On other hand, the literature on learning from failure has examined how organizations learn from prior failure to improve operational performance and productivity. While the first literature stresses on the constituents of innovation thereby ignoring the importance of organizational learning, the second has focused on the process of learning but ignored the context of innovation. Our work contributes to the intersection of these two literatures. We synthesize two perspectives to understand how learning from diverse failures help organizations combine knowledge to discover novel ideas.

While this research makes above-stated contributions, there are boundary conditions that influence its generalizability. First, the pharmaceutical industry is highly regulated which puts limitation on idea selection and executions. Second, given the steep resources needed for drug testing, complementary assets pay crucial role. This puts large organizations at an advantage compared with the small organizations. This may limit our understanding about the heterogeneity that exists in learning from failures. Finally, strong appropriability norms in the industry creates entry barriers and may prevent new entrants. This reduces the chances of technology disruption. Before extending the findings from this study to other industries, the above-stated boundary conditions need to be considered.

There are important normative implications for practicing managers. While failures are intrinsic part of experimentation, higher number of failures may not improve odds of innovation. Quality of failures in terms of their spread across knowledge domains will offer richer insights than failures that are concentrated. Further, it is also helpful to avoid mistakes in the known areas but encourage them in unknown areas.

There are several limitations to this research. First, we could not observe the role of failure-timing in deciphering causal mechanisms. As argued, late failures are very crucial to understanding causal mechanisms. Late failures are important specifically when multiple causal mechanisms form a system and early failures fail to shed light on all mechanisms that the system presents. In the pharmaceutical industry, late failures have significant implication from resources perspective as drug trials in the late stage can cost anywhere between US $ 500 million to US $ 1 billion. While failures at this stage may offer rich insights from learning perspective, it affects subsequent R&D projects allocation and thus may affect innovation output. Future research may shed more light on this by examining late failures that don not constrain resources.