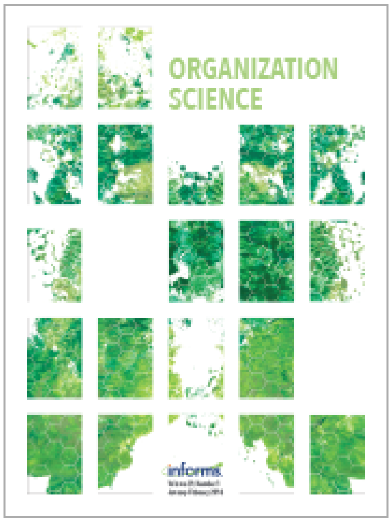


Organization Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Converging Tides Lift All Boats: Consensus in Evaluation Criteria Boosts Investments in Firms in Nascent Technology Sectors

Xirong (Subrina) Shen , Huisi (Jessica) Li , Pamela S. Tolbert

Published online in Articles in Advance 27 Sep 2021

. https://doi.org/10.1287/orsc.2021.1493

Published in Organization Science on September 27, 2021 as DOI: 10.1287/orsc.2021.1493.
This article has not been copyedited or formatted. The final version may differ from this version.

informs
http://pubsonline.informs.org/journal/orsc

ORGANIZATION SCIENCE
Articles in Advance, pp. 1–21
ISSN 1047-7039 (print), ISSN 1526-5455 (online)

Converging Tides Lift All Boats: Consensus in Evaluation Criteria Boosts Investments in Firms in Nascent Technology Sectors

Xirong (Subrina) Shen,[a] Huisi (Jessica) Li,[b] Pamela S. Tolbert[c]

[a] Department of Management, McCombs School of Business, University of Texas at Austin, Austin, Texas 78712; [b] Scheller College of Business, Georgia Institute of Technology, Atlanta, Georgia 30308; [c] Department of Organizational Behavior, School of Industrial and Labor Relations, Cornell University, Ithaca, New York 14853

Contact: xirong.shen@mccombs.utexas.edu, https://orcid.org/0000-0002-2769-6074 (X(S)S); jessica.li@scheller.gatech.edu, https://orcid.org/000-0002-7881-7809 (H(J)L); pam.tolbert@cornell.edu, https://orcid.org/0000-0001-9661-6528 (PST)

Received: September 14, 2019
Revised: June 6, 2020; December 15, 2020; March 23, 2021
Accepted: May 4, 2021
Published Online in Articles in Advance: September 27, 2021

https://doi.org/10.1287/orsc.2021.1493

Abstract. Although previous studies show that the emergence of evaluation criteria for a new technology improves the life chances of well-performing firms, we theorize that *consensus* in such criteria among technology experts increases investments to *all* firms in the new sector. We provide a variety of supportive evidence for this claim. First, in an experiment with 80 Chinese investors (Study 1), we provide evidence of a causal relation between evaluation consensus and investments. We follow this with a second experiment with 412 U.S. participants (Study 2), showing that evaluation criteria consensus increases participants' propensity to view a firm as technologically competent and to expect others to favor investing in the firm. Analyses of longitudinal archival data on investment in artificial intelligence technology firms in the United States (Study 3a) and China (Study 3b) support the generalizability of our findings. By exploring the social-cognitive processes that link evaluation criteria consensus to investors' decisions to invest in firms in nascent technology fields, this paper advances the scholarly understanding of the microfoundations of the institutionalization processes in new market sectors.

History: This paper has been accepted for the *Organization Science* Special Issue on Experiments in Organizational Theory.
Funding: This work was supported by Cornell University (Innovation, Entrepreneurship, and Technology Theme Research Grant, Jeffrey S. Lehman Fund for Scholarly Exchange with China).
Supplemental Material: The online appendix is available at https://doi.org/10.1287/orsc.2021.1493.

Keywords: evaluation criteria • evaluation criteria consensus • investment • nascent technology • micro-foundations of institutional theory • experimental methods

For investors, firms in nascent technology fields are thorny roses (Anderson and Tushman 1990, Aldrich and Fiol 1994). On the one hand, investing in a new firm whose technology may create market-changing opportunities can yield extraordinary returns. On the other hand, evaluating the firm's core technology may be challenging (Tushman 1992, Suchman 1995). Difficulties in evaluating firm technologies undermine investors' willingness to provide critical resources to firms in a nascent sector, limiting the growth of the sector as a whole (Benner and Zenger 2016). Past research has examined field-level evaluation institutions, such as organizations that award certifications (Sine et al. 2007, Graffin and Ward 2010, Lanahan and Armanios 2018), or rank firms on some chosen criteria (Rao 1994), and found that they have important consequences for firms. For example, the establishment of evaluation criteria has been shown to increase the survival rate of firms that perform well on the criteria (Rao 1994, Sine et al. 2007). Taking this a step further, Goldfarb et al. (2018) suggest that simply being evaluated may have a positive effect on perceptions of firms, independent of their performance on the criteria.

However, a critical neglect in existing work is how *consensus* (as opposed to disagreement) on the criteria for evaluating different technology solutions in a nascent sector affects firms' ability to acquire resources. Multiple evaluation criteria often exist in nascent fields (Lee and Sine 2012). For example, in the early automotive industry, different technology competitions used different criteria, such as speed, durability, and fuel efficiency, to assess entrants' technological performance (Rao 1994). Past studies alluded to the existence of multiple evaluation criteria in nascent technology fields (Rao 1994, Lee and Sine 2012, Grodal 2018), but have not examined how varying levels of consensus in criteria could shape firm- and field-level outcomes. Recognizing the importance of field-level consensus in nascent markets (Ozcan and

Santos 2015, Georgallis et al. 2018), this paper focuses on the effects of consensus (as opposed to disagreement) in evaluation criteria among technology experts on firms' resource acquisition.

In particular, we focus on how evaluation criteria consensus affects investors' assessment of and willingness to invest in firms in nascent technology fields. This focus allows us to shed light on the still mysterious process through which evaluation criteria influence firms. Prior works have implied that evaluation criteria affect firm survival by influencing resource providers' evaluation of firms. Yet, these works did not directly examine the evaluation processes of resource providers but rather inferred their reactions from the firms' survival or performance (Rao 1994, Sine et al. 2007, Goldfarb et al. 2018). This neglect is critical because individual resource providers' judgments of firms in a nascent sector are the "the micro motors" that coalesce to legitimate and direct resources to the sector (Powell and Colyvas 2008).

Addressing this gap, we unpack the socio-cognitive processes underpinning investors' evaluation of firms in nascent technology fields. In assessing the risk associated with a new technology, investors look to technology experts in both industry and academia (Fischhoff et al. 1982, Morgan 2014). For example, investors often attend criteria-setting technological competitions organized by experts at academic conferences (Malisiewicz 2015). Drawing on microinstitutional literature and the social psychological literature on judgment and decision making (Tversky and Kahneman 1974, Kahneman 2003, Tyler 2006, Bitektine 2011, Bitektine and Haack 2015, Bitektine et al. 2018, Haack et al. 2020), we theorize that field-level evaluation criteria consensus among experts promotes a collective perception that a nascent technology is accepted and appropriate (referred to as validity beliefs in Bitektine and Haack 2015, Haack et al. 2020). Specifically, we argue that evaluation criteria consensus encourages investment in individual firms in a nascent sector for two reasons. First, evaluation criteria consensus helps investors form more positive perceptions of the technological competence of firms in the sector (referred to as propriety perceptions in Bitektine and Haack 2015, Haack et al. 2020). In other words, it encourages investors to view the firms' technological solutions as more capable and reliable. Second, agreement on evaluation criteria increases investors' confidence that their peers will also hold a positive view of the focal firm (referred to as consensus perceptions in Haack et al. 2020), making it easier to justify investment decisions to peers and stakeholders. Both the perceptions of a firm's technological competence and others' positive evaluations of the firm should encourage individuals to invest in the firm.

We tested our predictions in two main experiments and two archival studies. In our first experiment, surveying 80 Chinese private equity investors, we found that expert consensus on the criteria for evaluating a new technology increases investors' willingness to invest in a firm using the technology. In a second experiment with 412 U.S. managers, we show that evaluation criteria consensus boosts investment because it increases both individuals' perceptions of firms' technological competence and their beliefs that others will also evaluate firms in the sector positively. To establish the external validity of our findings supporting Hypothesis 1, our third and fourth studies use longitudinal archival data on investment in artificial intelligence (hereafter referred to as AI) firms in the United States and China. We collected data from competitions held at major AI conferences, in which different technological solutions were evaluated. We operationalized evaluation criteria consensus as the overlap in evaluation metrics across different competitions. Both studies demonstrate that evaluation criteria consensus increases the likelihood of firms receiving investment.

This paper contributes both to the stream of work on the impacts of evaluation criteria in nascent technology sectors (Rao 1994, Sine et al. 2007, Lee and Sine 2012, Goldfarb et al. 2018), and to the microinstitutional literature that seeks to link microlevel individual evaluations and macro-level legitimation outcomes (Bitektine 2011, Bitektine and Haack 2015; Schilke 2018, Haack et al. 2020). First, we provide a microfoundation for the literature on the impacts of evaluation criteria in nascent sectors (Rao 1994, Sine et al. 2007, Lee and Sine 2012, Goldfarb et al. 2018). Existing studies imply that the development of evaluation criteria affects firm survival in nascent market sectors by influencing the judgments of key resource providers (Rao 1994, Sine et al. 2007). However, little work has directly theorized or empirically examined the socio-cognitive processes underpinning resource providers' judgment formation. Such neglect is partly due to the limits of conventional methodologies (e.g., macro-historical data analyses and case studies), which are not designed to capture individuals' perceptions or untangle multiple causal forces (Haack et al. 2020). We address this gap by leveraging a mixed-method approach. We focus on the consensus in evaluation criteria, a construct that has not been examined so far. Our experiments not only establish a causal impact of evaluation criteria consensus on investors' assessment of firms but also unpack two socio-cognitive mechanisms involved in their judgments: beliefs about the firm's technological competence and perceptions of others' assessment of the firm. Our two archival studies demonstrate that these microinstitutional processes coalesce to have macro-level consequences: in a field with high evaluation criteria consensus, perhaps even without consciously understanding the details of

Published in Organization Science on September 27, 2021 as DOI: 10.1287/orsc.2021.1493.
This article has not been copyedited or formatted. The final version may differ from this version.

the consensus, investors still accept it as a "social fact" (Zucker 1987), making them more willing to provide material support for the field and firms in the field.

Our analyses also extend recent theoretical work on the nature of legitimation processes by providing a rigorous empirical validation of the multilevel nature of the legitimation process (Bitektine 2011, Bitektine and Haack 2015, Schilke 2018, Haack et al. 2020). Prior work posits that individuals assess the legitimacy of an entity on their own as well as judge the existence and extent of shared agreement on its legitimacy. This multipronged conception of legitimacy offers key insights into how and when institutional change is likely to occur. By showing that investors' endorsement of a firm in a nascent technology sector is driven by *both* their own beliefs about the appropriateness and viability of the firm's technology (propriety beliefs) as well as anticipation of their peers' collective assessment of the firm (consensus beliefs), our work provides strong evidence for this theoretical conceptualization. This paper is also among the first to empirically bridge the micromacro link (Bitektine and Haack 2015, Haack et al. 2020). Our mixed-method approach provides a nuanced picture of how varying levels of consensus in evaluation criteria translate into systematic differences in resources across technology fields by shaping individual investors' judgment. As such, we echo Haack et al. (2020) in showing consensus as an important "micro motor" for institutionalization processes (Powell and Colyvas 2008).

Theory and Hypotheses

Evaluation of Firms in Emerging Technology Fields

A crucial factor that investors consider when deciding whether to invest in a technology-based firm is its technological competence or the reliability and capability of its core technology (Baum and Silverman 2004, Aggarwal et al. 2015). However, making such an assessment is inherently difficult in a nascent technology field because of two types of uncertainty, which Graffin and Ward (2010) refer to as technical uncertainty and performance standard uncertainty. The first arises from a lack of observable data and well-defined objective metrics for gauging the quality of performance. The latter is created by a lack of collective agreement on an acceptable performance level. When uncertainty in evaluation is high, as is common in nascent sectors, investors may withhold resources from firms or put a discounted price on them (Benner and Zenger 2016, Polidoro Jr 2020). Evaluation criteria address both technical and performance standard uncertainty (Graffin and Ward 2010). We define *evaluation criteria* for a technology as standards that define the observable features or functions on which firms' technological performances should be measured and the metrics for measuring these features and functions (Durand and Kremp 2016).

Past literature has shown that the development of evaluation criteria in a new sector can reassure investors of firms' technological competence. For example, in a study of competitions organized by automobile makers in the formative days of the industry, Rao (1994) showed that firms that won more competitions were more likely to survive. Extending this work, Goldfarb et al. (2018) found that it was not just the winning firms that were more likely to survive—those that placed second or third also enjoyed enhanced survival chances. Furthermore, they note that mere participation in the competitions could be as beneficial as winning. This suggests that simply being evaluated may affect audiences' perceptions of firms' viability (Khaire 2010). Similarly, several studies showed that the establishment of agencies that provided certifications to firms that met their criteria enhanced the survival rates of those firms, even if the criteria poorly measured firm performance (Sine et al. 2007, Lee and Sine 2012).

Although existing studies indicate that the presence of evaluation criteria in a new sector positively affects firm survival, they paid less attention to the consequences of coexisting and even competing evaluation criteria (Lee and Sine 2012, Chatterji et al. 2016, Lanahan and Armanios 2018, Noh and Tolbert 2019). The rapid growth of a nascent sector often attracts technology experts from diverse backgrounds with different interpretations of the defining features/functions of the nascent technology (Wry et al. 2011, Lee et al. 2017, Grodal 2018), resulting in multiple, competing evaluation criteria. Although some prior studies alluded to the existence of multiple criteria in a nascent sector (Wry et al. 2011, Lee et al. 2017, Grodal 2018), they have not explicitly measured evaluation criteria consensus nor theorized its effects on key resource providers, such as investors (for a relevant theoretical discussion, see Lanahan and Armanios 2018).

Evaluation Criteria Consensus and Investment to Firms

We conceptualize consensus in evaluation criteria for a nascent technology as the extent to which technology experts—academic and industry scientists with domain-specific knowledge of the technology—use similar criteria to compare different technology solutions in the field (Hsu et al. 2012).[1] Existing studies of nascent market sectors have examined the consequences of having other types of field-level consensus. For example, Ozcan and Santos (2015) showed that the lack of consensus on transactional arrangements among potential exchange partners from different industries prevented the formation of a global market for mobile payments. Similarly, Georgallis et al. (2018) found that country-level convergence in producers' identities in the new solar photovoltaic industry increased government policy support to the producers (for similar findings, see McKendrick et al. 2003, Negro et al. 2010, 2011).

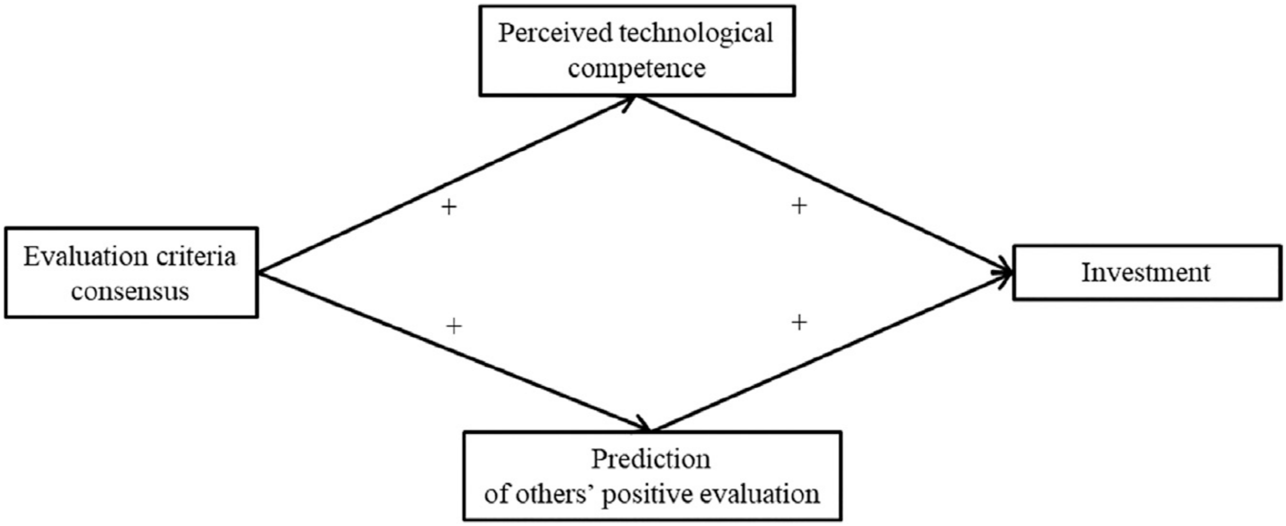
This paper builds on these insights and extends them. Our theorization provides a microfoundation to the existing historical studies by explicating the social-cognitive processes through which consensus on evaluation criteria affects resource providers' (i.e., investors') reactions to firms. As elaborated below, we argue that consensus among technology experts on evaluation criteria affects investment in a firm through two mechanisms: fostering investors' positive perceptions of the firm's technological competence and inducing them to predict positive evaluations of the firm from others. Figure 1 depicts our framework.

Perceived Technological Competence as a Mediator. The decision to invest in a firm specializing in a nascent technology depends on the investor's individual assessment of the firm's technological competence (Baum and Silverman 2004, Aggarwal et al. 2015). However, in a nascent sector, data on the technological competence of a firm are necessarily limited (Pollock and Gulati 2007, Plummer et al. 2016). As a result, investors rely on signals about the firm and its environment (MacMillan et al. 1985, MacMillan et al. 1987), such as the composition of the management team, its financial status, and its target market (Mason and Stark 2004). In line with this, we argue that investors' assessment of a firm is affected by their perception that the technology that the firm uses is acceptable and appropriate (referred to as validity beliefs in Haack et al. 2020).

Findings from the expert elicitation literature suggest that consistency in experts' opinions reduces nonexperts' perception of the risks associated with a new technology, thus making firms using the technology generally appear less risky (Lichtenstein et al. 1977, Slovic et al. 1985, Campbell 1998). Moreover, psychological research has shown that individuals tend to compensate for the lack of information on one judgment task (in our case, assessing the technological competence of a firm) by heuristically relying on other, more accessible information (e.g., evaluating the technology the firm uses) (Tversky and Kahneman 1974, Rosch 1999, Kahneman 2003). Therefore, when investors draw on evaluation criteria consensus as a cue to infer that the technology is reliable and appropriate, they are also likely to view firms using the technology as technologically competent. Because perceived technological competence of a firm should increase investment in it (Baum and Silverman 2004, Aggarwal et al. 2015), we propose that perceived technological competence mediates the positive relationship between evaluation criteria consensus and investment in a firm.

Prediction of Others' Positive Evaluation as a Mediator. Evaluation criteria consensus also helps to resolve investors' concerns about others' evaluations of the company—how their colleagues and other investors in the field will assess its value (referred to as perceptions of consensus in Haack et al. 2020). In assessing the appropriability and viability of a new entity, people usually look to their peers and consider what they do (DiMaggio and Powell 1983, Strang and Meyer 1993).

Figure 1. Summary of Theoretical Mechanisms



Published in Organization Science on September 27, 2021 as DOI: 10.1287/orsc.2021.1493.
This article has not been copyedited or formatted. The final version may differ from this version.

This tendency becomes even stronger in nascent markets (McDonald and Eisenhardt 2019). Moreover, resolving uncertainty in others' evaluations is particularly important for investors because they face pressures to communicate and justify their decisions to their colleagues and supervisors (Tetlock and Boettger 1989, Jensen 2006). More importantly, they are concerned about how to sell the company to other investors—either to syndicate partners in the short run or through investment exit in the long run (Cumming and Johan 2008). Investors would feel more comfortable investing in a firm when they predict that their peers and supervisors have a positive view of it.

Hypothesis 1. *Evaluation criteria consensus among technology experts in a nascent technology field increases investment in firms in the field.*

Hypothesis 2. *Investors' positive perception of firms' technological competence mediates the positive relationship between evaluation criteria consensus and investment.*

Hypothesis 3. *Investors' prediction of others' positive evaluations of a firm mediates the positive relationship between evaluation criteria consensus and investment.*

Scope Conditions. These arguments rest on some scope conditions. First, as stated in our title, our theoretical claims are applicable to nascent technology fields because uncertainty about a firm's potential performance is especially high in such fields. Under these conditions, investors are more likely to rely on experts' evaluation criteria consensus to infer the firm's technological competence and others' evaluations of the firm. Because evaluative uncertainty varies depending on how nascent a firm's technology and industry are, we reason that as the nascence level of a technology/industry varies, our hypothesized effects may change. For example, for firms that apply a mature technology in a well-established industry, there is less uncertainty involved in evaluating their technological competence. Thus, the effect of evaluation criteria consensus on investment would not be significant. We conducted an additional experiment to explore this scope condition. Please see Online Appendix A for details.

Similarly, because of limited data, the true technology competence of a young firm in a nascent field is hard, if not impossible, to assess. Nevertheless, the effect of consensus may vary depending on the firm's *reported* technological competence levels. For example, when the firm's reported technological competence is low, investors may disregard information from field-level evaluation, or if there is a high evaluation consensus, further discount the investment value of the firm because the clarity in evaluation criteria makes investors more certain that the firm is incompetent. Thus, we expect the effect of evaluation criteria consensus on investment to be weaker or not significant for firms that have low technological competence. Extrapolating this reasoning to the sector level, the positive effect of evaluation criteria should not hold for a nascent technology sector with a large proportion of technologically poor, "lemon" firms (Benner and Zenger 2016). We conducted an additional experiment to test this scope condition. Please see Online Appendix B for details.

Overview of Studies

To test our hypotheses, we conducted four main studies using two different methodologies (experimental and archival) across two national contexts (the United States and China). Combining experiments and archival studies allows us to examine the causal impact of evaluation criteria consensus on investor judgment, explore the underlying socio-cognitive processes, and reveal macro-level consequences of these judgments. Using data from both the United States and China ensures the generalizability of our findings. These two countries are not only home to firms that received most of the investments in the AI sectors (over 80% in 2019, CB Insights Research 2019), but also provide two different socio-economic contexts. In Study 1, we conducted an experiment with 80 experienced Chinese investors in the private equity sector to establish a positive causal relationship between evaluation criteria consensus and investment intentions (Hypothesis 1). Study 2 aimed to replicate the findings in Study 1 with a sample of 412 full-time employees in the United States and to extend our findings by testing the mediating effects of investors' perceptions of firm technological competence (Hypothesis 2) and predictions of others' positive evaluations (Hypothesis 3). In Studies 3a and 3b, we utilize archival data on AI firms in both the United States and China to establish the external validity of Hypothesis 1.

Study 1

Study 1 examines Hypothesis 1, positing a causal relationship between evaluation criteria consensus and investments. The design and analysis plans for this study were preregistered and can be accessed at https://osf.io/k8mbe/?view_only=23eba785b4d44e3386e61957d4563637.

Sample

A total of 80 Chinese investors in the private equity market were recruited from an online investor discussion group affiliated with the alumni association of a top research university in China to participate in our experiment. Participants in this group are verified alumni who invest in the private equity market and use the group to network and exchange information. Our survey was set up on Qualtrics (www.qualtrics.com), a widely used online survey platform. The survey link was posted by an active member of the discussion group. Participation was voluntary, anonymous, and without any monetary or other type of rewards. Of the 443 members of the group, 153 completed the study (response rate = 34.53%). As participants preregistered, we constrained the eligible participants to people with investment experience and excluded 17 participants who did not meet this criterion from our analysis. On average, the 136 (88.89%) eligible participants had 3.65 years of investment experience in the private equity market. Our study included simple attention checks (Meade and Craig 2012, Ward et al. 2017). (Online Appendix D provides details on the attention checks.) As preregistered and consistent with prior research (e.g., Hildreth et al. 2016), 56 participants who failed the attention checks (41.18%) were removed.[2] A total of 80 (58.82%) participants passed the attention checks and were included in our analysis. We preregistered to recruit 60 participants for Study 1. However, when the sample size reached 60, a power analysis results suggested that a sample of 80 was needed to get α = 0.05,1-β = 0.80. Thus, we collected a total of 80 participants to sufficiently power our analysis. Of these, 15 were female, 64 were male, and 1 unknown. The mean age was 33.27 (SD = 5.24). None of the participants were aware of our research purposes.

Design and Procedures

The task resembles the first-round screening process at venture capital firms (Tyebjee and Bruno 1984, MacMillan et al. 1987). Online Appendix C provides details on the task. Participants were instructed to imagine themselves as an investor who works in a team at a venture fund whose job included reading information about different technologies and firms. They were given a brief overview of a start-up firm that makes intelligent cameras, the firm's product, and the entrepreneur. This information was adapted from the description of a start-up on Indiegogo, a website that crowd-funds innovative products and features entrepreneurial campaigns.

Participants also read a description of the firm's core technology, image classification algorithms, and a paragraph about evaluation criteria for the technology: "Prominent academic and industry researchers at Stanford, MIT, Google, and Uber all have worked on assessing the algorithms. They have proposed various evaluation metrics for assessing the accuracy of image classification algorithms. The metrics include average precision, ROC-AUC, top-5 error, and others."

Participants were then randomly assigned to one of two conditions in a one-factor two-level (evaluation criteria consensus: high versus low) between-subject experimental design. The random assignment was automatically conducted by the Qualtrics platform. In the high (low) *evaluation criteria consensus* condition, participants read: "(However,) The academic and industry experts have agreed (do not agree) on the appropriate metrics. In other words, there is great (no) consensus about what metrics should be used for evaluating image classification algorithms." We manipulated *evaluation criteria consensus* as the agreement/disagreement among academic and industry experts because experts play an important role in shaping investor decision making in nascent technology fields. Past literature shows that experts' opinions influence investors' assessment of the risk associated with a new technology (Fischhoff et al. 1982, Morgan 2014).

Lastly, participants completed a survey with measures for the dependent variables, investment intention and investment amount (out of a possible Y2,000,000), attention checks, a manipulation check, and questions regarding their demographic information. Online Appendix D provides details on the measures.

Measures

Investment Intention. Three items adapted from Larsen and Newton-Smith (2001) measured intention to invest in the firm. Participants rated the extent to which they agreed or disagreed with statements such as, "I am willing to invest money in this firm." For all Likert scales in this paper, unless otherwise noted, 1 = strongly disagree, 2 = disagree, 3 = somewhat disagree, 4 = neither agree nor disagree, 5 = somewhat agree, 6 = agree, and 7 = strongly agree. α = 0.95.

Investment Amount. Participants were also asked, "How much of the ￥2,000,000 will you invest in this firm?" The number reported was the measure for investment amount (Kanze et al. 2018).

Manipulation Check. Three items measured participants' perceptions of evaluation criteria consensus, asking them to rate the extent to which they agreed or disagreed with statements such as, "There is consensus about the proper metrics to evaluate image classification algorithms." α = 0.89.

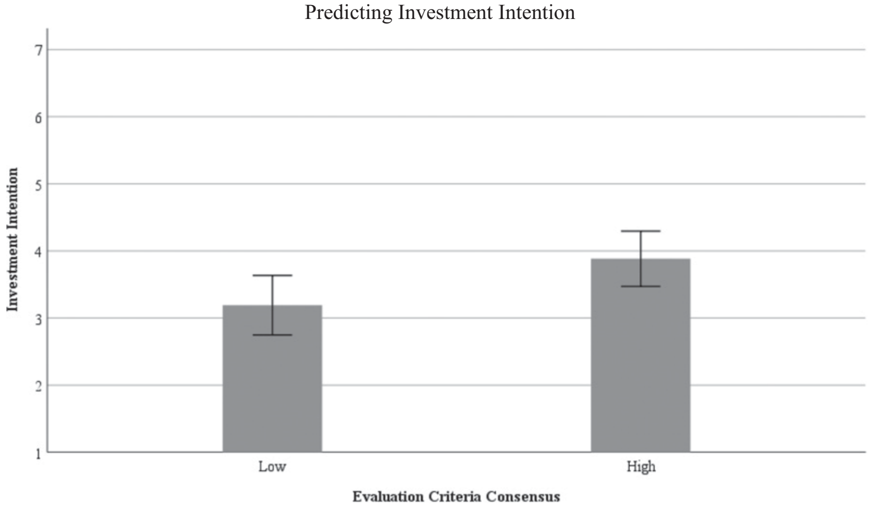
Results and Discussion

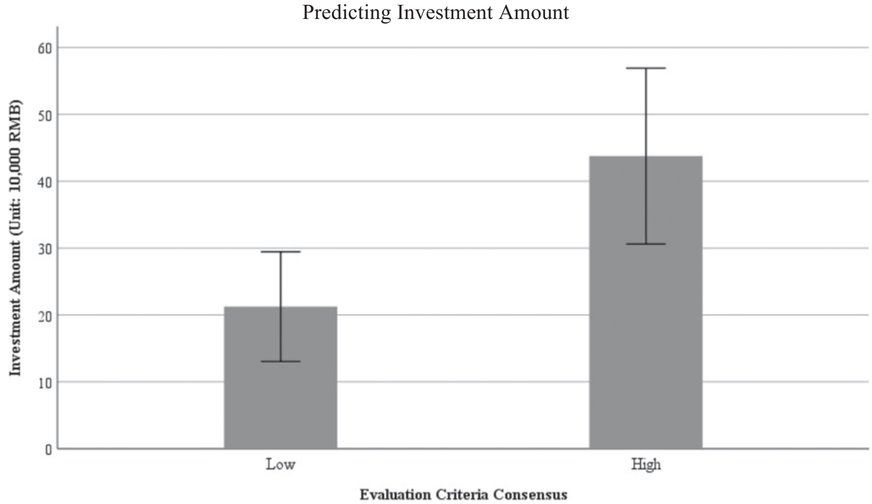
Manipulation Check. A *t*-test showed a significant difference in perceived evaluation criteria consensus between the high and the low conditions (Difference = 1.97, 95% CI = 1.35–2.59, d = 1.16, t (78) = 6.33, p < 0.001; $M_{high\text{-}consensus}$ = 4.68 and $M_{low\text{-}consensus}$ = 2.71), indicating that our manipulation worked as intended.

Hypotheses Testing. Supporting Hypothesis 1, as shown in Figure 2, a *t*-test revealed a significant

Published in Organization Science on September 27, 2021 as DOI: 10.1287/orsc.2021.1493.
This article has not been copyedited or formatted. The final version may differ from this version.

Figure 2. Study 1. Differences in (a) Investment Intention and (b) Investment Amount Cross High vs. Low Evaluation Criteria Consensus Conditions





Notes. $N = 80$. Error bars show 95% CI.

difference in *investment intention* between the high and low *evaluation criteria consensus* conditions (Difference = 0.69, 95% CI = 0.10–1.29, $d = 0.50$, $t\ (78) = 2.31$, $p = 0.02$; $M_{high\text{-}consensus} = 3.88$ and $M_{low\text{-}consensus} = 3.19$). Similarly, a *t*-test also suggested a significant difference in *investment amount* between the two *evaluation criteria consensus* conditions (Difference = 225,000.00, 95% CI = 71,900.00–378,040.00, $d = 0.63$, $t\ (80) = 2.94$, $p = 0.004$; $M_{high\text{-}consensus} = 437{,}500.00$ and $M_{low\text{-}consensus} = 212{,}500.00$). Thus, evaluation criteria consensus among experts increased both investors' investment intention and their investment amount. Results from this experiment provide causal evidence that evaluation criteria consensus increases investments.

Study 2

The aim of Study 2 was twofold. First, we sought to replicate the positive effect of evaluation criteria consensus on investment (Hypothesis 1) in another national context (the United States) and with a larger and more diverse sample. Second, we examined two mechanisms producing this effect: investors' perceptions of a firm's technological competence (Hypothesis 2) and their perceptions of others' positive evaluations of the firm (Hypothesis 3). The design and analysis plans were preregistered and can be accessed at https://osf.io/wnrd7/?view_only=5c458adf013c4b108e637f42f53c0a7d.

Sample

A total of 423 full-time employees with managerial experience in the United States drawn from the online platform Prolific participated in our experiment. Prolific connects researchers with their target participants, who earn rewards for participating in studies. Prolific is commonly used by management researchers because it enables fast, reliable, and high-quality data collection. We restricted participants to people with managerial experience because research shows such experience exposes people to entrepreneurship-relevant information and resources (Tonoyan et al. 2019), and thus prepares them for work related to equity investment. Each participant received a $0.60 reward, which is typical for participants on this and similar platforms. As participants preregistered, we included the same attention checks as those in Study 1 and excluded participants that failed any of the checks. A total of 412 (97.40%) participants passed the attention checks and were included in our analysis.[3] Of these, 178 were female, 226 were male, and 8 other (M_{age} = 37.14, SD = 10.62). We also replicated Study 2 with a sample of 559 online participants whose demographic profiles closely resemble those of professional investors—the participants in the General Social Survey (GSS; 1972–2018) who worked in the "securities, commodities, funds, and trusts" industry. The results are similar to those reported here.

Design and Procedures

As Study 1, this study was set up on Qualtrics. Participants were randomly assigned to one of two conditions in a one-factor two-level (evaluation criteria consensus: high versus low) between-subject experimental design. The scenario and manipulations for evaluation criteria consensus were the same as those in Study 1. Similarly, participants completed a survey that measured investment intention, investment amount, attention checks, a manipulation check, and participants' demographic information. In addition to the measures in the previous study, participants completed the scales of the two proposed mediators: perceived technological competence of the firm and prediction of others' positive evaluation of the firm. The mediators were perceptions of participants and were not manipulated. Online Appendix D provides details on the measures.

Measures[4]

Investment Intention. The items were the same as in Study 1 (α = 0.96).

Investment Amount. The question was the same as in Study 1 (available funds: $200,000).

Perceived Technology Competence. Participants rated the extent to which they agreed or disagreed with three scale items, such as, "This firm has demonstrated extraordinary technology competence" (α = 0.88).

Prediction of Others' Positive Evaluation. Participants also indicated the extent to which they agreed or disagreed with three scale items, such as, "Other people in my field (e.g., my boss, colleagues, and other investors) will be willing to invest money in this firm" (α = 0.95).

Manipulation Check. The items were the same as in Study 1 (α = 0.96).

Results and Discussion

Manipulation Check. A *t*-test showed a significant difference in perceived evaluation criteria consensus between the high and the low conditions (Difference = 3.93, 95% CI = 3.69–4.17, d = 1.68, t (410) = 31.68, $p <$ 0.001; $M_{high\text{-}consensus}$ = 5.93 and $M_{low\text{-}consensus}$ = 2.00), indicating that our manipulation worked as intended.

Main Effect of *Evaluation Criteria Consensus.* Replicating the results for Hypothesis 1 in Study 1, a *t*-test revealed a significant difference between the high and low consensus condition in *investment intention* (Difference = 0.60, 95% CI = 0.31–0.88, d = 0.40, t (410) = 4.14, $p <$ 0.001; $M_{high\text{-}consensus}$ = 4.76 and $M_{low\text{-}consensus}$ = 4.16) and in *investment amount* (Difference = 20,103.00, 95% CI = 11,674.07–28,530.95, d = 0.45, t (410) = 4.69, $p <$ 0.001; $M_{high\text{-}consensus}$ = 64,492.80 and $M_{low\text{-}consensus}$ = 44,390.02).

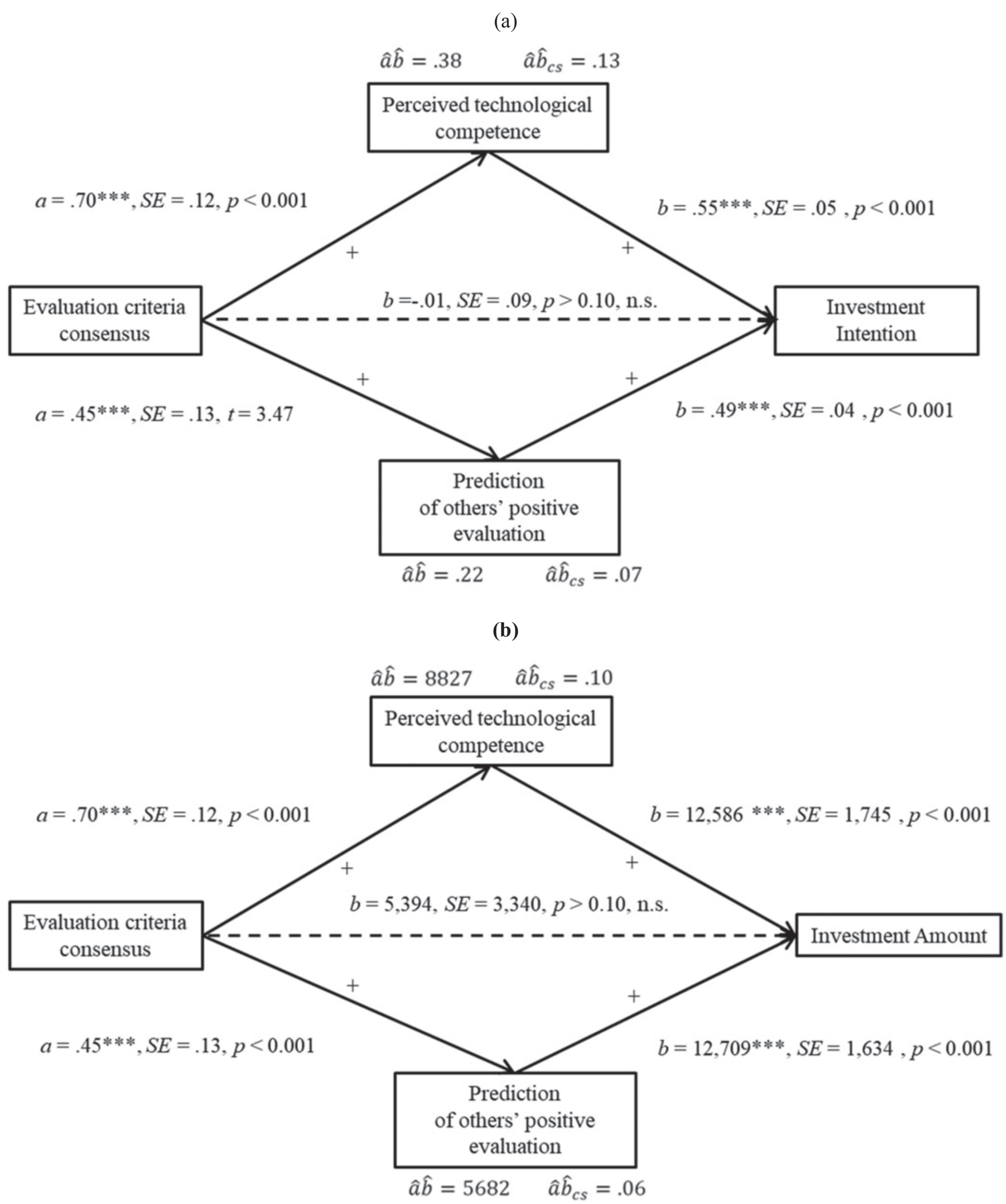
Mediation Analysis. To test *perceived technological competence* and *prediction of others' positive evaluation* as the mediators, we conducted mediation analysis with 5,000 bootstrap samples (MacKinnon and Dwyer 1993, MacKinnon et al. 2002, Preacher and Hayes 2004, Hayes 2017). Supporting Hypothesis 2, *perceived technological competence* mediated the effects of *evaluation criteria consensus* on both *investment intention* (indirect effect $\hat{a}\hat{b} = 0.63$, 95% CI = 0.41–0.84, $p <$ 0.001, standardized indirect effect $\hat{a}\hat{b}_{cs} = 0.21$) and *investment amount* (indirect effect $\hat{a}\hat{b} = 15{,}129.00$, 95% CI = 9,770.00–20,510.00, $p <$ 0.001, standardized indirect effect $\hat{a}\hat{b}_{cs} = 0.17$). Likewise, and in line with Hypothesis 3, *predictions of others' positive evaluation* also mediated the effects of *evaluation criteria consensus* on both *investment intention* ($\hat{a}\hat{b} = 0.37$, 95% CI = 0.17–0.57, $p <$ 0.001, $\hat{a}\hat{b}_{cs} = 0.12$) and *investment amount* ($\hat{a}\hat{b} = 9{,}170.00$, 95% CI = 4,170.00–14,270.00, $p <$ 0.001, $\hat{a}\hat{b}_{cs} = 0.10$). When we tested the mediation effects of both *perceived technology competence* and *predictions of* *others' positive evaluation* simultaneously, as shown in Figure 3, both mediating effects remained significant when predicting both *investment intentions* ($\hat{a}\hat{b} = 0.38$, 95% CI = 0.25–0.54, $\hat{a}\hat{b}_{cs} = 0.13$ and $\hat{a}\hat{b} = 0.22$, 95% CI = 0.10–0.35, $\hat{a}\hat{b}_{cs} = 0.07$; $p < 0.001$, respectively) and *investment amount* ($\hat{a}\hat{b} = 8,827.00$ and $\hat{a}\hat{b}_{cs} = 0.10$, 95% CI = 5,451.00–12,654.00 and $\hat{a}\hat{b} = 5,682.00$, $\hat{a}\hat{b}_{cs} = 0.06$, 95% CI = 2,604.00–9,116.00, $p < 0.001$, respectively).

In a different national context (i.e., the United States), and with a larger and more diverse sample of full-time employees with managerial experience, Study 2 replicated the results of Study 1 by showing that evaluation criteria consensus boosts investments. More importantly, Study 2 demonstrated that the positive effect of evaluation criteria consensus on investment is mediated by both investors' perceptions of a firm's technological competence and their predictions of others' positive evaluations of the firm.

Studies 3a and 3b

Studies 3a and 3b aimed to establish the external validity of our experimental findings supporting Hypothesis 1. Using two separate archival data sets on investments in Chinese and U.S. firms, we tested the relationship between evaluation criteria consensus

Figure 3. Study 2. Results of the Mediation Analysis. (a) Predicting Investment Intention; (b) Predicting Investment Amount



Notes. $N = 412$.
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

and investments in AI firms. To the extent that results from both the United States (Study 3a) and China (Study 3b) converge to support Hypothesis 1, we gain increased confidence in our experimental findings' robustness and generalizability.

Setting

These studies use data on new firms developing and using AI technologies. AI is a generic name for technologies that rely on computer programs modeled after human brains to draw inferences from big data (Poole et al. 1998; Legg and Hutter 2007a, b; Russell and Norvig 2016; WIPO 2019). These technologies have been used to develop products in various industrial contexts, such as security, healthcare, and finance. For example, computer vision is an AI technology that trains algorithms to recognize and classify different types of objects. It has been used to develop intelligent cameras as part of surveillance systems.

We chose this field as our research context because its potential investors often face high uncertainty in evaluating firms that develop and commercialize AI-based technologies. Since the early 2000s, as part of the effort to further the development of AI technologies, academic and industry experts have organized competitions at academic conferences aimed at comparing different AI-based programs tackling the same technological problem. The evaluation metrics proposed in these competitions influence both academia and industry. For example, an annual competition that assesses computer programs' performance in classifying objects started in 2010. In 2015, Microsoft Research cited their performance on a metric proposed in this competition (namely, top-5 test error) to claim that they have surpassed human capacity in classifying images (He et al. 2015).

Data and Sample

Our sample consisted of AI firms in the United States (Study 3a) and China (Study 3b), and we utilized data from several sources. First, we collected all evaluation metrics proposed at all 434 technology competitions at 19 top AI-themed academic conferences between 2003, when the first competition was held, and 2019. These competitions involved four major AI technologies: computer vision and pattern recognition, natural language processing, machine learning, and robotics. To create the list of top AI conferences, we first compiled a list of AI conferences that were ranked in the top 20 in each of the four categories according to their h5-index. The h5-index for a conference is the largest number (h) of published papers from a conference that have each been cited at least h times in the past five years. We then collected another list of "top-tier recommended conferences" under the AI category by the China Computer Federation. Our final list included the overlapping conferences from these two lists.

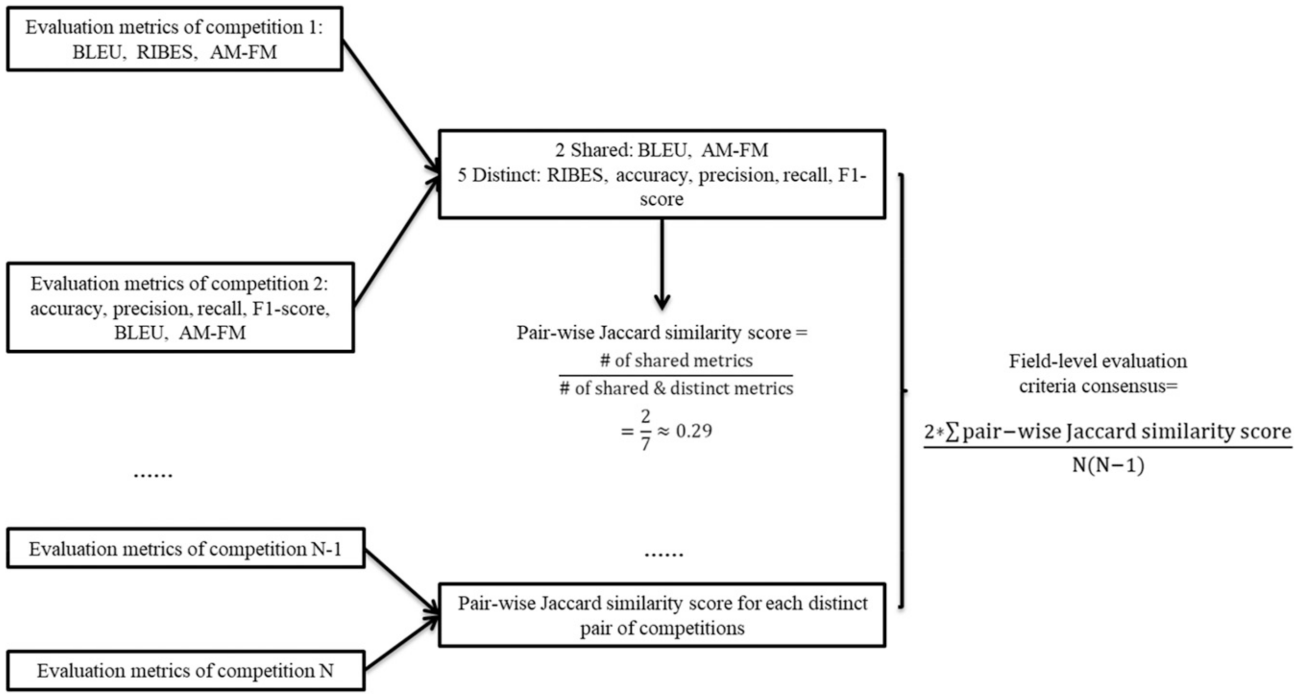
Second, we collected data on investment in AI firms from two major business information platforms: Crunchbase, for firms headquartered in the United States (Hallen et al. 2014, Nuscheler et al. 2019), and JuziIT, for firms in China. We chose Crunchbase because it contains more comprehensive information on startups, particularly the ones in the early-stage, than other databases (Hallen 2008, Gallagher 2013). Similarly, we chose JuziIT because it has been widely cited as the most complete and trustworthy information source for AI firms in China (Scmpnews 2017, Technode 2017). We included all firms identified as AI-focused by each platform. We obtained data on firm characteristics from both platforms. These data were collected by the platforms' data-scraping programs and contributed by firms and investors. They are regularly verified and updated by the platforms. Finally, we obtained data on firm patents from the U.S. Patents and Trademark Office database and the China National Intellectual Property Administration.

Measures

Investment. Following previous studies (Pontikes 2012, Wry et al. 2014), we measured investment in a firm with a binary variable, *investment*, with 1 indicating that the firm received investment(s) in a particular year and 0 otherwise. We decided not to use investment amount because these figures tend to be inaccurate and exaggerated (Wry et al. 2014).

Evaluation Criteria Consensus. Evaluation criteria consensus is operationalized as the overall similarity among criteria proposed for evaluating the same technology across different competitions. Following Hsu et al. (2012), we calculated this variable in four steps, as shown in Figure 4. First, we obtained the evaluation metrics proposed in all competitions. Two computer scientists with PhD training calibrated different expressions of the same metric. Second, we compared each competition with all other competitions involving the same technology in each year to calculate a pair-wise Jaccard similarity score (Niwattanakul et al. 2013, Bikard 2018, Bikard et al. 2019). This indicates the degree of overlap between the two sets of metrics proposed in the two competitions. Online Appendix E provides details about Jaccard similarity score calculations. Third, we summed all pair-wise Jaccard similarity scores and divided this by the number of pair-wise comparisons to get the average Jaccard score for each technology. Finally, for each firm, we calculated evaluation criteria consensus by summing Jaccard scores for the technology(ies) the firm identifies with and dividing the sum by the number of its claimed technology(ies). We matched each firm to one or more of the

Figure 4. Studies 3a and 3b. An Example of the Calculation of Evaluation Criteria Consensus



four technology(ies) based on the keywords identified by Crunchbase/JuziIT. To allow time for the market to respond, we lagged this variable for one year. Table 1 shows the keyword matching scheme.

Control Variables. We controlled for firm-level and field-level variables that may influence investments in a firm. At the firm level, we controlled for characteristics that may lead investors to favor a firm or make a firm self-select into a technology field. First, to capture signals about a firm's technological capability that may affect investment decisions (Busenitz et al. 2005, Conti et al. 2013, Hoenig and Henkel 2015), we controlled for the *number of patents logged* (Hsu and Ziedonis 2008),[5] *number of founders with a PhD degree*, and *number of founders with science or engineering degree(s)* (Hsu 2007, Eesley et al. 2014).

We also controlled for the prestige of founders' universities and past employers. The former is measured by *the number of founders from prestigious universities* (Tyebjee and Bruno 1984, MacMillan et al. 1985, Navis and Glynn 2011, Eesley et al. 2014). The latter is measured by *the number of founders from prestigious companies* to capture important status markers that have been found to influence investment decisions (Burton et al. 2002). Online Appendix F provides a list of prestigious universities and companies. Research has indicated that some investors favor firms founded by serial entrepreneurs (Wright et al. 1997, Hsu 2007). It is also possible that serial entrepreneurs are more experienced in identifying technology fields that attract

Table 1. Studies 3a and 3b. Matching Firm Technology with Technology Categories

Technology categories	Firm technology
Computer vision and pattern recognition	Computer vision, Image recognition, Image processing, Machine vision, Video, Facial recognition, Speech recognition
Machine learning	Machine learning, Deep learning, Big data, Data mining, Data analytics, Predictive analytics
Natural language processing	Natural language, Translation, Text analytics
Robotics	Robotics

Notes. To ensure that the keyword matching scheme is comprehensive, for the U.S. sample, we consulted three tenured professors and three PhD students in Computer Science, whose research focuses on AI, at a top research university in the United States. For the Chinese sample, we consulted the chair of the AI subdivision of the China Computer Federation, who is also a full professor at a top research university in China. Experts in both the United States and China verified the comprehensiveness of our matching scheme.

investors' interest. Hence, we included a dummy variable, *serial entrepreneur*, with 1 indicating at least one of the founders of the firm was a serial entrepreneur. We included a dummy variable, *C-round or after,* with 1 indicating firms at the C-round or after, as a firm's stage of development is associated with both its likelihood of obtaining investments and the development of its field (Jeng and Wells 2000, Davila et al. 2003). Finally, to account for industry idiosyncrasies, we controlled for the industrial sector(s) associated with a firm. Each of the two databases (Crunchbase and JuziIT) provided an industry classification system similar to the North American Industry Classification System (NAICS), which we used to create dummy variables.[6]

A large number of studies have provided evidence that resource providers often react to how easy it is to classify a firm into a distinct market or organizational categories (e.g., Zuckerman 1999, Hsu 2006, Pontikes 2012, Noh and Tolbert 2019). Pontikes (2012) specifically indicates that investors prefer firms that are associated with "fuzzy" categories (i.e., those comprised of firms that span multiple market categories). It could be that fuzzy categories are also likely to be characterized by low evaluation consensus. Given the potential correlation between category fuzziness and our independent variable—evaluation criteria consensus—we included a measure of a firm's category fuzziness as a control variable. Using Pontikes' methodology, we calculated the fuzziness of a technology category as the inverse of the sum of partial memberships divided by the number of members in the category. The partial membership of a firm to a category is the inverse of the total number of categories to which the firm belongs. For each firm, *category fuzziness* is measured by the sum of the fuzziness of all technology categories it claims divided by the total number of claimed categories.

Finally, at the field-level, evaluation criteria consensus is intricately intertwined with efforts to organize criteria-setting competitions, which could also affect investments in firms. Thus, we controlled for the *number of technology competitions*, measured by the number of competitions involving any of a firm's technologies in a year. To allow time for the market to respond, we lagged this variable for one year.

Estimation

Following past research (Wry and Lounsbury 2013, Hallen et al. 2014), we modeled the effects of evaluation criteria consensus on firms' likelihood of obtaining investments using Cox hazard rate survival models (Cox 1972, Cleves et al. 2008).[7] Our observation window begins in 1999, before which there had been fewer than five investments in any AI firms in the United States/China. Firms founded after 1999 enter the risk-set in their founding years. Firms leave the risk-set in the year they closed or were acquired. We analyzed the relationship between evaluation criteria consensus and investment from 2004 (the year following the first technology competition) to 2019. As noted above, we lagged our two competition-related variables, evaluation criteria consensus, and the number of technology competitions for one year. As the data set contains repeated observations of firms, we clustered standard errors by firms. We included year-fixed effects to account for temporal variation. There are 27,837 firm-year observations from 4,905 firms in the U.S. sample and 6,140 firm-year observations from 1,345 firms in the Chinese sample. We conducted our analyses using Stata 14.2 with the stcox command.

Results

Table 2 presents the descriptive statistics and correlations for the variables included in our analyses. Variance inflation factor (VIF) tests showed that none of the models we estimated had VIF scores over 10, the recommended threshold (Gujarati and Porter 2003).

Tables 3–5 present the results of our regression analyses. For ease of interpretation and comparison, standardized coefficients are presented. Models 1a and 2a in Table 3 present results from the Cox hazard rate regression analysis of investment in AI firms in the U.S. sample. Models 1b and 2b in Table 4 present results from the Chinese sample. Models 1a and 1b show the effects of control variables. Most of the firm-level controls have significant effects on investment in the expected directions in both samples. We did not find a significant positive effect of *category fuzziness* in either sample, contrary to Pontikes (2012). At the field level, the *number of technology competitions* has nonsignificant effects on *investment* in both samples. This suggests that more efforts in proposing evaluation criteria do not necessarily lead to more investments in firms.

Models 2a (in Table 3) and 2b (in Table 4) test Hypothesis 1, which predicts evaluation criteria consensus to have a positive effect on investment. In both samples, *evaluation criteria consensus* has a significant, positive effect on *investment*, net of all controls (β = 0.02, $p < 0.05$ in both samples).

Robustness Checks. We conducted additional analyses to ensure that our results remained robust to alternative explanations and specifications. First, we considered the possibility that the positive effect of evaluation criteria consensus is driven by a technology field's growth. This logic is in line with population ecologists' claim that growth in a nascent field increases its visibility and legitimacy, making all firms in the field more appealing to investors (Hannan and Carroll 1992,

Published in Organization Science on September 27, 2021 as DOI: 10.1287/orsc.2021.1493.
This article has not been copyedited or formatted. The final version may differ from this version.

Table 2. Studies 3a and 3b. Descriptive Statistics and Correlations of Variables

		Mean	SD	1	2	3	4	5	6	7	8	9	10
	Part A: The U.S. sample (N = 27,837)												
1	Investment (1 = Yes, 0 = No)	0.03	0.17										
2	Evaluation criteria consensus	0.33	0.37	0.04									
3	Number of founders from prestigious universities	0.18	0.48	0.12	0.04								
4	Number of founders from prestigious companies	0.13	0.40	0.09	0.01	0.34							
5	Number of founders with a PhD	0.05	0.24	0.08	0.03	0.39	0.17						
6	Number of founders with science/engineering degree(s)	0.13	0.41	0.08	0.03	0.40	0.34	0.22					
7	Serial entrepreneur (1 = Yes, 0 = No)	0.06	0.24	0.02	0.01	0.09	0.09	0.05	0.07				
8	Log (Number of patents)	0.10	0.53	−0.01	−0.01	0.06	0.08	0.01	0.04	0.00			
9	C-round or after (1 = Yes, 0 = No)	0.00	0.01	0.07	−0.00	0.01	0.00	0.02	0.00	0.00	0.00		
10	Category fuzziness	0.79	0.02	0.02	0.16	0.04	0.03	−0.02	0.06	0.02	0.00	0.00	
11	Number of technology competitions	6.97	9.61	0.00	0.21	0.01	0.03	0.04	0.00	0.02	−0.01	−0.01	0.16
	Part B: Chinese sample (N = 6,140)												
1	Investment (1 = Yes, 0 = No)	0.14	0.34										
2	Evaluation criteria consensus	0.23	0.37	0.15									
3	Number of founders from prestigious universities	0.11	0.62	0.20	0.03								
4	Number of founders from prestigious companies	0.08	0.50	0.24	0.05	0.19							
5	Number of founders with a PhD	0.06	0.30	0.26	0.07	0.48	0.29						
6	Number of founders with science/engineering degree(s)	0.02	0.15	0.15	0.06	0.20	0.25	0.29					
7	Serial entrepreneur (1 = Yes, 0 = No)	0.01	0.12	0.11	0.03	0.04	0.16	0.05	0.02				
8	Log (Number of patents)	1.61	7.18	0.02	0.07	0.00	−0.01	0.01	−0.01	−0.01			
9	C-round or after (1 = Yes, 0 = No)	0.03	0.16	0.06	−0.01	0.10	0.08	0.06	−0.02	0.01	0.02		
10	Category fuzziness	0.87	0.01	0.05	0.10	0.04	0.04	0.04	0.02	0.02	0.03	0.05	
11	Number of technology competitions	5.40	9.49	0.03	0.21	−0.02	0.00	0.02	0.03	−0.01	0.00	−0.03	0.08

Notes. For brevity, industry dummy indicators are omitted. All coefficients with an absolute value above 0.05 are significant at $p < 0.001$. Two-tailed test.

Table 3. Study 3a. Models of Investments in AI Firms in the U.S. Sample

Variables	Model 1a	Model 2a	Model 3a	Model 4a	Model 5a	Model 6a
Evaluation criteria consensus		0.02*	0.05***	0.04***	0.02*	12.52*
		(0.11)	(0.09)	(0.12)	(0.11)	(62.73)
Year			0.01			0.02*
			(0.01)			(0.01)
Evaluation criteria consensus * year						−12.47*
						(0.03)
Number of organizations using a focal firm's technology(ies) (Mean-centered)				0.09***		
				(0.00)		
Number of organizations using a focal firm's technology(ies) (Mean-centered, squared)				−0.09***		
				(0.00)		
Firm-level controls						
Number of founders from prestigious universities	0.05***	0.04***	0.05***	0.04***	0.04***	0.04***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Number of founders from prestigious companies	0.04***	0.04***	0.04***	0.04***	0.04***	0.04***
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Number of founders with a PhD	0.01*	0.01*	0.01*	0.01*	0.01*	0.01*
	(0.09)	(0.09)	(0.10)	(0.09)	(0.09)	(0.10)
Number of founders with science/engineering degree(s)	0.02**	0.02**	0.02**	0.02**	0.02**	0.02**
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Serial entrepreneur (1 = Yes, 0 = No)	0.01	0.01	0.01	0.01	0.01	0.01
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Log (Number of patents)	−0.00	−0.00	−0.01	−0.01	−0.01	−0.01
	(0.08)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)
C-round or after (1 = Yes, 0 = No)	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.34)	(0.33)	(0.30)	(0.27)	(0.31)	(0.30)
Category fuzziness	0.01	0.00	0.01	−0.00	0.00	0.01
	(2.40)	(2.46)	(2.44)	(2.47)	(2.47)	(2.45)

Table 3. (Continued)

Variables	Model 1a	Model 2a	Model 3a	Model 4a	Model 5a	Model 6a
Field-level controls						
Number of technology competitions	0.01	0.00	−0.02*	−0.01	0.00	−0.02
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	No	No	Yes	No
Log pseudolikelihood	−6,579.15	−6,576.11	−6,618.41	−6,609.50	−2,465.48	−6,200.30

Notes. N = 27,837. Model 1a–4a and Model 6a are Cox hazard rate models. Model 5a is a piecewise hazard rate model. In Model 5a, the time pieces are: 0–1 year since founding, 1–2 years since founding, 2–5 years since founding, 5–10 years since founding, and more than 10 years since founding. Industry dummies are included but omitted from the table for brevity. β coefficients are reported. Standard errors are in parentheses. Robust standard error was clustered by firm.

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. Two-tailed tests.

Hannan et al. 1995). We controlled for field growth in two different ways. In Model 3a in Table 3 (the U.S. sample) and Model 3b in Table 4 (the Chinese sample), we replaced year dummies with a continuous variable, *year,* to capture the passage of time as a proxy of field growth. Supporting Hypothesis 1, the positive effect of *evaluation criteria consensus* remained significant in both models (β = 0.05, p < 0.001 in the U.S. sample and β = 0.03, p < 0.001 in the Chinese sample). In Models 4a and 4b, we controlled for *the number of organizations using a focal firm's technology(ies) (mean-centered)* and the squared term of this variable to

Table 4. Study 3b. Models of Investments in AI Firms in the Chinese Sample

Variables	Model 1b	Model 2b	Model 3b	Model 4b	Model 5b	Model 6b
Evaluation criteria consensus		0.02*	0.03***	0.02*	0.01*	19.14***
		(0.10)	(0.09)	(0.11)	(0.10)	(65.37)
Year			0.13***			0.14***
			(0.01)			(0.01)
Evaluation criteria consensus * year						−19.11***
						(0.03)
Number of organizations using a focal firm's technology(ies) (Mean-centered)				0.11***		
				(0.00)		
Number of organizations using a focal firm's technology(ies) (Mean-centered, squared)				−0.04**		
				(0.00)		
Firm-level controls						
Number of founders from prestigious universities	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Number of founders from prestigious companies	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Number of founders with a PhD	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***
	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.04)
Number of founders with science/engineering degree(s)	0.01	0.00	0.02	0.00	0.00	0.01
	(0.14)	(0.14)	(0.14)	(0.13)	(0.14)	(0.13)
Serial entrepreneur (1 = Yes, 0 = No)	0.01***	0.01***	0.02***	0.02***	0.01***	0.02***
	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)
Log (Number of patents)	−0.00	−0.00	0.01	0.01*	−0.00	0.01
	(0.001)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
C-round or after (1 = Yes, 0 = No)	0.03***	0.03***	0.03***	0.02***	0.03***	0.02***
	(0.16)	(0.16)	(0.14)	(0.14)	(0.17)	(0.14)
Category fuzziness	0.03	0.03	0.03	0.02	0.03	0.03
	(12.97)	(12.31)	(11.12)	(7.10)	(12.23)	(11.81)
Field-level controls						
Number of technology competitions	0.01	0.02	−0.02*	−0.03**	0.01	−0.02*
	(0.002)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	No	No	Yes	No
Log pseudolikelihood	—5,196.53	—5,195.02	—5,313.38	—5,365.12	—1,384.52	—5,309.92

Notes. N = 6,140. Model 1b–4b and Model 6b are Cox hazard rate models. Model 5b is a piecewise hazard rate model. In Model 5b, the time pieces are: 0–1 year since founding, 1–2 years since founding, 2–5 years since founding, 5–10 years since founding, and more than 10 years since founding. Industry dummies are included but omitted from the table for brevity. β coefficients are reported. Standard errors are in parentheses. Robust standard error was clustered by firm.

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. Two-tailed tests.

Published in Organization Science on September 27, 2021 as DOI: 10.1287/orsc.2021.1493.
This article has not been copyedited or formatted. The final version may differ from this version.

capture more nuanced field-growth dynamics (Hannan et al. 1995). Because these two variables are highly correlated with the year dummies and yield VIF scores over 10 (Gujarati and Porter 2003), we removed year dummies. Again, supporting Hypothesis 1, the effect of *evaluation criteria consensus* remained significantly positive in both models ($\beta = 0.04$, $p < 0.001$ in the U.S. sample and $\beta = 0.02$, $p < 0.05$ in the Chinese sample).

One potential issue with the Cox hazard rate survival model is that it assumes the base rate of receiving investment to be constant over time. A piecewise exponential hazard rate model allows the base rate of funding to vary and thus does not require a strong assumption about the form of time dependence. Hence, in our third set of robustness checks, we replicated our analysis with piecewise exponential hazard rate models. Results are shown in Model 5a in Table 3 (for the U.S. sample) and Model 5b in Table 4 (for the Chinese sample). Again, speaking to the robustness of our effect, both models show significantly positive effects of *evaluation criteria consensus* on *investment* ($\beta = 0.02$, $p < 0.05$ in the U.S. sample and $\beta = 0.01$, $p < 0.05$ in the Chinese sample).

Additional Analyses. We also explored how the effect of evaluation criteria consensus might vary over time, among different investors and for different investment syndicates, using our archival data sets. Two processes may induce changes. One is the coevolution of evaluation criteria consensus and investment. Although investors did not directly participate in the development of technology evaluation criteria (academic and industry scientists did), increasing investments in a technology field may attract standard-setting efforts from technology experts with diverse backgrounds who propose competing evaluation criteria (Grodal 2018). This could result in reducing the correlation between investment and evaluation criteria consensus over time. A second process involves investor learning. As time goes by, investors may accumulate experience and domain-specific knowledge in nascent technology fields, feel less uncertain about how to assess firms in the sector, and rely less on field-level signals such as evaluation criteria consensus. In sum, both processes imply that the observed positive effect of evaluation criteria consensus would decline with time.

We present results of analyses examining these potential effects in Model 6a in Table 3 (for the U.S. sample) and Model 6b in Table 4 (for the Chinese sample). In both models, we used the continuous variable, *year*, to capture the passage of time. The negative, significant coefficients of the interaction term of *evaluation criteria consensus* and time in Models 6a and 6b indicate that the impact of expert evaluation criteria consensus on investors did decline with time ($\beta = -12.47$, $p < 0.05$ in the U.S. sample and $\beta = -19.11$, $p < 0.001$ in the Chinese sample). Hence, the impact of evaluation criteria consensus is the strongest when a technology field is newly formed.

The effect of evaluation criteria consensus may also vary for different types of investors (Benner and Ranganathan 2013, Benner and Zenger 2016, Theeke et al. 2018). Experienced investors are likely to have routines that help them quickly develop domain-specific knowledge about a nascent field (Levitt and March 1988, Cohen and Levinthal 1990, Cohen 1991). Similarly, professional venture capital firms are apt to have better routines for assessing the value of start-up firms than other types of investors (e.g., public institutions or individual investors) (Sykes 1990, Gompers and Lerner 2000). As experienced and professional investors are better at assessing venture firms, they may rely less on field-level signals such as evaluation criteria consensus. Therefore, the positive effect of evaluation criteria consensus should be weaker for more experienced investors and professional venture capital firms. We tested these predictions with the U.S. sample[8] and presented the results in Table 5. We looked at two types of investor experience: *general experience* and *AI-specific experience.* Both were continuous variables measuring the average years of general/AI-specific experience among actors investing in a focal firm in a given year. They were calculated by taking the difference between the year of observation and the founding year of the organization (for general experience) and the difference between the year of observation and the year the organization making its first investment in an AI firm (for AI-specific experience), respectively. *Professional VC firm* is a dummy variable with 1 indicating that there are one or more professional venture capital firms among the investors that invest in a firm in a year.

The significant, negative coefficients of the interaction terms for *evaluation criteria consensus* and *general experience* (in Model 7) and for *evaluation criteria consensus* and *AI-specific experience* (in Model 8) suggest that the effect of evaluation criteria consensus weakens as the general or AI-specific investment experience of investors increases ($\beta = -0.01$, $p < 0.01$ for both). Similarly, in Model 9, the significant, negative coefficient of the interaction term indicates that the investments by professional VC firms are less influenced by expert *evaluation criteria consensus* ($\beta = -0.02$, $p < 0.05$).

Third, we explored how the effect of evaluation criteria may vary for investment syndicates of different sizes. We theorized that investors' concerns over others' opinions were one of the reasons why they rely on field-level evaluation criteria consensus to make investments. In large investment syndicates, because more parties need to be persuaded, the concerns over others' opinions should be stronger. Therefore, we predict the effect of evaluation criteria consensus

Published in Organization Science on September 27, 2021 as DOI: 10.1287/orsc.2021.1493.
This article has not been copyedited or formatted. The final version may differ from this version.

Table 5. Study 3a. Contingent Effects of Evaluation Criteria Consensus in the U.S. Sample

Variables	Model 7	Model 8	Model 9	Model 10
Evaluation criteria consensus	0.04***	0.04***	0.05*	0.00
	(0.12)	(0.13)	(0.22)	(0.12)
General experience	0.07***			
	(0.00)			
Evaluation criteria consensus * general experience	−0.01**			
	(0.01)			
AI-specific experience		0.10***		
		(0.01)		
Evaluation criteria consensus * AI-specific experience		−0.01**		
		(0.01)		
Professional VC firm			0.25***	
			(0.13)	
Evaluation criteria consensus * Professional VC firm			−0.02*	
			(0.22)	
Syndicate size				0.06***
				(0.02)
Evaluation criteria consensus * Syndicate size				0.02***
				(0.02)
Controls	Yes		Yes	Yes
Industry dummies	Yes		Yes	Yes
Year fixed effects	Yes		Yes	Yes
Log pseudolikelihood	−6285.77		−5465.90	−6200.30

Notes. $N = 27{,}837$. All models are Cox hazard rate models. Control variables and industry dummies are included but omitted from the table for brevity. β coefficients are reported. Standard errors are in parentheses. Robust standard error was clustered by firm.

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. Two-tailed tests.

to be stronger for larger syndicates. We tested this prediction and presented the results in Model 10 in Table 5. *Syndicate size* is a continuous variable measuring the average syndicate size of the investments that a firm received in a year. The significantly positive coefficient of the interaction term of *evaluation criteria consensus* and *syndicate size* suggests that the effect of evaluation criteria consensus is stronger for larger syndicates ($\beta = 0.02$, $p < 0.001$), confirming our prediction.

Discussion

This paper investigated an understudied phenomenon in nascent technology fields: how consensus among experts on the criteria for evaluating new technologies affects investors' willingness to invest in firms in the field. We theorized that consensus on evaluation criteria increases willingness to invest by enhancing investors' perceptions of a firm's technological competence and their confidence that their peers also view the firm positively. We found support for these predictions using data from two experiments and two archival studies in the United States and China. The consistency of findings across different methodologies and national contexts boosts the validity and generalizability of our findings.

Theoretical Contributions

Our findings have implications for several on-going streams of organizational research, including work on nascent market sectors, microinstitutional theory, and technological evolution. First, this paper provides a microfoundation to the literature on evaluation criteria in nascent market sectors. Existing work on nascent market sectors has shown that individual/collective actions can reduce the inherent uncertainties of a nascent sector (Santos and Eisenhardt 2009, Dattée et al. 2018, Hannah and Eisenhardt 2018, Ozcan and Gurses 2018, McDonald and Eisenhardt 2019). In this context, a number of studies have documented the important role played by the development of evaluation criteria, particularly in industries and markets that rely on nascent technologies (Rao 1994, Sine et al. 2007, Lee and Sine 2012, Goldfarb et al. 2018). However, the processes through which the emergence of evaluation criteria affects resource providers, and thus the flow of resources to new organizations that comprise a fledgling sector, have not been fully explicated nor explored. This paper addresses this gap. Our findings establish a causal impact of evaluation criteria consensus on investors' positive evaluations of ventures in a nascent market sector. This effect holds except in cases where a firm has been specifically identified as having subpar performance. More importantly, our analyses also reveal the socio-cognitive mechanisms involved: consensus in evaluation criteria for a new technology enhances both investors' private beliefs of a firm's competence that uses the technology and their beliefs that other investors will view the firm positively.

Second, by examining the socio-cognitive mechanisms involved in investors' assessment of ventures in nascent technology fields, this paper extends a theoretical stream of work that seeks to link microlevel individual evaluation to macro-level legitimation outcomes (Bitektine 2011, Bitektine and Haack 2015, Schilke 2018, Haack et al. 2020). These studies have proposed separate components of the legitimation of an entity: at the microlevel, an individual forms private belief about the appropriateness of the entity (propriety); at the meso-level, the individual's reference group agrees or disagrees on whether characteristics of the entity are proper (consensus); at the macro-level, a collective, generalized perception emerges about the appropriateness of the entity for its social context (validity). These different components separately and jointly shape the legitimation and de-legitimation of entities (Haack et al. 2020). Our analyses provide rigorous empirical validation of such a multilevel legitimation process: investors' endorsement of a firm in a nascent technology sector is driven by both their own beliefs about the appropriateness and viability of the firm's technology (propriety beliefs) as well as their anticipation of peer investors' positive assessment of the firm (consensus beliefs). Answering the call to bridge the micromacro link (Bitektine and Haack 2015, Haack et al. 2020), our study is among the first to show how microlevel socio-cognitive processes coalesce to produce macro-level economic consequences: by shaping the judgments of individual investors, variations in consensus in evaluation criteria translate into systematic differences in resource acquisition across technology fields.

The literature on technological evolution and industry emergence has established that technology standards and dominant designs (Anderson and Tushman 1990, Suárez and Utterback 1995, Tassey 2000, Vakili 2016) reduce uncertainty in a nascent sector by coordinating transactions among market actors. Recently, organizational scholars have begun to examine how the development of shared definitions, such categories (Grodal 2007, Suarez et al. 2015, Zunino et al. 2019), technology frames (Kaplan 2008, Kaplan and Tripsas 2008), and field schema and labels affect uncertainty in a nascent technology field (Bingham and Kahl 2013). We contribute to this burgeoning literature by investigating how evaluation criteria consensus influences firm resource acquisition. Our findings indicate that experts reaching consensus in the technology's evaluation criteria constitutes a critical milestone in the development of a nascent technology field.

Limitations and Future Research Directions

This study has a few limitations that suggest future research opportunities. First, this paper focused on the consequences of consensus in evaluation criteria rather than its antecedents. Thus, we did not investigate how consensus (versus disagreement) in evaluation criteria occur. When multiple evaluation criteria coexist in a nascent technology field, there are usually different underlying technology frames and individual/organizational actors with different vested interests (Campbell 1998, Hoffman 1999, Kaplan 2008, Kaplan and Tripsas 2008). The history of technology competitions in AI illustrates such a process. From the mid-2000s until 2013, many early competitions were organized by a core group of academics and industry researchers who carefully inherited past evaluation criteria. In 2013, the number of competitions doubled. After that, technology firms organized a substantial number of competitions, many of which proposed new criteria. Such changes fueled disagreements over the criteria for evaluating technology solutions. We call for future research investigating how evaluation criteria consensus forms or changes as a result of the criteria-setting efforts of different types of actors. Another promising research direction is the coevolution of interorganizational networks and evaluation criteria consensus. Studying these questions would not only answer the intriguing question of how consensus arises but also unveil the coevolution of networks, power, and institutions (Padgett and Powell 2012, Powell and Oberg 2017).

Second, we focus on the effect of field-level consensus on technology evaluation criteria on investments in all firms in the field. We chose this focus because this paper, as the first theorization and test of evaluation criteria consensus, aims to document its effect on the entire nascent sector. Nevertheless, we tested factors that explain within-field variance in the effects of evaluation criteria consensus. For example, in our archival studies (Studies 3a and 3b), we found that the effect varies over time and across investors and investment syndicates. Moreover, in our experimental Study 4 (reported in Online Appendix A), we explored how the effects of evaluation criteria consensus on investment vary across sectors with different levels of technology and industry nascence. In our experimental Study 5 (reported in Online Appendix B), we found that information about a firm's reported technological competence moderates the effects of evaluation criteria consensus on investment. These analyses point to how characteristics of investors, technologies, firms, and industries cause heterogeneity in the effects of evaluation criteria consensus. We encourage future researchers to test other potential moderators. For example, the substantial differences in resources, competencies, and legitimacy between *de alio* and de novo firms may influence the extent to which they benefit from a field-level consensus in evaluation criteria (Carroll et al. 1996, Helfat and Lieberman 2002, Khessina and Carroll 2008). Due to our focus on start-up firms, we did not examine this question. On a separate note, the consensus in evaluation criteria resolves two types of uncertainties in a nascent technology field—technical uncertainty (i.e., what features to measure) and standard uncertainty (i.e., what threshold to use) (Graffin and Ward 2010). The two types of uncertainties do not always covary. We encourage future research on how these two types of uncertainties moderate the effect of evaluation criteria consensus.

Conclusion

The past decade has witnessed an outpouring of scholarly interest in understanding the forces that enable the institutionalization of new industries and markets (e.g., Kennedy 2008, Padgett and Powell 2012, Grodal et al. 2015). Within this context, a substantial amount of work has examined the impacts of evaluation institutions on organizational survival and performance (Rao 1994, Sine et al. 2007, Lee and Sine 2012, Goldfarb et al. 2018). Although being insightful, these studies leave open how the emergence of evaluation criteria affect resources providers, and thus the flow of resources to firms that comprise a nascent sector. Our paper makes important steps toward further understanding the operation of evaluation institutions by unpacking the sociocognitive processes underlying investors' assessment of organizations in nascent technology sectors. Specifically, we propose a new construct, namely the field-level consensus in the criteria used to evaluate technological solutions, and show that such consensus boosts investment in firms in nascent sectors. By demonstrating how "converging tides lift all boats," we hope to stimulate future research that digs deeper into the processes and consequences of evaluation institutions in nascent fields.

Acknowledgments

The authors acknowledge research support from the Innovation, Entrepreneurship, and Technology Theme research grant and Jeffrey S. Lehman Fund for Scholarly Exchange with China at Cornell University. The authors thank Chris Marquis, Glen Dowell, Rodolphe Durand, Wesley Sine, Yuan Shi, and Shon Hiatt, who have provided helpful comments on earlier versions of this paper. The authors are appreciative of the guidance and suggestions from special issue editors Olenka Kacperczyk, Sheen Levine, Oliver Schilke, and Lynn Zucker and two anonymous reviewers.

Endnotes

[1] Per our definition, in order to have a high level of consensus in evaluation criteria, it is not necessary that experts agree on one dominant criterion; there could be multiple criteria. As long as experts agree on the relative importance of each one that a competent technology solution needs to satisfy, there is high consensus.

[2] The failure rate was relatively high for a few reasons: (1) the subjects were professional investors with a busy schedule, and as a result, their participation could be easily interrupted; (2) participation was voluntary, anonymous, and without monetary or other type of reward. Testing Hypothesis 1 on the full sample of participants, including those who failed the attention checks, yielded consistent results.

[3] The included and excluded respondents do not differ significantly in their age, investment experience, or gender. Including participants who failed the attention checks in our analysis does not change our results, speaking to the robustness of our findings.

[4] We conducted confirmative factor analysis using Mplus to validate that all items load onto their respective constructs as intended. The measurement model exhibited excellent fit: χ^2 (36) = 79.94, comparative fit index (CFI) = .986, root mean square error of approximation (RMSEA) = .095, and SRMR = .021. All item loadings were significant at $p < .001$ on their respective factors.

[5] Following the advice of an anonymous reviewer, we alternatively controlled for the number of invention patents for the Chinese sample and found consistent results.

[6] The sectors in Crunchbase are security, finance, health, education, home, marketing, law, customer service, automotive, manufacturing, logistics, agriculture, environment, tourism, media and culture, business-to-business, electricity and energy, architecture and real estate, insurance, aircraft and fashion. In JuziIT, they are security, finance, health, education, home, marketing, law, customer service, automotive, manufacturing, logistics, agriculture, environment, tourism, media, business-to-business, electricity and energy, architecture and real estate, social network, robots, apps for life, insurance, aircraft, and fashion.

[7] We chose Cox-proportional hazard rate survival models based on the assumption of no significant yearly change in the baseline hazard rates of receiving investment. Nevertheless, to ensure that potential changes in the baseline hazard rates do not bias our results, we ran piece-wise exponential hazard rate models and found consistent results.

[8] Detailed data on investor experience and characteristics were not available for the Chinese data, so we tested this set of predictions only using the U.S. data.

References

Aggarwal R, Kryscynski D, Singh H (2015) Evaluating venture technical competence in venture capitalist investment decisions. *Management Sci.* 61(11):2685–2706.

Aldrich HE, Fiol CM (1994) Fools rush in? The institutional context of industry creation. *Acad. Management Rev.* 19(4):645–670.

Anderson P, Tushman ML (1990) Technological discontinuities and dominant designs: A cyclical model of technological change. *Admin. Sci. Quart.* 35(4):604–633.

Baum JA, Silverman BS (2004) Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology startups. *J. Bus. Venturing* 19(3):411–436.

Benner MJ, Ranganathan R (2013) Divergent reactions to convergent strategies: Investor beliefs and analyst reactions during technological change. *Organ. Sci.* 24(2):378–394.

Benner MJ, Zenger T (2016) The lemons problem in markets for strategy. *Strategy Sci.* 1(2):71–89.

Bikard M (2018) Made in academia: The effect of institutional origin on inventors' attention to science. *Organ. Sci.* 29(5):818–836.

Bikard M, Vakili K, Teodoridis F (2019) When collaboration bridges institutions: The impact of university–industry collaboration on academic productivity. *Organ. Sci.* 30(2):426–445.

Bingham CB, Kahl SJ (2013) The process of schema emergence: Assimilation, deconstruction, unitization and the plurality of analogies. *Acad. Management J.* 56(1):14–34.

Bitektine A (2011) Toward a theory of social judgments of organizations: The case of legitimacy, reputation, and status. *Acad. Management Rev.* 36(1):151–179.

Bitektine A, Haack P (2015) The "macro" and the "micro" of legitimacy: Toward a multilevel theory of the legitimacy process. *Acad. Management Rev.* 40(1):49–75.

Bitektine A, Lucas J, Schilke O (2018) Institutions under a microscope: Experimental methods in institutional theory. Bryman A, Buchanan DA, eds. *Unconventional Methodology in Organization and Management Research* (Oxford University Press, Oxford, UK), 147–167.

Burton MD, Sørensen JB, Beckman CM (2002) Coming from good stock: Career histories and new venture formation. *Social Structure and Organizations Revisited* (Emerald Group Publishing Limited, Bingley, UK), 229–262.

Busenitz LW, Fiet JO, Moesel DD (2005) Signaling in venture capitalist—new venture team funding decisions: Does it indicate long–term venture outcomes? *Entrep. Theory Pract.* 29(1):1–12.

Campbell JL (1998) Institutional analysis and the role of ideas in political economy. *Theory Soc.* 27(3):377–409.

Carroll GR, Bigelow LS, Seidel MDL, Tsai LB (1996) The fates of de novo and de alio producers in the American automobile industry 1885–1981. *Strateg. Management J.* 17(S1):117–137.

CB Insights Research (2019) China is starting to edge out the US in AI investment. *CB Insights Research* (December 2), https://www.cbinsights.com/research/china-artificial-intelligence-investment-startups-tech/.

Chatterji AK, Durand R, Levine DI, Touboul S (2016) Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strateg. Management J.* 37(8):1597–1614.

Cleves M, Gould W, Gould WW, Gutierrez R, Marchenko Y (2008) *An Introduction to Survival Analysis Using Stata* (Stata Press, College Station, TX).

Cohen MD (1991) Individual learning and organizational routine: Emerging connections. *Organ. Sci.* 2(1):135–139.

Cohen WM, Levinthal DA (1990) Absorptive capacity: A new perspective on learning and innovation. *Admin. Sci. Quart.* 35(1): 128–152.

Conti A, Thursby M, Rothaermel FT (2013) Show me the right stuff: Signals for high-tech startups. *J. Econom. Management Strategy* 22(2):341–364.

Cox DR (1972) Regression models and life-tables. *J. R. Stat. Soc. B.* 34(2):187–202.

Cumming D, Johan SAB (2008) Preplanned exit strategies in venture capital. *Eur. Econom. Rev.* 52(7):1209–1241.

Dattée B, Alexy O, Autio E (2018) Maneuvering in poor visibility: How firms play the ecosystem game when uncertainty is high. *Acad. Management J.* 61(2):466–498.

Davila A, Foster G, Gupta M (2003) Venture capital financing and the growth of startup firms. *J. Bus. Venturing* 18(6):689–708.

DiMaggio PJ, Powell WW (1983) The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *Amer. Sociol. Rev.* 48(2):147–160.

Durand R, Kremp P-A (2016) Classical deviation: Organizational and individual status as antecedents of conformity. *Acad. Management J.* 59(1):65–89.

Eesley CE, Hsu DH, Roberts EB (2014) The contingent effects of top management teams on venture performance: Aligning founding team composition with innovation strategy and commercialization environment. *Strategic Management J.* 35(12): 1798–1817.

Fischhoff B, Slovic P, Lichtenstein S (1982) Lay foibles and expert fables in judgments about risk. *Amer. Statistician* 36(3b):240–255.

Gallagher B (2013) How CrunchBase data compares to other industry sources. *Tech Crunch* (July 23), http://social.techcrunch.com/2013/07/23/how-crunchbase-data-compares-to-other-industry-sources/.

Georgallis P, Dowell G, Durand R (2018) Shine on me: Industry coherence and policy support for emerging industries. *Admin. Sci. Quart.* 64(3):503–541.

Goldfarb B, Zavyalova A, Pillai S (2018) Did victories in certification contests affect the survival of organizations in the American automobile industry during 1895–1912? A replication study. *Strategic Management J.* 39(8):2335–2361.

Gompers P, Lerner J (2000) The determinants of corporate venture capital success: Organizational structure, incentives, and complementarities. *Concentrated Corporate Ownership* (University of Chicago Press, Chicago, IL), 17–54.

Graffin SD, Ward AJ (2010) Certifications and reputation: Determining the standard of desirability amidst uncertainty. *Organ. Sci.* 21(2):331–346.

Grodal S (2007) The emergence of a new organizational field: Labels, meaning and emotions in nanotechnology. Unpublished doctoral dissertation, Stanford University, Palo Alto, CA.

Grodal S (2018) Field expansion and contraction: How communities shape social and symbolic boundaries. *Admin. Sci. Quart.* 63(4): 783–818.

Grodal S, Gotsopoulos A, Suarez FF (2015) The coevolution of technologies and categories during industry emergence. *Acad. Management Rev.* 40(3):423–445.

Gujarati DN, Porter DC (2003) *Basic Econometrics*, 4th ed. (McGraw-Hill, New York).

Haack P, Schilke O, Zucker L (2020) Legitimacy revisited: Disentangling propriety, validity, and consensus. *J. Management Stud.* 58(3):749–781.

Hallen BL (2008) The causes and consequences of the initial network positions of new organizations: From whom do entrepreneurs receive investments? *Admin. Sci. Quart.* 53(4):685–718.

Hallen BL, Bingham CB, Cohen S (2014) Do accelerators accelerate? A study of venture accelerators as a path to success? *Academy of Management Proceedings* (Academy of Management, Briarcliff Manor, NY), 12955.

Hannan MT, Carroll GR (1992) *Dynamics of Organizational Populations: Density, Legitimation, and Competition* (Oxford University Press, Oxford, UK).

Hannah DP, Eisenhardt KM (2018) How firms navigate cooperation and competition in nascent ecosystems. *Strategic Management J.* 39(12):3163–3192.

Hannan MT, Carroll GR, Dundon EA, Torres JC (1995) Organizational evolution in a multinational context: Entries of automobile manufacturers in Belgium, Britain, France, Germany, and Italy. *Amer. Sociol. Rev.* 60(4):509–528.

Hayes AF (2017) *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (Guilford Publications, New York).

He K, Zhang X, Ren S, Sun J (2015) Delving deep into rectifiers: Surpassing human-level performance on imagenet classification *Proc. IEEE Internat. Conf. Comput. Vision* (ICCV, Santiago, Chile), 1026–1034.

Helfat CE, Lieberman MB (2002) The birth of capabilities: market entry and the importance of pre-history. *Indust. Corporate Change* 11(4):725–760.

Hildreth JAD, Gino F, Bazerman M (2016) Blind loyalty? When group loyalty makes us see evil or engage in it. *Organ. Behav. Human Decision Processes* 132:16–36.

Hoenig D, Henkel J (2015) Quality signals? The role of patents, alliances, and team experience in venture capital financing. *Res. Policy* 44(5):1049–1064.

Hoffman AJ (1999) Institutional evolution and change: Environmentalism and the US chemical industry. *Acad. Management J.* 42(4): 351–371.

Hsu G (2006) Jacks of all trades and masters of none: Audiences' reactions to spanning genres in feature film production. *Admin. Sci. Quart.* 51(3):420–450.

Hsu DH (2007) Experienced entrepreneurial founders, organizational capital, and venture capital funding. *Res. Policy* 36(5):722–741.

Hsu DH, Ziedonis RH (2008) Patents as quality signals for entrepreneurial ventures. *Academy of Management Proceedings* (Academy of Management, Briarcliff Manor, NY), 1–6.

Hsu G, Roberts PW, Swaminathan A (2012) Evaluative schemas and the mediating role of critics. *Organ. Sci.* 23(1):83–97.

Jeng LA, Wells PC (2000) The determinants of venture capital funding: Evidence across countries. *J. Corporate Finance* 6(3):241–289.

Jensen M (2006) Should we stay or should we go? Accountability, status anxiety, and client defections. *Admin. Sci. Quart.* 51(1):97–128.

Kahneman D (2003) A perspective on judgment and choice: mapping bounded rationality. *Amer. Psych.* 58(9):697.

Kanze D, Huang L, Conley MA, Higgins ET (2018) We ask men to win and women not to lose: Closing the gender gap in startup funding. *Acad. Management J.* 61(2):586–614.

Kaplan S (2008) Framing contests: Strategy making under uncertainty. *Organ. Sci.* 19(5):729–752.

Kaplan S, Tripsas M (2008) Thinking about technology: Applying a cognitive lens to technical change. *Res. Policy* 37(5):790–805.

Kennedy MT (2008) Getting counted: Markets, media, and reality. *Amer. Sociol. Rev.* 73(2):270–295.

Khaire M (2010) Young and no money? Never mind: The material impact of social resources on new venture growth. *Organ. Sci.* 21(1):168–185.

Khessina OM, Carroll GR (2008) Product demography of de novo and de alio firms in the optical disk drive industry, 1983–1999. *Organ. Sci.* 19(1):25–38.

Lanahan L, Armanios D (2018) Does more certification always benefit a venture? *Organ. Sci.* 29(5):931–947.

Larsen JD, Newton-Smith R (2001) Using surveys of investment intentions. *Bank England Quart. Bull.* Summer 2001.

Lee BH, Sine WD (2012) Certifying the harvest: Entry dynamics and standards-based certification organizations in a nascent market. Preprint, written August 15, https://dx.doi.org/10.2139/ssrn.2340796.

Lee BH, Hiatt SR, Lounsbury M (2017) Market mediators and the trade-offs of legitimacy-seeking behaviors in a nascent category. *Organ. Sci.* 28(3):447–470.

Legg S, Hutter M (2007a) A collection of definitions of intelligence. *Frontiers Artificial Intelligence Appl.* 157:17–24.

Legg S, Hutter M (2007b) Universal intelligence: A definition of machine intelligence. *Minds Mach.* 17(4):391–444.

Levitt B, March JG (1988) Organizational learning. *Annual Rev. Sociol.* 14(1):319–338.

Lichtenstein S, Fischhoff B, Phillips LD (1977) Calibration of probabilities: The state of the art. *Decision Making and Change in Human Affairs,* 275–324.

MacKinnon DP, Dwyer JH (1993) Estimating mediated effects in prevention studies. *Eval. Rev.* 17(2):144–158.

MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V (2002) A comparison of methods to test mediation and other intervening variable effects. *Psych. Methods* 7(1):83.

MacMillan IC, Siegel R, Narasimha PS (1985) Criteria used by venture capitalists to evaluate new venture proposals. *J. Bus. Venturing* 1(1):119–128.

MacMillan IC, Zemann L, Subbanarasimha P (1987) Criteria distinguishing successful from unsuccessful ventures in the venture screening process. *J. Bus. Venturing* 2(2):123–137.

Malisiewicz T (2015) *Venture Pitch Contest at CVPR 2015 in Boston* (Tombone's Computer Vision Blog, MA).

Mason C, Stark M (2004) What do investors look for in a business plan? A comparison of the investment criteria of bankers, venture capitalists and business angels. *Internat. Small Bus. J.* 22(3):227–248.

McDonald RM, Eisenhardt KM (2019) Parallel play: Startups, nascent markets, and effective business-model design. *Admin. Sci. Quart.* 65(2):483–523.

McKendrick DG, Jaffee J, Carroll GR, Khessina OM (2003) In the bud? Disk array producers as a (possibly) emergent organizational form. *Admin. Sci. Quart.* 48(1):60–93.

Meade AW, Craig SB (2012) Identifying careless responses in survey data. *Psych. Methods* 17(3):437.

Morgan MG (2014) Use (and abuse) of expert elicitation in support of decision making for public policy. Clark WC, ed. *Proc. Natl. Acad. Sci.* (National Academy of Science, Cambridge, MA), 7176–7184.

Navis C, Glynn MA (2011) Legitimate distinctiveness and the entrepreneurial identity: Influence on investor judgments of new venture plausibility. *Acad. Management Rev.* 36(3):479–499.

Negro G, Hannan MT, Rao H (2010) Categorical contrast and audience appeal: Niche width and critical success in winemaking. *Indust. Corporate Change* 19(5):1397–1425.

Negro G, Hannan MT, Rao H (2011) Category reinterpretation and defection: Modernism and tradition in Italian winemaking. *Organ. Sci.* 22(6):1449–1463.

Niwattanakul S, Singthongchai J, Naenudorn E, Wanapu S (2013) Using of Jaccard coefficient for keywords similarity. *Proc. Internat. Multiconfer. Engineers Comput. Scientists* (IMECS, Hong Kong), 380–384.

Noh S, Tolbert PS (2019) Organizational identities of US art museums and audience reactions. *Poetics* 72:94–107.

Nuscheler D, Engelen A, Zahra SA (2019) The role of top management teams in transforming technology-based new ventures' product introductions into growth. *J. Bus. Venturing* 34(1):122–140.

Ozcan P, Gurses K (2018) Playing cat and mouse: Contests over regulatory categorization of dietary supplements in the United States. *Acad. Management J.* 61(5):1789–1820.

Ozcan P, Santos FM (2015) The market that never was: Turf wars and failed alliances in mobile payments. *Strategic Management J.* 36(10):1486–1512.

Padgett JF, Powell WW (2012) *The Emergence of Organizations and Markets* (Princeton University Press, Princeton, NJ).

Plummer LA, Allison TH, Connelly BL (2016) Better together? Signaling interactions in new venture pursuit of initial external capital. *Acad. Management J.* 59(5):1585–1604.

Polidoro F Jr (2020) Knowledge, routines, and cognitive effects in nonmarket selection environments: An examination of the regulatory review of innovations. *Strategic Management J.* 41(13):2400–2435.

Pollock TG, Gulati R (2007) Standing out from the crowd: The visibility-enhancing effects of IPO-related signals on alliance formation by entrepreneurial firms. *Strategic Organ.* 5(4):339–372.

Pontikes EG (2012) Two sides of the same coin: How ambiguous classification affects multiple audiences' evaluations. *Admin. Sci. Quart.* 57(1):81–118.

Poole D, Mackworth A, Goebel R (1998) *Computational Intelligence: A Logical Approach* (Oxford University Press, Oxford, UK).

Powell WW, Colyvas JA (2008) Microfoundations of institutional theory. Greenwood R, Olive C, Suddaby R, Sahlin-Andersson K, eds. *The Sage Handbook of Organizational Institutionalism* (Sage Publishers, Thousand Oaks, CA), 276–298.

Powell WW, Oberg A (2017) Networks and institutions. Greenwood R, Oliver C, Lawrence TB, Meyer RE, eds. *The Sage Handbook of Organizational Institutionalism* (Sage Publishers, Thousand Oaks, CA), 446–476.

Preacher KJ, Hayes AF (2004) SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behav. Res. Methods Instrum. Comput.* 36(4):717–731.

Rao H (1994) The social construction of reputation: Certification contests, legitimation, and the survival of organizations in the American automobile industry: 1895–1912. *Strategic Management J.* 15(S1):29–44.

Rosch E (1999) Principles of categorization. Margolis E, Laurence S, eds. *Concepts: Core Readings* (The MIT Press, Cambridge, MA), 189–206.

Russell SJ, Norvig P (2016) *Artificial Intelligence: A Modern Approach* (Pearson Education Limited, Malaysia).

Santos FM, Eisenhardt KM (2009) Constructing markets and shaping boundaries: Entrepreneurial power in nascent fields. *Acad. Management J.* 52(4):643–671.

Schilke O (2018) A micro-institutional inquiry into resistance to environmental pressures. *Acad. Management J.* 61(4):1431–1466.

Scmpnews (2017) China turns to artificial intelligence to boost its education system. *South China Morning Post* (October 14), https://www.scmp.com/tech/science-research/article/2115271/china-wants-bring-artificial-intelligence-its-classrooms-boost.

Sine WD, David RJ, Mitsuhashi H (2007) From plan to plant: Effects of certification on operational start-up in the emergent independent power sector. *Organ. Sci.* 18(4):578–594.

Slovic P, Fischhoff B, Lichtenstein S (1985) Rating the risks: The structure of expert and lay perceptions. Covello VT, Mumpower JL, Stallen PJM, Uppuluri VRR, eds. *Environmental Impact Assessment, Technology Assessment, and Risk Analysis* (Springer, Berlin, Germany), 131–156.

Strang D, Meyer JW (1993) Institutional conditions for diffusion. *Theory Soc.* 22(4):487–511.

Suárez FF, Utterback JM (1995) Dominant designs and the survival of firms. *Strategic Management J.* 16(6):415–430.

Suarez FF, Grodal S, Gotsopoulos A (2015) Perfect timing? Dominant category, dominant design, and the window of opportunity for firm entry. *Strategic Management J.* 36(3):437–448.

Suchman MC (1995) Managing legitimacy: Strategic and institutional approaches. *Management Rev.* 20(3):571–610.

Sykes HB (1990) Corporate venture capital: Strategies for success. *J. Bus. Venturing* 5(1):37–47.

Tassey G (2000) Standardization in technology-based markets. *Res. Policy* 29(4-5):587–602.

Technode (2017) China vs US: Who is winning the big AI battle? *Technode China* (October 22), https://technode.com/2017/10/22/china-vs-us-ai/.

Tetlock PE, Boettger R (1989) Accountability: A social magnifier of the dilution effect. *J. Pers. Soc. Psych.* 57(3):388.

Theeke M, Polidoro F Jr, Fredrickson JW (2018) Path-dependent routines in the evaluation of novelty: the effects of innovators' new knowledge use on brokerage firms' coverage. *Admin. Sci. Quart.* 63(4):910–942.

Tonoyan V, Strohmeyer R, Jennings JE (2019) Gender gaps in perceived start-up ease: Implications of sex-based labor market segregation for entrepreneurship across 22 European countries. *Admin. Sci. Quart.* 65(1):181–225.

Tushman ML (1992) Organizational determinants of technological change: toward a sociology of technological evolution. *Res. Organ. Behav* 14:311–347.

Tversky A, Kahneman D (1974) Judgment under uncertainty: Heuristics and biases. *Sci.* 185(4157):1124–1131.

Tyebjee TT, Bruno AV (1984) A model of venture capitalist investment activity. *Management Sci.* 30(9):1051–1066.

Tyler TR (2006) Psychological perspectives on legitimacy and legitimation. *Annual Rev. Psych.* 57:375–400.

Vakili K (2016) Collaborative promotion of technology standards and the impact on innovation, industry structure, and organizational capabilities: Evidence from modern patent pools. *Organ. Sci.* 27(6):1504–1524.

Ward M, Meade AW, Allred CM, Pappalardo G, Stoughton JW (2017) Careless response and attrition as sources of bias in online survey assessments of personality traits and performance. *Comput. Human Behav.* 76:417–430.

WIPO (2019) *World Intellectual Property Organization Technology Trends 2019: Artificial Intelligence* (World Intellectual Property Organization).

Wright M, Robbie K, Ennew C (1997) Venture capitalists and serial entrepreneurs. *J. Bus. Venturing* 12(3):227–249.

Wry T, Lounsbury M (2013) Contextualizing the categorical imperative: Category linkages, technology focus, and resource acquisition in nanotechnology entrepreneurship. *J. Bus. Venturing* 28(1):117–133.

Wry T, Lounsbury M, Glynn MA (2011) Legitimating nascent collective identities: Coordinating cultural entrepreneurship. *Organ. Sci.* 22(2):449–463.

Wry T, Lounsbury M, Jennings PD (2014) Hybrid vigor: Securing venture capital by spanning categories in nanotechnology. *Acad. Management J.* 57(5):1309–1333.

Zucker LG (1987) Institutional theories of organization. *Annual Rev. Sociol.* 13(1):443–464.

Zuckerman EW (1999) The categorical imperative: Securities analysts and the illegitimacy discount. *Amer. J. Sociol.* 104(5):1398–1438.

Zunino D, Suarez FF, Grodal S (2019) Familiarity, creativity, and the adoption of category labels in technology industries. *Organ. Sci.* 30(1):169–190.

Xirong (Subrina) Shen is an assistant professor of management at the McCombs School of Business at the University of Texas at Austin. She received her PhD from the Johnson Graduate School of Management at Cornell University. Her research interests focus on entrepreneurship and innovation in nascent technology sectors.

Huisi (Jessica) Li is an assistant professor of organizational behavior at the Scheller College of Business at the Georgia Institute of Technology. She received her PhD from the Johnson Graduate School of Management at Cornell University. Her research interests include social hierarchy, diverse teams, and employee turnover.

Pamela S. Tolbert is the Lois Gray Professor of Industrial Relations and Social Sciences in the ILR School at Cornell University. She received her PhD in sociology from University of California, Los Angeles. Her general areas of research include organizations and social stratification, organizational decision-making, and processes of institutional change.