SMARTER SYSTEMS:
How Tweaking Your Diligence Process Can Unlock Overlooked Opportunities

RESEARCH METHODOLOGY

village capital

RESEARCH PARTNERS

IFC International Finance Corporation
WORLD BANK GROUP
Creating Markets, Creating Opportunities

we-fi
WOMEN ENTREPRENEURS
FINANCE INITIATIVE

SPONSORS

VISA Foundation

MOODY'S

USAID FROM THE AMERICAN PEOPLE

ANDE ASPEN NETWORK OF DEVELOPMENT ENTREPRENEURS
aspen institute

SPF SASAKAWA PEACE FOUNDATION

SGB EVIDENCE FUND

Part 1: Background

Part 2: Methodology — Experimental Programs

Acknowledgements

Part One:
Background

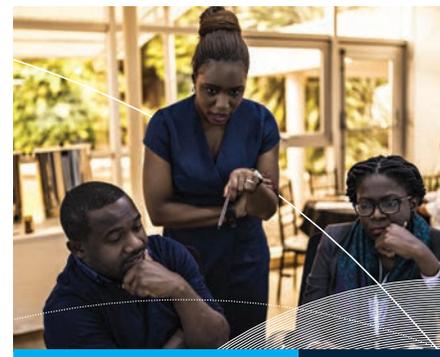# The Role of Accelerators in the Gender Financing Gap

In 2019, Village Capital partnered with International Finance Corporation (IFC), the World Bank Gender Innovation Lab, and the Women Entrepreneurs Finance Initiative (We-Fi) to develop "*Venture Capital and the Gender Financing Gap: The Role of Accelerators*," a study on the role of accelerators in the gender financing gap that leverages data from the Global Accelerator Learning Initiative (GALI).

We found that, while acceleration removes the financing disadvantage women-led startups face when raising debt, it has little to no effect on their ability to raise equity. In fact, men-led startups raise 2.6 times more equity post-acceleration than do women-led startups.

Venture Capital and the Gender Financing Gap: The Role of Accelerators
RESEARCH SNAPSHOT | FEBRUARY 2020

IN PARTNERSHIP WITH
we-fi | VillageCapital | IFC International Finance Corporation

This financing gap could not be attributed to any quantifiable aspect of either startup or founder differences. These included founder characteristics (such as education level or experience) and startup characteristics (such as intellectual property, sector of operation, geography, or revenue generated). Moreover, we found no clear accelerator design elements that could overcome this gap.

Consequently, we hypothesized that investor biases may play a role in explaining the gap.

Based on this hypothesis, we sought to better understand investor behavior and identify effective interventions that could be employed by both accelerators and investment organizations to mitigate discrepancies in evaluation processes that disadvantage women founders during investor decision-making.

Reducing Disparities in Evaluations

To identify which interventions we could ultimately test, we first reviewed existing research on investor behavior, conducted a lab- in-the-field experiment, and interviewed a global sample of 30 investors. Our learnings from this process later informed the interventions that we designed and tested.

Review of Existing Research

We turned to existing research to better understand investor behavior when evaluating startups. We identified three characteristics that later informed the strategies we tested.

1. The gender of the entrepreneur influences funding decisions.[1] After evaluating data from pitch competitions, one study found that men-led startups were 60% more likely to receive funding than were women-led startups.[2] When presented identical pitches, differentiated only by the gender of the voice narrating the pitch, "68.33% of participants chose to fund ventures pitched by a male voice, while only 31.67% of participants chose to fund the ventures pitched by a female voice."[3]

2. Women-led startups receive more difficult questions from investors. One study found that, while investors ask women entrepreneurs more risk-focused questions such as "What potential regulatory issues or future government mandates are possible?", they ask men entrepreneurs more growth-focused questions such as "How much can sales increase?"[4]

3. We found three theories that could explain why investors prefer to invest in men entrepreneurs:

   ▶ *Pattern-matching based on previous successes:* Investors typically invest in entrepreneurs within their networks, which tend to be the same gender and come from similar backgrounds.[5] With the purpose of replicating past successes, investors may also often seek out entrepreneurs who are similar to those with whom they have had success in the past. This leads investors to favor men over women entrepreneurs.

- *Gender role congruity theory*: Based on this theory, investors possibly evaluate women less favorably because they perceive attributes typically associated with women as incongruent with those required to be a competent entrepreneur who shows potential for success.[6]

- **Evaluating potential vs performance:** Research has found that, when evaluating employee leadership potential, women are consistently scored lower than men despite being scored equally or higher for their performance.[7] After the initial evaluation among those with equal scores for potential and performance, women still outperform men.[8] The promotion gap is closed when evaluations of potential take into account how well the employee scored in their performance.[9] Research also suggests that evaluating how well someone performs at something results in more equitable and objective hiring decisions that are not clouded by gender bias.[10]

Together, these three learnings suggested that there are discrepancies in how investors evaluate men-led and women-led startups, possibly driven by gender bias. This could reduce the accuracy and objectivity of evaluations. For example, asking women more risk-focused questions and men more growth-focused questions could result in investors overlooking potential for risks or growth opportunities.

Lab-in-the-Field Experiment

We conducted a lab simulation with 150 external investors to test the effectiveness of two strategies to mitigate discrepancies in investor decision-making. Participants were randomized into control and treatment groups. In our setting, providing investors with legitimating information (that the startup had received an investment from a fund with a diversity mandate) did not influence scores. Directional effects suggested that only men-led startups benefitted from this treatment. However, changing the organizational evaluation frameworks investors used to assess startups affected the likelihood that an investor would be interested to take a startup through due diligence. We also observed that investors sought risk- and growth-related information from startups depending on founder gender, suggesting that the follow-on diligence process would be more difficult for women-led startups than for those led by men.

These results suggested that how investors evaluated — and how organizations asked them to evaluate — could meaningfully affect the scores given to startups.

Investor Interviews

We next examined the evaluation strategies used by investors to identify potential areas in which we could intervene. We interviewed a global sample of over 30 investors to better understand how investors typically evaluate startups. We found that investors often use well-known criteria to evaluate a startup venture, highlighted in light blue in the diagram below.

However, investors also shared that they used criteria to evaluate the founding team, beyond basic information about the team's qualifications. This aligns with a large-scale investor survey, where **95% of VC firms cited founders as being one of the most important factors when deciding to invest in a startup.**[11] Given that early-stage startups cannot provide ample information on their business trajectory, it is likely that investors focus heavily on evaluating the founding team's potential to grow the startup in order to determine if an investment should be made.
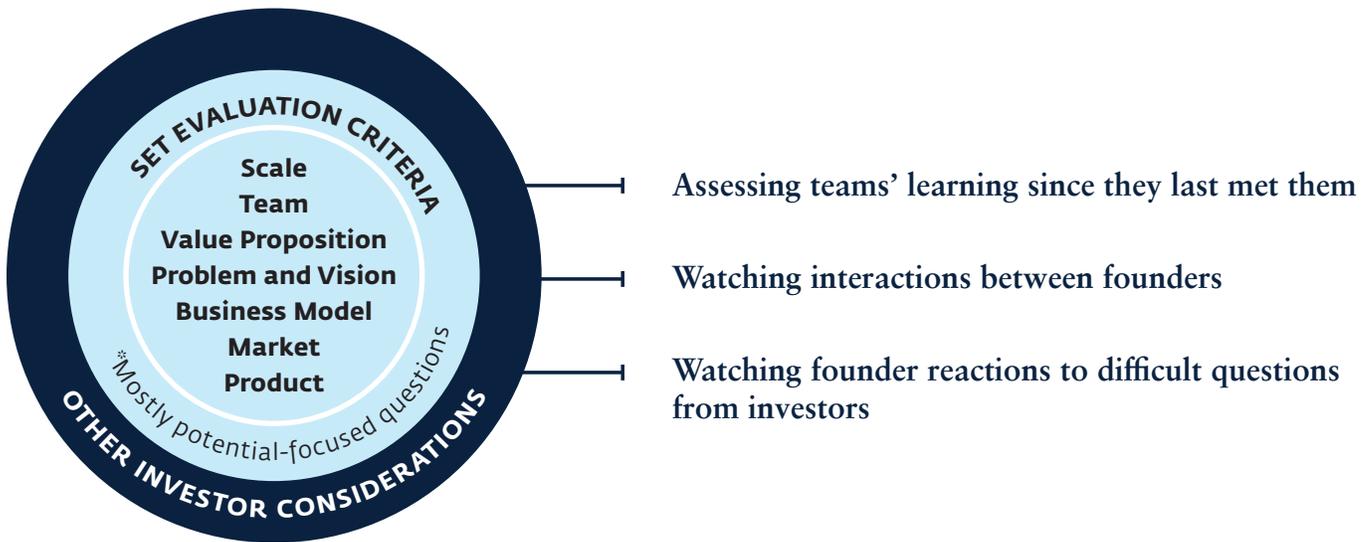
For example, one investor stated: "Our investment thesis is built strongly around the team— which is more intangible. That means evaluation tends to be a little fuzzy, less specific, and less able to concretize into many evaluation models."

Through these interviews, we found that **evaluations lack consistent criteria and standards with which to evaluate founding teams, even though investors saw them as important.** Given that this process lacks data, investors rely on their "gut instinct" to evaluate the founders' "potential."

This could lead to inconsistent evaluations, as assessing potential without relying on much data leaves space for evaluations to be influenced by gender bias.[12] More specifically, this could result in more favorable outcomes for men than for women, likely due to the influence of gender bias.

Interventions Tested In the Experimental Programs

Our research review, lab-in-the-field experiment, and investor interviews all helped us identify what inconsistencies to target in the evaluation process and how to do so. We found that evaluation processes that lack structure and consistency — which reduce the accuracy of assessments and, therefore, lead to investors overlooking promising startups. As a result, we added three steps in the evaluation process to reduce gender disparities and increase the accuracy of all startup assessments by making them more consistent, comprehensive, and data-driven.

Step #1

Pre defining evaluation criteria

In the treatment group, the investor had to predefine how much weight they would give to each criterion when scoring. In other words, they had to predefine which criteria would most heavily determine their scores.

This step was designed based on our learnings from the lab-in-the-field experiment.

Reasoning:

Research has also found that evaluators adjust the characteristics they initially wanted to see in a successful candidate to fit the characteristics displayed by candidates of their preferred gender.[13] Predefining the weight applied to each criterion prompts the investor to commit to evaluating all startups consistently, preventing the investor from redefining the criteria for success based on the gender of the founders.[14]

Step #2

Consistently seek risk and growth opportunity information from all startups

During each evaluation round, investors in the treatment group were asked two additional questions: "What additional information would you want on this venture's potential for growth?" and "What additional information would you want on how this venture will mitigate risks?"

This step was designed based on existing research, which has found that women are asked more risk-focused questions.[15]

Reasoning:

While startups warrant different degrees of focus on their risks and growth opportunities, investors have been found to focus disproportionately on one of the two, depending on the gender composition of the founding team. This decreases the accuracy of evaluations, as investors may focus too much or too little on one of the two criteria and thus overlook key risks or growth opportunities that could impact their assessment.

Explicitly prompting investors to think about both risk- and growth-related questions sought to increase consistency and accuracy of evaluations by preventing investors from focusing too much or too little on a single one. In doing so, it sought to prevent investors from overlooking promising startups that are not as risky as perceived and/or overestimating startups that pose more risks than perceived.

Step #3

Assess a team's potential by evaluating how much they have demonstrated an ability to improve their startup

In the treatment group, two additional categories were added to the evaluation framework to evaluate a founding team's demonstrated ability to improve in understanding and executing its growth and risk mitigation strategies.

This step was designed based on our learnings from the investor interviews.

Reasoning:

Research suggests that evaluating how well someone performs at something results in more equitable and objective hiring decisions.[16]

A startup has potential if it seems likely they will be able to grow. To do so, the startup must be able to continually make improvements that allow it to grow.

Consequently, evaluating how a founding team improves their startup in the short-term helps the investor make a more accurate, performance-based assessment of the startup's future potential, by creating new data to assess how well the team will be able to continue making improvements to their company strategies in the future.

We avoided focusing on individual investors because research has found that targeting change in organizational processes and structures is more effective at producing individual behavioral change.[17] These steps, however, can also be adopted by individual investors.

We also avoided strategies that can result in unintended consequences, such as reinforcing gender stereotypes by portraying women-led startups as needing help or creating backlash. The purpose of these steps is also not to explicitly train investors or those involved in evaluating startups to be less biased, as research has found that bias training is not always effective and, in fact, can be counterproductive in some circumstances.[18]

Rather, we designed these steps to improve the evaluation process by reducing discrepancies in two parallel ways so that investors could evaluate ventures of similar quality equally. First, they provided more structure and consistency to areas that lacked them. Second, they focused investor attention on proven entrepreneur competencies to reduce their reliance on their gut instincts when evaluating the founding team.

Part Two:
Methodology

Experiment Setting

We leveraged Village Capital's selection model to test our experiment on "trainee investors." Village Capital uses a peer selection model to allocate investments into two startups in each program, usually over a period of three months. Under this model, within each regional, industry-specific program, each startup evaluates and scores its peers using Village Capital's VIRAL (Venture Investment-Readiness and Awareness Levels) framework. At the end of each program, the two companies with the highest scores in the last evaluation round receive an investment.

Village Capital designed VIRAL after carefully studying the areas most commonly evaluated by investors and the different levels of maturity startups experience through each funding stage. It is composed of eight evaluation categories: team, problem and vision, value proposition, product, market, business model, scale, and investor exit. Investors use two criteria — growth opportunity and investment opportunity — to assess the startups' performance in these categories. The table below provides sample questions typically used during peer due diligence:

	Product	Business Model
Growth Opportunity	How confident are you that this company's product will solve its customers' problems and delight users?	How confident are you that the company's business model is viable and that it can make money?
Investment Opportunity	How confident are you that the product can expand to multiple offerings and outpace the market on innovation?	How confident are you that the business model is resilient enough to withstand strategic and market risks, and adapt as necessary to provide a meaningful return to its investors?

At the same time, the framework outlines the nine levels that startups go through over the lifetime of the firm, beginning with "establishing the founding team" and ending with an "exit in sight." It also indicates the type of funding that usually takes place at each level. Based on the startups' maturity in each of the eight evaluation categories, the framework locates the startup on its corresponding level of investment-readiness. In this way, VIRAL facilitates communication among investors and entrepreneurs by creating a common language to talk about progress and investment expectations.

As a result, when entrepreneurs use VIRAL both for their self-assessment and to evaluate their peers, they are able to adopt the mindset of an investor. We are also confident that these entrepreneurs can positively evaluate their peers due to their experience in the relevant industry and region. A previous study, *Flipping The Power Dynamics: Can Entrepreneurs Make Successful Investment Decisions?*, found that the final peer-selected rank accurately reflected the future ability of ventures to raise capital.[19]

We were also able to leverage this setting because we could change how Village Capital trained investors to evaluate, and we could observe investors' evaluations over time. This allowed us to isolate the effect of adjusting something within the evaluation process. We measured changes in the scores to determine the effectiveness of the strategies in producing more equitable evaluations.

Setting up the Randomized Controlled Trial (RCT)

We conducted a randomized controlled trial (RCT) in eight Village Capital accelerator programs across four geographies. Participants were selected using Village Capital's rigorous selection framework to ensure comparable quality and level of investment readiness across all startups. They all shared similar characteristics in terms of sector, region, and stage or maturity, and they were all startups that Village Capital would invest in. We also randomized the startups into a control or treatment group, stratifying by region and gender. In each group, we ensured parity across founder gender, startup characteristics, and program participation.

Randomizing into control and treatment groups helped us to identify any differences among the groups as being driven by our interventions. In other words, any difference in scores between the treatment and control groups is very unlikely to be driven by any other factor than our treatments. The RCT tested the effectiveness of three strategies at improving the objectivity and accuracy of the evaluation process in a way that ventures of similar quality are evaluated more equally, regardless of the gender composition of the founding team.

Setting and Sample

65 startups (or trainee investors) participated in our eight accelerator programs. Of these, almost half were women-led startups, meaning that they had at least one woman on the founding team.

In our experiment, investors evaluated startups four times over three months on a variety of startup criteria. This resulted in a total of over 30,000 scores over the course of 1,503 investor-startup evaluations. These evaluations resulted in 510 investment decisions.

Investors scored startups using Village Capital's VIRAL framework — designed to guide the investor in consistently collecting objective data about the startup—on our online platform. This framework evaluates categories that many investor evaluations typically include (see table below); investors assign a score for each of the eight categories. The top two ventures with the highest scores in the last evaluation round were peer selected for an investment of $20,000.

Criteria	Company Breakdown Categories							
Growth Opportunity	Team	Vision	Value Proposition	Product	Market	Business Model	Scale	Investor Exit
Investment Opportunity								

Each evaluation round, investors assigned a score between 1 and 4 for each of the eight categories in the control group, as well as each of the 12 categories in the treatment group.

The platform automatically converts these scores into z-scores, which indicate how much the original score deviates from the mean score. A positive z-score indicates that the original score is above the mean, while a negative z-score indicates that the original score is below the mean. The z-score is the average of all z-scores calculated per investor (i.e., across all their decisions per round), which helps control for investors who score generously or harshly. We measured changes in the z-scores to determine the effectiveness of the interventions.

Incorporating Our Three Steps Into The Evaluation Framework

In the first evaluation (our baseline), investors were simply asked to score each startup, indicating how inclined they would be to initiate due diligence on the venture. In the three evaluation rounds that followed, both control and treatment groups used Village Capital's VIRAL framework. In the treatment group, we incorporated our three new steps into the framework.

Step #1   **Pre defining evaluation criteria**

Prior to each evaluation round, treated investors were asked to pre define how much each evaluation criterion would determine their scores. To do this, investors had to answer the following question:

> **Please think about how you will make your decisions and weight the criteria below with percentages of how much weight you would place on each criterion. (Please make sure it adds up to 100%!)**
>
> | Growth Opportunity (%) | Investment Opportunity (%) | Improvement Made During Program (%) | Other (%) |

Step #2   **Consistently seek risk and growth opportunity information from all startups**

During each evaluation round, investors were asked to submit the additional questions they had regarding each venture's growth opportunities and risks. These answers were shared with the cohort. Specifically, investors were asked the following two questions:

> **What additional information would you want on [startup name]'s potential for growth?**
>
> **What additional information would you want on how [startup name] will mitigate risks?**

We incorporated four new categories into VIRAL to assess how much a startup had improved, as a whole, in understanding and executing its growth and risk mitigation strategies.

Similar to how investors assessed each of the original eight VIRAL categories using two criteria (growth opportunity and investment opportunity), investors were asked to evaluate how much a founding team had shown improvement in the four new categories throughout the program. To do so, they were provided with a guiding question to help guide their assessment of each category. The table below shows the categories added to VIRAL that were used by investors in the treatment group:

Whole company improvement categories	Understanding potential for *growth*	Demonstrating potential for *growth*	Understanding *risks*	Demonstrating *risk mitigation*
Guiding question	*How much has this company **improved in understanding** its path to growth?*	*How much has this company **improved in executing** its path to growth?*	*How much has this company **improved in understanding** its risks?*	*How much has this company **improved in executing** on risk mitigation?*

Research Design and Analysis Plan

We pre-registered our hypothesis on the AEA RCT trial website and, before running the experiment, we received feedback on the research design from academics, World Bank researchers, the Innovation Growth Lab, and research leads in our partner organizations.

We have put an overview of the methodology below. To find out more, please contact Amisha Miller (amisham@bu.edu).

Our dependent variable is score. Following McKenzie (2012), we pre-registered an Ancova regression not only to assess the difference in scores from the beginning to the end of the program, but also to increase our power. This means that we put the baseline score on the right-hand side of the regression.

Our independent variables are female, treatment applied, and an interaction between female and treatment. This interaction is our interest — what is the effect of the treatment on female-led startups? We report these results in the toolkit.

We clustered all standard errors by investor to make sure that we assessed each investor decision as correlated. In other words, we did not treat each investor decision as totally independent of their other decisions.

To reduce the effect of other factors, we also included fixed effects for region and time of evaluation. We also controlled for startup characteristics (but these controls did not affect our main results).

We conducted two mechanism analyses and had evidence that supported both mechanisms:

1. Assessed the effect of our seeking information treatment on the number of risk questions asked. (We simply placed "number of risk questions asked" as the dependent variable).

2. Assessed the effect of our improvement treatment by evaluating whether treated investors evaluated improvement compared to the control (through interviews and by assessing the weights they placed on their criteria).

We also included robustness checks by changing how we measured the dependent and independent variables, and we did not find that this changed our main results.

ENDNOTES

1   Alison Wood Brooks , Laura Huang , Sarah Wood Kearney and Fiona E. Murray, "Investors prefer entrepreneurial ventures pitched by attractive men," Proceedings of the National Academy of Sciences of the United States of America (PNAS), February 2014, https://gap.hks.harvard.edu/investors-prefer-entrepreneurial-ventures-pitched-attractive-men

2   Brooks, "Investors prefer entrepreneurial ventures pitched by attractive men," https://gap.hks.harvard.edu/investors-prefer-entrepreneurial-ventures-pitched-attractive-men

3   Brooks, "Investors prefer entrepreneurial ventures pitched by attractive men," https://gap.hks.harvard.edu/investors-prefer-entrepreneurial-ventures-pitched-attractive-men

4   Dana Kanze, Laura Huang, Mark A. Conley, and E. Tory Higgins, "Male and Female Entrepreneurs Get Asked Different Questions by VCs — and It Affects How Much Funding They Get," Harvard Business Review, June 27, 2017, https://hbr.org/2017/06/male-and-female-entrepreneurs-get-asked-different-questions-by-vcs-and-it-affects-how-much-funding-they-get

5   Paul Gompers, Will Gornall, Steven N. Kaplan, and Ilya A. Strebulaev, "How venture capitalists make decisions," Harvard Business Review, March 2021, https://hbr.org/2021/03/how-venture-capitalists-make-decisions

6   Alice A. Eagly and Steven J. Karau, "Role congruity theory of prejudice toward female leaders," Psychological Review, 2022, https://www.women-unlimited.com/wp-content/uploads/prejudice_against_women.pdf

7   Alan Benson, Danielle Li, and Kelly Shue, "Potential" and the Gender Promotion Gap," June 22, 2022, https://danielle-li.github.io/assets/docs/PotentialAndTheGenderPromotionGap.pdf

8   Benson, "Potential" and the Gender Promotion Gap," https://danielle-li.github.io/assets/docs/PotentialAndTheGenderPromotionGap.pdf

9   Benson, "Potential" and the Gender Promotion Gap," https://danielle-li.github.io/assets/docs/PotentialAndTheGenderPromotionGap.pdf

10 Claudia Goldin and Cecilia Rouse, "Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians," American Economic Review, September 2000, https://www.aeaweb.org/articles?id=10.1257/aer.90.4.715

11 Gompers, "How do Venture Capitalists Make Decisions?" https://hbr.org/2021/03/how-venture-capitalists-make-decisions

12 Benson, "Potential" and the Gender Promotion Gap," https://danielle-li.github.io/assets/docs/PotentialAndTheGenderPromotionGap.pdf

13 Eric Luis Uhlmann and Geoffrey L. Cohen, "Redefining Merit to Justify Discrimination," Psychological Science, June 1, 2005, https://doi.org/10.1111/j.0956-7976.2005.01559.x

14 Uhlmann, "Redefining Merit to Justify Discrimination," https://doi.org/10.1111/j.0956-7976.2005.01559.x

15 Kanze, "Male and Female Entrepreneurs Get Asked Different Questions by VCs — and It Affects How Much Funding They Get." https://hbr.org/2017/06/male-and-female-entrepreneurs-get-asked-different-questions-by-vcs-and-it-affects-how-much-funding-they-get

16 Goldin, "Orchestrating Impartiality: The Impact of 'Blind' Auditions on Female Musicians," https://www.aeaweb.org/articles?id=10.1257/aer.90.4.715

17 Dr. Kristen Liesch, "How is Venture Capital Failing to Fund a Gender Equal Future?" Tidal Equality, https://www.tidalequality.com/blog/how-is-vc-failing-to-fund-a-gender-equal-future

18 Dr. Kristen Liesch, "Don't do unconscious bias training." Tidal Equality, https://www.tidalequality.com/blog/dont-do-unconscious-bias-training

19 Allie Burns, Robert Tashima, and Heather Matranga, "Flipping the Power Dynamics: Can Entrepreneurs Make Successful Investment Decisions?" Village Capital, January 2019, https://newsandviews.vilcap.com/reports/flipping-the-power-dynamics-can-entrepreneurs-make-successful-investment-decisions