

Project RISE: Cultivating Collaborative Cultures

Summary of Key Findings

April 2022



Report prepared by Audrey Aday,
Hilary Bergsieker, Grace Denney,
Toni Schmader, & Jessica Trickey



ENGENDERING SUCCESS IN STEM

Table of Contents

Click any title below to jump to that section of the report.

	p.
<u>Message from Project Leads</u>	3
<u>Executive Summary</u>	4
<u>Project Overview</u>	6
<u>Sample Overview</u>	7
<u>Results</u>	8
<u>Workshop Feedback & Outcomes</u>	8
<u>Leadership & Bias Recognition</u>	9
<u>Implicit Stereotypes & Awareness</u>	10
<u>Beliefs About Bias & Allyship</u>	11
<u>Enacting Allyship</u>	12
<u>Enacting Leadership</u>	13
<u>Fit & Engagement</u>	14
<u>Key Findings at a Glance</u>	15
<u>Glossary</u>	16
<u>Definitions</u>	16
<u>Symbols & Abbreviations</u>	17
<u>Interpretation Guide</u>	18
<u>Tables & Graphs</u>	18
<u>Effect Sizes</u>	19
<u>Further Reading</u>	20
<u>Project RISE Team</u>	21
<u>Thank You to Our ESS Partners</u>	22
<u>Appendices</u>	23
<u>Measures: Items, Range, Sample Size Over Time</u>	24
<u>Full Sample: Baseline Levels</u>	26
<u>Full Sample: Short-Term Outcomes</u>	27
<u>Full Sample: Over-Time Outcomes</u>	28

Message from Project Leads

Organizations spend millions of dollars annually on employee diversity and inclusion training. Typically, the effectiveness of this training is not measured. Some training programs are effective at changing beliefs and attitudes, but have minimal or only short-term effects on behaviour. Research is needed to identify when and how such training can be effective, especially over time.

Project RISE: Cultivating Collaborative Cultures is a research partnership among social scientists, STEM experts, and organizational partners committed to fostering gender diversity and inclusion in **science, technology, engineering, and math (STEM)**. The project used a **randomized control trial** design, the gold standard in research, to assess effects of an in-person, interactive half-day workshop on **Inclusive Innovation** compared with a parallel workshop on **Influential Leaders**.

In partnership with WinSETT, the RISE team designed the **Inclusive Innovation** workshop to reflect the best available science on organizational values, the cognitive science of implicit bias, the harm caused by gender stereotypes, and how supportive allyship actions can help mitigate that harm and cultivate greater inclusion for women and other marginalized groups. This training was designed to be evidence based, minimize reactance (backlash), and respect participants of different genders, backgrounds, and intersecting identities. Opportunities for in-depth dialogues aimed to foster greater understanding of people's lived experiences of gender in the workplace.

The **Influential Leaders** workshop introduced participants to the science of status and influence, challenges in managing teams, four distinct leadership strategies that are effective in STEM, and ideas for how to increase one's everyday leadership. Our goal in this workshop was to provide attendees with a similarly positive and evidence-based experience to those in the Inclusion workshop, with content relevant to organizational outcomes, but without bearing directly on issues related to equity, diversity, or inclusion.

In line with best practices in research, we pre-registered our hypotheses and analysis plan. We **predicted distinct outcomes** for each workshop at several time points over **nearly two years**. This report summarizes key findings on outcomes related to learning, attitudes, behaviour, and flourishing. Whenever possible, we controlled for or display changes in these outcomes from baseline, that is, prior to participating in a workshop.

Project RISE: Cultivating Collaborative Cultures was a truly collaborative effort. The dedicated facilitators at WinSETT helped to shape workshop content and co-facilitated sessions alongside Drs. Bergsieker and Schmader. Members of the Project RISE research team sourced content, created surveys, managed materials, and coded and analyzed results. Our STEM experts provided advice and input at critical points in the project's execution. Most importantly, our partners provided invaluable insight into the research development and invested their time to recruit employees. Finally, **we thank the nearly 300 employees** who participated. Only with their involvement can we better understand possible benefits of cultivating collaborative cultures.



Dr. Toni Schmader
Director of Engendering Success in STEM
University of British Columbia



Dr. Hilary Bergsieker
Project RISE co-lead
University of Waterloo

Executive Summary



Immediate Workshop Impacts

The Leadership and Inclusion workshops had a parallel structure and were rated to be similarly positive. Both were highly recommended, particularly the Inclusion workshop.

That said, immediate workshop effects were distinct:

- Leadership participants gained appreciation for leadership as a skill, but were not more motivated than Inclusion participants to enact leadership.
- Inclusion participants were more motivated to enact allyship and felt more self-efficacy for both allyship and (unexpectedly) leadership behaviors.

Recognizing Bias and Leadership

Each workshop trained participants to accurately recognize how the content applies in everyday settings.

- Inclusion participants were more accurate at identifying instances of implicit gender bias.
- Leadership participants were more accurate at identifying distinct instances of leadership.

Implicit Stereotypes and Awareness

A key learning objective of the Inclusion workshop, was to raise awareness of people's own implicit stereotypes, an automatic tendency to "think STEM, think male."

- Across workshops, participants (men in particular) exhibited evidence of this implicit stereotype.
- The Inclusion workshop increased men's awareness of their own gender stereotypes.
- Women's awareness was already high and not elevated by participating in the Inclusion workshop.

Executive Summary



Bias a Problem; Allyship a Solution

Some participants entered the study more skeptical about gender bias as a problem in STEM. One concern is that such skeptics might show reactance to Inclusion content.

Instead, the Inclusion workshop was most effective at boosting the belief that gender bias is a problem among those most skeptical at baseline. This effect persisted throughout the study, up to 20 months post-workshop.

The Inclusion workshop also fostered the belief that supportive allyship actions from men can be uniquely effective at foster gender inclusion.

Allyship and Leadership Behaviour

Two weeks after the workshops, only women in Inclusion reported engaging in more allyship actions, either in general, or specific to the action plan they had developed.

However, men enacted their specific allyship action plan increasingly over time, closing this gap by 18-20 months.

Leadership participants engaged in newer, less autocratic leadership strategies, an effect that persisted over time.

Fit and Engagement

The Inclusion workshop aimed to change people's beliefs and behaviours in ways that would have longer term benefits for participants' sense of fit and engagement.

Only Leadership participants showed declining levels of fit and engagement over the 18-20 month study period.

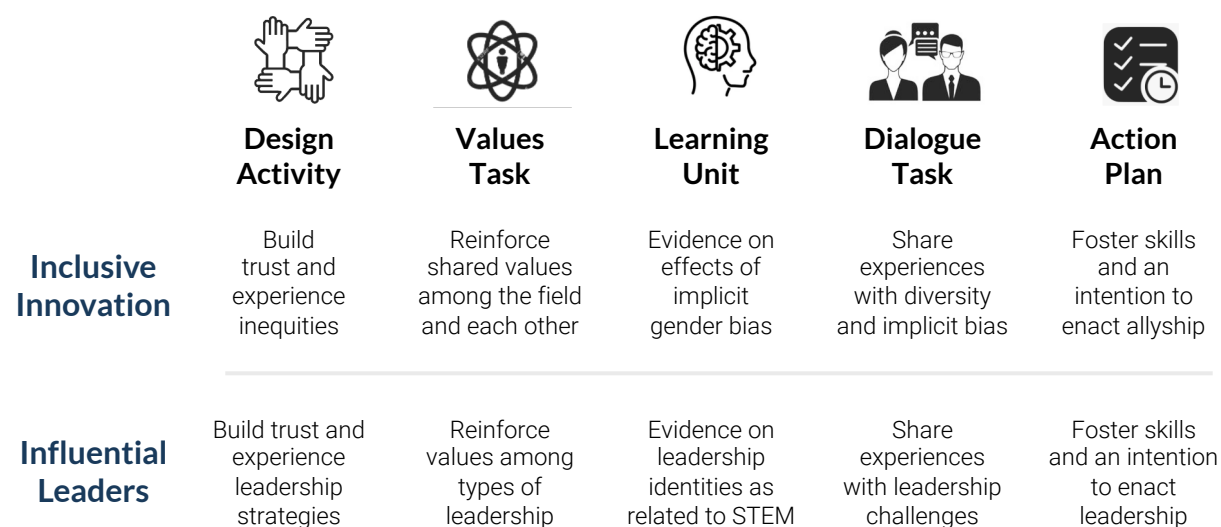
At 18-20 months, Inclusion participants (women in particular) reported more fit and engagement than those in the Leadership workshop.

Project Overview

Project RISE conducted a longitudinal study of employees in STEM (science, technology, engineering, and math) fields. The project used a **randomized control trial** to test short- and long-term effects of participating in one of two face-to-face evidence-based workshops.

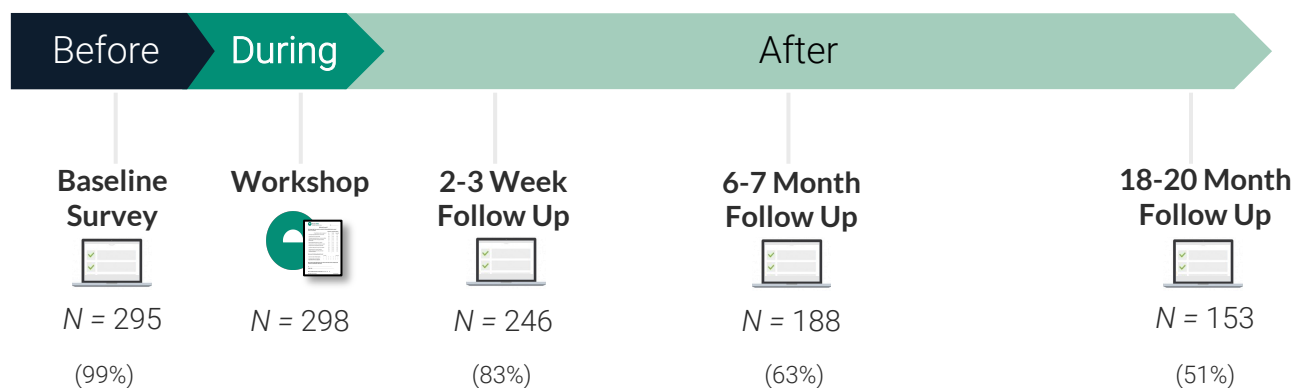
Workshop Architecture

Participants were randomly assigned to attend one of two workshops: **Inclusive Innovation** or **Influential Leaders**. The workshops followed a parallel structure with topic-tailored content.



Project Timeline

Participants completed surveys 1-2 weeks **before, during, and after** the workshop, across 2 years. Over half (51%) of the 298 workshop participants completed the final follow-up survey.

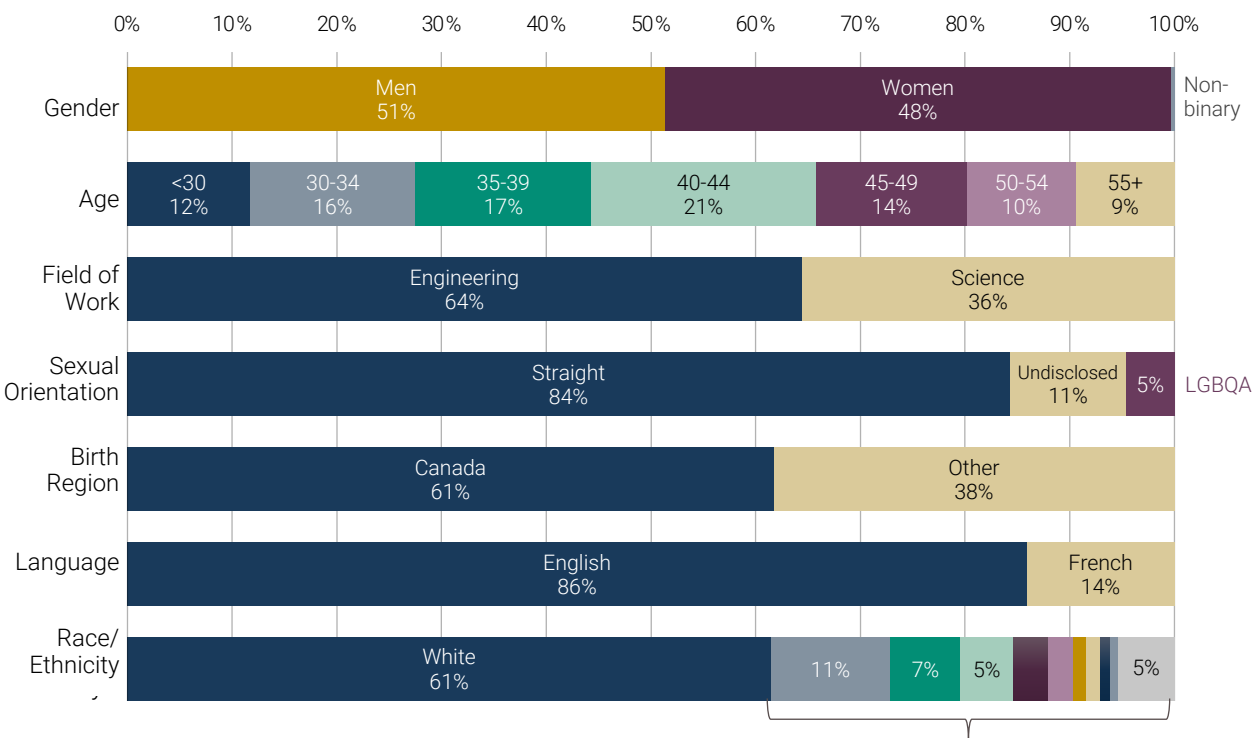


Sample Overview

298 employees from 7 organizations attended a 4.5-hour professional development workshop. Participant demographics were **similar** across the two workshop topics.

Participant Backgrounds

The average participant was **40-44 years old**, had worked for **5 years** in their organization (and **14 years** in the field), and reported having **mid-level status** within their organization.



Gender by Workshop Assignment

Similar gender distributions for **Inclusion** and **Leadership**.

	Men	Women	Nonbinary	Total
Inclusion	79	78	1	158
Leadership	74	66	0	140
Total	153	144	1	298

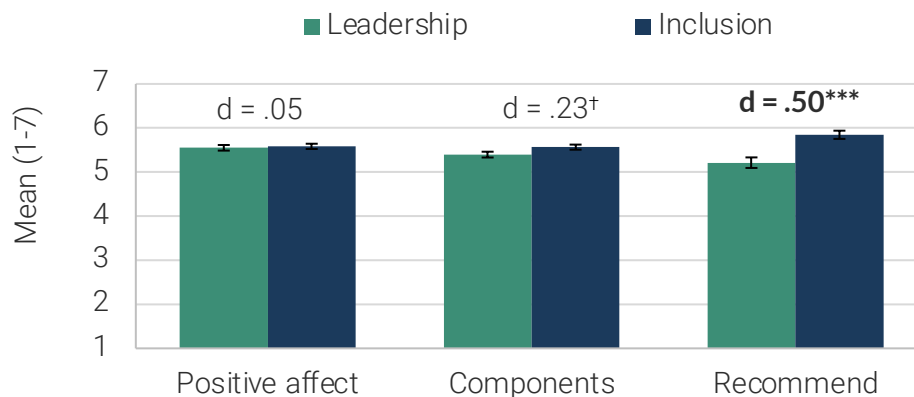
Note. For data privacy, no nonbinary results are reported ($n = 1$).

East Asian	11%
Middle Eastern	7%
South Asian	5%
Multi-racial	3%
Black	2%
Latinx	1%
Southeast Asian	1%
Indigenous	1%
Another Option	1%
Undisclosed	5%

Workshop Feedback & Outcomes

Both Workshops Were Rated Positively

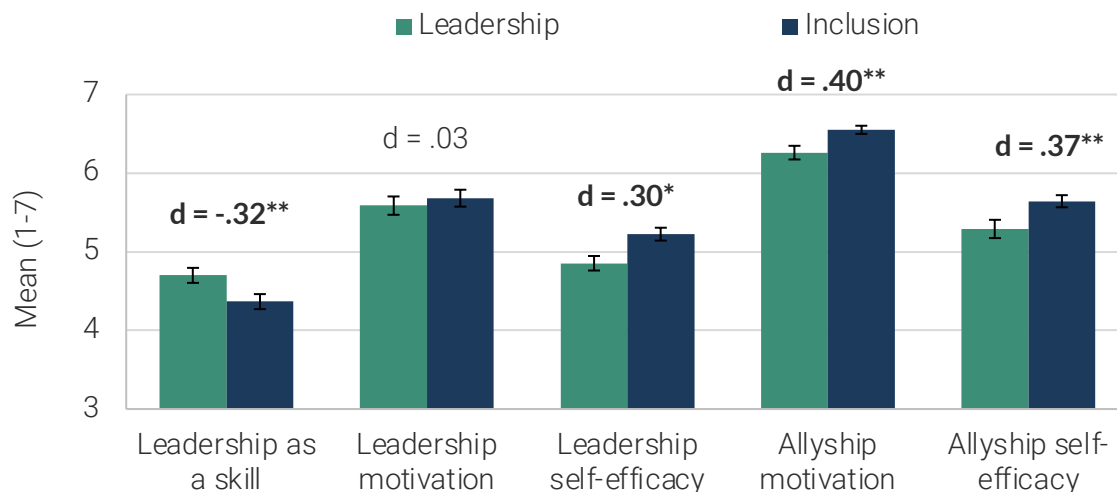
Participants **felt positive** (e.g., inspired) after both workshops and rated **components** favourably. Both workshops were **highly recommended**, particularly the Inclusion workshop.



Immediate Workshop Impacts

End-of-workshop ratings showed workshop effects on content-related beliefs and motivation. Tests of workshop differences control for participants' baseline levels of each outcome.

- Leadership participants gained more appreciation for **leadership as a skill** (vs. trait).
- Inclusion participants reported more **allyship motivation** and **allyship self-efficacy**.
- Unexpectedly, Inclusion participants also reported more **leadership self-efficacy**.



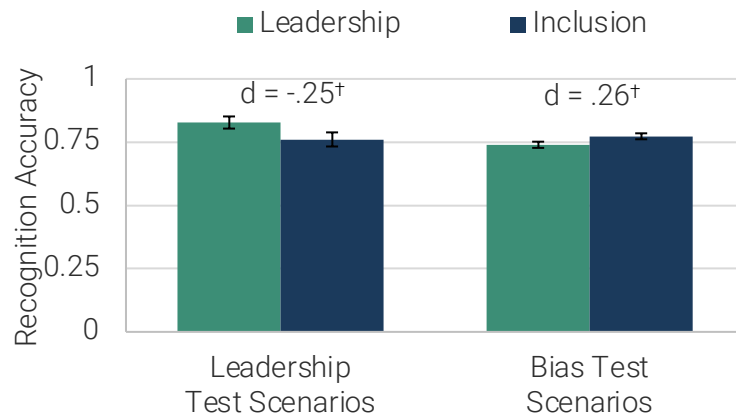
Leadership & Bias Recognition

Key Findings

- We tested participants' ability to **accurately recognize** leadership strategies and instances of bias.
- Recognition is key to **enacting** effective leadership and **confronting** bias.
- Leadership workshop increased accuracy in **identifying leadership strategies**.
- Inclusion workshop increased accuracy in **recognizing subtle gender bias**.
- Better bias recognition involved realizing **not all negative outcomes** women experience reflect gender bias.

Each Workshop Improved Accuracy for Related Content

The leadership workshop boosted accuracy in identifying leadership strategies and the Inclusion workshop improved accuracy in recognizing subtle gender bias.



Interaction (workshop x accuracy type), $p = .011$

Accuracy Test Scenarios

All participants saw scenarios depicting varied leadership strategies or presence vs. absence of subtle gender bias.

Your coworker Adam learns about a new system that he believes might help a project that Mary is spearheading in another department. He offers to put Mary in touch with another colleague using the system so that she can assess if it would be a useful resource.

Adam's behaviour is an example of which type of leadership?

Your coworker Jason mentions he is having difficulty on a project. You suggest he ask your other coworker, Samantha, for help because she's dealt with a similar situation. Jason scoffs and says Samantha is not exactly the most skilled person in the office and then rolls his eyes and laughs.

How likely is it that Jason's behaviour reflects gender bias?

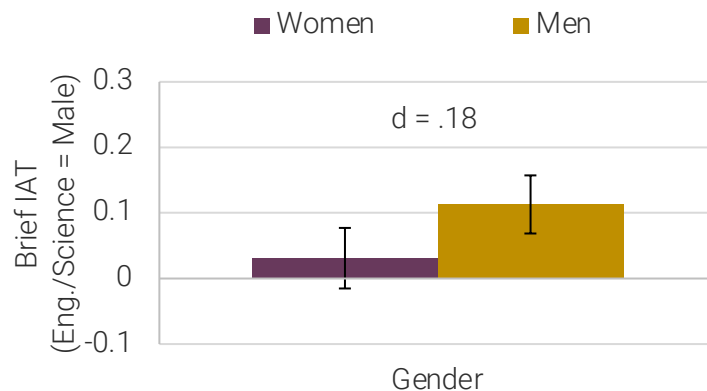
Implicit Stereotypes & Awareness

Key Findings

- **Men** had largest STEM = male implicit gender stereotype.
- **No workshop effects** on BIAT scores were predicted or found.
- **Engineers** tended to hold stronger implicit stereotypes than scientists ($p = .032$).
- Men held slightly stronger implicit gender stereotypes, yet women were more aware of their own stereotypes ($p = .014$).
- For men, the Inclusion workshop increased **awareness of own gender stereotypes**.

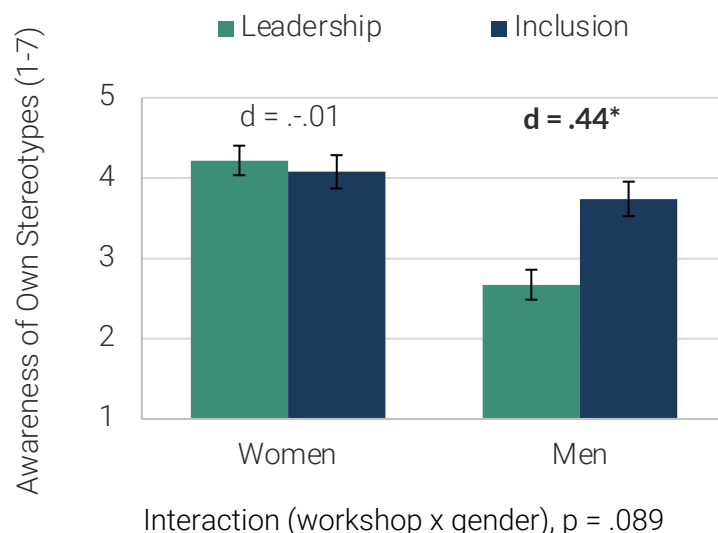
STEM = Male Implicit Association Observed, Especially Among Men

The Brief Implicit Association Test (BIAT) measured the automatic tendency to “think STEM, think male.” This implicit STEM = male association differed significantly from zero in the sample overall ($p = .036$), but primarily among men ($p = .012$) and not women ($p = .503$).



Inclusion Workshop Raises Men's Awareness of Own Stereotypes

The Inclusion workshop increased awareness of one's own implicit gender stereotypes among men, but not women (who started out higher in baseline awareness).



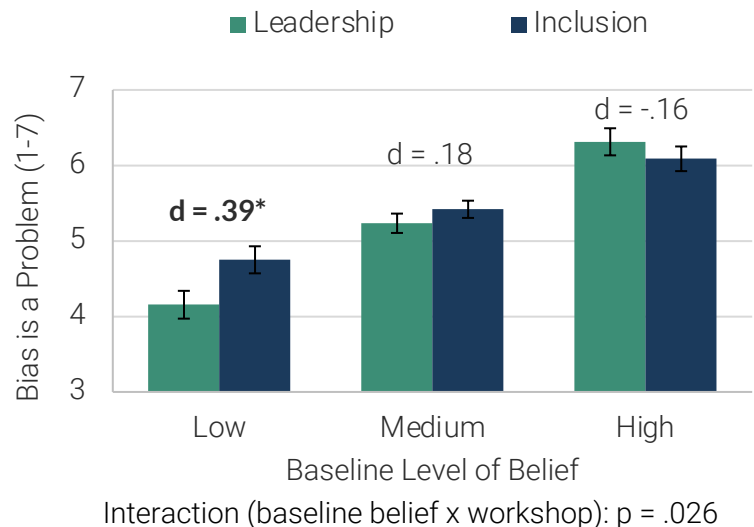
Beliefs About Bias & Allyship

Key Findings

- Inclusion workshop boosted the **belief that gender bias is a problem**, particularly among initial **skeptics**.
- Skeptics' increased awareness of bias as a problem suggests no evidence of **backlash**.
- Inclusion workshop **demonstrated** how allyship can increase women's inclusion.
- Participants viewed men's allyship as **more effective** after Inclusion workshop.

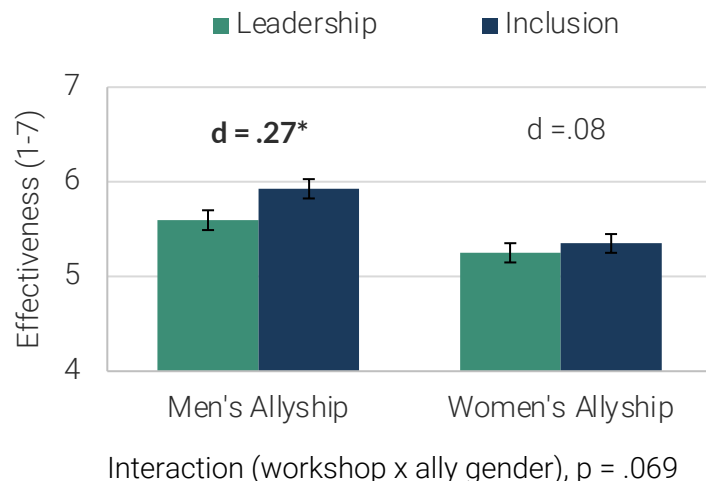
Inclusion Workshop Boosts Skeptics' Belief That Gender Bias is a Problem

For "skeptics" who initially saw bias as less problematic (1 SD below the mean), the Inclusion workshop boosted this belief 2-3 weeks ($d = .31^*$), 6-7 months ($d = .61^{**}$), and 18-20 months ($d = .39^*$, graphed below) later.



Inclusion Workshop Fosters Belief That Men's Allyship Works

Inclusion workshop helped participants see that allyship done by men (more so than women) is effective for making STEM workplaces more supportive for women.



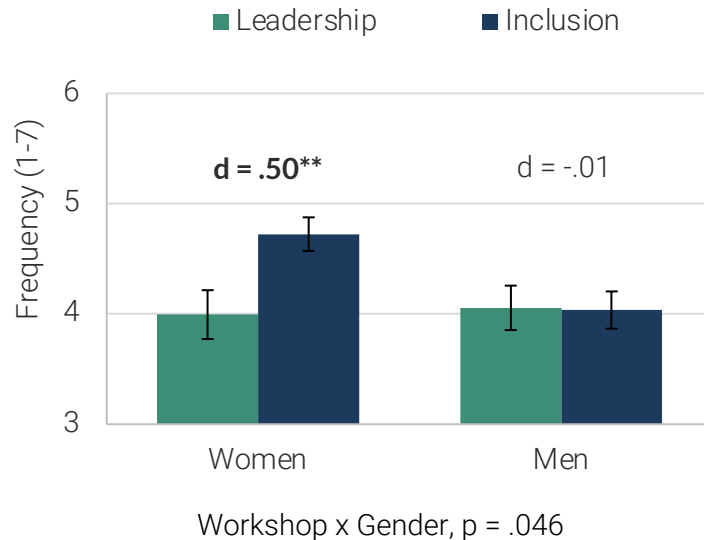
Enacting Allyship

Key Findings

- **Allyship to women in STEM** was assessed in general and for a personal action plan.
- Inclusion workshop boosted **women's general allyship** 2-3 weeks later.
- There were no workshop effects for **general allyship later** (see Appendices).
- After 2-3 weeks, women enacted their **personal allyship action plan** more often than did men.
- Men took their planned action **increasingly often over time**, closing the gender gap by 18-20 mos.

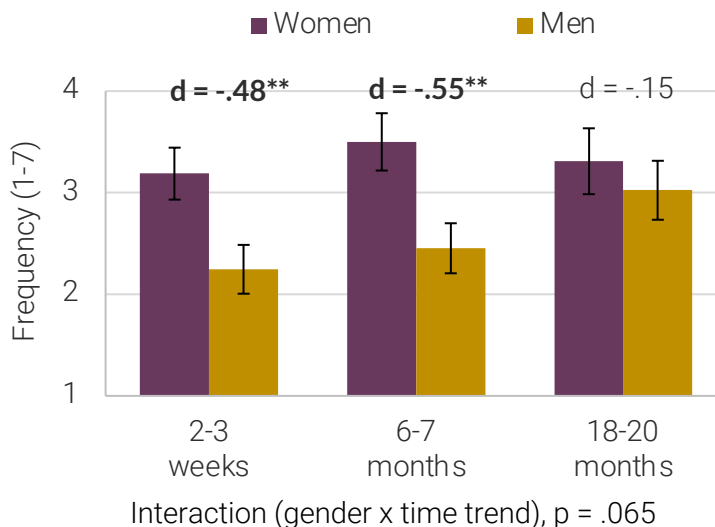
Women Acted as Allies More Often Than Men 2-3 Weeks Later

The Inclusion workshop increased women's (not men's) allyship actions in the subsequent 2-3 weeks.



For Specific Allyship Action Plans, the Gender Gap Narrows Over Time

Men enacted their specific allyship action plan increasingly often over time (trend $d = 0.46^{**}$), closing the gender gap.



Enacting Leadership

Key Findings

- Leadership workshop boosted belief in **technical mastery** as an **effective form** of leadership.
- Leadership workshop increased **use of new leadership strategies** (vs. autocratic ones).
- This shift in leadership strategy **persists across time** points.
- Workshop-based shift in strategy slightly **stronger among men**.
- Women and men enacted their personal **leadership action plans** similarly often (no over-time change).

Focus on Leadership Strategies

Leadership workshop contrasted traditional autocratic vs. newer STEM leadership strategies (Rottmann et al., 2015).



Autocratic Direction

Delegation + structure



Technical Mastery

Technical expertise + mentorship



Collaborative Optimization

Process optimization + team catalyst



Organizational Innovation

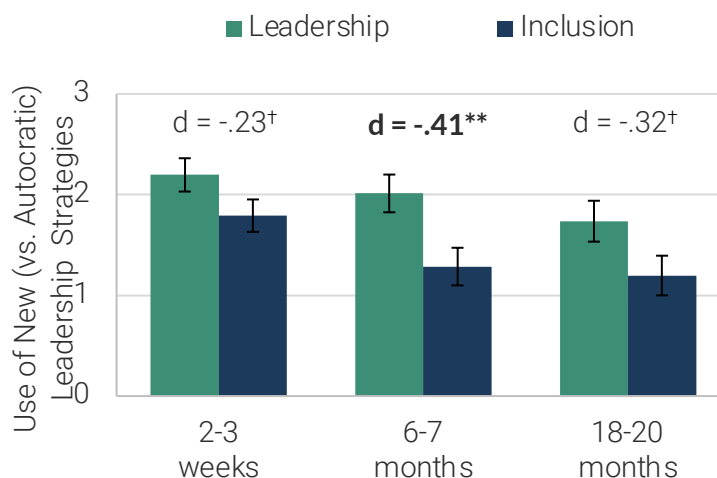
Innovation + realization

Predict having more nimble and innovative solutions

Leadership Increased Use of Newer Leadership Strategies

A behavioural shift to using these new strategies more than autocratic direction persisted across time points.

This workshop-driven shift was larger for men ($d_s \approx -.50^*$).



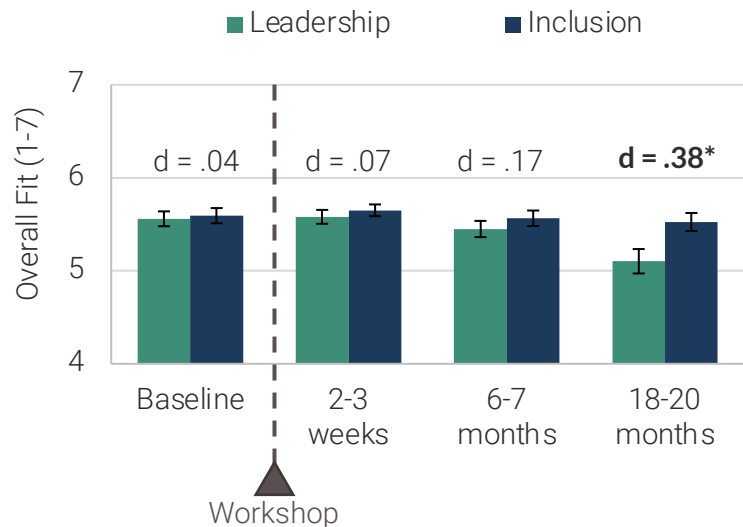
Fit & Engagement

Key Findings

- Inclusion workshop buffered against declines in **fit** over 1.5 years later.
- Inclusion workshop led to maintenance of **engagement** (i.e., less burnout) over time.
- Long-term workshop effects were **strongest for women's** fit and engagement.
- Women's engagement and fit were strongly **positively correlated**.

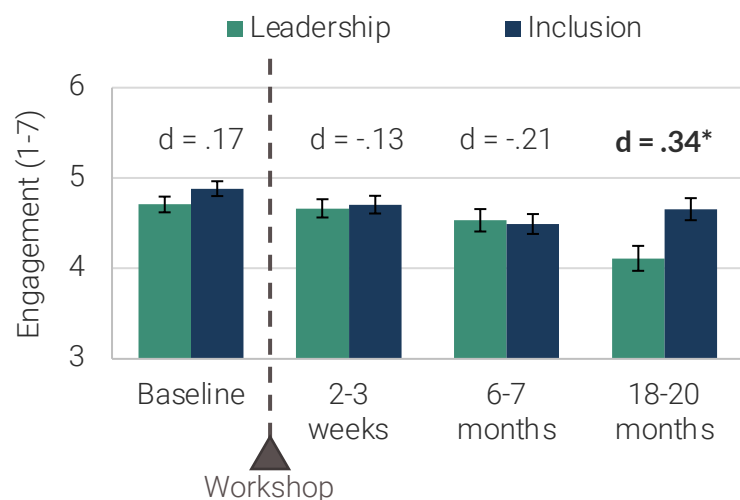
Inclusion Buffers Against Decrease in Fit Over Time

Inclusion sustained sense of fit 18-20 months later, especially among women ($d = .54^*$).



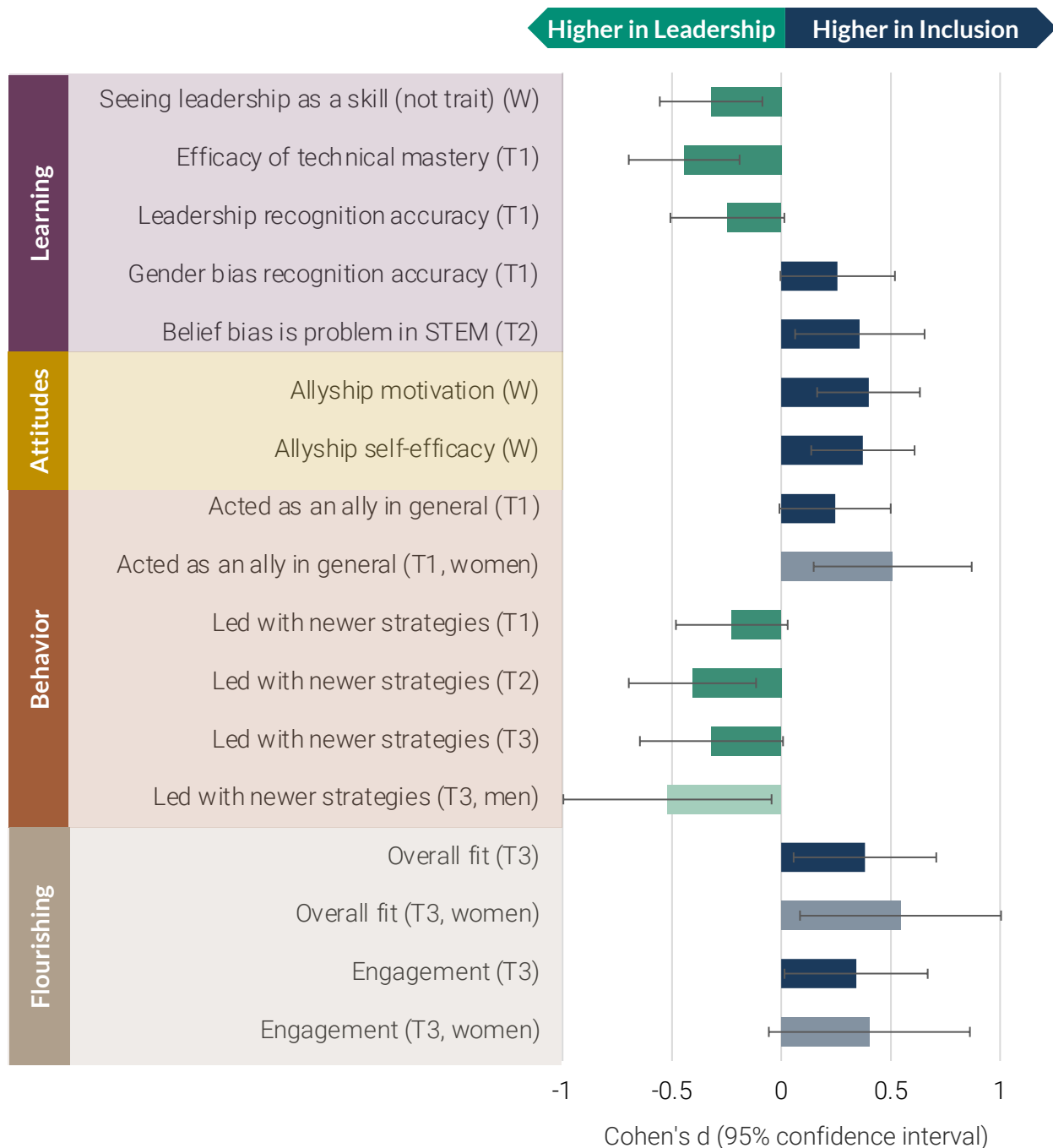
Inclusion Maintains Engagement

Inclusion also led to better maintenance of engagement (i.e., less burnout, greater commitment to organization) over time, especially for women ($d = .40^+$).



Key Findings at a Glance

This summary shows **effect sizes** for outcomes during the workshops (**W**), 2-3 weeks later (**T1**), 6-7 months later (**T2**), and 18-20 months later (**T3**), highlighting some workshop effects that were strongest among **men** or **women** specifically.



Glossary



Definitions

Gender Bias: Unequal treatment toward someone based on their gender (e.g., dismissing a woman's suggestion because she isn't expected to have technical expertise)

Implicit Gender Stereotypes: An automatic tendency to associate STEM more with men than with women.

Allyship: Taking action to support those who might otherwise be or feel excluded

- **Reactive Allyship:** Reacting to bias when we see it (e.g., confronting disrespectful, biased remarks)
- **Proactive Allyship:** Proactive efforts to increase inclusion (e.g., inviting women onto key projects)

Leadership Strategies: Strategies for influencing people

- **Autocratic Direction:** Initiating structure and delegating tasks
- **Technical Mastery:** Technical expertise and related mentoring
- **Collaborative Optimization:** Process optimization and catalyzing team skills
- **Organizational Innovation:** Innovation and shifting paradigms

Engagement: Experiencing work as meaningful, feeling less burnout and greater commitment to organization

Fit: Feeling well suited to one's environment based on self-concept match, goal alignment, and social belonging

Glossary



Symbols & Abbreviations

N – Sample size (number of participants)

n – Subgroup size

M – Mean (average)

SD – Standard deviation

SE – Standard error of mean

d – Cohen's d (measure of effect size)

p – indicator of statistical significance

[†] – marginal statistical significance at $p < .10$

^{*} – statistical significance at $p < .05$

^{**} – statistical significance at $p < .01$

^{***} – statistical significance at $p < .001$

Interpretation Guide: Tables & Graphs

Tables and graphs give **summary statistics**: the average score for participants in a given group.

Results of **inferential tests** (e.g., *d* and *p* values) additionally control for **variation across gender and workshop** and **baseline levels** of the same variable (where measured pre-workshop).

	Leadership	Inclusion	<i>d</i>
Name of first measure	4.70 (1.13)	4.37 (1.20)	-0.32**
Name of second measure	0.74 (0.13)	0.77 (0.13)	0.26†

Mean:
Group average

Standard Deviation:
Spread of typical person's score from group average

Cohen's *d*:

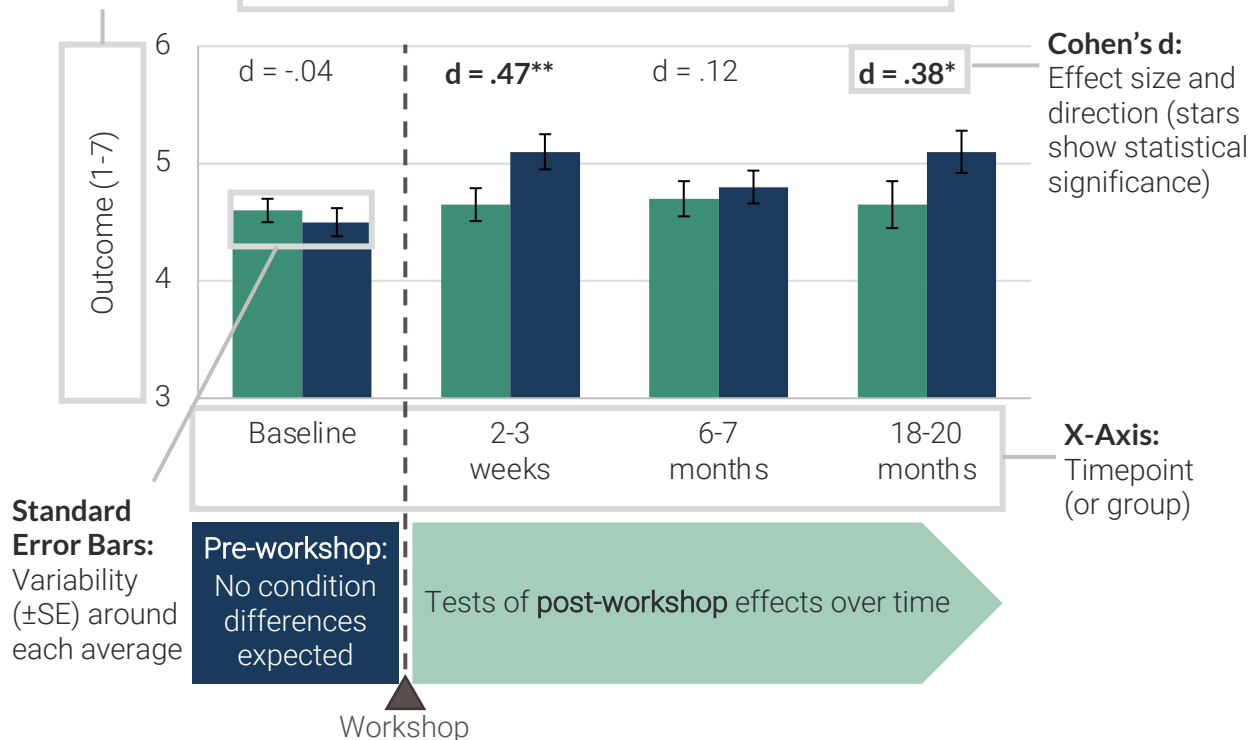
Effect size and direction (stars show statistical significance)

Example Graph

Below is an example of a graph showing results across time by workshop.

Y-Axis:

Outcome measured and its scale



Interpretation Guide: Effect Sizes

Definitions

Cohen's d: A standardized mean difference effect size metric expressing size of the difference between two groups in standard deviation units; unlimited range

Common language (CL): Probability that a score randomly sampled from one group will be larger than a score randomly sampled from the other group.

Interpreting Cohen's d

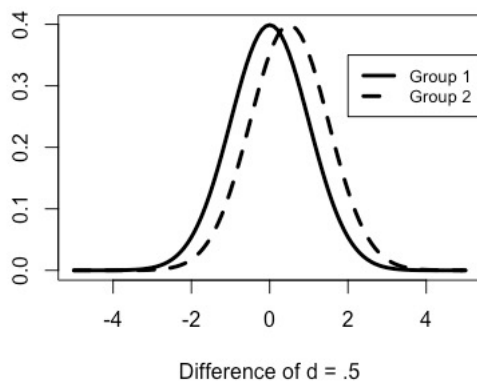
The **sign** of d indicates the **direction** of an effect.

Negative d = Leadership > Inclusion or Women > Men

Positive d = Leadership < Inclusion or Women < Men

Cohen's d	CL	Pearson's r	Interpretation (Funder & Ozer, 2019)
.00	50%	.00	No effect/association
.20	56%	.10	Small effect/association
.41	59%	.20	Medium effect/association
.63	65%	.30	Large effect/association
.87	71%	.40	Very large effect/association

Example: d = .5



Graph: Degree of overlap between two groups for d = .5, assuming normally distributed groups, each with a standard deviation of 1.

Interpretation: If d = .5, there is a 63.8% chance that a random member of Group 2 will score higher than someone from Group 1.

Further Reading

- Chang, E. H., et al. (2019). The mixed effects of online diversity training. *Proceedings of the National Academy of Sciences*, 116, 7778-7783.
- Devine, P. G., & Ash, T. L. (2022). Diversity training goals, limitations, and promise: a review of the multidisciplinary literature. *Annual Review of Psychology*, 73, 403-429.
- De Souza, L., & Schmader, T. (2021). The misjudgment of men: Does pluralistic ignorance inhibit allyship? *Journal of Personality and Social Psychology*, 122, 265-285.
- Hall, W. M., et al. (2015). Engineering exchanges: Daily social identity threat predicts burnout among female engineers. *Social Psychological and Personality Science*, 6, 528-534.
- Kalev, A., et al. (2006). Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. *American Sociological Review*, 71, 589-617.
- Moss-Racusin, C. A., et al. (2018). Reducing STEM gender bias with VIDS (video interventions for diversity in STEM). *Journal of Experimental Psychology: Applied*, 24, 236-260.
- Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A review and assessment of research and practice. *Annual Review of Psychology*, 60, 339-367.
- Rottmann, C., et al. (2015). Engineering leadership: Grounding leadership theory in engineers' professional identities. *Leadership*, 11, 351-373.
- Schmader, T., et al. (2020). Cracking the Culture Code. *Forgas, JP, Crano, WD, Fiedler, K*, 334-355.
- Schmader, T., et al. (2021). Why anti-bias interventions (need not) fail. doi.org/10.31234/osf.io/5rxk8
- Sriram, N., & Greenwald, A. G. (2009). The brief implicit association test. *Experimental Psychology*, 56, 283-294.

Visit us online to stay up-to-date on our latest findings



@ESSConsortium



@ess_consortium



successinSTEM.ca

Project RISE Team

Team Leaders



Hilary Bergsieker, PhD
University of
Waterloo



Toni Schmader, PhD
University of
British Columbia

Research Team



Audrey Aday, MA
University of
British Columbia



Emily N. Cyr, MA
University of
Waterloo



Lucy De Souza, MA
University of
British Columbia



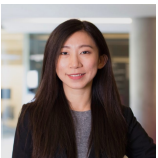
Tara C. Dennehy, PhD
University of
British Columbia



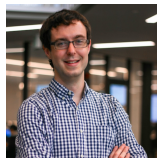
Grace Denney, BSc
University of
Waterloo



Will Hall, PhD
Brock
University



Joyce He, PhD
University of California,
Los Angeles



Erik Jansen, BA (Hon)
University of
Waterloo



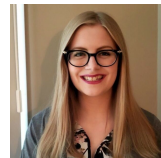
Sonia Kang, PhD
University of
Toronto



Lesley Shannon, PhD
Simon Fraser
University



Sheryl Staub-French, PhD
University of
British Columbia



Jessica Trickey, MA
University of
Waterloo

Workshop Facilitation



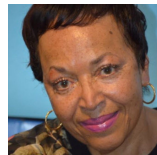
Susan Hollett, MSc



Cheryl Kristiansen, PEng

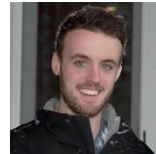


Erica Lee-Garcia, PEng

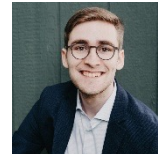


Helen Wilson, MA

Research Staff



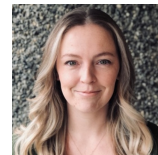
Tyler Hartwig, BSc



Seth Mahon, BSc



Trisha Smith, MA



Keltie Sommer, MSc

Thank You to Our ESS Partners

ENGINEERING & RESEARCH ORGANIZATIONS

City of Vancouver*
City of Port Moody*
City of New Westminster*
District of West Vancouver*
Metro Vancouver*
Cloud Army Network Inc.
Magnitude | Simba*
McElhanney Consulting Services*
National Research Council*
Teck Resources*
TRIUMF*
Bstro

PROFESSIONAL ASSOCIATIONS

Engineers Canada
Engineers and Geoscientists BC
MiHR

SCIENCE EDUCATION

Actua
Eng-Cite
Engineering Science Quest
Geering Up
Science Alive
Science World

NON-PROFIT CHANGE AGENTS

Engineering Change Lab
Gender and the Economy
WinSETT Centre
NSERC Chairs for Women in Science and Engineering
Ontario Network for Women in Engineering
Society for Canadian Women in Science and Technology
Canadian Commission for UNESCO

UNIVERSITIES

University of British Columbia
University of British Columbia, Faculty of Applied Science* & Faculty of Science*
Simon Fraser University
University of Toronto, Department of Psychology

University of Toronto, Faculty of Applied Science & Engineering
University of Toronto, Institute for Management & Innovation
University of Waterloo
McMaster University, Faculty of Engineering

This research made possible by:

SSHRC  **CRSH**

Social Sciences and Humanities Research Council of Canada
Conseil de recherches en sciences humaines du Canada

* RISE Data Collection Partners

Appendices



Measures: Items, Range, Sample Size Over Time

Subsets of measures were collected 1-2 weeks prior (**T0**), during the workshops (**W**), 2-3 weeks later (**T1**), 6-7 months later (**T2**), and 18-20 months later (**T3**). Sample items are provided below.

		Sample Size by Time					
Sample Survey Item(s) / Stimuli		Range	T0	W	T1	T2	T3
Workshop evaluations							
Enjoyment/positive affect	After participating in today's workshop, I feel [hopeful, inspired, empowered, valued - frustrated, guilty, frustrated, drained].	1-7	—	292	—	—	—
Rating of components	How did each of the following elements contribute to your experience in the workshop? [evidence, discussions, surveys...]	1-7	—	289	—	—	—
Recommendation likelihood	Are you likely to recommend this workshop to other coworkers?	1-7	—	292	—	—	—
Utility for personal leadership	Is this workshop likely to help you achieve your personal leadership goals?	1-7	—	293	—	—	—
Utility for organization D&I goals	Is this workshop likely to help your organization achieve its goals for diversity and inclusion?	1-7	—	293	—	—	—
Leadership beliefs							
Seeing leadership as a skill	To what extent is leadership a trait people are born with or a skillset that people can develop?	1-7	287	297	246	—	—
Bias-related beliefs & attitudes							
Support for gender inclusion	I personally support initiatives to increase women's participation in science.	1-5	284	—	246	—	—
Awareness of own gender bias	To what extent do you believe that you hold implicit gender stereotypes?	1-5	283	—	246	—	—
Belief bias is problem in STEM	To what extent do you believe that implicit gender bias towards women is a problem in STEM fields?	1-5	289	—	246	185	147
Motivation to address bias	When it comes to gender bias, I focus on how I can improve things.	1-7	288	—	246	185	146
Motivation to avoid bias	When it comes to gender bias, I want to avoid doing the wrong thing.	1-7	289	—	246	185	146
Motivation & efficacy							
Leadership motivation	I am motivated to learn and use new leadership strategies.	1-7	287	295	246	188	151
Leadership self-efficacy	I feel like I know how to be an effective leader in my organization.	1-7	287	291	246	188	151
Seeing self as a leader	Being a leader is an important part of my self-image.	1-7	287	—	246	188	151
Allyship motivation	I want to be an ally to female [scientists/engineers].	1-7	288	294	246	187	151
Allyship self-efficacy	I feel like I know how to be a strong ally to female scientists/engineers].	1-7	289	290	246	187	151
Perceived organizational climate							
Inclusive workplace culture	My organization has a culture of valuing the contributions that women and men make in the workplace.	1-7	245	292	246	185	152
% of men acting as allies	Please estimate the percentage of men at your organization who act as allies to female [scientists/engineers] in practice.	0-100	277	—	243	186	150
% of women acting as allies	Please estimate the percentage of women at your organization who act as allies to female [scientists/engineers] in practice.	0-100	277	—	243	187	150
Social identity threat	Over the past two weeks at work, I have felt very aware of my gender.	1-7	289	290	247	185	147
Organizational fit & engagement							
Overall fit	Composite: Average of self-concept, value, and social fit	1-7					
Self-concept fit	My job at my organization suits the way I see myself.	1-7	294	—	247	288	153
Value fit	The values of my organization are a good fit to my own personal values.	1-7	294	—	247	288	153
Social fit	I feel respected by people who work in my organization.	1-7	294	—	247	288	153
Engagement	I feel "emotionally attached" to this organization.	1-7	293	—	247	188	152

Measures: Items, Range, Sample Size Over Time (Continued)

Subsets of measures were collected 1-2 weeks prior (**T0**), during the workshops (**W**), 2-3 weeks later (**T1**), 6-7 months later (**T2**), and 18-20 months later (**T3**). Sample items are provided below.

		Sample Size by Time					
Sample Survey Item(s) / Stimuli		Range	T0	W	T1	T2	T3
Accuracy & understanding							
Leadership recognition accuracy	[Character's] behaviour is an example of which type of leadership? (Cross-scenario average)	0-1	—	—	233	—	—
Gender bias recognition accuracy	Difference score: Bias-present scenarios - bias-absent scenarios (rescaled to range 0 to 1)	0-1	—	—	231	—	—
Bias detection when present	How likely is it that [character's] behaviour reflects gender bias? [Averaged of scenarios with bias]	1-7	—	—	231	—	—
Bias detection when absent	How likely is it that [character's] behaviour reflects gender bias? [Averaged of scenarios without bias]	1-7	—	—	231	—	—
Implicit gender stereotype (BIAT)							
STEM = male implicit association	Faster association of male (vs. female) names with STEM words (testing, technology, design, math)	N/A	—	—	226	—	—
Efficacy perceptions (allyship)							
<i>How effective are the following practices for making workplaces more inclusive for women in STEM?</i>							
Men's reactive allyship	Men speaking up and reacting to instances of subtle bias when it occurs.	1-7	—	—	246	—	—
Men's proactive allyship	Men making proactive efforts to show respect for women's contributions	1-7	—	—	246	—	—
Women's reactive allyship	Women speaking up and reacting to instances of subtle bias when it occurs	1-7	—	—	246	—	—
Women's proactive allyship	Women making proactive efforts to show respect for women's contributions.	1-7	—	—	246	—	—
Efficacy perceptions (leadership)							
<i>To what extent do you think the following strategies are effective forms of leadership?</i>							
Autocratic direction	Autocratic direction: Adopts sole responsibility for setting organizational structure and delegating tasks.	1-7	—	—	246	—	—
Technical mastery	Technical mastery: Models technical expertise and mentors others on technical content.	1-7	—	—	246	—	—
Collaborative optimization	Collaborative optimization: Builds strong links and communication by matching project objectives, workers' skills, resources.	1-7	—	—	246	—	—
Organizational innovation	Organizational innovation: Promotes innovation toward the realization of visionary ideas that have impact.	1-7	—	—	246	—	—
Leadership & allyship behavior							
Acted as an ally in general	Over the past two weeks I have acted as an ally to women in STEM.	1-7	—	—	244	187	151
Action plan "if" scenario occurred	Since the workshop, how often has the "IF" situation occurred?	1-7	—	—	245	188	150
Took action plan "then" action	Since the workshop, how often have you taken your "then" action?	1-7	—	—	244	188	150
Led with newer strategies	Difference score: Average TM, CO, and OI strategy use - AD use <i>Over the past two weeks, I have led via...</i>	N/A	—	—	242	188	150
Used autocratic direction	...autocratic direction (e.g., giving directives to subordinates)	1-7	—	—	242	188	150
Used technical mastery	...technical mastery (e.g., technical mentoring)	1-7	—	—	243	188	150
Used collaborative optimization	...collaborative optimization (e.g., promoting effective team communication)	1-7	—	—	244	188	150
Used organizational innovation	...organizational innovation (e.g., helping teams see the big picture)	1-7	—	—	244	188	150

Full Sample: Baseline Levels

Measures assessed 1-2 weeks prior to the workshop. No workshop differences were expected.

	Leadership	Inclusion	d
Leadership beliefs			
Seeing leadership as a skill	4.62 (1.27)	4.52 (1.20)	-0.09
Bias-related beliefs & attitudes			
Support for gender inclusion	3.82 (0.82)	3.95 (0.67)	0.16
Awareness of own gender stereotypes	3.20 (1.60)	3.45 (1.67)	0.14
Belief bias is problem in STEM	4.53 (1.75)	4.79 (1.50)	0.14
Motivation to address bias	5.47 (1.49)	5.53 (1.30)	0.03
Motivation to avoid bias	5.76 (1.52)	5.73 (1.42)	-0.02
Motivation & self-efficacy			
Leadership motivation	5.95 (0.90)	6.00 (0.95)	0.05
Leadership self-efficacy	4.61 (1.41)	4.77 (1.42)	0.12
Seeing self as a leader	5.39 (1.28)	5.52 (1.29)	0.11
Allyship motivation	6.09 (1.18)	6.01 (1.16)	-0.08
Allyship self-efficacy	4.94 (1.55)	4.79 (1.48)	-0.11
Perceived organizational climate			
Inclusive workplace culture	5.54 (1.16)	5.40 (1.10)	-0.11
% of men acting as allies	62.0 (23.3)	56.6 (24.8)	-0.22 ⁺
% of women acting as allies	73.5 (21.2)	70.8 (20.8)	-0.13
Social identity threat	3.21 (1.73)	3.12 (1.76)	-0.08
Organizational fit & engagement			
Overall fit	5.56 (0.95)	5.59 (1.01)	0.04
Self-concept fit	5.61 (1.11)	5.51 (1.21)	-0.08
Value fit	5.59 (1.00)	5.68 (1.13)	0.09
Social fit	5.47 (1.26)	5.60 (1.27)	0.10
Engagement	4.71 (0.95)	4.88 (1.08)	0.17

Full Sample: Short-Term Outcomes

These measures were assessed **during** the workshop (*Workshop evaluations* and *End-of-workshop outcomes*) or **2-3 weeks later**. Analyses control for baseline levels (if available) and use a pooled full-sample error term, so estimated workshop effect sizes may differ in size or direction from directly comparing descriptive means.

	Overall			Women			Men		
	Leadership	Inclusion	d	Leadership	Inclusion	d	Leadership	Inclusion	d
Workshop evaluations									
Enjoyment/positive affect	5.55 (0.74)	5.58 (0.74)	0.05	5.52 (0.81)	5.72 (0.67)	0.28 ⁺	5.58 (0.67)	5.45 (0.78)	-0.19
Rating of components	5.40 (0.77)	5.57 (0.72)	0.23 ⁺	5.47 (0.68)	5.67 (0.65)	0.27	5.33 (0.84)	5.47 (0.77)	0.19
Recommendation likelihood	5.21 (1.40)	5.85 (1.16)	0.50***	5.25 (1.41)	5.91 (1.12)	0.52**	5.18 (1.39)	5.78 (1.21)	0.47**
Utility for personal leadership	4.92 (1.10)	5.15 (1.31)	0.19	4.98 (1.12)	5.18 (1.38)	0.17	4.86 (1.09)	5.12 (1.25)	0.21
Utility for organization D&I goals	4.13 (1.44)	5.58 (0.98)	1.21***	3.98 (1.42)	5.69 (0.98)	1.43***	4.27 (1.45)	5.46 (0.96)	1.00***
End-of-workshop outcomes									
Seeing leadership as a skill	4.70 (1.13)	4.37 (1.20)	-0.32**	4.82 (1.24)	4.48 (1.22)	-0.35*	4.59 (1.02)	4.25 (1.17)	-0.30 ⁺
Leadership motivation	5.59 (1.37)	5.68 (1.35)	0.03	5.52 (1.28)	5.95 (1.23)	0.19	5.64 (1.46)	5.42 (1.42)	-0.13
Leadership self-efficacy	4.85 (1.08)	5.22 (1.02)	0.30*	4.69 (1.14)	5.30 (0.95)	0.41*	5.00 (1.01)	5.15 (1.08)	0.19
Allyship motivation	6.26 (1.02)	6.55 (0.65)	0.40***	6.32 (0.95)	6.68 (0.55)	0.46**	6.21 (1.09)	6.43 (0.71)	0.34*
Allyship self-efficacy	5.29 (1.35)	5.64 (0.96)	0.37**	5.10 (1.41)	5.84 (0.87)	0.67***	5.46 (1.29)	5.46 (1.00)	0.09
Inclusive workplace culture	5.41 (1.23)	5.44 (1.00)	0.13	5.02 (1.26)	5.30 (1.12)	0.31	5.77 (1.10)	5.57 (1.84)	-0.04
Social identity threat	3.30 (1.93)	3.07 (1.80)	-0.11	3.98 (1.82)	3.39 (1.82)	-0.34 ⁺	2.68 (1.84)	2.74 (1.72)	0.11
Accuracy & understanding									
Seeing leadership as a skill	4.91 (1.08)	4.53 (1.22)	-0.34**	5.02 (1.14)	4.75 (1.15)	-0.24	4.80 (1.01)	4.31 (1.26)	-0.44*
Leadership recognition accuracy	0.83 (0.26)	0.76 (0.30)	-0.25 ⁺	0.81 (0.26)	0.80 (0.28)	-0.04	0.84 (0.26)	0.72 (0.32)	-0.46*
Gender bias recognition accuracy	0.74 (0.13)	0.77 (0.13)	0.26 ⁺	0.75 (0.13)	0.78 (0.13)	0.28	0.73 (0.13)	0.76 (0.13)	0.24
Bias detection when present	4.92 (1.31)	4.98 (1.27)	0.04	5.12 (1.25)	5.07 (1.30)	-0.03	4.73 (1.36)	4.88 (1.25)	0.11
Bias detection when absent	2.05 (1.13)	1.70 (0.99)	-0.33*	2.13 (1.24)	1.65 (1.00)	-0.45*	1.97 (1.02)	1.75 (0.99)	-0.21
Implicit gender stereotype (BIAT)									
STEM = male implicit association	0.04 (0.45)	0.10 (0.47)	0.12	0.05 (0.49)	0.03 (0.40)	-0.05	0.03 (0.42)	0.17 (0.52)	0.30
Bias-related attitudes									
Awareness of own gender stereotypes	3.20 (1.60)	3.45 (1.67)	0.21	3.97 (1.47)	3.64 (1.71)	-0.01	2.51 (1.38)	3.27 (1.62)	0.44*
Support for gender inclusion	3.75 (0.77)	3.96 (0.67)	0.22 ⁺	3.96 (0.65)	4.03 (0.67)	0.05	3.55 (0.83)	3.89 (0.68)	0.40*
Efficacy perceptions									
Men's reactive allyship	5.53 (1.39)	5.87 (1.20)	0.26*	5.75 (1.33)	5.97 (1.27)	0.18	5.33 (1.42)	5.76 (1.11)	0.34 ⁺
Men's proactive allyship	5.65 (1.23)	5.99 (1.20)	0.28*	5.88 (1.19)	6.11 (1.30)	0.19	5.43 (1.24)	5.87 (1.09)	0.37*
Women's reactive allyship	5.04 (1.43)	5.14 (1.29)	0.08	4.88 (1.50)	5.06 (1.34)	0.14	5.20 (1.36)	5.23 (1.23)	0.02
Women's proactive allyship	5.46 (1.10)	5.56 (1.15)	0.08	5.68 (0.97)	5.59 (1.19)	-0.08	5.25 (1.18)	5.52 (1.11)	0.25
Autocratic direction	3.23 (1.42)	3.19 (1.63)	-0.03	3.17 (1.39)	3.22 (1.69)	0.03	3.30 (1.45)	3.16 (1.58)	-0.09
Technical mastery	5.32 (1.11)	4.76 (1.38)	-0.44***	5.36 (0.98)	4.78 (1.27)	-0.46*	5.28 (1.23)	4.74 (1.50)	-0.43*
Collaborative optimization	6.32 (0.80)	6.30 (0.75)	-0.02	6.41 (0.72)	6.41 (0.66)	0.00	6.23 (0.86)	6.19 (0.83)	-0.05
Organizational innovation	5.62 (1.01)	5.77 (1.13)	0.14	5.81 (0.96)	5.98 (0.97)	0.16	5.43 (1.04)	5.55 (1.25)	0.12

Full Sample: Over-Time Outcomes

These measures were assessed at each post-workshop survey timepoint. All analyses control for baseline, except leadership and allyship behaviour (baseline not collected), and use a pooled full-sample error term, so estimated workshop effect sizes may differ in size or direction from directly comparing descriptive means.

Action plan content varied by workshop, so these outcomes are not directly comparable.

	2-3 Weeks Post-Workshop			6-7 Months Post-Workshop			18-20 Months Post-Workshop		
	Leadership	Inclusion	d	Leadership	Inclusion	d	Leadership	Inclusion	d
Bias-related beliefs									
Belief bias is problem in STEM	4.85 (1.73)	5.19 (1.50)	0.17	4.96 (1.62)	5.39 (1.17)	0.36*	5.30 (1.50)	5.47 (1.18)	0.16
Motivation to address bias	5.68 (1.19)	5.85 (1.00)	0.24 [†]	5.59 (1.56)	5.78 (1.11)	0.30*	5.62 (1.58)	5.74 (1.24)	0.21
Motivation to avoid bias	5.92 (1.17)	6.09 (1.07)	0.21	5.75 (1.48)	6.18 (1.08)	0.44**	5.75 (1.54)	5.86 (1.31)	0.11
Motivation & efficacy									
Leadership motivation	5.74 (0.97)	5.84 (1.01)	0.12	5.54 (1.06)	5.63 (1.09)	0.16	5.27 (1.22)	5.60 (1.21)	0.44**
Leadership self-efficacy	4.93 (1.17)	4.88 (1.29)	-0.17	4.88 (1.12)	5.02 (1.31)	-0.02	4.75 (1.47)	4.93 (1.34)	-0.11
Seeing self as a leader	5.13 (1.31)	5.34 (1.31)	0.09	5.13 (1.37)	5.50 (1.19)	0.25 [†]	4.93 (1.35)	5.61 (1.26)	0.57***
Allyship motivation	6.18 (1.00)	6.30 (0.89)	0.18	6.24 (1.11)	6.39 (0.84)	0.22	6.33 (1.06)	6.45 (0.79)	0.13
Allyship self-efficacy	5.16 (1.26)	5.25 (1.11)	0.09	5.38 (1.20)	5.28 (1.06)	-0.03	5.52 (1.06)	5.48 (1.01)	-0.05
Leadership & allyship behaviour									
Acted as an ally in general	4.03 (1.62)	4.39 (1.31)	0.25 [†]	4.36 (1.55)	4.12 (1.46)	-0.16	4.62 (1.56)	4.42 (1.49)	-0.12
Action plan "if" scenario occurred	3.67 (1.58)	2.26 (1.56)	-0.91***	3.93 (1.79)	2.61 (1.61)	-0.80***	4.32 (1.65)	2.87 (1.52)	-0.92***
Took action plan "then" action	3.87 (1.91)	2.73 (2.01)	-0.60***	4.02 (1.94)	2.97 (1.98)	-0.54***	4.33 (1.91)	3.17 (1.99)	-0.59***
Led with newer strategies	5.16 (1.26)	5.25 (1.11)	-0.23 [†]	5.38 (1.20)	5.28 (1.06)	-0.41**	5.52 (1.06)	5.48 (1.01)	-0.32 [†]
Used autocratic direction	2.71 (1.73)	2.73 (1.77)	0.01	2.77 (1.65)	3.30 (1.94)	0.30*	3.03 (1.67)	3.42 (1.87)	0.22
Used technical mastery	4.88 (1.53)	4.28 (1.83)	-0.36**	4.84 (1.55)	4.40 (1.68)	-0.28 [†]	4.48 (1.49)	4.46 (1.70)	-0.02
Used collaborative optimization	5.39 (1.23)	5.04 (1.42)	-0.27*	5.34 (1.32)	4.99 (1.52)	-0.24 [†]	5.24 (1.47)	5.00 (1.51)	-0.16
Used organizational innovation	4.43 (1.79)	4.27 (1.75)	-0.09	4.17 (1.70)	4.38 (1.80)	0.12	4.58 (1.57)	4.40 (1.64)	-0.12
Perceived organizational climate									
Inclusive workplace culture	5.60 (1.12)	5.53 (1.00)	0.05	5.59 (1.20)	5.57 (1.03)	0.12	5.25 (1.17)	5.39 (1.00)	0.17
% of men acting as allies	62.6 (24.4)	56.4 (23.1)	-0.14	63.4 (23.8)	59.5 (19.9)	-0.19	62.9 (24.5)	61.7 (20.6)	-0.02
% of women acting as allies	74.6 (19.5)	67.3 (22.6)	-0.38**	74.3 (19.8)	71.5 (18.9)	-0.09	74.6 (17.8)	71.5 (19.9)	-0.25
Social identity threat	2.64 (1.65)	3.02 (1.73)	0.28*	2.66 (1.69)	2.55 (1.73)	-0.02	3.02 (1.84)	2.67 (1.63)	-0.16
Organizational fit & engagement									
Overall fit	5.58 (0.82)	5.65 (0.71)	0.07	5.45 (0.80)	5.57 (0.86)	0.17	5.10 (1.08)	5.53 (0.89)	0.38*
Self-concept fit	5.43 (1.06)	5.67 (0.98)	0.32*	5.30 (0.98)	5.47 (1.18)	0.25 [†]	4.93 (1.33)	5.49 (1.05)	0.50**
Value fit	5.67 (0.88)	5.77 (0.87)	0.08	5.52 (1.00)	5.60 (0.95)	0.05	5.01 (1.35)	5.55 (1.01)	0.40*
Social fit	5.64 (1.08)	5.51 (1.07)	-0.23 [†]	5.53 (0.99)	5.63 (0.96)	0.09	5.37 (1.17)	5.53 (1.10)	0.04
Engagement	4.67 (1.09)	4.71 (1.10)	-0.13	4.53 (1.09)	4.49 (1.15)	-0.21	4.10 (1.19)	4.66 (1.07)	0.34*