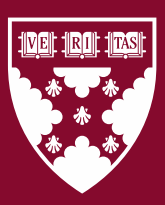
Working Paper 25-034

Crossing the Design-Use Divide: How Process Manipulation Shapes the Design and Use of AI

Rebecca Karp



Harvard Business School

Crossing the Design-Use Divide: How Process Manipulation Shapes the Design and Use of AI

Rebecca Karp
Harvard Business School

Working Paper 25-034

Funding for this research was provided in part by Harvard Business School.

Crossing the Design-Use Divide: How Process Manipulation Shapes the Design and Use of AI

Abstract
Existing literature often separates research on the design of innovations from their implementation and use, neglecting the role of selection—how organizations choose which innovations to implement. Although scholars suggest scientific approaches for selecting novel technologies, there is limited research on how these methods are practically employed in decision-making. This study addresses this gap by examining how organizations decide which innovations to implement and how the selection process influences their design and use. Drawing on a two-year ethnographic study, the research explores how 13 dyadic pairs of entrepreneurial firms and health system committees piloted AI-based medical diagnostic innovations. Committees, composed of members with polarized views on AI, formed coalitions reflecting these views. Dominant coalitions engaged in "process manipulation," strategically altering the piloting process to achieve self-interested outcomes while maintaining an appearance of rigor. Coalitions enthusiastic about AI scoped pilots to test basic uses, ensuring success, while skeptical committees tested advanced uses, hoping for failure. This manipulation constrained entrepreneurs' ability to advocate for their innovations and demonstrate market differentiation. The paper highlights the dynamics of process manipulation and its impact on AI innovation development and use.

Word count: 185

INTRODUCTION

The implementation of a novel technology or innovation within an organization can trigger broad scale and transformative change to work processes (Barley, 1986; Beane, 2020), roles (Bechky, 2020; Pachidi et al., 2021), and even occupational identity, such as when librarians began using internet search technologies (Nelson & Irwin, 2014). Yet little research on implementation examines how innovations, with the power to catalyze substantive change, appear within a given organization at the start. Rather, scholars operate across a divide. On one side, students of innovation and technological evolution focus on how firms, communities, and varying social groups organize to design new technologies, stopping short of investigating how a discrete technology is adopted and used within a consuming organization (Pinch & Bijker, 1984; Kline & Pinch, 1996; Suarez & Utterback, 1995; Rosenkopf & Tushman, 1998; Hargadon & Douglas, 2001; Anderson &

Tushman, 2018; Benner & Tripsas, 2012). On the other side, scholars interested in implementation and use focus on how new technologies alter, or reinforce, work tasks, roles, and organizational structures, without examining how innovations arrive within an adopting organization (Barley, 1987; Orlikowski, 1993; Leonardi 2009; Edmondson et al., 2001; Nelson & Irwin, 2015; Beane, 2020; Pachidi et al., 2021; Anthony & Beane, 2023). While both literatures offer valuable insights regarding how innovations and organizations evolve, they neglect to explain how adopting organizations select innovations for use. This is an important omission to rectify, as the way organizations select innovations may very well shape their design and use (Thomas, 1994).

Selection, "choosing an innovation for use," involves the set of activities organizational members engage in when determining whether to implement or reject a given technology (Rogers, 1995). Although not an entirely linear process, scholars contend that selection occurs after design, but before an innovation is implemented and used within an organization (Leonardi & Barley, 2010; Leonardi, 2007). To select novel innovations, scholars posit that organizations may run a "scientific-like" process that involves piloting an idea, process, or innovation before implementing it (List, 2022; Thomke, 2003). For example, Edmondson et al. (2001) show how hospitals piloted minimally invasive surgery, a very novel technique, before broadly rolling it out. However, this technical or scientific perspective ignores a more behaviorally plausible model of selection, where some organizational members might coalesce to advocate for certain innovations over others, to reinforce their own power or status (March 1994). Pilots may be more subjectively executed than appreciated and organizational members may not design tests with an eye towards selecting the most proficient innovation. Power plays and persuasion from both designers and organizational members might be as likely a determinant of why an innovation is selected compared with the unbiased results of any given pilot. However, without directly observing how the selection process unfolds, it is challenging to understand how certain innovations come to be implemented within an organization while others do not. As Leonardi and Barley (2010: 38) explain, "when studies begin with little insight into why technologies were designed as they were, why one technology was chosen over another … it is impossible to determine whether patterns of use are shaped in important ways by dynamics of power."

The process of selection does not just effect how organizations use innovations downstream, it also likely shapes how innovations are designed upstream as selection screens the features and capabilities a designer can encapsulate within an innovation. Buoyed by desires to see their products adopted, designers may be influenced by powerful members of potential adopting organizations, to develop products that promote some features over others (Karp, 2023) or that benefit certain members at others' expense (Myers, 2023). Designers can create products that even temper performance in an appeal to special interests. For example, Noble (1984) showed how powerful groups shaped the design of numerical control systems in ways that diminished technical performance. As Noble asks: "The best technology? Best for whom? Best for what?" (2011: 240). While scholars often advocate for user-centered designs (von Hippel, 1986; Norman, 2011), it is the organizational members involved in selection that may have outsized influence on design choices, rather than users, as designers hope to convince those involved in selection to advocate for and pick their innovations. Users' and decision-makers' interests may not align. Yet little research explores how the process of selection advances or tempers the development of an innovation, nor how the will of those in charge of selecting a novel innovation interacts with the intentions of designers.

Examining selection takes on heightened importance when considering innovations built on AI (Bailey & Barley, 2020; Anthony, Bechky, & Fayard, 2023; Kellogg, 2022). Innovations that leverage AI create "constant change, invisibility, and inscrutability [which] call into question the applicability of prior findings about technology" (Anthony et al., 2023: 1674). AI is inherently modular and decomposable (Baldwin, 2023), as features can easily be turned on or switched off to suit the preferences of those in decision-making positions (Karp, 2023). Thus, two organizations might select to pilot very different instantiations of the same innovation, which is less plausible with other types of novel standalone innovations. Those with power to make these decisions have become increasingly important agents in not just gating but also shaping innovations, as they select not only if an innovation will be used, but also what aspects of an innovation and for what uses. At the point of selection, organizations make a host of visible decisions about how they might use AI, and AI technologies are granted access to organizational databases and data. Access to such sources of data are key differentiators in the quest to train and develop AI-based

innovations. Thus, observing the decisions organizational members make regarding the selection of AI provides a unique opportunity to bridge the gap and explore the relationship between design and use.

This paper draws on two years of ethnographic data to unpack: How do organizations decide which innovations to implement and how does this process shape the design and use of an innovation? Specifically, I examine how 13 overlapping dyadic pairs of three health systems and six entrepreneurial firms attempted to pilot AI-driven innovations for medical diagnostics and treatment. Paying attention to the role internal organizational politics plays (Selznick, 1953; March, 1994; Thomas, 1994), I reveal how selection can be manipulated when committees form coalitions based on strong outside priors about an innovation's capabilities. Dominant coalitions wary of AI designed pilots to test only advanced capabilities of entrepreneurs' innovations in hopes that the pilots would fail. In contrast, dominant coalitions enthusiastic about AI designed pilots focused on testing basic capabilities. These pilots succeeded, but limited entrepreneurs' ability to prove that their innovations were differentiated in the market. Very few innovations piloted to test advanced capabilities were successful. However, conditional on a successful pilot, these innovations were more likely to be used than innovations introducing basic capabilities. I introduce the concept of process manipulation: working within an accepted process but strategically engineering select steps or micro-decisions to achieve specific results - to explain how dominant coalitions altered the process of piloting to achieve their own self-interested ends. I show how process manipulation can alter the trajectory of what is tested and subsequently designed and used.

SELECTION AND CROSSING THE DESIGN-USE DIVIDE

Scholars have long called for studies that cross the "implementation line" (Leonardi, 2007) and account for the relationship between the design and use of novel innovations (Thomas, 1994; Leonardi, 2007; Leonardi & Barley, 2010; Bailey & Barley, 2020; Anthony et al., 2023). As scholars contend, understanding this relationship is crucial as design and use are not "discontinuous events" (Leonardi, 2007), but rather part of a reflexive cycle, where the way organizations use an innovation informs an innovation's design. Designers' intentions and development decisions may in turn constrain or enable any organizational change triggered by the implementation of an innovation and its use (Forsythe, 2001). These calls have

reached a fever pitch in the era of AI as these technologies are "radically different from prior technologies in [their] potential to transform the landscape of work" Bailey and Barley (2020: 1). Though often neglected, a useful way to explore the relationship between design and use is to focus on how organizations select novel innovations. During selection, organizations make decisions about how to use an innovation, and designers are open to adapting and modifying their innovations as they learn of what it takes to convince adopting organizations to choose their innovations.

Historically, selecting a novel innovation involved fewer decision-makers and a minimal set of decisions (Rogers, 1983). Since innovations were more standalone and less decomposable, decision-makers made a choice, if to adopt an innovation, not what aspects of an innovation to use (Baldwin, 2023). Consider Barley's (1987) CT scanner: it is difficult to imagine that hospital decision-makers could choose to rapidly adapt, limit, or modify the functionality of the CT scanner at the time. Staff in Barley's study were exposed to CT scanners with the same set of embedded capabilities. In contrast, when an innovation is digital or built upon AI, organizational decision-makers ability to locally shape and adapt an innovation is amplified (Kellogg, 2022), which means two of the same "scanners" can offer the user a different capability set. Today, an organization might quickly change or modify software to adjust the way technology is used on-site. Decision-makers often have latitude to turn on and off the multitude of features embedded within a digital or AI-based innovation to suit their own needs (Karp, 2023). Organizational members at all levels may be involved with decision-making processes, as organizations seek out the opinions and grant authority to a variety of different staff (Lee, 2024). The combination of increased decisions and decision-makers introduces degrees of freedom into the selection process. This flexibility further blurs the line between design and use and elevates the role selection plays in shaping both.

A Scientific Approach to Selection

Selection may be a simple process to undertake when innovations are familiar (Pisano, 1994). For example, Leonardi (2007) described how ITSM, a help-desk queuing application already in use at the SkyLabs organization was implemented in a new group of networking engineers with no internal deliberation

and in only a few weeks' time. However, to manage the complexity of multiple decision-makers and multiple decisions, scholars posit that organizations use a technical or "scientific" approach (Merton, 1973) to decision making by experimenting and testing innovative products, ideas, and practices before implementing them (Pisano, 2001; Thomke, 2003). For example, in contrast to SkyLabs, all 16 hospitals in Edmondson et al.'s (2001) study "piloted" a novel form of minimally invasive cardiac surgery before integrating this method into regular use. Piloting is a type of experiment that involves testing the performance of an innovation on a representative sample of tasks or use cases, usually in a live or quasi-live environment, and then comparing the pilot's performance with that of traditional ways of working (List, 2022). As pilots unfold, adopting organizations and designers can learn how an innovation might work in situ and make modifications in response to real-time feedback. Scholars advocating for a scientific approach to decision-making implore that because piloting and experimentation are scientific methods, they quell latent or acknowledged biases that would otherwise cloud decision-making (Wuebker, Zenger & Felin, 2023; Thomke, 2003; Ries, 2010; Camuffo et al., 2020; Koning, 2020; List, 2023). As such, a scientific approach can help organizations generate data to analyze decisions before they are broadly enacted and hopefully diminish the challenge of managing differing voices and choice sets.

Yet, determining how to design a pilot is not a straightforward feat. Ideally, when designing a pilot scientifically, scholars suggest that organizations look to balance learning about the radical or most advanced capabilities of an innovation with gaining information about the generalizable uses of that innovation (Eisenmann, 2010; Gans, 2023). Innovations are viewed as more generalizable or scalable when they offer benefit to a wider set of users (Thomas, 1994). Offering a greater set of users benefit usually means testing more basic aspects of an innovation; aspects that are less specialized and that a greater quantity of individuals can understand and use (Bresnahan, 2010). For instance, the application Wix offers users templated website designs, no coding required. While most individuals can utilize this application with ease, it limits the types of customizations or advanced and specialized capabilities a user can build in. Thus, to satisfy both goals—learning and generalizability—organizational decision-makers manage tradeoffs, focusing pilots on aspects of an innovation that balance between advanced and basic capabilities. Following this logic, it is unlikely that

organizations will design pilots testing the tails of an innovation's capabilities: the advanced and radical capabilities of an innovation or the most incremental, benign capabilities. The implication for the design and use of novel innovations is that if organizations make decisions based on a scientific approach, they will select and implement innovations with capabilities that balance technical proficiency with generalizability. From these pilots, designers likely also learn how a their innovations perform across a range of uses.

Yet enacting science is rarely "value free" (Merton, 1973). Organizations may use quite a bit of intuition and judgement when designing pilots, which might deviate from the purely scientific. For example, Gans (2023) recently theorized that organizations may design pilots to reaffirm their views on the market, rather than gain information about an innovation's capabilities. Gaining information may not be the only goal. Drawing from March (1994), organizational decision-making often occurs as part of a political process where action is strategic, rather than scientific, and organizational members band together and form alliances or coalitions to see their preferences come to fruition. Much literature has focused on the role these power dynamics play in shaping who is granted decision-making authority within organizations (Mechanic, 1962; Pfeffer & Salanick, 1974; Lackman, 1989; March, 1994; Goldstein & Hays, 2011; Friedkin, 2011; Truelove & Kellogg, 2016). Rather than make decisions based on a technical process, decision-makers with power may instead select innovations that help them maintain their advantage. Those without power might form alliances to advocate for their own interests. These power dynamics may very well shape how organizations select novel innovations, yet they are neglected by a scientific approach to decision-making.

A Behavioral Approach to Selection

Instead of making decisions purely using a scientific lens, organizational members might take a more strategic or behavioral approach to selection, and use their authority or power to assert decisions. Following March (1962; 1994), organizational members accomplish this task generally through one of two ways: (1) by leveraging their structural position within an organization; (2) by using their skill at impression management and issue selling (Fligstein, 2001; Dutton & Ashford, 1993). For example, Pfeffer & Salanick (1974) show how committees tasked with allocating university resources granted powerful departments more than their fair share of graduate fellowships. Howard-Grenville (2010) explains how issue sellers in relative positions of

disadvantage compared with issue recipients were able to advocate for their interests and shift Chipco towards a more environmentally friendly manufacturing process. While the mechanism by which members acquire power differ, one structural, the other skill-based, the implication is the same: those with power, or those savvy enough to acquire it, determine whether an innovation is selected and for what purposes. Yet how these dynamics might play out and the consequences of these dynamics on both design and use are less clear. For example, Thomas (1994), shows how mid-level engineers engaged in issue selling and convinced senior managers to select surface mount technology to improve their status within the organization. But selection took a long time. Implementation took a long time as well and did not advance the design of SMT in any notable way, nor the status of engineers. Further, advocacy may originate from designers, who may attempt to find powerful decision-makers, be they positionally powerful or good at issue selling, to champion their innovations even if those innovations benefit certain organizational members over others (Howell & Higgins, 1990).

It is also unlikely that organizations completely abstain from scientific approaches to selection given their popularity. Piloting and experimentation are commonly used techniques and likely do inform decision-making processes (List, 2022). Thus, organizational members might balance a scientific approach with a more behavioral one, to select novel innovations. For example, in this study, all of the adopting organizations designed and executed pilots to test out novel innovations. However, those with power used pilots to engage in process manipulation—working within an accepted process but strategically engineering select steps or micro-decisions to achieve specific results. Process manipulation varied from other forms of more assertive power (March, 1962; 1994), as it involved maintaining the veneer of a scientific process by making the right claims and involving all decision-makers. I explain how process manipulation occurs within organizations and how it shapes both the design and use of novel innovations in material ways.

METHODS

In 2016, I engaged in a broad field study to examine how entrepreneurs innovating in healthcare commercialized their products and services. I interacted with three regulatory organizations, more than 50 entrepreneurial firms and 100 experts in the healthcare industry, and 10 large organizations that might

ultimately license and use entrepreneurial firms' innovations. Ethnographic interviews with the latter organizations revealed customers' outsized interest in the digital technologies—mainly machine learning and artificial intelligence—underpinning some of the entrepreneurs' innovations. Some of these customers nevertheless feared that deployment of these technologies, purported to improve both the quality and efficiency of healthcare delivery, might displace work tasks or routines within their organizations and shift organizational power structures in politically unpalatable ways. Entrepreneurs were well aware of such concerns.

Because machine learning, or "AI," technologies were spoken of so ubiquitously and frequently, I made them the focus of my efforts to better understand the decisions that determined whether and how these innovations were implemented within customer organizations. Because the selection process was infrequently observed, I leveraged an inductive field research design, which was particularly well suited to developing an understanding of less researched settings (Bailey & Barley, 2020; Edmondson & McManus, 2009). The use of machine learning for medical diagnostics has been hotly and publicly debated due in large part to scholars' and practitioners' concerns about its role in displacing work (Ghassemi et al., 2019; Jamison & Goldfarb, 2019; Topol, 2019; Brynjolfsson et al., 2020; Leibowitz et al., 2021). Although the use of artificial intelligence to support differential diagnoses is not a novel concept, recent advances in machine learning have enabled algorithms to replace, rather than merely support, diagnostic work performed by doctors and nurses (Esteva et al., 2017; Topol, 2019). A qualitative field research design offered me the space through observations and interviews to discern and probe upon informants' experiences, which could have been lost through surveying or other quantitative means (Becker, 1998).

All the organizations in my study tested how to both roll out and use an innovation. To gain more in-depth knowledge of the piloting process, I focused data collection efforts from 2017 to 2019 on three large healthcare systems that agreed to pilot the AI-based innovations of six entrepreneurial firms. These three customer organizations and six entrepreneurial firms formed overlapping dyads. In all cases, at least two customer organizations agreed in principle to pilot each entrepreneurial firm's innovation. Analysis at the dyadic level helped clarify whether selection varied because of idiosyncratic behaviors of the customer

organization or entrepreneurial firm. Analysis at the dyadic level allowed some generally unobservable alternative explanations to be more observable, such as a healthcare system's or entrepreneurial firm's lack of adequate implementation capabilities. This concern would be revealed over a set of multiple dyadic pairs if, for example, one entrepreneurial firm was unable to successfully pilot its innovation with any customer or any customer was unable to successfully pilot with any entrepreneur.

Research Context and Sample Selection

All six entrepreneurial firms in this study used AI as foundational inputs to performing medical diagnostic work. AI is a "form of computational statistics, [and] is based on algorithms that use data to generate predictions" (Jamieson & Goldfarb, 2019: 778), which improve automatically as an algorithm encounters more data, enabling machines to eventually perform work with minimal or no human intervention. This inherent learning process may be particularly important in shaping the innovative trajectory of a given application (Fraser & Ozcan, working paper). Access to one type of data versus another may influence an application's future capabilities and determine how well it competes with comparable market offerings. For example, an application trained on diverse sets of data may be better able to detect illness in minority populations and therefore be more valuable than comparable applications not equivalently trained.

AI technologies are generally used to perform medical diagnostic work in one of three ways: (1) by capturing or counting known medical irregularities difficult to observe with the human eye or existing technologies, (2) identifying irregularities not previously known to be indicators of an illness, and (3) offering potential treatment plans for confirmed diagnoses. MAMMO, one of the firms in my sample, leveraged machine learning in two of these ways. Its application could limit false positives by more accurately determining when mammograms contained no known malignancies compared with radiologists and could identify novel indicators of cancerous tumors in patients whose mammograms were previously diagnosed as clinically negative. No intervention by doctors was required to make these diagnoses.

Sampling strategy. Over the course of my time in the field, I engaged with three large health systems (Red Hospital, Blue Hospital, and Community Hospital) that were interested in purchasing and using

AI-based innovations. Two were formed around large academic hospitals. One health system was originally formed around a community hospital. These organizations shared an overlapping interest in licensing and using the offerings of six entrepreneurial firms (MAMMO, SENSOR, INFECTION, SKIN, DIAB, WHITE). The attestations of these organization members were not merely lip service, as all three organizations had paid to pilot entrepreneurial firms' innovations and established internal committees tasked with designing and executing pilots. Piloting committees were diverse, composed of individuals who served in various roles (e.g., developers, doctors, nurses, administrators) and at different levels (e.g., directors, managers, coordinators) within each health system. Committee members varied within organization and members were selected by a mix of stakeholders: heads of innovation, departments heads, and technology executives. Although diverse, all of the informants I spoke to in these organizations were well versed in the "gestalt" of AI and knew the arguments for and against implementing this technology. People were not shy about sharing their opinions. The individuals that selected the committees were sometimes part of the committees themselves. When members were asked about why they were chosen to participate on the committee, rationale was mixed and ranged from "I am diligent" and "not afraid to share my opinions respectfully" to "not really sure why" and "I guess I did something right." A head of innovation shared, "Committees were formed by people who had a stake in the use of this innovation, and I tried to pick people who could think carefully about the innovation." Surprisingly, members or leaders never commented about being selected or selecting members based on their views on AI.

Although I engaged with many entrepreneurial firms focused on innovating in healthcare using AI, I selected the six identified above as they had pilot agreements in place with at least two of the large customer organizations in my sample. All had proof that their innovations could work in situ and were looking to expand upon that evidence in ways that could differentiate their innovations in the market. They also exhibited marked differences, varying in team experience, targeted medical specialty, and geographic proximity to investors or academic centers. Scholars show that entrepreneurial firm performance is influenced by prior team experience (Kor, 2003), sector focus (Agarwal & Gort, 2002), and propinquity to sources of capital or knowledge (Powell et al., 2005). The diversity of this sample offered the theoretical

range needed to illuminate common aspects of the selection process and highlight differences regardless of firm experience, sector focus, or geography (Lawrence & Lorsch, 1967; Harris & Sutton, 1986; Santos & Eisenhardt, 2008). Table 1 profiles the six entrepreneurial firms in my sample.

Insert Table 1

From this set of entrepreneurial firms and customer organizations, I constructed a sample of 13 overlapping dyadic pairs. Table 2 provides an overview of each dyad.

Insert Table 2

Data collection. Although the present study draws primarily from ethnographic observations conducted between 2017 and 2019, I supplemented this data with a series of structured interviews and entrepreneurial and customer firm data that included strategic planning documents, news articles, blog posts, and scientific publications.

Ethnographic observations. Observations occurred in meetings between entrepreneurial firms and customer organizations and at numerous public events. Over the course of the study, I conducted over 1,000 hours of observations. After each day of observation, I recorded field notes along with any emerging insights.

Interviews. Throughout the course of the study, I conducted 40 formal interviews, ranging in duration from one to two hours, involving multiple members of each entrepreneurial firm and decision-makers at different levels in each of the three customer organizations.

Entrepreneurial and customer firm data. I collected three types of firm data: (1) news articles and blog posts on both entrepreneurial firms and customer organizations; (2) scientific reports and journal articles detailing results of any pilots or trials run by the entrepreneurial firms; and (3) self-reported progress reports detailing any roadblocks to piloting encountered by the entrepreneurial firms. Progress reports were collected in person from each entrepreneurial firm. I collected progress reports five times during the course of the study. From these data, I was able to construct an understanding of how entrepreneurial firms and customer organizations interacted to engage in the process of piloting an innovation.

Data analysis. I first compiled and reviewed my field notes (Emerson et al., 1995; Locke, 2002) by dyadic pair of customer organization and entrepreneurial firm. My notes revealed that to design and execute

pilots, customer organizations stood up formal committees of five to eight members that spanned levels, roles, and occupational groups. Although agreements to pilot entrepreneurs' innovations occurred before pilot committees were formed, the committees exerted significant power over shaping pilots and ensuring that they got off the ground. Informants shared that committees could "kill pilots." I became interested in how these seemingly powerful committees might shape the trajectory of entrepreneurs' technologies and what leverage, if any, entrepreneurs had in directing these pilots towards their own interests, if they veered off course. Analysis proceeded in five phases.

Phase 1: Mapping the piloting process. After ordering and reviewing my field notes, I realized that piloting seemed to involve three stages: scoping, integrating, and running and implementation. I arrived at this conclusion through several means. First, committees generally talked about the piloting process in three stages, with clear goals and milestones for each stage. At Red Hospital, members talked about picking the capabilities, a stage or step that involved determining what exactly a pilot would test. At Community Hospital, members talked about how before they ran pilots they had to "develop the pilot." I moved away from leveraging either of these more emic terms, as "developing or picking" had clarity in context but failed to communicate the essence of this step in general. Thus, I termed this stage "scoping," during which entrepreneurial firms and pilot committees determined which aspects of an innovation to test, and the number and types (e.g., doctors and nurses) of users that would participate in the pilot. I knew this stage had ended when committees mostly stopped discussing what a pilot might test and began developing materials and altering systems to support the pilot. This stage, the integrating stage, involved determining how entrepreneurs' innovations would interoperate with customers' technical systems and workflow and modifying an innovation or operational processes in response. MAMMO, for example, changed its data security protocol to comply with Red Hospital's requirements. When conversation about integration ended, committees moved on to execute the pilots and discuss their progress. I termed this the running and implementation stage, where customers actually utilized an innovation among selected test user groups, monitored results, and then attempted to implement these innovations, if pilots were successful. I validated

these steps among entrepreneurs, asking many of them, "Can you tell me about how the pilot process works?".

Phase 2: Categorizing committee preferences and building coalitions. Committee members held strong priors about AI. Informants were legitimately dogmatic about AI, sharing statements like "It will be the death of medicine if we use this ML stuff this way," but also "Machine learning will revolutionize medicine" and "This innovation will literally save the hospital." Consequently, I coded and categorized committee members' beliefs and sentiment about entrepreneurs' innovations. I grouped committee members beliefs about AI by theme and valence (positive, neutral, or negative). Although I had planned to code a range of responses from negative to neutral to positive, no one was neutral about AI. Thus, committee members fell into one of two categories: apocalyptic and evangelist. When an informant shared negative concerns regarding the use of an entrepreneurial firm's innovation, I coded this person as apocalyptic. I coded members as evangelistic when they expressed excitement for an innovation or highlighted beneficial features of an innovation. I then aggregated the views of each member by committee to understand if committees were dominated by evangelists or apocalyptics. In all 13 cases, committees were formed of members that expressed differences in sentiment. There were two formations: committees were majority apocalyptic or evangelist with at least two committee members holding an opposing view. Dominant coalitions formed around members sharing the majority view on AI.

Like Vuori and Huy (2016), I was concerned that a senior member's expressions of wariness or enthusiasm about machine learning might color the sentiment of the entire committee, even if most members held opposing views. I consequently coded the role and level of committee members in the majority and minority. Committees were composed of business leaders, such as heads of nursing, chief medical officers, administrators, IT representatives, and chief information or data security officers, whose groups or direct work might be affected by an entrepreneur's innovation, and frequently included users (line doctors, nurses, assistants, and administrators) as well. Hierarchy and status did not appear to play a role in preferences for or against machine learning. Because committee members were so vocal about AI in meetings, I parsed my observations of committee meetings from private interviews I had with informants to

understand if informants "preformed" in committee meetings, sharing views on AI that they thought others wanted to hear. Informants shared consistent views across all points of contact, or in other words, what was said in private was also said in public. For example, one informant shared in a meeting, "I am extremely concerned about using this technology to check really sensitive data. I mean where could the data go?" He also said privately, "There are serious privacy issues associated with using these types of technology. Once someone like an entrepreneur, who is totally motivated by getting their business going, has it, what will he do with it?"

Phase 3. Analyzing dominant coalition activities by phase. I next analyzed the actions dominant coalitions engaged in during each phase of the piloting process. During the scoping phase, dominant coalitions (e.g., an apocalyptic on an apocalyptic committee or evangelistic on an evangelistic committee) attempted to use claims of learning and generalizability to justify their decisions to narrow the scope of entrepreneurs' pilots. For example, when a nurse on a pilot committee at Community Hospital shared that they were going to limit the scope of an entrepreneur's pilot to ensure they could really learn how to use an innovation, I coded this as claiming learning. Dominant coalitions also controlled the flow of discussion in committee meetings through two actions: Folding and Temporally loading discussion. Folding discussion involved following up comments made by minority members in a committee meeting with a differing opinion or statement. When dominant coalitions Temporally loaded discussion, they either pulled forward conversation that usually occurred later in the scoping process or pushed topics back. For example, when committees engaged in discussion about integration with scoping decisions, I coded this activity as temporally loading conversation.

During the integrating stage, committees ensured that an innovation could interoperate with the necessary systems and processes used in each hospital and assigned responsibility for conducting that integration work. When committees took primary responsibility for integration, I coded this as claiming responsibility. For example, when a member of WHITE shared that the hospital was managing integration, and that there was very little to change about their innovation on their end at this point, I coded this as the hospital claiming responsibility. In contrast, some committees also disavowed responsibility. When the co-founder of MAMMO shared that the hospital was doing very little to help with integration, and that he

would need to do "a ton of work and design a whole new process for the hospital" to test their innovation, I coded this as the committee disavowing responsibility for integration.

During the running and implementation stage, committees and entrepreneurs executed pilots and tracked whether the pilot's results met, exceeded, or underwhelmed compared to a back-up or baseline method. I coded for the dominant coalitions' responses to the results of each pilot and whether they were contested. I also coded each pilot's outcome including whether it ran to completion and customer organizations agreed to continue using the innovation. I then assessed whether innovations were further implemented after pilots were completed.

Phase 4. Analyzing entrepreneurial actions and actions of committee members in the minority. I traced entrepreneurial firms' responses to actions taken by their counterpart committees in customer organizations. During scoping, entrepreneurial firms always adjusted the functional scope of pilots in response to committees' prompting. The direction of a pilot's functional scope could be shifted in one of two ways: towards testing more *basic* uses of innovations, such as diagnosing known and understood abnormalities, or towards testing more *advanced* uses, such as diagnosing "unknown" abnormalities or suggesting treatments not yet incorporated in doctors' and scientists' work. For example, when SKIN shifted their pilot to diagnosing abnormalities often detected by doctors, I coded this as shifting functional scope toward testing *basic* uses. I also analyzed how entrepreneurs responded to actions taken by pilot committees during the integration and run phases of piloting. For example, I assessed whether they shifted resources to support a pilot. I also analyzed how committee members in the minority responded to actions by the dominant coalition. I coded for whether they were silenced, spoke up, if they tried to argue with decisions, or attempted to bring in others across the organization for support.

I was surprised that the committee members in the minority were so vocal about AI, but more stayed in their responses to actions taken by the dominant coalition. Thus, I scrutinized what motivated these responses. Informants shared concerns about disrupting the providence of the pilot process, or looking like "bad citizens." Informants in general shared concerns about reputational risk if their behavior on a committee, or the committee itself, was viewed as unreliable. Members also shared desires to preserve the

committee process as in the future they might be given the chance to join a dominant coalition and see their goals achieved. Entrepreneurs and minority members were aware that dominant coalition members were manipulating pilots but did not stop their actions.

Phase 5. Explaining how and why pilots were manipulated to shape the design and implementation of AI. I compared what I observed with how the literature described coalition and committee interactions (March, 1994) and identified differences. First, the formation of coalitions did not relate to expertise, role, department affiliation, position within the organization, or skill (March, 1994; Salanick & Pfeffer, 1974; Young-Hyman, 2017; Truelove & Kellogg, 2016; Fligstein, 2001; Howard-Grenville, 2010). Instead, coalitions were formed around common preferences for AI. Second I compared the actions dominant committees took to assert control with those portrayed in the literature. Here, too, I found differences, as minority members did not try to draw out allies (Truelove & Kellogg, 2017), or partner with others outside the focal committee or organization (Davis & Thompson, 1984).

Thus, I developed a model to help explain the way power manifested in this setting and the role it played in shaping the design and use of innovations underpinned by AI. Dominant coalitions leveraged process manipulation - working within an accepted process but strategically engineering select steps or micro-decisions to achieve specific results – to assert power and achieve their objectives (to kill a pilot or ensure it succeeded). Dominant coalitions could not directly push their agenda, but instead manipulated pilots in the hopes that such tactics would help them achieve their ends.

ESTABLISHING COALITIONS AND ENGAGING IN PROCESS MANIPULATION

Committees kicked off the selection process with a meeting where members met, often for the first time. When asked to share initial ideas about how to select innovations, committee members were quick to pause such discussions and instead interject their personal views on AI. Right off the blocks, it became quite clear who was a true believer in AI and who thought "AI is like the Anti-Christ. It will ruin humanity and everything good about what we do," shared a nurse at Red Hospital. Preferences were never neutral. There was no hiding an opinion away about AI and members quickly recognized who shared their views and as a result, who was on their side. Interviews and observational data never indicated that members met privately

to coordinate their efforts. Yet, coalitions formed around common views—evangelistic (for AI) or apocalyptic (against AI). This bifurcation tangibly and quickly appeared. Within two meetings, members with like opinions came to sit next to each other: like in pews at a church wedding, apocalyptics on one side, evangelists on the others. Coalitions could have formed along different lines: "All the IT people together, the nurses versus the doc. You know typical stuff," explained Tom, a committee member. But they did not. Coalitions formed around strong shared views about AI, with a dominant coalition coalescing around members holding the majority view.

Despite having stridently bifurcated views on AI, never did members talk over each other, coerce agreement, leave those with opposing opinions out of meetings, or issue sell to change minds (Howard-Granville, 2010; Kellogg, 2019). This does not mean that dominant coalitions did not attempt to affirm their will on the piloting process and fight for their entrenched views. But battle gear was hidden rather than outwardly donned. Members in dominant coalitions quickly realized they had a numerical advantage over those in the minority. Yet, unlike in typical departmental settings, or fixed subunits, they could not just discount dissenting voices. John, the CIO of Blue Hospital, explained:

> Here, you can sort of ignore some dissent if you have the authority. But these committees are temporary and filled with people from different places in the organization. The purpose is super important and certainly elevates participants for a period of time, So, the process has to be unquestioned, and you have to be viewed as appropriate. Otherwise, you get an organization-wide reputation as either an idiot or an asshole. What would your department think about that?"

Members viewed roles on committees as true opportunities to shape the organization. At the same time, participation exposed members to broad reputational risk if committees went awry. Thus, "When you have the numbers, you have to play the game. Fight for your views but do so in a backhanded way," explained Tom.

On one hand, dominant coalitions attempted to maintain consensus to avoid looking like "assholes" or "idiots," coloring within the process to avoid disrupting the providence of the piloting process. On the other, they colored with muddied shades as they attempted to achieve their own ends—ensuring a pilot would work or effectively killing it. Entrepreneurs had difficulty reading and responding to this process, as it was unclear who they could align with to see their own agenda promoted. This study proceeds by explaining how pilots unfolded between committees and entrepreneurial firms by examining the practices dominant

coalitions used to assert power by pilot stage (scoping, integrating, running and implementation) and how entrepreneurs and members in the minority responded to these practices.

Scoping Stage

According to the literature, organizations scope pilots to (1) produce results that can generalize and scale across an organization, and (2) learn about an innovation's capabilities (Gans, 2023). The challenge is that these two goals are not often aligned. Striving for generalizability may limit the use cases or features an organization pilots. For example, an organization that tests an innovation with the user groups most inclined to adopt it might limit learning about how to implement the innovation with less enthusiastic groups. Thus, organizations have a lot of discretion in how they attempt to achieve these goals when designing pilots. On the surface, members of pilot committees shared this sentiment about discovering the best and most robust uses of an innovation. When asked initially, scoping was characterized by an IT executive as "getting the experiment right" and by a nurse as "seeing how much we can learn." Although committees referenced and claimed learning and generalizability as goals, pilots were rarely scoped to accomplish them. For instance, when asked about learning from pilots, Tom explained: "That's what they all say [name]. That's sort of naïve. It, [piloting], is really about being clear on what you think will come to pass and making that your reality." I relay the experiences of Red Hospital and Blue Hospital working with INFECTION to scope a pilot. Tables 4a, provides additional representative data of the actions taken by dominant coalitions and responses of entrepreneurs and members in the minority.

Insert Table 4a

Apocalyptic committees. Apocalyptic committees were open about their views on AI. At Blue Hospital, for example, committee members debated the value of investing time and money in AI, which most did not believe to be a viable substitute for human decision-making. They discussed quite openly whether they should "kill the pilot," shared Steve, CIO of INFECTION. Blue Hospital, he said, "had a million general questions about the efficacy of using machine learning, but hardly any questions were about how we specifically use it." Steve fully expected the committee to terminate the pilot. But those in power did not. Instead, the apocalyptic committee engaged in two practices to assert their power and manipulate the process: claiming learning and folding the discussion. These practices helped the dominant coalition narrow the scope of INFECTION's pilot, to test the most advanced capabilities underpinning their innovation, capabilities that were previously untested and would be very difficult to prove out.

Claiming learning: Members in the dominant coalition engaged in a practice entitled claiming learning, which involved designing a pilot to generate the perception that it would maximize learning. In this case, apocalyptics wanted to kill INFECTION's pilot, but could not forcefully do so for fear that it would make them look like "assholes" or create questions about the rigor of the piloting process. Committee members in the minority who viewed AI favorably voiced their opinions as scoping progressed. Denise, a coordinator "straight out of college" and an avowed evangelist, was encouraged to share her views. "I may not have the main views here, but I believe in the power of machine learning and am going to let the rest of the committee know," she said. To appease members enthusiastic about the potential of machine learning, the committee agreed to carry on with the pilot. The head of nursing explained the rationale: "We are a consensus-oriented group," and "We like to find a solution that everyone can live with. [Blue Hospital CIO] John really liked INFECTION, and so we carved out a way."

However, the committee did narrow the pilot's scope. INFECTION's innovation could diagnose bacterial strains likely to be resistant to typical antibiotic treatments as well as suggest treatment for such infections. The committee focused the pilot on treatment, a capability that was thought to be quite advanced. "INFECTION is the only company I know of that says it can do this," explained John. The head of nursing explained in a committee meeting that the scope would "ensure they would gain significant learning from the pilot." However, privately, the head of nursing shared that the committee had narrowed the scope of the pilot to effectively kill it.

> We picked this focus because no one person really owns it. Doctors yes, of course, treat patients, but we don't really have a set process for treating resistance. We do it by trial and error today. So we at least had a hard time figuring out what … the pilot could look like. And to be honest, in my experience, if there is no real owner and you can't visualize how it will work, it will likely die anyway. But if [INFECTION] figures it out, well, it will provide something valuable, I am sure.

Directing the scope of the pilot towards advanced activities without clear support processes was a way for apocalyptic members of the committee to appease their colleagues, while setting the pilot up to "likely die anyway." The committee used the veneer of learning, a key goal of piloting, to justify a shift in scope that was really about passively killing the pilot.

Entrepreneurial response. This rationale was completely apparent, though never stated outwardly to the team at INFECTION. "This was the compromise," reflected Christina, INFECTION's CEO. "They agreed on treatment. I swear they did this because it was a hard test for our application." But the team was not upset

with the change in scope. Rather, the scope change was seen as a massive opportunity to test their innovation's advanced capabilities, a test needed to gain further market traction. "This is quite the opportunity," explained Christina, "but wow, this is pretty advanced stuff. We have never tested just these capabilities live." The committee pushed the pilot into new territory.

Entrepreneurs like Christina were usually willing to "re-litigate their innovations" in response to shifts in pilot scope. The challenge for Christina and the INFECTION team was how to get the information necessary to make their pilot work given the new, and previously never tested, scope. "Testing something like this is going to be really hard. We don't have all the information about how to do this. We need help. But I'm not even sure who to ask. The people advocating for this are the haters. The others, they seem to lack power." Seeking information to improve the likelihood of the innovation meeting the requirements of a more advanced pilot, Christina asked the committee questions like "How does the treatment piece work today?" and "What systems will be important for us to hook into?" But these questions were brushed off as the committee attempted to rapidly conclude conversations about the pilot's scope. Not wanting to sow doubt about the innovation's capability or ask too many questions for fear the pilot would be canceled outright, Christina stopped searching for information.

Minority response. Those in the minority were also aware of the dominant group's tactics. As Denise, the coordinator, stated, "We know what they are doing. They are rigging this and setting up the pilot for failure." But members of the minority were reluctant to mention their concerns during committee meetings or to others in the organization for fear that doing so would affect the validity of the pilot process and change the rules for structuring pilot committees in the future. When asked why not share her concern, Denise explained:

> Denise: [Authors name], why should I say something?
>
> Author: Well, I mean, you did kinda mention that you think they are rigging the pilot.
>
> Denise: I don't think it. I know it. But that doesn't change anything.
>
> Author: Why not?
>
> Denise: We need to keep the pilot intact. We were given this responsibility and well, it would look bad if we started raising red flags without even seeing this thing scoped and run.
>
> Author: Why would it look bad?
>
> Denise: Because the hospital is committed to this process. And honestly, I don't want the process to change. See, in the future, things might go a different way. If I do a good job, maybe I'll be

> on a committee again and my opinion might be the majority opinion. In that case, I want to advance my ideas.

Members in the minority explained how it was in their best interests to agree on the narrowed scope as this would keep "the pilot intact" and preserve opportunities to be in the dominant coalition in the future.

Folding. Members in dominant coalitions, regardless of being apocalyptic or evangelistic, also engaged in a practice entitled folding. Folding involved stacking the discussion so that comments made by members in the minority coalition were sandwiched between comments made by members in the dominant coalition, with last words by members in the dominant coalition. For example, the team discussed sharing data and systems information early in the scoping process. Members of the minority wanted to share this information early on as they believed it might be an important input to scoping. "If it would be difficult to share certain sets of data that were needed to run the pilot, the pilot would not run very well," explained John. Members of the dominant coalition were loath to have this discussion as doing so might trigger the committee to change the scope of the pilot in a way that improved its chance for success. John (an evangelist) commented:

> Maybe we should think about what systems are involved in running this pilot. For example, do we need to share access with EPIC, or our internal disease control data?
>
> Lynn, a member of the innovation team and an apocalyptic responded:
> We should save these conversation for later. When we get into the details of next phase.
>
> Denise retorted:
> But this is sort of important to settle on now. Right, we don't want to scope something that can't be done.
>
> Mike parried back:
> Well, we will discover if this can be done by running the pilot. We want to make sure we set up the pilot to learn the most possible.

Because the minority coalition had fewer members, they also had fewer voices to continue the argument. Rarely did committee members voice opinions about a discussion point multiple times, as that, John described, "was in poor taste and looks like you are being pushy." Thus, folding allowed the dominant coalition to keep control of conversation, even without an appointed leader.

Entrepreneurs' response. Entrepreneurs found this practice frustrating, as it limited what they saw as productive conversations about how to move their pilots successfully forward. Yet they too were careful to not interject frequently and be seen as siding with the minority. As Christina explained, "When you see the clients going back and forth, you want to just let them go through it. You can try to interject here and there but you do not want to back up the wrong person. This pilot is so precarious. So, we have to just let it play

out and not make a mess of it by piling on with [CIO of Blue Hospital] and being in the minority." While entrepreneurs could have engaged in these conversations, bolstering the minority, they mostly kept quiet.

Minority response. Members in the minority coalition were not surprised that the dominant coalition attempted to control the flow of conversation as doing so helped them assert and maintain power in the absence of formal authority. "That is what I would do," explained John. "It is an interesting situation, you want to control things, but you don't have the stated ability to, so you control the way we engage." However, John and others in the minority coalition were unsettled by the way Christina and the INFECTION team kept silent when they attempted to help them gain useful information about systems and data. Denise shared, "Why didn't they [INFECTION] speak up? We were trying to get info for them. I'm not going to put my neck out for others if they are going to advocate for themselves." Denise explained how despite wanting to see the pilot work, she was unlikely to go out of her way for INFECTION. She would help if they directly asked but would not push to see them gain an advantage. In this way, the quiet response by entrepreneurs to folding chilled the minority's interest in supporting their efforts to successfully run a pilot. The pilot's scope remained fixed on testing advanced capabilities.

Evangelistic committees. INFECTION also engaged with Red Hospital where the team was enthusiastic about AI. Evangelist committees also utilized a set of practices to manipulate the scope of pilots and manage the flow of conversation: claiming generalizability and temporally loading conversation. However, these actions led to different ends in terms of how pilots were scoped. Red Hospital's pilot originally planned to test multiple aspects of INFECTION's innovation, but the committee narrowed its scope to test the innovation's basic capabilities rather than more advanced capabilities. I review how these committees utilized claiming generalizability and temporally loading conversation to scope pilots.

Claiming generalizability. Just like when committees were apocalyptic, evangelist committees listened to the minority. For example, when apocalyptics, now in the minority, voiced concerns about running a pilot, the dominant coalition responded and narrowed the scope of their pilot. In contrast to the apocalyptic committee at Blue Hospital, the committee at Red Hospital shifted the pilot scope to areas that the application could easily accomplish, towards basic capabilities. Red Hospital agreed to pilot INFECTION's diagnostic capabilities, with emphasis on diagnosing resistant strains that were well known to doctors. Explained Mark, the director of innovation:

> We wanted to be sensitive to members of the committee who were worried that some of the more innovative aspects of INFECTION's application might not work. So, we directed the pilot at an area that was well documented, pretty discrete, and that few people could argue over so it would be easy for us to show the pilot worked.

The committee validated this shift by making claims that narrowing the pilot scope on basic capabilities would help show how INFECTION's innovation could scale or "generalize" across relevant groups in the organization. "They said it would be easy to repeat it," said Christina at INFECTION of the committee's public justification to narrow the scope of the pilot on basic capabilities. However, all parties knew that claiming generalizability as the rationale for narrowing the scope was a red herring; the scope was narrowed to make the pilot more likely to succeed.

Entrepreneurs' response. INFECTION was disappointed that Red Hospital did not offer the opportunity to test the aspects of their innovation that were differentiating. INFECTION's CTO Steve explained why: "Lots of people can do this. Lots of people can't do the treatment stuff. Treatment gets us a better valuation, a better chance on making it." Entrepreneurs lost opportunities to test advanced capabilities and also access to data that might help entrepreneurs refine these applications. "If we don't get a live environment to test our best stuff with, we can't robustly train our algo," shared Steve. Despite their disappointment, entrepreneurs also explained how they had little recourse to broaden the scope of their pilots. They were not sure who they needed to convince within the customer's organization. Steve expounded, "It's weird. The CTO [not on the committee] won't talk to us about it [changing the scope of the pilot]. The committee is it. And on the committee, the guys that were for us are also limiting us. We can't really figure it out." Evangelists, with their enthusiasm for AI, were on paper entrepreneurs' natural advocates. Yet entrepreneurs did not view evangelists as fighting for their interests, as they limited the scope of the pilot. Entrepreneurs could not ask apocalyptics for help as "they rarely want to talk with us," explained Christina.

Minority response. Apocalyptics on the committee were neutral on the change in scope. Jane, a doctor, shared, "Oh, they are setting it up so that it works. That is obvious. But this might not work to their advantage in the long run." Minority members refrained from "doing much" to change the scope. First, they worried about creating questions about the validity of the pilots and pilot process, which was not to their advantage, as they too believed in preserving the process to improve their chances of seeing their interests fulfilled in the future. Second, even if the pilot worked, they were not sure how the broader organization

would use or accept entrepreneurs' innovations. "Of course they will believe the results of the pilot. But will people actually use this thing, if it is basic and doesn't change much?" asked Jane.

Temporally loading conversation. Dominant coalitions also engaged in a practice entitled temporal loading. Temporal loading involved moving discussions topics earlier or later in the course of the piloting process to maintain control and improve a pilot's chances of either succeeding or failing. As Red Hospital set the scope of INFECTION's pilot, the committee requested that entrepreneurs share technical details and requirements for integrating their innovation with existing systems and process. This task usually would occur later, during the integration phase of a pilot, but "We wanted to understand everything," explained Maria, CIO of Red Hospital. "The code, what type of data they were drawing on. We wanted to make sure their data sources were robust to diverse populations. We wanted to see it all." Evangelist committees continued to adjust the scope of pilots, removing features or capabilities from the pilot that might prove difficult to integrate. For example, INFECTION went back and forth with the team at Red Hospital to understand critical features to test. Features that used data the Hospital might have difficulty supporting were removed from the pilot. Maria explained the rationale for temporally loading:

> We wanted to make sure that issues did not come up later when we tried to connect the innovation that could create reasons for the naysayers to try to cancel the pilot and say no to the innovation. This pilot has to work. We need machine learning. So we will do what it takes.

Moving discussion of integration earlier in the process helped evangelists control the flow of discussion and prevent "naysayers" from later interjecting ideas or doubts that could upend the pilot, much after the scope was settled.

Entrepreneurs' response. Frustration ensued from entrepreneurs who spent significant time hammering out the details with committees but saw the extent of their pilots continuously shrink. "What exactly are we even testing?" Steve relayed. "How are we going to get anything out of this?" complained Christina. Not sharing these details was also not an option, as doing so might prevent the pilot from running altogether. Entrepreneurs felt they had little influence over the continued narrowing of their pilots.

Minority coalitions' response. Members in the minority attempted to push conversation about integration back to its perceived rightful place, after the scoping was complete. For example, Jane explained, "I tried a few times to stop the conversation about integration. It is better to cleanly set the scope and then see if we can make that scope work. But others were like no. No good designing a pilot

that won't even work." Jane was skeptical of the dominant coalition's motives. "I mean, I think they are doing it because they truly are worried about designing a pilot that doesn't work. But it also is sort of convenient. Learn about a tech spec, narrow the scope, learn about something else, narrow the scope." Jane started to wonder what the team was even going to test. "I mean, what is the use case now? It is super benign." While Jane was annoyed that the majority used control over discussion to improve the odds of the pilot working, observational and interview data suggests that she never raised this concern to any member on the committee.

Integrating Stage

During integration, entrepreneurs attempted to get their innovations to interoperate with customer organizations' systems and processes. Among critical questions debated and resolved during integration were what aspects of a customer's operating environment required adjustment to accommodate a pilot, what aspects of an innovation needed to be adapted to the customer organization, and who was responsible for these adaptations? Literature suggests that integration requires mutual participation (Karp & O'Mahony, 2024). However, little collaboration occurred between committees and entrepreneurs. Apocalyptic committees generally assigned responsibility for integration to entrepreneurs, leaving entrepreneurs the task of figuring out how to make their innovations interoperate. Evangelistic committees, in contrast, assumed most of the tasks associated with integration. Committees justified their decision to assume or disavow responsibility for integration by claiming their choices generated learning or helped to ensure a pilot's results could generalize across relevant groups in their organization. Table 4b relays the activities dominant coalitions, entrepreneurs and miniority members undertook as they engaged in integration across both committee types, compared with what is suggested by existing research. I detail the experiences of Red Hospital and Blue Hospital working with SKIN.

Insert Table 4b.

Apocalyptic committees. When SKIN encountered a committee mostly wary of AI technologies at Red Hospital, they faced a score of time-consuming challenges getting their pilot integrated into Red Hospital's work environment. The pilot was scoped to test advanced capabilities related to diagnosing

cancerous skin abnormalities difficult for humans to physically observe. The committee asked SKIN a plethora of questions about integration but rarely provided help gaining answers.

Disavowing responsibility. Apocalyptic committees peppered entrepreneurs with questions such as: "How do we double-check [an innovation]?" and "When should double checking take place?" and "How can we update our systems with your data?" but disavowed responsibility for figuring out the answers. Committees were unwilling to directly work through these questions with entrepreneurs. "It is [SKIN's] innovation," remarked Bethany, a coordinator on the committee. "They agreed to the pilot scope and now have to figure out how to make it work here. How can we say that [name of application] can scale across the organization if [the leadership team of SKIN] cannot get this pilot to work?" Apocalyptic committees believed "making it work" to be the entrepreneurs' responsibility and disavowed ownership, even joint ownership, of integration. Explained a chief of medicine, "If the pilot can't work here, that's not on us. It's on [the entrepreneurs]." The dominant coalition claimed that disavowing responsibility would ensure that the results of the pilot could generalize more broadly across the organization. But this also made an impossible task even more impossible.

Entrepreneurs' response. To satisfy the committee's concerns and integrate their innovations, SKIN took on a heavy workload. "There was not a clear plan in place for integration," SKIN's CEO Tina bemoaned. "And if we wanted the pilot to work, we were going to put a lot of effort into it." SKIN devoted a full-time person to the pilot (a non-trivial action given that it was a firm of few people). SKIN added a new data set to their database, based on concerns that their innovation did not represent diverse populations. They customized their innovation to meet the unique security requirements of Red Hospital. "Because we were doing something new," explained Tina, "they had this crazy thing they wanted us to do, too, to absolutely ensure data privacy. … We built it in."

SKIN not only revised technical aspects of their innovation but also developed new protocols and procedures for hospital staff to employ during the pilot. Because it was difficult for doctors and nurses to evaluate the malignancies that SKIN's application surfaced, using existing tools, it was challenging to double-check the results produced by the application. "We had to create all these workflows and write protocols for who checks the output of the machine learning algo," recalled Tina. "First the algo checks itself, then a nurse checks it and forwards it to an oncologist. It was complicated." To get it going, SKIN had to "pilot the pilot."

> We built into our application three or four different routing paths between different specialists and nurses, three different workflows, and I still don't know if it is going to work. To do [this pilot], we had to stop working on some of our other efforts. … We had to stop sourcing some new business leads.

Surprisingly, all the entrepreneurs in the sample were willing to dedicate similar effort to pilots testing advanced capabilities. All reallocated resources or stopped pursuing other leads, and even other pilots, to bolster their efforts. When asked why, a representative of WHITE explained, "This is a huge opportunity, to test out these capabilities." Proving such capabilities could create differentiation in the market. Tina recalled, "It was a total slog. Horrible. But they also really pushed us, they pushed for [the application] to be better." Apocalyptic committees pushed entrepreneurs to improve the most innovative aspects of their innovation.

Minority response. Evangelistic members in the minority were reluctant to help out. In part, they believed that "aggravating the majority is never a good idea." One evangelistic committee member said, "We can work around the edges . . . but the entrepreneurs have to get their applications up and running." Further, these members remembered having gone to bat for entrepreneurs in the past, during the scoping phase. Entrepreneurs neglected to back them up. Tina shared, "I think committee members felt a bit burned when they spoke up in prior meetings and we didn't say anything. These guys weren't going to lift a finger in our defense." While the minority understood that holding back support might impact entrepreneurs' pilots, they were not "going to put their goodwill on the line," reflected Tina.

Evangelistic committees. The integration stage consumed less effort and time when entrepreneurial firms engaged with evangelistic committees that focused the scope of their pilots on more basic capabilities. Entrepreneurs at this stage made minimal changes to their innovations and customers made few changes to their systems or processes. As explained by Tina, who also worked with a team at Blue Hospital whose members were generally enthusiastic about machine learning, "Much of the hard work had been done." During the scoping phase, SKIN had augmented and customized their innovation to appease Blue Hospital's questions about how their innovation worked. Consequently, most issues were "taken care of already and it felt like our product just fit in."

Asserting responsibility. A key difference among pilots managed by apocalyptic committees and evangelistic committees was that evangelistic committees took over responsibility of the pilots. As the lead nurse from Blue Hospital explained, she "took over a few of the steps" required to run the pilot. As one CIO on an evangelistic committee explained, "We wanted to show we were learning. If we can do it, that is strong

proof, and helps eliminate complaints later, like, oh, the entrepreneurs are the only ones who can do it, or, oh, that thing needs a lot of support." Whereas apocalyptic committees drew on notions of generalizability to disavow responsibility, evangelistic committees embraced learning as a reason to assert responsibility, as running the pilots would teach the organization how to use the application—though only a small sliver of it.

Entrepreneurs' response. From the perspective of entrepreneurs, when committees asserted responsibility, it limited their learning and opportunity to build new capabilities. Because Blue Hospital was "responsible," SKIN was less frequently invited to participate in conversations about integration and rarely consulted when changes or adaptations were made on the ground, even to their own innovation. Blue Hospital made adaptions to SKIN's innovation, building workarounds and customizations (Tyre & Orlikowsi, 1994) to improve operability, but never mentioned these changes to SKIN. "We were limited in what we could see about how to improve [the application]," explained SKIN's CIO. While many of the changes made by Blue Hospital were minor and content-specific, knowledge of these changes may have provided SKIN with useful information.

Minority response. Apocalyptic members of the committee were annoyed that those in the dominant coalition had taken on additional workload that they would have to support. "I really don't want to work on fixing [SKIN's] application. It will take a ton of my own time for something I am generally against," shared Susan, an administrator on the committee. Yet members in the minority did not push back and accomplished the tasks assigned to them to the best of their abilities. Susan shared how she did not want to be found responsible for "breaking the pilot" and thus would do her best, despite her frustration. Time and time again, pilot committee members were careful not to disturb the providence of the piloting process by raising concerns about the process inside the committee or with others outside the committee.

Running and Implementation Stage

The running and implementation stage involved releasing innovations into customer work environments, where assigned teams used those innovations in practice. And then, if pilots worked, implementing an innovation with additional user groups. When organizations ran a pilot, they tracked and interpreted the results of the pilot compared with a controlled test or with performance of the prior process. Scholars observe that tensions can arise at the end of a pilot as organizational members interpret results (Adner, 2020). Outcomes may be discounted in the face of doubts about the validity of the piloting process (e.g., Lind et al., 1993). In no instance were the rigor of the process or pilot outcomes debated.

When pilots were scoped around basic capabilities, few adjustments were made, and most pilots were successfully completed (70%). Yet, pilots did not always make attempts to implement an innovation easier. Of these pilots, 60% went on to be further implemented. When pilots were scoped around advanced capabilities, only approximately 30% were successful (2 of 6). However, both of these pilots were later implemented. SKIN was fully rolled out at Red Hospital. Table 4c provides representative data explaining how the running and implementation stage unfolded. I explain the experiences of Red Hospital working with SKIN and Community Hospital working with SENSOR.

Insert Table 4c

Apocalyptic committees. When committees were apocalyptic, pilots did not run smoothly. Despite developing new workflows and procedures during the integration stage, when entrepreneurs focused on testing advanced, radical capabilities, pilots usually fell short. Users often struggled to understand their roles and navigate new processes. Nevertheless, sometimes they were successful, but entrepreneurs had to work it out mainly on their own.

Denying support but supporting results. Committees took on a limited role running pilots. Denying help or support for running a pilot was a useful way to once again attempt to make a pilot fail. Apocalyptics claimed that taking on a limited role running pilots was important because it would help support greater belief in pilot results. If committees intervened, others might question if an innovation would work when fully rolled out. Maria of Red Hospital explained:

> We need people to ultimately agree with the pilots, or we will have to go through a vetting process over and over again with each department or group we try to roll out with. It is very easy for people to make an excuse of the pilot, and say, oh oh we don't believe the results. They will never actually use the innovation.

Of course, it was sensible to support the results of a pilot, when the pilot failed, which was often in the case for apocalyptic committees. Yet, these committees also supported results when pilots succeeded. "Even though I hate it," conceded one committee member, "I can't kill it at this point. They showed it can actually work." Since the apocalyptics supported the results of these pilots, evangelists could more easily make a case that SKIN's innovation should be implemented within the organization. Maria, CIO of Red Hospital, explained that it was difficult for committee members to speak out negatively against SKIN's innovation, because the pilot worked. "Everyone got in line and spoke about how it was surprising it worked and wow, we should be using it all over if we can."

Entrepreneurs' response. To make their pilot work, the team at SKIN had to rewrite their code and help Red Hospital develop entirely new processes and procedures to support the pilot. "We made it work, but barely," recalled Tina. "I am not sure we could do this kind of intensive support for every pilot. We wouldn't survive." Apocalyptic committees pushed entrepreneurs to improve how their applications could work in practice, by developing new protocols, processes, workflows, and technical functionalities. "If they didn't hate machine learning so much and as a result didn't push us so hard," acknowledged Tina, "we would have never rebuilt aspects of our application, aspects that were actually critical to delivering new capabilities in the health system." Committees with strong negative priors about AI triggered entrepreneurs to test and, in some cases, prove out advanced uses for their applications. All the entrepreneurial firms put more effort into piloting with apocalyptic committees, committing more resources in the form of time and personnel, than when piloting with evangelistic committees.

When pilots were unsuccessful, which was usually the case, entrepreneurs walked away with valuable information, but conceded that they could not put this type of effort in all of the time. Entrepreneurs talked about being caught between a "rock and hard place" as these "advanced" pilots were necessary but utilized a lot of resources. As the CEO of SENSOR shared, "We learned a ton even if the pilot did not work. Plus we got to train our algo on their data sets. But we can't do this all the time at no success." These pilots always created opportunities to learn but when they failed did not help generate differentiation. "There is a point in which we sort of learn enough and need to execute," stated the SENSOR CEO.

Minority response. In the case of SKIN, members of the minority were thrilled that the pilot worked. But members of the minority were aware that the SKIN outcome was not usual. "SKIN worked really hard, and we got kind of lucky. This probably won't happen again," explained a committee member. When pilots were successful, minority members of the committee became more vocal in their support of an innovation. As the CEO of SKIN explained, "After the pilot, things were different. People really talked us up with others outside the committee. I think this helped us get a deeper footing in the company." With the results of the pilot in hand, members in the minority felt more secure getting behind SKIN, and ensuring their preferences were acted upon across the organization.

Evangelistic committees. Entrepreneurial firms that engaged with evangelistic committees focused on piloting more incremental aspects of their innovations. Approximately 70% of these pilots were successfully completed. The focus on incremental capabilities simplified the piloting process by limiting the

need to develop new protocols, processes, and tasks. When basic pilots were completed successfully, organizations paid to roll out these innovations more broadly throughout their organizations. However, only 60% (3 of 5) were actually used and more broadly diffused. Maria explained that even though the pilots showed that these innovations were superior to existing ways of works, users did not perceive them as an improvement in practice. "People felt they would have to still learn how to do something that seemed just a bit better. And then, interestingly, people wondered if there would be better future iterations, that they would again have to relearn how to do everything for." This was exactly what entrepreneurial firms feared when limited from showing how their innovations could compete at the edge of what was possible—that the door would be left open for others to show more advanced capabilities.

Delimiting support of the results. Dominant coalitions supported the results of the pilot but just the results and not the expanded use of the innovation for other more advanced purposes. For example, SENSOR's pilot with Community Hospital focused on monitoring the mental health of elderly patients admitted to the hospital. Nurses were alerted to assess existing assigned patients (already in their wards) and, if necessary, involve a doctor to more fully determine if a patient was suffering from an undiagnosed mental health challenge. SENSOR's device neither added additional patients to nurses' workloads nor altered the work arrangements of doctors, nurses, or other occupational groups within the hospital. The pilot worked well, and the dominant coalition of evangelists were excited about implementing the innovation within their organization. Yet, they were reluctant to push to run more pilots testing other aspects of SENSOR's device, or expand their use of SENSOR's device, beyond what was tested in the pilot. The CEO of SENSOR explained, "We asked if we could run another pilot and they were like no. We just spent all these resources on the first pilot. Let's see how [SENSOR's device] works and maybe next year we can do something more." With all the resources just expended on a pilot, hospitals were reluctant to run additional trials.

Entrepreneurs' response. SENSOR showed how the basic capability of its application could work in practice, but was unable to demonstrate more advanced capabilities. "The pilot with Community [Hospital] was a win; I don't want to take away anything from that," acknowledged SENSOR's CEO. "But we sort of have shown that we can do this stuff before. We didn't get a chance to push the limit. And, well, we need to push the limit if we want to beat our competitors and really show customers and investors how valuable we are." Community Hospital paid to roll out SENSOR's device with three groups within the hospital. Yet, according to the CEO of SENSOR, these groups did not find their device more useful than what was already

in place: "Nurses were like, we already have good altering systems, I don't think this is that much better." Community Hospital did not order any additional devices. The SENSOR team was frustrated that they did not get the chance to show advanced capabilities or fully penetrate Community Hospital.

Minority response. Members in the minority were disappointed when pilots worked successfully but supported the results of these pilots. These committee members also understood that because the pilots only tested basic capabilities, that basic capabilities were what the organization would roll out. Other organization members might view these capabilities as highly incremental, and not worth using over existing methods of working. According to SENSOR's CEO, the committee members that were fearful of AI "insisted on rolling our innovation out exactly how it was piloted. Because that was what worked and was proven."

Outcomes

Table 5 shows the outcomes of pilots in terms of whether pilots were deemed successful and if innovations were subsequently implemented within the customer organizations that piloted the innovation. Across all pilots, evangelistic committees increased the chances that an entrepreneurs' innovation was successfully piloted compared with pilots designed by apocalyptic committees. In total, more innovations piloted by evangelistic committees were implemented and used within organizations than innovations piloted by apocalyptics (3 to 2 ratio). However, given a successful pilot, innovations tested by apocalyptic committees were more likely to be implemented within customer organizations. Evangelistic committees increased the total number of innovations organizations were willing to implement. Apocalyptic committees limited the number of innovations introduced to their organizations by designing pilots that mostly failed, but increased the likelihood that innovations would be implemented if a pilot succeeded. In this way, both types of committees achieved their goals. Evangelist committees enabled more AI innovations to be introduced and implemented; apocalyptic committees ensured fewer innovations were implemented and used. Taken together, no committee tested both basic and advanced uses, or attempted to find some middle ground. This suggests that well before implementation, dominant coalitions, by engaging in process manipulation, filtered the types of innovations introduced into organizations and placed guardrails on how those innovations were later used to achieve a vision in line with their preferences on AI. This gating process is often neglected by research on implementation, but important to acknowledge. All of the advanced instantiations of innovations were implemented; the basic instantiations were not.

HOW PROCESS MANIPULATION SHAPES THE USE AND DESIGN OF AI

To better understand the mechanisms that might explain the observed results, I investigated a number of theoretically grounded explanations. For example, scholars frequently suggest that entrepreneurial firms with more abundant capital endowments are generally able to engage in more risky endeavors (Agarwal, Gans, & Stern, 2021). The present study did not find an association between funding or revenue level and the innovative directionality of a pilot or pilot completion. Firms that had accrued more funding or generated more revenue at the start of the study were not more likely to focus on advanced capabilities or complete their pilots. Similarly, founder gender, proximity to capital sources, and market segment focus did not seem to influence whether firms completed pilots. I next explored whether particular organizations or entrepreneurs might simply be ineffective at running pilots by reexamining each dyad to assess whether either side was more or less likely to complete a pilot. All of the customers and entrepreneurs in the study completed at least one pilot. Lastly, I examined whether the power structure embedded in customer committees influenced the observed results. Extant scholarship shows that the expressed opinions of those in senior roles frequently shapes the group's decisions, as others may be fearful of explicating a different opinion or may view the senior member as an expert or authority (Thomas, 1994; Edmondson, 2001; Vuori & Huy, 2015). I evaluated whether more senior members of committees (determined by a member's title) had outsized influence on how pilots were scoped and evolved. I found dissent to be expressed by junior and mid-level as well as senior members of committees. One informant shared that senior committee members were not "really that influential. Especially because this is about trying something new, piloting. Sometimes they have very specific views, but rarely act as the only authority."

None of these alternative explanations could account for the results I observed. Thus, I developed a grounded theoretical model that helps explain how organizations select novel innovations. The model, as shown in Figure 1, details how committees engage in process manipulation to shape the use and design of AI. When committees have strong priors about technologies such as AI, they form coalitions around those priors. Under these conditions, coalitions with a numerical or majority advantage (the most members whose priors align) become the dominant coalition. Dominant coalitions will control decision-making, exerting power to determine the design of a pilot and what is tested. However, the dominant coalition cannot directly act upon their interests by overruling dissenting members or dismissing them from the committee. Because appointments on committees are temporary, but important, members will worry about reputational risk if

identified as personally inappropriate or if the pilot process is deemed unreliable and dubious by others in the organization. Thus, dominant coalitions must carefully manipulate the pilot process, by folding and temporally loading conversation, narrowing the scope of pilots, and disavowing or claiming responsibility for pilot design and execution, to achieve their goals. Committees appear to run a scientific process on the surface but manipulate decisions within the process in ways that are beneficial to the dominant coalition. The actions of dominant coalitions will waylay entrepreneurs from engaging and working with natural allies or organizational "champions" during the piloting process.

INSERT FIGURE 1 HERE

During the scoping phase of piloting, committee members will agree to run a pilots, but those with majority views will narrow the scope of pilots towards more advanced and challenging or basic and simple tasks. Committees largely wary or skeptical of AI will scope a pilot to test more advanced features or capabilities in hopes that the pilot will fail. Committee members excited and enthusiastic about an innovation will scope a pilot in the opposite direction, to test basic features or capabilities in hopes that the pilot will succeed. Dominant coalitions will do this by making claims that such actions help satisfy more scientific goals of generating learning or generalizability. However, they will also use practices like folding and temporally loading conversation to help restrict the flow of information provided to entrepreneurs and other dissenting members of the committee to control discussion and the scope of pilots. As the model also shows, to maintain the facade of a well-run process, committees will claim that their decisions to shift scope either improve robust learning of an innovation's capabilities or help engender results that can be generalized, but these adjustments to scope really help dominant coalitions achieve their goals.

Entrepreneurs will find it difficult to determine who might help them achieve their own agenda, which is to conduct a robust pilot so that they can adapt and revise their innovation in ways that create competitive advantage. When committees are apocalyptic, entrepreneurs will temper the tone and quantity of questions asked in meetings in fear that pushing too much will ruin chances for a pilot. This will prevent entrepreneurs from advocating for themselves and getting information that might advance their innovations. When committees are mostly evangelistic, entrepreneurs will still temper their commentary in meetings, but for different reasons. Entrepreneurs will find it difficult to ask evangelists to expand the scope of pilots, given it was their idea to limit them to basic capabilities in the first place. Committee members in the minority will also limit their own pushback for fear it might make them look inappropriate. These committee

members attempt to preserve the perception that the piloting process is pristine as they desire future opportunities to manipulate pilots if on the dominant side.

Throughout integration, pilot committees will continue to make claims of generalizability or learning to justify their decisions to disavow or claim responsibility for a pilot. Committees that are wary of an innovation or technology will typically disavow responsibility for integration to diminish the changes a pilot will work. These committees will use claims of generalizability to justify their decision. Conversely, evangelistic committees use calls to learn as justification for claiming responsibility of a pilot. Evangelistic committees will claim responsibility for integration to ensure that a pilot successfully runs.

When committees disavow responsibility for integration, entrepreneurs will struggle to create new processes or modify existing organizational procedures to ensure that their innovations interoperate with customers' work environments. Though challenging and resource-intensive, these situations allow entrepreneurs to learn new information about customers' work environments and develop new features and technical capabilities for their innovations. Committees that claim responsibility manage the work of integration primarily on their own. This will put less stress on entrepreneurs' resources, but limits entrepreneurs' learning about how their innovations can work in situ. Committee members in the minority will be frustrated by their inability to help entrepreneurs or the expectations that they will take on the responsibility of integrating entrepreneurs' innovations. In both cases, members of the minority will not share their concerns within or outside of committee meetings, as they do not want to degrade the piloting process.

During the running and implementation stage, organizations will deploy pilots and measure their outcome. When pilots test an innovation's advanced capabilities they will generally fail. When pilots test basic capabilities, they will generally succeed. Committees will accept pilot results on the basis that the process was legitimate, but only the results produced by the pilots. When pilots testing advanced capabilities are successful, the innovations are implemented within organizations. Innovations piloted for basic capabilities are less likely to be implemented compared with innovations tested for advanced capabilities. However, because more innovations are successfully piloted when pilots test basic capabilities, more of these innovations will be implemented in total compared to innovations focused on advanced uses.

Throughout the piloting process, entrepreneurs are afforded little room to alter the outcomes or contest the validity of pilots. However, when a pilot tests an innovation's advanced capabilities successfully, it

creates an opportunity for entrepreneurs to claim that their innovations are differentiated, offering a potential competitive advantage in the market. When pilots test basic capabilities successfully, they provide revenue opportunities, but limit entrepreneurs from proving that their innovations are differentiated. When AI or machine learning technologies constitute a fundamental input to innovations, which capabilities are tested may also have important implications for the innovations' subsequent development. Because AI learns and adapts as it encounters new sets of data, innovations that can access and use more specialized data earlier than others may gain a useful temporal advantage in the market.

DISCUSSION

Calls to investigate the design and use of AI have risen to a crescendo as "organizations (private and public) are heavily investing in developing, acquiring, and implementing AI, and we need to develop better and richer understandings of AI, how it is designed, implemented, and used." (Anthony et al, 2023:1690). Novel innovations built on technologies like AI have the potential to transform organizational structures and systems in ways that may differ dramatically from what has come before, by creating new mechanisms for data collection, analysis, and monitoring that may be more difficult to alter once in place (Brynjolfsson & McAfee, 2011; Anthony et al, 2023; Kellogg et al, 2021; Bailey & Barley, 2020). Research that spans the lifecycle of an innovation becomes all the more urgent to undertake if decisions about how to design and use intelligent technologies can quickly and durably become "encoded with the systems themselves" (Forsythe, 2001:12). Who is in power to select intelligent technologies for use inside organizations? How are these decisions made and with what consequences to the use and design of novel technologies?

This study heeds these calls by examining the process by which organizations select novel innovations. This is an important juncture as during selection, organizational decision-makers consider whether and how they might use a novel innovation. Designers can learn from the feedback decision-makers share and adapt their innovations in response – bridging the gap between design and use. I examined how 13 committees tasked with selecting novel innovations designed and piloted entrepreneurial firms' AI innovations. When committee members had strong priors about AI, those priors fractured more typical or theorized ways organizations make decisions – through a scientific approach, or where powerful subunit members exert power (Pfeffer & Salanick, 1974; March, 1994), or those skilled at persuading can convince

(Dutton & Ashford, 1993; Howard-Grenville, 2007). Rather, committee members gained power based on their preferences regarding AI and new alliances emerged. Committee decision-makers in dominant power coalitions engaged in process manipulation, which impacted the design and use of novel innovations in surprising ways. While dominant coalitions had power, their authority was nested within the larger organizational structure. Dominant coalitions could not just act strictly in their own interests but had to maintain and promote the perception that the pilot process was legitimate. As a result, dominant coalitions manipulated the pilot process to achieve their own interests. I show how when decision-makers engage in process manipulation it can alter the trajectory of what is tested and subsequently designed and used.

Process Manipulation, Selection, and a Scientific Approach to Decision Making

Recent research and theorizing have highlighted the utility of adopting a technical or "scientific" approach to decision making through experimentation and piloting, as these methods can help facilitate robust learning and discovery (Camuffo et al, 2024; Wuebker, Zenger & Felin, 2023; Merton, 1973; Thomke, 2003; Camuffo et al., 2020; Koning et al., 2022; Gans, 2023; List, 2022; Shelef, Wuebker & Barney, 2020). For example, Edmondson et al. (2001) argue that such approaches create environments where "new beliefs, new skills, and new collaborative routines are simultaneously developed" (p.697). A scientific approach can aid entrepreneurs, organizational decision-makers, and policymakers in generating data and evidence, which in turn supports more informed and less biased decisions. The benefits of these methods have been well documented. For instance, Koning et al. (2022) demonstrated that experimentation improves startup performance, while Thomke (2003) illustrated how testing has driven evolution in the semiconductor and pharmaceutical industries. Companies like Amazon use A/B testing to make decisions about webpage design. However, despite the theoretical support for these methods, there is limited research on how organizations actually employ them in practice to make decisions of consequence.

This research suggests that the ability of organizations to use a scientific approach and design pilots or experiments that generate unbiased information may be dubious at best. In this study, pilots did not create venues for developing new routines or skills. Instead, in all 13 pairings of entrepreneurial firms and customer organizations, committees manipulated the pilot process to achieve desired outcomes. Rather than generating unbiased information, these committees engaged in process manipulation—working within an accepted process but strategically engineering select steps or micro-decisions to achieve specific results. For example, dominant coalition members agreed to pilot novel innovations and abide by the pilot results. However, they

manipulated aspects of the piloting process, controlling the flow of information by folding and temporally loading the discussion, and offering carefully crafted claims of learning and generalizability to justify decisions that narrowed the scope of pilots. Dominant coalitions used process manipulation to achieve results aligned with their preferences while minimizing the reputational risk associated with running a process that might be questioned or doubted by others.

Process manipulation emerged as a powerful tool, particularly for those lacking traditionally theorized sources of power, such as positional or skill-based authority. Contrary to scholarship on decoupling, when organizational members engaged in process manipulation, they did not separate policy from practice or means from ends (Bromley and Powell, 2012). Instead, they immersed themselves in the process and strategically used it to achieve self-interested goals. This behavior suggests that organizational members at all levels may be more strategic than ceremonial. This finding has significant implications for scholars interested in technical or scientific approaches to decision making, as it suggests that while organizations claim to utilize these methods, in surveys or when asked, they may do so in ways that confirm biases rather than eliminate them.

Perhaps most importantly, the use of process manipulation influenced what was tested. Scholars advocating for a scientific approach typically assume that when constructing a pilot or experiment, organizational members will optimize for learning and generalizability (Wuebker, Zenger & Felin, 2023; Thomke, 2003; Camuffo et al., 2020; Koning et al., 2022; List, 2022; Shelef, Wuebker & Barney, 2020.). Organizations are likely to test moderate or mid-range elements of an innovation, rather than the most advanced or basic capabilities. However, this research indicates the opposite may occur. I found that organizations only piloted an innovation's advanced or basic capabilities. Gans (2023) theorizes that instead of optimizing for learning and generalizability, established organizations may instead design "high-bar" experiments to test the most complex ideas or the advanced capabilities of an innovation, to confirm their market views. Building on Gans (2023), I found that organizations do indeed construct pilots to confirm their views, but this does not mean they only design high-bar experiments. Committees that were fearful of AI constructed pilots in the manner Gans (2023) describes. But in contrast, committees that were enthusiastic about AI, developed pilots to test "low-bar" experiments, as a means of manipulating pilots to assert power in a situation where decision-makers had varying views. This is an important addendum to consider, that organizations are composed of individuals with varying perspectives, to theories on how

organizations design experiments and make decisions using a scientific approach. In this study, where selection was managed by a committee with bifurcated opinions about the utility of AI, some negotiation occurred and perhaps altered the way pilots would have otherwise been designed. Without apocalyptics, evangelistic committees would have likely designed high-bar experiments. Apocalyptic committees would have killed pilots – thus only high bar experiments would be run by adopting organization.

Surprisingly, many committee members in this study were medical professionals who, although familiar with randomized control trials and other experimental practices, were nevertheless willing to manipulate the piloting process. Future work would benefit from testing whether there are conditions under which organizations leverage pilots and experiments as scientists or "technicians" (e.g., Merton, 1973), aiming to generate unadulterated information. Given the prevalence of manipulation in this setting, further research investigating how those with scientific training execute experiments and pilots would be particularly useful. A cynic might suggest that once organizational members learn how to use these methods, they will bend them to their own purposes.

Process Manipulation, Selection and Use

Process manipulation also impacted what innovations were ultimately implemented and used within organizations. Surprisingly, in this study, those most likely to advocate for and convince others of an innovation's promise tempered how these innovations were tested and later used within organizations. Evangelists or champions, rather than advancing innovations, often attenuated their progression. Considerable research has investigated the benefits for external vendors and lower-status employees in mobilizing champions to sell an issue within an organization (Howard-Grenville, 2007) or to promote the adoption of a new idea, practice, or product (Schon, 1963; Rogers, 1983). Prior literature has emphasized the usefulness of piloting new ideas with groups that champion them (Edmondson et al., 2001), as champions tend to take bigger risks and more frequently use persuasion tactics, and impression management to sell their agenda (Howell and Higgins, 1990). However, this research suggests that champions may not always push for innovations to be used in ways that align with the goals of those seeking support. Champions were able to pilot innovations more successfully. As a result of conducting more successful pilots, more innovations testing basic capabilities were implemented and used by adopting organizations. However, innovations offering basic capabilities were more likely to be rejected during implementation compared with those that offered advanced capabilities and champions were not able to convince others to test more advanced

capabilities. When it came to using entrepreneurs' innovations, organizational members often failed to see the benefit, as the innovations offered features that were not sufficiently "better" than what was already in place, even if they improved performance.

Contrary to the existing literature, champions may limit the exposure of a novel innovation within an organization to protect it and ensure its uptake, even if restrained. Evangelistic champions may consequently temper learning about radical or advanced aspects of an innovation, whereas naysayers may promote advancement. Furthermore, in contrast to Howell and Higgins (1990), champions and naysayers used influence tactics with similar frequency to gain traction on their respective agendas. In this context, naysayers were not neutral parties in their organizations; they too had agendas to pursue and skillfully used manipulation to achieve their objectives. Given that organizations are frequently composed of individuals willing to push for their own agenda, champions may have to restrain their influence tactics to achieve their objectives. This restraint may limit the use of novel innovations within consuming organizations.

Committees formulated of apocalyptics were successful in stemming the total number of AI innovations implemented within adopting organizations. However, when these innovations were successfully piloted, they were more likely to be implemented and used, than innovations offering basic capabilities. Truelove and Kellogg (2016) have shown how when implementing radical technology organizational members can ally to develop more moderate options. This did not appear to be the case in this study, as organizations implemented what was successfully piloted. In this study, we can observe how committees were formed of diverse individuals who could not use brute strength or coercive power (e.g., Alder and Borys, 1996) due to concerns about how this would reflect on the process as a whole, they had to instead engage in process manipulation. These actions allowed committee members to achieve broader buy in, early in the process. This may have quelled organizational members from offering more moderate solutions. While process manipulation did indeed temper the adoption of radical or advanced uses of entrepreneurs' innovations, it also created a pathway for buy-in and implementation. Further research would do well to compare when and how more radical innovations are implemented within organizations.

Process Manipulation, Selection and the Design of AI

Scholars have hypothesized how designers' choices might impact organizations that adopt those innovations. As Bailey and Barley (2020:2) explain, "those who design and promulgate technologies have visions of what work is and what it should be like. Such visions shape the outcomes of technological

change." For example, Beane (2019) relays how the implementation of novel surgical robots, designed with small screens situated away from operating tables, limited peripheral participation by residents in the beginning and end stages of surgery. Some residents found operating with these robots difficult and subsequently did not use them, while others benefited. While designers' choices are not negligible, as designers can create innovations with the potential to upend existing ways of work (Forsythe, 1993), in this case, it was organizational members, not designers, who had control over how innovations were used and for what purposes.

Research has demonstrated how organizational members can locally adapt to fit with organizational users' distinct needs innovations (Desantis et al., 1992; Tyre and Orlikowski, 1994). Similarly, I show how a significant amount of design shaping occurs after innovations are introduced into organizations. These adjustments can have critical effects when they occur during selection. It is during the selection process, as organizations pilot innovations, that access to customers' systems and processes is granted. When this occurs, designers learn about how their innovations work in situ. This has important consequences for innovations built upon AI, as these are the moments when such applications can begin to train on customer data sets (Ozcan & Fraser, working paper). Thus, the decisions made by adopting organizations regarding how they expose an innovation to their work environment have implications for what that innovation might be capable of doing in the future. In a race to differentiate, entrepreneurs and designers who can gain this access earlier stand to win. While entrepreneurs were more likely to see their innovations implemented when pilots tested basic capabilities, these short-term gains may not help ultimately create a differentiated market position. When access to data is a critical differentiator in the market, enablers or facilitators of that access become critical partners.

Limitations

Although this study involved a range of organizations and entrepreneurial firms, the process of piloting is likely still contextual. For instance, these findings may not generalize to organizations implementing innovations they are already familiar with and currently using. While all the organizations in this study utilized committees to make selection decisions, not all organizations use this governance structure, and these findings may not generalize to hierarchical organizations. Another limitation is the size and maturity of the firms studied: the entrepreneurial firms were immature and small, while the customers were

mature and large. Entrepreneurial firms usually do not command market power over such customers and thus have limited influence. Larger, more equivalent firms may have more power to direct how customers test their new products. Future research would benefit from investigating the role of market power in shaping the way novel innovations are used and tested. Firms that have gained market power might be expected to push back on customer organizations with greater force and successfully push to pilot more robust uses of their innovations. Lastly, this study focused on how organizations select innovations that leverage AI—a polarizing technology. While it is possible that future policies, technologies, and ideas will create similarly bifurcating effects, not all will. Therefore, the findings of this study, particularly regarding the use of process manipulation, may not extrapolate to settings where decisions are less contested and decision-makers are less polarized.

Managerial Implications

This research has important implications for managers and entrepreneurs looking to introduce and use novel digital innovations. First, for managers, it is crucial to recognize that employees and leaders in this study formed strong opinions about AI technologies before even engaging with an entrepreneur's innovation. While this study does not shed light on how organizational members formed these opinions, it is reasonable to assume that the near-ubiquitous use of social media and access to news influenced their perspectives (e.g., Aral). Although some technologies may not provoke such strong and polarizing responses, it is reasonable to believe that members of organizations will form strong priors about many promising innovations in the future. These strong and polarized priors set the stage for process manipulation, which has significant implications for how organizations structure and design processes to engage with novel innovations. The use of committees and consensus-oriented governance structures is commonly employed to make decisions of importance within organizations. However, their use in selecting and testing novel innovations might be detrimental. This study suggests that committee structures may temper or limit the radicalness of what is considered. Other types of governance structures may encourage engagement with more novel innovations. For example, a structure with fewer decision makers may enable more experimentation, and minimize manipulation.

For entrepreneurs, this study underscores the importance of understanding the decision-making context within organizations. Entrepreneurs in this study did not attempt to shift the balance of power, despite being aware of committees' efforts to stifle their advancement. Entrepreneurs should consider

aligning with those who appear "out of power" to see if they can alter the direction or manipulative tactics used to pilot their innovations. In a world where organizational members have access to new types of information from public sources of varying quality, entrepreneurs should expect decision makers to have strong opinions about their innovations even before sharing them. Thus, new strategies beyond sharing their narratives or using traditional impression management techniques may be necessary to advance the design of their innovations.

References

Adler, Paul S., and Bryan Borys. "Two types of bureaucracy: Enabling and coercive." *Administrative science quarterly* (1996): 61-89.

Adner, Ron. "Sharing value for ecosystem success." *MIT Sloan Management Review* 63.2 (2022): 85-90.

Agarwal, Rajshree, and Michael Gort. "Firm and product life cycles and firm survival." *American Economic Review* 92.2 (2002): 184-190.

Amabile, Teresa, and Steven Kramer. *The progress principle: Using small wins to ignite joy, engagement, and creativity at work*. Harvard Business Press, 2011.

Anteby, Michel, Curtis K. Chan, and Julia DiBenigno. "Three lenses on occupations and professions in organizations: Becoming, doing, and relating." *Academy of Management Annals* 10.1 (2016): 183-244.

Anthony, Callen, Beth A. Bechky, and Anne-Laure Fayard. ""Collaborating" with AI: Taking a system view to explore the future of work." *Organization Science* (2023).

Bailey, Diane E., and Stephen R. Barley. "Teaching-learning ecologies: Mapping the environment to structure through action." *Organization Science* 22.1 (2011): 262-285.

Bailey, Diane E., and Stephen R. Barley. "Beyond design and use: How scholars should study intelligent technologies." *Information and Organization* 30.2 (2020): 100286.

Baldwin, Carliss Y. "Design rules: past and future." *Industrial and Corporate Change* 32.1 (2023): 11-27.

Barley, Stephen R. "Technology as an occasion for structuring: Evidence from observations of CT scanners and the social order of radiology departments." *Administrative science quarterly* (1986): 78-108.

Barley, William C. "Anticipatory work: How the need to represent knowledge across boundaries shapes work practices within them." *Organization Science* 26.6 (2015): 1612-1628.

Barrett, Michael, et al. "Reconfiguring boundary relations: Robotic innovations in pharmacy work." *Organization Science* 23.5 (2012): 1448-1466.

Beane, Matthew. "Shadow learning: Building robotic surgical skill when approved means fail." *Administrative Science Quarterly* 64.1 (2019): 87-123.

Beane, Matthew, and Callen Anthony. "Inverted Apprenticeship: How Senior Occupational Members Develop Practical Expertise and Preserve Their Position When New Technologies Arrive." *Organization Science* (2023).

Becker, Howard S. *Tricks of the trade: How to think about your research while you're doing it*. University of Chicago press, 2008.

Benner, M. J., & Tripsas, M. (2012). The influence of prior industry affiliation on framing in nascent industries: The evolution of digital cameras. *Strategic Management Journal*, *33*(3), 277-302.

Brynjolfsson, Erik, Seth Benzell, and Daniel Rock. "Understanding and addressing the modern productivity paradox." *Massachussetts Institute of Technology. Available online: https://workofthefuture. mit. edu/research-post/understanding-and-addressingthe-modern-productivity-paradox/(accessed on 9 May 2022)* (2020).

Camuffo, Arnaldo, et al. "A scientific approach to entrepreneurial decision making: Evidence from a randomized control trial." *Management Science* 66.2 (2020): 564-586.

DeSanctis, Gerardine, et al. "Using computing to facilitate the quality improvement process: The IRS-Minnesota project." *Quality control and applied statistics* 37.7 (1992): 363-366.

Edmondson, Amy C., Richard M. Bohmer, and Gary P. Pisano. "Disrupted routines: Team learning and new technology implementation in hospitals." *Administrative science quarterly* 46.4 (2001): 685-716.
Edmondson, Amy C., and Stacy E. McManus. "Methodological fit in management field research." *Academy of management review* 32.4 (2007): 1246-1264.
Fiol, C. Marlene. "Consensus, diversity, and learning in organizations." *Organization science* 5.3 (1994): 403-420.
Forsythe, D. E. (1993). Engineering knowledge: The construction of knowledge in artificial intelligence. *Social studies of science*, *23*(3), 445-477.
Forsythe, D. (2001). *Studying those who study us: An anthropologist in the world of artificial intelligence*. Stanford University Press.
Fraser, Jack Lewis, and Pinar Ozcan. "From Flexibility to Rigidity: The Interplay of Actors and Technology in Machine Learning Ventures." *Academy of Management Proceedings*. Vol. 2022. No. 1. Briarcliff Manor, NY 10510: Academy of Management, 2022.
Friedkin, N. E. (2011). A formal theory of reflected appraisals in the evolution of power. *Administrative Science Quarterly*, *56*(4), 501-529.
Gans, J. S. (2023). Experimental Choice and Disruptive Technologies. *Management Science*.
Gans, Joshua S., Scott Stern, and Jane Wu. "Foundations of entrepreneurial strategy." *Strategic Management Journal* 40.5 (2019): 736-756.
Goldstein, N. J., & Hays, N. A. (2011). Illusory power transference: The vicarious experience of power. *Administrative Science Quarterly*, *56*(4), 593-621.
Harris, Stanley G., and Robert I. Sutton. "Functions of parting ceremonies in dying organizations." *Academy of Management journal* 29.1 (1986): 5-30.
Huy, Q., and T. Vuori. "Who Killed Nokia." *Nokia Did, Available online at https://knowledge. insead. edu/strategy/who-killed-nokia-nokia-did-4268, INSEAD Knowledge* (2015).
Gans, J. S. (2023). Experimental Choice and Disruptive Technologies. *Management Science*.
Howard-Grenville, J. A. (2007). Developing issue-selling effectiveness over time: Issue selling as resourcing. *Organization Science*, *18*(4), 560-577.
Howell, Jane M., and Christopher A. Higgins. "Champions of technological innovation." *Administrative science quarterly* (1990): 317-341.
Jamieson, Trevor, and Avi Goldfarb. "Clinical considerations when applying machine learning to decision-support tasks versus automation." *BMJ quality & safety* 28.10 (2019): 778-781.
Jick, Todd D. "Mixing qualitative and quantitative methods: Triangulation in action." *Administrative science quarterly* 24.4 (1979): 602-611.
Karp, Rebecca. "Gaining Organizational Adoption: Strategically Pacing the Position of Digital Innovations." *Academy of Management Journal* 66.3 (2023): 773-796.
Kellogg, Katherine C. "Local adaptation without work intensification: experimentalist governance of digital technology for mutually beneficial role reconfiguration in organizations." *Organization Science* 33.2 (2022): 571-599.
Kellogg, Katherine C., Melissa A. Valentine, and Angele Christin. "Algorithms at work: The new contested terrain of control." *Academy of Management Annals* 14.1 (2020): 366-410.
Koning, Rembrand, Sharique Hasan, and Aaron Chatterji. "Experimentation and start-up performance: Evidence from A/B testing." *Management Science* 68.9 (2022): 6434-6453.
Kor, Yasemin Y. "Experience-based top management team competence and sustained growth." *Organization science* 14.6 (2003): 707-719.
Lawrence, Paul R., and Jay W. Lorsch. "Differentiation and integration in complex organizations." *Administrative science quarterly* (1967): 1-47.
Leonardi, Paul M. "Activating the informational capabilities of information technology for organizational change." *Organization science* 18.5 (2007): 813-831.
Leonardi, Paul M. "Crossing the implementation line: The mutual constitution of technology and organizing across development and use activities." *Communication Theory* 19.3 (2009): 278-310.
Leonardi, Paul M., and Stephen R. Barley. "Materiality and change: Challenges to building better theory about technology and organizing." *Information and organization* 18.3 (2008): 159-176.
Lind, E. A., Kulik, C. T., Ambrose, M., & de Vera Park, M. V. (1993). Individual and corporate dispute resolution: Using procedural fairness as a decision heuristic. *Administrative science quarterly*, 224-251.
List, John A. *The voltage effect: How to make good ideas great and great ideas scale*. Currency, 2022.

March, J. G. (1962). The business firm as a political coalition. *The Journal of politics*, *24*(4), 662-678.
March, J. G. (1994). *A primer on decision making: How decisions happen*. Simon and Schuster.
Mechanic, David. "Sources of power of lower participants in complex organizations." *Administrative science quarterly* (1962): 349-364.
Merton, Robert K. *The sociology of science: Theoretical and empirical investigations*. University of Chicago press, 1973.
Myers, Jenna Elizabeth. "When Big Brother is Benevolent: How Technology Developers Navigate Power Dynamics among Users to Elevate Worker Interests." *Academy of Management Discoveries* ja (2023).
Rogers, E. M. (1983). Diffusion of. *Innovations (Glencoe, IL: The Free Press, 1962)*.
Pachidi, S., Berends, H., Faraj, S., & Huysman, M. (2021). Make way for the algorithms: Symbolic actions and change in a regime of knowing. *Organization Science*, *32*(1), 18-41.
Pisano, Gary P. "Knowledge, integration, and the locus of learning: An empirical analysis of process development." *Strategic management journal* 15.S1 (1994): 85-100.
Podolny, Joel M. "Market uncertainty and the social character of economic exchange." *Administrative science quarterly* (1994): 458-483.
Salancik, G. R., & Pfeffer, J. (1974). The bases and use of power in organizational decision making: The case of a university. *Administrative science quarterly*, 453-473.
Schon, Donald A. "Champions for radical new inventions." *Harvard business review* 41 (1963): 77-86.
Stinchcombe, Arthur L. *Sugar island slavery in the age of enlightenment: the political economy of the Caribbean world*. Princeton University Press, 1995.
Szulanski, Gabriel. "The process of knowledge transfer: A diachronic analysis of stickiness." *Organizational behavior and human decision processes* 82.1 (2000): 9-27.
Thomas, Robert J. *What machines can't do: Politics and technology in the industrial enterprise*. Univ of California Press, 1994.
Thomke, Stefan H. *Experimentation matters: unlocking the potential of new technologies for innovation*. Harvard Business Press, 2003.
Topol, Eric J. "High-performance medicine: the convergence of human and artificial intelligence." *Nature medicine* 25.1 (2019): 44-56.
Truelove, Emily, and Katherine C. Kellogg. "The radical flank effect and cross-occupational collaboration for technology development during a power shift." *Administrative Science Quarterly* 61.4 (2016): 662-701.
Tyre, Marcie J., and Wanda J. Orlikowski. "Windows of opportunity: Temporal patterns of technological adaptation in organizations." *Organization science* 5.1 (1994): 98-118.
Valentine, M., and R. Hinds. *Algorithms and the org chart*. Working Paper. Stanford University, 2019.
Wuebker, R., Zenger, T., & Felin, T. (2023). The theory-based view: Entrepreneurial microfoundations, resources, and choices. *Strategic*
Vuori, Timo O., and Quy N. Huy. "Distributed attention and shared emotions in the innovation process: How Nokia lost the smartphone battle." *Administrative Science Quarterly* 61.1 (2016): 9-51.

1: Entrepreneurial firm characteristics

Firm Pseudonym	Funding ($000)	Revenue ($000)	Geography	CEO Gender	CEO Background	Diagnostic area	ML capabilities diagnose
MAMMO	$500	$550	Boston	Male	M.D. & Ph.D	Breast Cancer	known and unknown abnormalities
WHITE	$1,000		Europe	Male	Ph.D	Immunity	known and unknown abnormalities
DIAB	$150	$450	Latin America	Male	Ph.D	Diabetes	known and unknown abnormalities and treatment
SENSOR	$1,500		Boston	Female	Ph.D	Brain Injury & Mental Health	known and unknown abnormalities
SKIN	$400	$500	New York	Female	MBA	Derm. diseases	known and unknown abnormalities
INFECTION	$225	--	San Francisco	Female	M.D	Bacterial diseases	known and unknown abnormalities and treatment

Table 2: Entrepreneurial firm and Customer Dyads

Customer	Entrepreneurial firm	Pilot opportunity	Dominant coalition	Committee Size
Blue health System	MAMMO	Mammogram scans for known abnormalities	67% Evangelistic	6
	SENSOR	Patient mental health monitoring in hospital	63% Apocalyptic	8
	INFECTION	Diagnosis & treatment of resistant bacteria	67% Apocalyptic	6
	SKIN	Skin cancer diagnosis	60% Evangelistic	5
Red Health System	SENSOR	At home elderly mental health monitoring	67% Apocalyptic	6
	INFECTION	Diagnosis of resistant bacteria strains	71% Evangelistic	7
	SKIN	diagnosis of unknown skin abnormalities	75% Apocalyptic	8
	DIAB	Diabetes diagnosis	60% Evangelistic	5
	WHITE	known indicators of immune system strength	67% Evangelistic	6
Community Health System	MAMMO	Mammogram scans for unknown abnormalities	71% Apocalyptic	7
	SENSOR	In hospital patient monitoring	57% Evangelistic	7
	WHITE	Known indicators of immune system strength	67% Evangelistic	6
	DIAB	Diabetes diagnosis and treatment	60% Apocalyptic	5

3: Customer committees' issues and concerns regarding entrepreneurial firms' pilots

Type of issue	Issue	Definitions	Representative data
Benefits of Machine Learning Technology	Reduce loads	Decreases the number of patient cases reviewed or worked on by a health professional	"It was a pretty clear benefit, it [MAMMO's innovation] could decrease loads by like 80% in the long run, because 80% of what we see is pretty generic" (CIO)
	Improves quality	Improves quality of patient care	"If [Skin's innovation] can really see stuff that doctors miss, it will change how we take care of patients. It is a cheaper model for health, because we can catch things earlier with less labor, and a better model for health because we can catch things we might otherwise miss." (Chief of Medicine)
	Opportunity to learn	Creates an opportunity to learn about machine learning technologies and potentially a platform for career advancement	"Part of the reason we do this, pilot, is to learn about the tech. We can take it apart, look at it, implement it. People won't say this, but they want to take it apart. If it is the next thing in health, well you want to know about it. You want to be an expert in it." (head of innovation)
Concerns about Machine learning technology	Efficacy of machine learning	Concerns about the efficacy of Machine learning in solving complex medical issues	"One thing we heard, is that diagnosing is multifactorial, you can't just focus on one thing and spit out the answer. So, your head hurts, you have a stomachache, and a bunch of other things. A machine can't possible understand that It is too complex of an algorithm." (CIO of SENSOR)
	Ethicality	Concerns about sharing data necessary to train and maintain machine learning algorithm	"Oh ya. Someone in every meeting says, but what about ethics. I get it, it is an issue. But honestly, machines are just a reflection of society. So how is it less ethical than what goes on in society. If you are concerned with the robustness of the data to serve diverse communities, give us more data. Oh but wait, you don't have that data really, because you don't really do good studies on people who don't' go to the doctor and don't have readily available healthcare. (VP at DIAB)
	Workplace impact	Concerns machine learning will cause deskilling or role replacement	"We were very concerned about what machine learning might do to jobs in our health system. Yes, it could create efficiencies, but what happens to the people?" (Doctor at Blue hospital)
Benefits of an entrepreneurial firm's innovation	Technologically proficient	Innovation is well architected and technically constructed	"Their application was really good. Like super smart. I hadn't seen someone with an application that was of high quality" (IT representative at Community health system)
	reduce loads	Decreases the number of patient cases reviewed or worked on by a health professional	"I'm always wary of applications that say they reduce workload, but this one, I feel more confident about" (IT representative at Community health system)
Concerns about an entrepreneurial firm's innovation	Implementation	Innovation will not easily interoperate with existing organization's systems or processes	"Last time we did something with an entrepreneur, it did not work out. Their stuff is just never far enough along that it can just plug into our systems or work with our protocols." (Doctor at Red health system)
	Efficacy of an innovation	Innovation will not work as purported in delivering expected health or performance outcomes	"I just think they are vaporware at this point. It is not going to work like they say it will." (IT representative at Red health system)
	Workplace impact	Concerns that an innovation will cause deskilling or role replacement	"If we do this, people in my department will lose their jobs in the long-run" (Nurse at Red hospital)

Table 4a: Scoping Stage –Comparison of Scoping process in Apocalyptic committees compared with Evangelistic committees

	Activities		**Scope**
Literature on Experimentation & Piloting: (Edmondson et al, 2001; Thomke, 2003; Camuffo et al, 2020; Adner, 2020; List, 2022; Gans, 2023)	**Designing scope to optimize Learning and Generalizability** – extant literature describes how pilots enable organizations to gain information about the underlying capabilities of an innovation while also generating information about how to deploy that innovation within an organization.		Should test a mix of capabilities or moderate capabilities
Dominant coalition **Apocalyptic committee**	**Claims - Justifying actions with claims of learning**	**Folding discussion**	**Tested advanced capabilities**
	Committee meeting - Doctor notes the importance of learning what the app can really do through testing treatment. Multiple heads nod, and no one disagrees**.**	"We can each say our piece but because we have more people, we get to speak last. That is important, because we can then turn the conversation, kill topics that are out of bounds" (Doctor Apocalyptic)	"I am really doubtful about AI, so I was like look, let's make the pilot hard. We have to keep the pilot to make everyone happy but Let's really test it and if it falls, so what." (Doctor Apocalyptic)
Entrepreneurs' response	"We knew what they were up to. This was not about learning. It was about killing us" (Founder DIAB)	We really wanted them to talk through how the pilot would go before we settled on the scope. But they were like, we will get to that. We couldn't push them anymore. (Founder DIAB)	"I could tell the nurse that is our buddy was disappointed. But he was like, something is better than nothing" (Founder DIAB)
Minority coalition response	"I wish we could do more for [company name]. But we got the pilot going at least". (Chief of medicine Evangelist)	"We tried a few times to get our opinion in, but they always had someone else saying why it was not a good idea." (Chief of medicine Evangelist)	"This was the compromise. Making it hard. This is the way it goes, sometimes you have the numbers and sometimes you don't." (Coordinator Evangelist).
Dominant coalition **Evangelistic committee**	**Claims - Justifying actions with claims of generalizability**	**Temporally loading discussion**	**Tested basic capabilities**
	We want this to work, so we have to make sure it is bullet proof. We wanted to pick a test that we knew our systems could execute. (CIO Red Hospital)	"Skip the predictive stuff on expectation on the aggressiveness. That will be hard to work out a process to support that. So, let's take that out of scope. Let's just leave it at identifying and turning over to docs." **(**Doctor Evangelist)	"Let's pick a use-case that is simple enough that it will work. Think MVP" (IT coordinator Evangelist)
Entrepreneurs' response	"I'm not sure what they our we learn from this [the pilot]?." (CIO White)	"Every time the talked about integration, they scoped more stuff out. Oh, the API will be hard to use, let's not test that. Our scope got smaller and smaller." Founder White	"We understood the thinking. Make the test work, but that kind of screws us over. We don't get the chance to see what our stuff can do." Founder White.
Minority coalition response	I'm sure this will generalize." (Nurse Apocalyptic)	If they want to talk about integration and take stuff of the table. don't mind that at all." (Nurse Apocalyptic)	"We aren't really testing much. We are confirming." (Nurse Apocalyptic)

Table 4b: Integrating Stage –Comparison of Scoping process in Apocalyptic committees compared with Evangelistic committees

	Activities	Claims
Literature on Experimentation & Piloting: **(Edmondson et al, 2001; Thomke, 2003; Adner, 2020)**	**Developing new beliefs, skills and collaborative routines** – scholars suggest that as organizations develop a pilot, they will develop new beliefs and build collaborative routines.	**Mutual adjustment** – Integration is often depicted as requiring adjustment by all parties (e.g., Long-Lingo & O'Mahony, 2010) and thus entrepreneurs would make changes to their innovations and organizations to their operating environments.
Dominant coalition Apocalyptic Committee	**Disavowing responsibility** – pushing responsibility for integration upon other partiers. "It's [SENSOR's] pilot. They have to figure this out." (Nurse Apocalyptic)	**Justifying with claims of generalizability** "It has to work across lots of uses right? So we need to show it can." (Nurse Apocalyptic)
Entrepreneurs' response	"They would hardly help. They were like, here is the manual with our systems. Can you get your stuff to comply?" (Founder SENSOR) "Not sure if the lack of help was about proving the pilot or about decreasing our chances for success." (Founder SENSOR)	"They would hardly help. They were like, here is the manual with our systems. Can you get your stuff to comply?" (Founder SENSOR)
Minority coalition response	Committee meeting note:, no one offered SENSOR support to understand who are the right people to connect with. They talked about connecting with geriatric nurses, and all agreed it a good idea but no one had any contacts of the person who was in charge.	"I don't really believe their clams. But we are stuck here. I don't want to rock the boat" IT associate
Dominant coalition Evangelistic Committee	**Claiming responsibility** – taking on responsibility for integration "We had to know how to work the app. So we took control. Running the pilot required some new process maps and detailed plans about how psych docs would find about patients admitted for cardio problems. The team here built them." (IT director Evangelistic)	**Justifying with claims of learning** **IT director:** "What do companies typically do? They run the pilot right? They have to learn about what is going on. You can't leave that to chance and later realize it doesn't work and have people not use or stop using it. That would look bad. **Author:** How would it look bad? **IT director:** Well, everyone would start in on machine learning.
Entrepreneurs' response	"It seemed like every day, they would do more and we would do less to get the pilot up and running. I don't even really know all the stuff they did. (Founder SENSOR)	"We are being shut out of learning anything. But I guess they get to really see if they can do it." (Founder SENSOR)
Minority coalition response	I have so much more work, because they have limited the entrepreneurs' work. Frustrating. (Nurse Apocalyptic)	"It is a smart move to say this is about what we need to learn." (Doctor Apocalyptic)

Table 4c: Running and Implementation Stage –Comparison of Scoping process in Apocalyptic committees compared with Evangelistic committees

	Activities	Outcomes
Literature on Experimentation & Piloting: (Edmondson et al, 2001; Thomke, 2003; Camuffo et al, 2020; Adner, 2020; List, 2022; Gans, 2023)	**Tracking pilot results** – collecting evidence of how pilot unfolded compared with "control" test or "back up" method **Interpreting pilot results** Scholars contend that contestation can occurs as organizational members interpret and analyze pilot results as results may be tested on groups that are not representative enough of other relevant populations in the organization or the data is informative, but one of many data points used to evaluate a new idea or product, and thus it is ruled out (Adner, 2020; Thomke, 2003).	If an innovation is piloted successfully, it will likely be implemented, especially if the results are not contested
Dominant coalition Apocalyptic Committees	**Accepted results** **Author:** "Are you surprised everyone agreed with the results of the pilot? **Head of Nursing:** "We had a good process. Everyone had their say. We paid attention to everybody's concerns. So, no. Why would anyone complain. That would look really petty."	**~33% of pilots were successful** **All innovations that were successfully piloted were implemented and used**
Entrepreneurs' response	"They set it up for failure and well, it failed. It sucks. We tried super hard, spent a lot of time, but that is the way it goes." (CIO INFECTION)	"When these pilots worked, it was amazing. We got to prove out our technologies in ways that created real differentiation" (CEO Skin)
Minority response	"The pilot didn't work well. The doctors didn't really know when to use the suggestions, or they looked at them to double check their own analysis. So, it didn't add efficiency or accuracy. It was hard to dispute it and so we stopped working with [INFECTION]." (Coordinator Evangelist)	"We were hoping the pilot would worked. We would push through the most advanced, state of the art parts of this technology." Doctor (Evangelist)
Dominant coalition Evangelistic committees	**Accepted results** "The pilot proved out and so we went forward with them. We are going to broadened the teams using their application." Director of Innovation	**~ 60% of pilots were successful** **60% of innovations that were successful piloted were implemented and used**
Entrepreneurs' response	"We will make some money from those guys. No one really could push back on what we did." (CEO, INFECTION)	"People didn't really think our innovation offered much more than what was already in place." (CEO Sensor)
Minority response	"Some people may not be happy, but they bought into the process." (Director of Innovation)	"It was sort of great. We did a pilot, it worked but then it was just not that different than the existing way we do stuff. So people didn't use it." (Doctor Apocalyptic)

Table 5: Outcomes

Committee sentiment	Customer	Entrepreneurial firm	Successful Pilot	Innovation Implemented
Evangelistic	Red Health System	INFECTION	x	x
	Community Health System	WHITE	x	x
	Blue health System	MAMMO	x	
	Red Health System	DIAB	x	x
	Blue Health System	SKIN	x	
	Red Health System	WHITE		
	Community Health System	SENSOR		
Apocalyptic	Red Health System	SKIN	x	x
	Community Health System	MAMMO	x	x
	Red Health System	SENSOR		
	Blue health System	INFECTION		
	Blue health System	SENSOR		
	Community Health System	DIAB		

FIGURE 1: Explaining How Process Manipulation Shapes the Innovative Trajectory of Novel Innovations