Impact of a Women in STEM Conference on Two Indicators of Career Persistence: Evaluation Design and Results

Vivian Meng[*a1], Jennifer Pelletier[b], Rebekah Parker[c], and Elizabeth Croft[b]

[a] Department of Statistics, University of British Columbia, 3182 Earth Sciences Building, 2207 Main Mall, Vancouver, BC, Canada, V6T 1Z4
[b] Department of Mechanical Engineering, University of British Columbia, 2054-6250 Applied Science Lane, Vancouver, BC, Canada, V6T 1Z4
[c] Department of Curriculum & Pedagogy, University of British Columbia, 2125 Main Mall, Vancouver, BC, Canada, V6T 1Z4

*Corresponding author. E-mail: vivian.meng@mail.mcgill.ca

Introduction

Workshops and conferences are common venues for dissemination of information and are often cast as interventions aimed at a specific outcome for the target audience, e.g., gaining knowledge in a particular field, or developing a particular skill. While the outcomes of very focused training sessions can be tested using traditional means, the effectiveness of broader-topic, short-duration events (e.g., conferences for: women in science, technology, engineering and math (STEM), aboriginal educational leadership, men in nursing) to change outcomes for participants is not well documented.. Without measurable proof of an intervention's impact, it is difficult to create sponsor and participant buy-in for future activities. Measurement, therefore, is an important yet neglected component of these types of interventions.

In the past few decades, studies have presented the business case for gender diversity, citing benefits for organisations including better economic performance (Council of Canadian Academies, 2012; Catalyst, 2004, 2007, 2011; Adler, 1999, 2009; Orser, 2000), improved governance (Mateos de Cabo *et al.*, 2012; Brown *et al.*, 2002; Boulanta, 2013), increased innovation (Woolley and Malone, 2011; Torchia *et al.*, 2011; Diaz-Garcia *et al.*, 2013), and recruiting from a wider talent pool (Grosvold, 2011). Many traditional workplaces currently emphasise linear career paths, a large amount of in-person time at the workplace, and long hours; characteristics that do not reflect the challenges of many highly qualified women, 60% of whom have non-linear careers (Hewlett, 2007). Engineering, specifically, has a highly competitive culture which can increase stress levels for women, (Dryburgh, 1999), dissuade employees from taking advantage of inclusive policies (Lee *et al.*, 2010), and does not promote inclusivity for women in the industry (Cheryan, 2012; Diekman *et al.*, 2010). Shifting traditional workplace

[1] Present address: Department of Mathematics and Statistics, McGill University, Burnside Hall Room 1032, 805 Sherbrooke W., Montreal, QC, H2A 0B9, Canada

climates towards inclusive practices is essential for retaining a diverse workforce, and an inclusive climate can reduce relationship and task conflict, and turnover in gender-diverse groups (Nishii, 2013). Given the need to replace a rapidly aging engineering workforce (Engineers Canada, 2015), gender equality in the workforce is worth pursuing.

The participation of women in Science, Technology, Engineering, and Mathematics (STEM) remains low despite achieving overall gender representation in the workforce. In Canada, women make up 47.4% of the total workforce in 2006, but only 21.9% of the paid workforce in science and engineering occupations (Statistics Canada, 2006). For engineering specifically, only 11.7% of Canadian Professional Engineers (P.Eng.) are women (Engineers Canada, 2014). Similar figures exist in the US; in 2010 women held 28% of science and engineering positions, and only 13% of engineering positions (National Science Foundation, 2013).

Research on the barriers to women's participation in STEM is prolific, as summarised in reports initiated by national bodies, including *Why so few?* (AAUW, 2010) supported by the American National Science Foundation, and *The Gender Dimension* (Council of Canadian Academies, 2012) published in response to a request by the Canadian Minister of Industry. Both reports list perceived interest, workplace environment, implicit bias towards women and family responsibility as major barriers for women's participation in STEM.

These investigative reports and other literature also provide suggested practices to increase the recruitment and retention rates of women in STEM. Some documented practices/interventions include mentorship (Rutz and Shafter, 2011), workshops (Lawrence and Mancuso, 2012; Rutz and Shafer, 2011; Weavers *et al.*, 2011), and leadership development programs. Other suggestions include workplace interventions such as adopting family friendly policies, changing workplace culture, and increasing the presence of women in higher positions (Servon and Visser, 2010).

Conferences on the topic of women in STEM are increasingly common, aimed at raising awareness of barriers to and best practices for supporting women's involvement in STEM. This type of event has the potential to influence a large number of people. Examples of high profile conferences include: nationally, the Canadian Coalition for Women in Engineering, Science, Trades and Technology (CCWESTT) Conference, and the Gender and STEM Conference in the Netherlands, and internationally, the Gender Summit, and the Anita Borg Institute's Grace Hopper Celebration. These conferences share similar mandates along three themes: (1) increase women's involvement and persistence in STEM; (2) build support networks and mentorship; and (3) advertise career opportunities (VHTO, n.d.; Canadian Coalition of Women in Engineering, Science, Trades and Technology, 2014; Portia Ltd., 2013; Anita Borg Institute, n.d.). However, there is a lack of public reporting on conference evaluation, with the few available statistics focusing on participant satisfaction (Anita Borg Institute, n.d.) as a measure of success, rather than reflecting on the conference goals.

Program evaluation is considered by experts as crucial to program success and sustainability (Caffarella, 2002; George-Jackson and Rincon, 2012); it allows us to understand

and communicate the value of the intervention to the stakeholders, and the public. The lack of program evaluation in STEM conferences may be a reason why major reports such as *The Gender Dimension* have not mentioned conference activities as an important intervention in the multifaceted effort towards increasing women's representation and persistence in STEM.

This paper investigates the value and impact of conferences on women's representation and career persistence in STEM. The objective of this paper is twofold. First, we propose a framework of program and evaluation planning that allows us to understand the value of conferences to women's representation and persistence. Our focus is on deciding *what* and *how* to evaluate at a conference. We develop our approach with the case study of the regional Creating Connections Conference 2013 (CC2013) held in Vancouver, BC.

The second objective of this paper is to demonstrate that conferences, generally of short duration, impact participants by inducing changes in the participants that are sustained in the short- and medium-term. This is highlighted through the results from CC2013, where we captured short-term changes with a post-event measurement immediately following the conference, and medium-term changes with a follow-up measurement at six months post-event. A statistical test of significance is used to conclude a non-zero change in participants on average.

We note that this paper marks the first attempt at not only explicitly linking the impact of conference events to women's persistence in STEM, but also in the use of validated instruments at a Canadian conference for women in STEM. The subsequent sections account the inter-dependence between conference and evaluation design decisions; thus, this paper provides an exemplary demonstration of the importance of hand-in-hand program and evaluation planning (Caffarella, 2002).

Methods

Selecting appropriate measures of success

The importance of embedding program objectives in evaluation is seen in popular approaches to evaluation, including: the "Levels of Evaluation" approach (Kirkpatrick, 1998; Guskey, 2000), objective-based approach (Caffarella, 2002), and the "Accountability Planner" approach (Vella *et al.*, 1998). We adopt the objective-based approach in our design; in this approach, "the purpose, design, and criteria for the evaluation are all drawn from [the] objectives" (Caffarella, 2002, p. 249).

As previously mentioned, many conference organizers, including the authors, consider *increasing* the involvement and career persistence for women in STEM as an important goal for intervention activities. While women's increased involvement and persistence in STEM can be measured through a longitudinal study of gender ratio and career attrition rates, it is unrealistic to expect to relate a single event to career outcomes. Rather we propose to measure proxy targets, namely attitudes or conditions that would be precursors to involvement and persistence for women in STEM. We reviewed psychology literature to find psychological constructs that can

reflect changes incited by interventions in a timely fashion. We later define the changes required of these target constructs as the program objectives.

The awareness of the benefits of gender diversity in the workplace (awareness-BGD) was identified as an appropriate construct for evaluation because it is related to an important barrier to women's participation in STEM: devaluation in the work environment. In the Panel analysis for *The Gender Dimension* on a Survey of Canadian Research Chairs (Council of Canadian Academies, 2012), devaluation by colleagues and superiors was the most frequently cited barrier to women's participation in STEM. Literature suggests that raising the consciousness (synonymous to awareness) of a problem behavior is the first step towards creating behavioral (Prochaska et al., 1992) and social (Kloos et al., 2011) change. Based on research by Smith and Petty (1996) on efficacy of messaging, we chose to target a positive message (the awareness of the value/benefit of gender diversity), rather than a negative message (the consequences of devaluation). We adopted the 18-item, validated awareness-BGD instrument developed by the Authors (Author citation, forthcoming) for measuring awareness-BGD at CC2013. The instrument was tested for construct validity through factor analysis and "known-group" approach, internal-consistency through the Crochbach's alpha coefficient, test-retest reliability through a two week separation between tests, and sensitivity to change through a controlled experiment.

Career self-efficacy is the second construct we chose to target and measure in our conference. Self-efficacy is an important target of intervention because research shows that it is unequal amongst genders (Concannon and Barrow, 2012) and a recognized barrier (Orser *et al.*, 2012) for women in STEM. Furthermore, it positively correlates with career persistence (Hackett and Betz, 1981; Ballout, 2009). We measured the change in self-efficacy to correlate the impact of our conference with change in career persistence. We adopted the six-item career self-efficacy instrument developed by Rigotti *et al.* (2008) for measuring career-self-efficacy at the conference.

Conference design

Creating Connections 2013 was a 300-person regional conference held in Vancouver, BC, Canada at the University of British Columbia. The conference was open to all, but focused on issues related to gender diversity in STEM. It was part of an established bi-annual conference series, namely, the Building Communities Symposium in 2007, and Creating Connections Conferences in 2009 and 2011. Based on the literature review above, conference organizers established two main objectives for the conference: (1) increasing participants' awareness of the benefits of gender diversity in STEM, and (2) having a positive effect of career persistence for women participants.

The format of the intervention was a one-evening, one-day conference. This allowed the use of multiple learning formats and styles, ensuring that participants would stay engaged (Ravn, 2007; Louw and Zuber-Skerritt, 2011), could select sessions that best suited their learning preferences (Haley *et al.*, 2009), and allowed for a larger number of participants than could be

accommodated at a workshop or seminar. Efforts were made to ensure barriers to participation were eliminated (childcare provisions, sponsorships for students, travel funding). The conference included a wide range of topics beyond gender diversity, addressed through a diversity paradigm, to attract a broader audience (Hoyt and White, 2011; Mair and Thompson, 2009; Briziarelli, 1996).

The conference structure included formal and informal learning, mentoring, and reflection. Formal learning took place through keynote lectures, workshops, and panel discussions. Informal learning included networking, world café discussions (loosely guided brainstorming), and idea exchanges (informal discussion circles). Providing both formal and informal learning opportunities allowed participants to both obtain and exchange knowledge.

Participants were guided through mentoring and reflection through purposeful framing of the event by the organizers in the program and in introductions to conference content (Ravn, 2007; Ravn and Elsborg, 2011), and Connect and Reconnect sessions - small groups that gathered at the beginning and end of the day. The reflections were designed to ensure participants could explicitly identify key experiences and knowledge. The questions for these sessions are listed in Table 1. Reflection is important to clarify an experience and link it to other domains.

Bandura (1977) specified four factors contributing to self-efficacy: (1) performance accomplishments, (2) modeling, (3) encouragement and support, and (4) reduced anxiety. Conference programming targeted two of these factors. Modeling was present in the panel discussions, keynote lectures, and some parallel sessions. Encouragement and support was provided through networking, Connect and Reconnect, and informal learning sessions.

Table 1. Questions for Reflection during the Creating Connections conference 2013

Framing Messaging	We leave you with three goals to accomplish today: 1. Connect with three new people and learn their stories 2. Discuss the value gender diversity brings to our organization 3. Let yourself be inspired
Connect Group Questions	1. Icebreaker: Introduce yourself: who are you, how do you spend your time, and why do you relate to this interest group? 2. Why did you decide to come to Creating Connections? Name one thing that you want to take away from today's event.

Reconnect Group Questions	1. What is the key thing that you are taking away from your experience today? 2. What one thing will you do or change as a result of your experience today?

Evaluation design

Survey format

All survey questionnaires used in this study contain four major components: study description and informed consent, self-generated identifier code (Yurek *et al.*, 2008), self-efficacy survey items, followed by awareness-BGD survey items.

The self-efficacy survey items were taken from Rigotti *et al.* (2008), composed of six questions and were used without modification. The awareness-BGD survey items were taken from (Author citation, forthcoming), a total of eighteen questions, and were used without modification.

The post-event survey contained an extra page at the end, which solicited the participant's consent to be contacted for the follow-up survey. Participants were asked to provide contact details if they consented to being contacted.

Evaluation logistics

Conference evaluation was administered at three points: immediately pre-conference (pre), immediately post-conference (post), and at six months after the conference (follow-up). The pre-conference evaluation consisted of paper-based surveys. At the time of registration before the plenary event, participants were given the pre-conference survey and asked to fill it out. Participants were also approached by staff to encourage participation of the survey. Responses to pre-conference surveys were accepted from the time registration opened until the end of the first plenary event. When respondents exited the lecture hall they were asked to return the completed pre-conference survey, or were reminded to do so in the next five minutes. In total there was roughly 1.5 hours for participants to complete and submit their pre-conference survey.

The post-conference evaluation also consisted of paper-based surveys. Surveying began prior to the final plenary session. Participants were given the post-conference survey before they entered the lecture hall for the closing plenary. Survey responses were collected after the plenary talk, when participants exited the hall. Participants were also reminded during the plenary talk to fill out and return the survey when exiting the hall had they not done so. The survey officially closed after all participants left the event venue.

The follow-up event was administered via an online survey system, six months after the event. All respondents who consented to be contacted were emailed the survey link, and were given two weeks to complete the follow-up survey.

Data Matching

Prior to performing statistical analysis, we matched surveys based on anonymized participant-generated identifiers on every returned survey using Levenshtein distance – a probability based matching scheme recommended in Schnell *et al.* (2010). We use the R package "RecordLinkage" to perform matching by Levenshtein distance. The author of this package recommended using the Levenshtein similarity function (Borg, 2013) which produced a value between 0 and 1 as an indication of degree of similarity. As an example of the matching process, we compared a self-generated code from the post-event surveys against all self-generated identifiers from the pre-event surveys using the Levenshtein similarity function; the pre-event identifier that had the highest similarity value was taken as the match. To avoid false matches, we required a minimum similarity of 0.6 before a match was declared. The minimum similarity requirement was strict enough that a manual inspection of declared matches showed no sign of false matches.

Hypothesis testing

We tested a total of four hypotheses, that, on average, attendees of CC2013 had (1) a short-term improvement in self-efficacy, (2) a short-term improvement in awareness-BGD, (3) a medium-term improvement in self-efficacy, and, (4) a medium-term improvement in awareness-BGD. To show short-term change we studied paired responses from the pre- and the post-surveys. To show medium-term change we studied the pre- and follow-up surveys. The self-efficacy and awareness-BGD components were examined separately, each on the basis of *a total score* of survey items (Boone and Boone, 2012).

To test for a change, we used a two-sided, paired-t-test. The t-test allowed us to conclude if on average, a person's score at a later time differed from a person's score at an earlier time. We used a finite population correction (FPC) corrected t-test whenever we matched a large proportion of the total number of participants, to account for the representativeness of the result captured, similar to Curtis and Keeves (2000).

We set our overall significance level to 5%. This corresponded to a chance of 1/20 of erroneously concluding statistical significance. Since we identified four primary effects of interest (short-term self-efficacy, medium-term self-efficacy, short-term awareness, medium-term awareness), we required the p-value for *each* test to be less than 0.0125 for statistical significance, according to the Bonferroni correction as advocated by Bland and Altman (1995).

Results

A total of 316 people registered for the Creating Connections Conference 2013. However, the actual participation at the conference was projected to be between 200-250 people. The demographics of registered participants can be found in Table 2.

Table 2. Participant Demographics at CC2013

by Sector	n	% total (exclude n/a)	Note
Engineering	124	43%	
Science	106	37%	
Education	31	11%	
Technology	13	4%	
Other	15	5%	1 not-for-profit, 1 consulting, 1 media, 1 commerce, 1 family services, 1 political science, 2 psychology, 7 presenters
n/a	27		
total	**316**		

by Role	n	% total (exclude n/a)	Note
Graduate Student	66	23%	
Undergraduate students	52	18%	
Junior industry and Academe	46	16%	
Senior Industry and Academe	38	13%	
Industry (unknown level)	14	5%	
Junior Management	14	5%	
Currently seeking Employment	13	4%	
Executive	12	4%	
Senior Management	10	3%	
Teacher/Counsellor/Advisor	8	3%	
Highschool student	4	1%	
HR and Admin	4	1%	
Other	12	4%	1 not-for-profit, 1 citizen journalist, 3 conference staff, 7 presenters
n/a	23		
total	**316**		

The result from four hypothesis tests is summarized in

Table 3. The columns lists the two psychometric measures being evaluated, and row indexes either short term or medium term change. Detailed item by item results are found in Appendix A.

Table 3. summary of changes in self-efficacy and awareness-BGD

	Self-efficacy	**Awareness-BGD**
Short term (post – pre)	Paired-t-test with FPC (n = 135, $\bar{\Delta} = 1.27(2.28)$) **p-value << 0.01**	Paired-t-test with FPC (n = 113, $\bar{\Delta} = 0.97(3.84)$ **p-value < 0.01**
Medium term (follow-up – pre)	Paired-t-test (n = 33, $\bar{\Delta} = 0.42(2.55)$ **p-value= 0.34**	Paired-t-test (n = 32, $\bar{\Delta} = 1.46(3.41)$ **p-value= 0.02**

Discussion

The results of our study show that, overall people who attend CC2013 experienced positive changes in self-efficacy and awareness-BGD in the short term that are statistically significant. Recall that self-efficacy is correlated with career persistence, and awareness-BGD addresses an important barrier for women in STEM in general; our result indirectly suggest that CC2013 had a positive impact on women's participation and persistence in STEM.

After six months, however, there was some evidence of improved awareness-BGD, but no evidence of lasting improvement in self-efficacy. This is likely due to competing confounding factors that we cannot control, such as obstacles and challenges in participants' daily lives that diminished the effects of increased self-efficacy. This observation has an important implication in the design of interventions for women in STEM. Much of the existing research on the effects of STEM interventions tests the hypothesis that interventions improve self-efficacy for women in STEM. These studies focus on the effects of one-shot intervention on short term self-efficacy, e.g. immediately post intervention (Betz and Schifano, 2000; Dawes *et al.*, 2000), or at four weeks' time (Luzzo *et al.*, 1999). However, our study is one of the first to explore the persistence of increased self-efficacy six months post intervention. The lack of persistence after six months may indicate that the one-shot intervention format, on its own, is not well suited to increasing self-efficacy, and that other intervention targets are more suitable for long-term outcomes of these type of events.

Of course, it is not possible to show a causal relationship between the attendance at a STEM conference and the measured changes. We examine a number of potential explanations for the observed changes, also known as confounding factors. First, we believe that during the conference the participants did not experience *external* influences, e.g. a competing intervention, given the short time-frame of eight hours that separated the pre and post surveys. However, a phenomenon known as "response-shift bias" - the change in the participants' metric for answering questions from the pre-test to the post-test due to a new understanding of a concept being taught (Klatt and Taylor-Powell, 2005) - cannot be ruled out completely. The use of a

retrospective pre-test design could prevent response-shift bias (Howard and Dailey, 1979), but incurs further problems such as: recall bias, social desirability, effort justification, and cognitive dissonance (Colosi, 2006). Had we chosen a retrospective-design, the problem of social desirability and effort justification could not be ignored. Some research reports that response-shift bias in pre-post design led to more conservative p-values and estimated effects, compared to a retrospective design (Rohs and Langone, 1997; Rohs *et al.*, 2001). In our study context, we believe conservatism on concluding statistical significance to be a benign error, and thus the pre-post design remained the better option.

We believe the chance of volunteerism bias occurring in the measured short-term changes is also low. First, our post-test sample was highly representative of the whole participating population (~50%) given its voluntary nature. Secondly, our staff deduced that the reduced participation in the post-event survey was likely due to other personal commitments at the end of the day, as the conference was held on a Saturday. This factor is not related to self-efficacy, nor awareness-BGD, thus we do not have an indication of volunteerism bias. While a larger drop-out in participation was observed for the follow-up survey, 33 respondents completed a follow-up survey out of 80 who consented to be contacted – an excellent participation rate. Furthermore, we found that whether a person responded to the follow-up survey or not did not correlate with the self-efficacy score (p-value = 0.24), nor with the awareness-GBD score (p-value = 0.73) at the closing of the intervention (see Appendix B).

As the follow-up survey was administered over the internet, we had no means of guaranteeing that participants had intentionally learned the correct answers to the awareness-BGD knowledge testing questions for the purpose of survey completion. However, we note that the majority of questions that were unrelated to awareness-BGD knowledge also showed positive changes in the follow-up surveys (see Appendix A).

We note that the evaluation framework itself potentially offered value-added for our conference aside from enabling better reporting to stakeholders. The use of the awareness-BGD instrument, which in part measures a person's intention and belief toward advancing gender diversity in the technical workplace, potentially creates a mere-measurement effect on participants' future behavior. The mere-measurement effect is a change in the future behavior of a person who is asked to self-report on either attitude, intention or belief towards a subject (Sprott *et al.*, 2006; Chapman, 2001; Godin *et al.*, 2008). In the context of the awareness-BGD survey, the effect of mere-measurement may result in a changed behavior relating to the attendance of events related to gender diversity, participation in advocacy, and public support for actions increasing gender diversity. We are unable to confirm nor reject this conjecture as it is outside of the scope of this study, but recognize the potential for future work.

Lastly, positive results achieved at CC2013 serve as evidence in support of using STEM conference as a tool for increasing gender diversity in STEM. Furthermore, as a tool for creating change, we believe the evaluation framework detailed in this paper could serve well as a

measurement standard for future STEM conference since it utilizes a collection of standardized procedures, instruments, tests for establishing impact. An adoption of this framework in future conferences will enable a more consistent comparison on the effects, necessary for identifying best practices.

Conclusion

In this paper, we present the first documented evaluation framework for a STEM conference type intervention and demonstrate its impact on women's participation and persistence in STEM. We did this through targeting and measuring two proxy measures, career self-efficacy and the awareness of the benefit of gender diversity (awareness-BDG) in the workplace, which are correlated to women's participation and persistence based on theory and empirical evidence. We targeted self-efficacy and awareness-BGD through embedded messaging, and measure them with a pre-post-follow-up evaluation design to capture both short term and medium term effects while minimizing the chance of bias.

We applied this evaluation framework to a conference, Creating Connections 2013, and found that it succeeded in positively impacting both self-efficacy and awareness-BGD in the short term, confirming the value of our intervention to women's participation and career persistence. Furthermore, despite being a short-term event, our conference produced an increase in awareness-BGD that persisted at least six months after the conference ended. This suggests that short-term interventions may create sustained impact after its conclusion for certain constructs.

This study is a first investigation on the value of conference to women's in STEM based on two measures only, which is by no means an exhaustive exploration. Future work lies in identifying and evaluation other constructs related to women's participation and persistence in STEM. In particular we hope to compare the sensitivity to change and duration of impact sustained by the various constructs, to identify the most effective set of programming goals for a conference type interventions. We also hope extend the comparison to other interventions to select the most effective programming target for other types of intervention activities.

Furthermore, program design can also largely influence the sensitivity to change and duration of impact sustained by different constructs being evaluated. In our conference we addressed the selected conference objectives with an embodied learning approach and without any comparison cases. We hope to test out other approaches to program designs in the future to understand of how program design affects the duration of conference impact. Only with an understanding of both "what to target in an intervention" and "how to target" can we ensure that resources are spent efficiently for the best sustained effect to increase women's participation and persistence in the field of science, technology, engineering, and mathematics.

Appendix A

Table 4. Summary of survey results on self-efficacy

	Short Term Change		Medium Term Change	
Career Self-efficacy items	**M**	**SD**	**M**	**SD**
I can remain calm when facing difficulties in my job because I can rely on my abilities.	0.26	0.60	0.21	0.93
When I am confronted with a problem in my job, I can usually find several solutions.	0.10	0.61	0.06	0.83
Whatever comes my way in my job, I can usually handle it.	0.14	0.65	-0.21	0.72
My past experiences in my job have prepared me well for my occupational future.	0.18	0.91	0.12	0.82
I meet the goals that I set for myself in my job.	0.22	0.65	0.21	0.65
I feel prepared for most of the demands in my job.	0.36	0.83	0.03	0.98
TOTAL	1.27	2.28	0.42	2.55

Table 5. Summary of survey results on Awareness-BGD

		Short term *change*		Medium term *change*	
Awareness-BGD items	**category**	**M**	**SD**	**M**	**SD**
I am likely to attend gender diversity workshops in the future.	a	0.08	1.13	0.18	1.16
Companies should not actively promote gender diversity in the workplace.	e	0.20	1.51	0.36	1.17

Companies should spend more resources toward creating a gender diverse workplace.	e	0.04	1.25	0.03	0.68
Within my current knowledge, I know where to find information on how I can help advance gender diversity.	a	0.52	1.31	0.54	1.2
Gender diversity in the technical workplace benefits society on a:					
personal level	e	0.13	0.93	0.03	0.88
interpersonal level (e.g. when interacting with people around us)	e	0.04	0.88	-0.15	0.91
Corporate level	e	-0.03	0.90	-0.03	0.92
National level	e	-0.03	0.95	-0.03	0.81
Are the following items a result of increased gender diversity in the workforce for technical industries?					
Access to foreign markets	k	-0.23	0.81	0.00	0.62
Access to a broader talent base	k	0.08	0.30	0.06	0.24
Increase in innovation potential	k	0.04	0.27	0.03	0.17
Increase in cash reserves	k	-0.03	0.77	-0.09	0.46
Increase in cost of staffing	k	0.17	1.03	0.06	0.70
Solution to skill shortages	k	0.05	0.69	0.09	0.52
Enhanced market development	k	0.11	0.60	0.24	0.44
Stronger financial performance	k	0.07	0.81	0.39	0.61
Greater return on human resource investment	k	-0.03	0.72	0.00	0.56
Weighted* TOTAL		0.97	3.84	1.46	3.41

* Weighted total is calculated by rescaling each of the 3 subcomponents (action (a), empathy (e) and knowledge (k)) such that each contribute 10 marks to the total score.

Appendix B

In Appendix B we provide a summary of the observed correlation between responding to the follow-up survey and self-efficacy score, and for the correlation between responding to the follow-up survey and awareness-BGD score. Sample correlation coefficients were calculated, along with 95% confidence interval on the true correlation coefficients and the p-value from the test for Pearson's correlation coefficient on the hypothesis that the true correlation is 0. This result is summarised in Table 6.

Table 6. Correlation between a respondent's participation in the follow-up survey and his/her psychometric measurements at post-event

	Self-efficacy at post-event	**Awareness-BGD at post-event**

Status as participating in the follow-up survey (yes/no)	n= 152;
	r = -0.096;
	95%CI [-0.25, 0.06];
	p-value* = 0.24
		n=147;
		r= -0.34;
		95%CI [-0.19, 0.13];
		p-value* = 0.73

*Test of Pearson's product moment correlation coefficient

References

Adler, R. D. (1999), "Women in the executive suite correlate to high profits. European Project on Equal Pay", available at: http://www.lonelotsarna.se/equalpay/docs/en/adler_web.pdf (accessed 31 July 2015).

Adler, R. (2009, February), "Profit, thy name is… woman? *Pacific Standard*", available at: http://www.psmag.com/navigation/business-economics/profit-thy-name-is-woman-3920/ (accessed 31 July 2015).

American Association of University Women (2010), "Why so few: Women in science technology engineering and mathematics", available at: http://www.aauw.org/resource/why-so-few-women-in-science-technology-engineering-and-mathematics/ (accessed 31 July 2015).

Anita Borg Institute (n.d.), "*Grace Hopper Celebration | About*", available at: http://gracehopper.org/about/ (accessed 02 August 2015).

Author citation (forthcoming), "Short instrument on the public awareness of the benefits of gender diversity for assessing the level of valuation of gender diversity".

Ballout, H.I. (2009), "Career commitment and career success: moderating role of self-efficacy", *Career Development International*, Vol. 14 No. 7, pp. 655-670.

Bandura, A. (1977), "Self-efficacy: toward a unifying theory of behavioral change", *Psychological Review*, Vol. 84 No. 2, pp. 191-215.

Betz, N.E. and Schifano, R.S. (2000), "Evaluation of an intervention to increase realistic Self-Efficacy and Interests in College Women", *Journal of Vocational Behavior*, Vol. 56 No. 1, pp. 35-52.

Bland, J.M., and Altman, D.G. (1995), "Multiple significance tests: The Bonferroni method", *British Medical Journal*, Vol. 310 No. 6973, pp. 170.

Boone, H.N., and Boone, D.A. (2012), "Analyzing likert data", *Journal of Extension*, Vol. 50 No. 2, pp. 1-5.

Borg, A. and Sariyar, M. (2013, August 29), "*Reference Manual: Package 'RecordLinkage'*", available at: http://cran.r-project.org/web/packages/RecordLinkage/RecordLinkage.pdf (accessed 02 August 2015).

Boulanta, I. (2013), "Hidden connections: the link between board gender diversity and corporate social performance", *Journal of Business Ethics*, Vol. 113 No. 2, pp. 185-197.

Briziarelli, G., and Tice, T.L. (1996), "Anxiety-free conference design", *Training and Development*, Vol. 50 No. 6, pp. 47-50.

Brown, D.A.A., Brown, D.L., and Anastasopoulos, V. (2002), "Women on boards: Not just the right thing ... not the "bright" thing", The Conference Board of Canada, available at: http://www.conferenceboard.ca/e-library/abstract.aspx?did=374 (accessed 02 August 2015).

Caffarella, R.S. (2002), *Planning programs for adult learners: A practical guide for educators, trainers, and staff developers*, Jossey-Bass, San Francisco, California.

Canadian Coalition of Women in Engineering, Science, Trades and Technology (2014), "*Welcome to CCWESTT 2014*", available at: http://www.ccwestt2014.ca/ (accessed 13 May 2014).

Catalyst (2004), "The bottom line: Connecting corporate performance and gender diversity", available at: http://www.catalyst.org/knowledge/bottom-line-connecting-corporate-performance-and-gender-diversity (accessed 02 August 2015).

Catalyst (2007), "The bottom line: Corporate performance and women's representation on boards", available at: http://www.catalyst.org/knowledge/bottom-line-corporate-performance-and-womens-representation-boards (accessed 02 August 2015).

Catalyst (2011), "The bottom line: Corporate performance and women's representation on boards (2004–2008)", available at: http://www.catalyst.org/knowledge/bottom-line-corporate-performance-and-womens-representation-boards-20042008. (accessed 02 August 2015).

Chapman, K.J. (2001), "Measuring intent: There's nothing "mere" about mere measurement effects", *Psychology & Marketing*, Vol. 18 No. 8, pp. 811-841.

Cheryan, C. (2012), "Understanding the paradox in math-related fields: Why do some gender gaps remain while others do not?", *Sex Roles*, Vol.66, pp. 184-90.

Colosi, L., and Dunifon, R. (2006), "What's the difference: "Post then Pre" & "Pre then Post." Prepared for Cornell Cooperative Extension", available at: http://www.citra.org/Assets/documents/evaluation%20design.pdf (accessed 05 May 2014).

Concannon, J.P. and Barrow, L. (2012), "A reanalysis of engineering majors' self-efficacy beliefs", *Journal of Science Education and Technology*, Vol. 21 No. 6, pp. 742-753.

Council of Canadian Academies (2012), "Strengthening Canada's research capacity: The gender dimension", available at: http://www.scienceadvice.ca/en/assessments/completed/women-researchers.aspx (accessed 02 August 2015)

Curtis, D.D., and Keeves, J.P. (2000), "The course experience questionnaire as institutional performance indicator", *International Education Journal*, Vol. 1 No. 2, pp. 73-82.

Dawes, M.E., Horan, J.J., and Hackett, G. (2000), „Experimental evaluation of self-efficacy treatment on technical/scientific career outcomes", *British Journal of Guidance and Counselling*, Vol. 28 No. 1, pp. 87-99.

Diaz-Garcia, C., Gonzalez-Moreno, A., and Saez-Martinez, F.J. (2013), "Gender diversity within R&D teams: Its impact on radicalness of innovation", *Innovation: Management, Policy, & Practice*, Vol. 15 No. 2, pp. 149-160.

Diekman, A.B., Brown, E.R., Johnston, A.M. and Clark, E.K. (2010), "Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers", *Psychological Science*, Vol. 21 No. 8, pp. 1051-1057.

Dryburgh, H. (1999), "Work hard, play hard: women and professionalization in engineering - adapting to the culture", *Gender & Society*, Vol. 13 No. 5, pp. 664-82.

Engineers Canada (2014), "2013 membership survey", available at: http://www.engineerscanada.ca/national-membership-report/ (accessed 02 August 2015).

Engineers Canada (2015), "Engineering Labour Market in Canada: Projections to 2025", available at: https://www.engineerscanada.ca/labour-market-report/ (accessed 02 August 2015).

George-Jackson, C.E., and Rincon, B. (2012), "Increasing Sustainability of STEM Intervention Programs through Evaluation", in Furst-Bowe, J., Padro, F. and Veenstra, C. (Eds.), *Advancing the STEM Agenda: Quality Improvement Supports STEM*, ASQ Quality Press, Milwaukee, pp. 249-266.

Godin, G., Sheeran, P., Conner, M., and Germain, M. (2008), "Asking questions changes behavior: mere measurement effects on frequency of blood donation", *Health Psychology*, Vol. 27 No. 2, pp. 179.

Grosvold, J. (2011), "Where are all the women? Institutional context and the prevalence of women on the corporate board of directors", *Business & Society*, Vol. 50 No. 3, pp. 531-555.

Guskey, T.R, (2000), *Evaluating professional development*, Corwin Press, Thousand Oaks, CA.

Hackett, G., and Betz, N.E. (1981), "A self-efficacy approach to the career development of women" *Journal of Vocational Behaviour*, Vol. 18 No. 3, pp. 326-339.

Haley, K.J., Wiessner, C.A., Robinson, E.E. (2009), "Encountering new information and perspectives: Constructing knowledge in conference contexts", *The Journal of Continuing Higher Education*, Vol. 57 No. 2, pp. 72-82.

Hewlett, S.A. (2007), *Off-Ramps and on-ramps*, Harvard Business School Press, Boston, MA.

Howard, G.S., and Dailey, P.R. (1979), "Response-shift bias: A source of contamination of self-report measures", *Journal of Applied Psychology*, Vol. 64 No. 2, pp. 144-150.

Hoyt, J.E., and Whyte, C. (2011), "Increasing the quality and value of conferences, seminars, and workshops", *The Journal of Continuing Higher Education,* Vol. 59 No. 2, pp. 97-103.

Jackson, D. (2013), "Making the Connection: the Impact of Support Systems on Female Transfer Students in Science, Technology, Engineering, and Mathematics (STEM)", *Community College Enterprise*, Vol. 19 No. 1, pp. 19-33.

Kirkpatrick, D.L. (1998), *Evaluation training programs - the four levels*, Berrett-Koehler, San Francisco, CA.

Klatt, J., and Taylor-Powell, E. (2005), "Synthesis of literature relative to the retrospective pretest design", Paper presented at American Evaluation Association, at Toronto, ON.

Kloos, B., Hill, J., Thomas, E., Wandersman, A., and Elias, M. (2011), *Community psychology: Linking individuals and communities*, Cengage Learning, Independence, KY.

Lawrence, D.A., and Mancuso, T.A. (2012), "Promoting girls' awareness and interest in engineering", *Technology & Engineering Teacher*, Vol. 72 No. 1, pp. 11-16.

Lee, L., Falkner, W., and Alemany, C. (2010), "Turning good policies into good practice: Why is it so difficult?", *International Journal of Gender, Science and Technology*, Vol. 2 No. 1, pp. 90-99.

Louw, I., and Zuber-Skerritt, O. (2011), "The learning conference: Knowledge creation through participation and publication", *The Learning Organization*, Vol. 18 No. 4, pp. 288-300.

Luzzo, D.A., Hasper, P., Albert, K.A., Bibby, M.A., and Martinelli Jr, E.A. (1999), "Effects of self-efficacy-enhancing interventions on the math/science self-efficacy and career interests, goals, and actions of career undecided college students", *Journal of Counseling Psychology*, Vol. 46 No. 2, pp. 233.

Mair, J., and Thompson, K. (2009), "The UK association conference attendance decision-making process", *Tourism Management*, Vol. 30 No. 3, pp. 400-409.

Mateos de Cabo, R., Gimeno, R., and Nieto, M.J. (2012), "Gender diversity on European banks' board of directors", *Journal of Business Ethics*, Vol. 109 No. 2, pp.142-162.

National Science Foundation (2013), "Employed scientists and engineers, by occupation, highest degree level, and sex: 2010 (updated)" [Data file], available at: http://www.nsf.gov/statistics/wmpd/2013/tables/tab9-5_updated_2013_11.xls (accessed 02 August 2015).

Nishii, L.H. (2013), "The benefits of climate for inclusion for gender-diverse groups", *Academy of Management Journal*, Vol. 56 No. 6, pp. 1754-1774.

Orser, B. (2000), "Creating High - Performance Organizations: Leveraging Women's Leadership," The Conference Board of Canada, available at: http://www.conferenceboard.ca/e-library/abstract.aspx?did=128 (accessed 02 August 2015).

Orser, B., Riding, A., and Stanley, J. (2012). Perceived career challenges and response strategies of women in the advanced technology sector. *Entrepreneurship & Regional Development*, *24*(1/2), 73-93.

Portia Ltd (2013), "*Home - Gender Summit*", available at: http://www.gendersummit2013-na.com/ (accessed 02 August 2015).

Prochaska, J.O., DiClemente, C.C., and Norcross, J.C. (1992), "In search of the structure of change", in Klar, Y., Fisher, J.D., Chinsky, J.M. and Nadler, A. (Eds.), *Self Change*, Springer New York, pp. 87-114.

Ravn, I. (2007), "The learning conference", *Journal of European Industrial Training*, Vol. 31 No. 3, pp. 212-222.

Ravn, I., and Elsborg, S. (2011), "Facilitating learning at conferences", *International Journal of Learning and Change*, Vol. 5, pp. 84-99.

Rigotti, T., Schyns, B., and Mohr, G. (2008), "A short version of the occupational self-efficacy scale: Structural and construct validity across five countries", *Journal of Career Assessment*, Vol. 16 No. 2, pp. 238-255.

Rohs, F.R., and Langone, C.A. (1997), "Increased accuracy in measuring leadership impacts", *Journal of Leadership & Organizational Studies*, Vol. 4 No. 1, pp. 150-158.

Rohs, F.R., Langone, C.A., and Coleman, R.K. (2001), "Response shift bias: a problem in evaluating nutrition training using self-report measures", *Journal of Nutrition Education*, Vol. 33 No. 3, pp. 165-170.

Rutz, E., and Shafer, M. (2011), "Impact of an engineering case study in a high school pre-engineering course", *Journal Of STEM Education: Innovations & Research*, Vol. 12, pp. 26-34.

Schnell, R., Bachteler, T., and Reiher, J. (2010), "Improving the use of self-generated identification codes", *Evaluation Review*, Vol. 34 No. 5, pp. 391-418.

Schull, P.J. and Weiner, M. (2002), "Thinking inside the box: self-efficacy of women in engineering", *International Journal of Engaging Education*, Vol. 18 No. 4, pp. 438-446.

Servon, L. and Visser, M.A. (2010), "Progress hindered: the retention and advancement of women in science, engineering and technology careers", *Human Resource Management Journal*, Vol. 21 No. 3, pp. 272-284.

Smith, S.M., and Petty, R.E. (1996), "Message framing and persuasion: A message processing analysis", *Personality and Social Psychology Bulletin*, Vol. 22, pp. 257-268.

Sprott, D.E., Spangenberg, E.R., Block, L.G., Fitzsimons, G.J., Morwitz, V.G., and Williams, P. (2006), "The question–behavior effect: What we know and where we go from here", *Social Influence*, Vol. 1 No. 2, pp. 128-137.

Statistics Canada (2006), "*Occupation - National Occupational Classification for Statistics 2006 (720), Class of Worker (6) and Sex (3) for the Labour Force 15 Years and Over of Canada, Provinces, Territories, Census Metropolitan Areas and Census Agglomerations, 2006 Census - 20% Sample Data"*, (Catalogue number 97-559-XCB2006011), available at: http://www12.statcan.gc.ca/census-recensement/2006/dp-pd/tbt/Rp-ENG.cfm?TABID=1&LANG=E&A=R&APATH=3&DETAIL=1&DIM=0&FL=A&FREE=0&GC=01&GID=837928&GK=1&GRP=1&O=D&PID=92104&PRID=0&PTYPE=

88971,97154&S=0&SHOWALL=0&SUB=0&Temporal=2006&THEME=74&VID=0&VNAMEE=&VNAME (accessed 13 February 2014).

Torchia, M., Calabrò, A., and Huse, M. (2011), "Women directors on corporate boards: From tokenism to critical mass" *Journal of Business Ethics*, Vol. 102 No. 2, pp. 299-317.

Vella, J., Bernardinelli, P., and Burrow, J. (1998), *How do they know they know? Evaluating adult learning*, Jossey-Bass, San Francisco, CA.

VHTO (n.d.), "*Network Gender & Stem: Conference 2014*", available at: http://www.genderandstem.com/conference-2016/conference-2014.html (accessed 02 August 2015).

Weavers, L., Bautista, D., Williams, M., Moses, M., Marron, C., and La Rue, G. (2011), "Assessing an engineering day camp for middle-school girls", *Journal of Professional Issues in Engineering Education and Practice*, Vol. 137 No. 3, pp. 127–134.

Woolley, A., and Malone, T., (2011), "What makes a team smarter? More women", *Harvard Business Review*, Vol. 89 No. 6, pp. 32-33. Retrieved from http://hbr.org/2011/06/defend-your-research-what-makes-a-team-smarter-more-women/ar/1

Yurek, L.A., Vasey, J., and Sullivan Havens, D. (2008), "The use of self-generated identification codes in longitudinal research", *Evaluation Review*, Vol. 32 No. 5, pp. 435-452.