Journal of MultiDisciplinary Evaluation
Volume 15, Issue 33, 2019

JMDE
Journal of MultiDisciplinary Evaluation

ISSN 1556-8180
http://www.jmde.com

Using Program Theory to Evaluate a Graduate Student Development Program

Shannon Guss
University of Denver

Chelsea E. Bullard
Oklahoma State University

Ashley Keener
Oklahoma State University Center for Health Sciences

Sarah Gordon
Arkansas Tech University

Background: The "3 Minute Presentation" is a graduate student competition based off the more popular "3 Minute Thesis" competition. The program aims to help graduate students learn to inform others of their research in a quick and accessible manner. Programs to engage graduate students more deeply in their education require evaluation to determine if they are useful and effective at meeting their intended goals. Evaluation literature in graduate educational programs is currently limited, but increasingly needed for both the field and the students served.

Purpose: Development and testing of a program-theory evaluation to understand participation, recruitment, preparation, training, skills, and confidence of graduate students engaging in a "3 Minute Presentation" competition at a state university.

Setting: Institution of higher education.

Intervention: 3 Minute Presentation competition.

Research Design: Mixed-method program-theory evaluation.

Data Collection and Analysis: Direct observations and closed-ended survey analyzed through qualitative coding, descriptive statistics, group comparisons, and correlation analysis.

Findings: Overall, the program evaluation found, with a possible lack of diversity in participants, that the program components of recruitment, preparation, and skill development work as expected. Additionally, engagement in preparation was associated with competition scores and the perceived helpfulness of preparation was related to students' confidence in their presentation skills. This evaluation was deemed useful for program improvement and capacity building in the program's continuation at the university.

Keywords: *graduate student engagement; self-determination theory; 3 Minute Thesis; program-theory evaluation.*

Introduction

Efforts to engage college students more deeply in their education generates programs focused on student needs. These programs can benefit from program evaluation through testing of assumptions, clarification of design, and questioning of practices. Assessment and evaluation efforts for programs on college campuses are important and timely, especially as accountability and improvement drive decision-making for funding, policy, research, staffing, and student learning opportunities. In this study, a program theory-driven approach was applied to gain clarity about the stakeholder perceptions of the operations of a program and to tailor an evaluation to the unique aspects of the specific university context.

The program evaluated in this study was the Three Minute Presentation (3MP), which is a program run by the Graduate College at our institution (a large doctoral-granting land-grant university in the Midwest). The 3MP is based on the 3-Minuite Thesis (3MT) competition that originated at the University of Queensland in 2008. The 3MT competition "cultivates students' academic, presentation, and research communication skills" by increasing student capacity to effectively and briefly explain their research "in a language appropriate to a non-specialist audience" (University of Queensland, 2017, para. 1). We approached the evaluation of the 3MP, an adaptation of the 3MT, with program theory-driven evaluation.

The goals of this paper are to 1) provide an example of a theory-based evaluation in a student-focused higher education program, using our Graduate College's Three Minute Presentation (3MP) program as a case study and 2) disseminate results of an evaluation of a graduate student program to help build knowledge about the programs' efficacy. While the study sample size is small, this outlined method proved useful for elucidating the program theory to stakeholders and evaluation of the program in the early years of its development. This method of evaluation could prove useful in future evaluations of the 3MP or its more widely utilized counterpart, the 3MT competition. Following a description of the program, we discuss the program evaluation plan, provide visual diagrams of the program theory, discuss evaluation results, consider lessons learned, and provide recommendations for future efforts in higher education that may use this program evaluation approach.

Literature Review

Program Description

The Three Minute Presentation (3MP) is based on the 3-Minuite Thesis (3MT) competition originating at the University of Queensland in 2008. The 3MT allows graduate-level students to present their completed or ongoing research projects in three minutes with a single static PowerPoint slide. The 3MT competition "cultivates students' academic, presentation, and research communication skills" by increasing student capacity to effectively and briefly explain their research "in a language appropriate to a non-specialist audience" (University of Queensland, 2017, para. 1). Benefits of participating in the 3MT program include publicizing the research and the researcher, prize money, increased potential employment opportunities, media attention, and opportunities to write about research for the media (3MT, 2015).

The 3MP was adapted from the 3MT model (changing the 'T'—thesis— to 'P'—presentation) to be more inclusive of students who did not pursue thesis options for their graduate degree but still engaged in research and/or creative components; it also allowed for doctoral students who wrote dissertations to participate. Like the 3MT program, the 3MP program is also structured as a competition amongst graduate students. Students must create a three-minute presentation to discuss their research with a single static slide as their only visual. Those eligible to participate in the 3MP include currently enrolled students seeking a non-thesis master's degree, specialist, or graduate certificate program. Prior institutional winners of the 3MP and the related 3MT competition are not eligible to participate. The 3MP program is coordinated by an associate dean of the Graduate College. The program is advertised to students, graduate coordinators, and advisors via email.

Students in certain degree programs are required to participate in the 3MP program (as determined by their academic program faculty), while students from other degree programs may elect to participate. Additionally, students may elect to attend an optional training/preparation meeting offered by the coordinator of the 3MP.

3MP presentations are judged by non-faculty community members and staff of the university. The presentations are scored based on the ability for a lay audience to comprehend the topic covered and the ability of the speaker to create a compelling and engaging presentation. Various preliminary rounds determine who will compete at the additional levels, with cash prizes at each of the three levels: 1. Preliminary, 2. University Finals, and 3. President's Fellows. The 3MP is a new program at the University, beginning in 2015. The program stakeholders requested for the new program to be evaluated for effectiveness and potential improvements.

Evaluation Method

While evaluation has always been a part of higher education practice at some level, it was not until the 1980s and 1990s that it came to the forefront as an essential element of practice for higher education professionals (Schuh, 2009). Theory-based evaluation, or theory-driven evaluation, can be especially useful because it necessitates a discussion about the components of a program that are critical to achieving the desired outcomes (Donaldson, 2007; Fitz-Gibbon & Morris, 1975). Program-theory evaluations seek not to know just if a program works, but how and why a program works by developing theories that identify the relationships between the problems a program aims to solve, the conditions program components/processes operate within, and how the program plans to solve those problems (Bickman 1987; Donaldson 2007). Program-theory based evaluation was chosen to investigate the 3MP program due to its ability to addresses system-level assumptions and goals articulated by stakeholders and how program elements are thought to impact outcomes.

Methods

The following steps outline the undertaken process of developing the program theory, developing and then testing grounded evaluation questions.

Program Theory Identification

The evaluators first met with the key stakeholder to better understand the program from the stakeholder's perspective (Donaldson, 2007). The stakeholder explained the program to the evaluators and the evaluators reflected back their understanding of the program to ensure consensus and clarity. Importantly, the evaluators were able to understand the stakeholder's view of the program and how it met intended goals. The implicit assumptions underlying the program (Weiss, 2000) were: 1) the program was created in part as a way to provide graduate students an opportunity to be engaged (as supported by research on student involvement/engagement); 2) the program would help them in their academic and research endeavors (as promoted by the 3MT program the 3MP is modeled after); and 3) the components of the program worked together to help students build their skills.

Specifically, the process of engaging and developing students' skills occurred through the process of choice of participation, preparation, and presentation in the competition. Some students were recruited through email, while others were in programs which mandate participation. The program offered an opportunity for students to engage in their area of study by preparing and presenting their research. While training was not mandatory for all participants, the training was expected to assist students in preparing a concise and engaging presentation. Students could also elect to prepare for the 3MP independently. Regardless of how students chose to prepare, the preparation process was seen as part of the student engagement opportunity that the competition provided. Stakeholders believed that students gained skills through preparation, both through deepening engagement in their area of study and pursuing a professional development

activity that would be useful to students as they enter their chosen careers.

Following interviews with stakeholders, ethnographic observation of training lectures, and examination of documents related to the program, evaluators then reviewed literature related to student development, involvement, and engagement. These concepts served as the foundation for the stakeholder's conception of the primary program aims. Evaluators then incorporated literature from student involvement theory and self-determination theory. These literatures informed the tools that were chosen to measure program components and the expected relationships among them.

Student Involvement Theory

One of the most widely supported and accepted theories in college student development literature is Student Involvement Theory (Renn & Reason, 2013). Involvement is defined as "the amount of physical and psychological energy that the student devotes to the academic experience" (Astin, 1984, p. 297). Also referred to as "engagement", the theory proposes that meaningful student involvement leads to cognitive complexity, learning, and development (Renn & Reason, 2013). While there is extensive research on the effect of involvement and engagement among undergraduates (Hartnett, 1965; Kuh et al., 1991; Pascarella & Terenzini, 2005; Terenzini & Pascarella, 1991) less is known about graduate student involvement. While there were approximately 1.78 million students enrolled in graduate programs in the United States as of the Fall 2015, research on graduate student involvement and its potential to support student outcomes is sparse (Okahana, Feaster, & Allum, 2016).

While some research points to similarities between undergraduate and graduate student involvement and persistence as related to time spent in clubs, organization, and other campus-related activities (Gardner & Barnes, 2007; Thomas, Clewell, & Pearson, 1992; Tinto 1993), research on graduate student involvement was found to be particularly important for intellectual development, development of skills needed for thesis/doctoral completion, (Tinto, 1993) and socialization into the professional academic

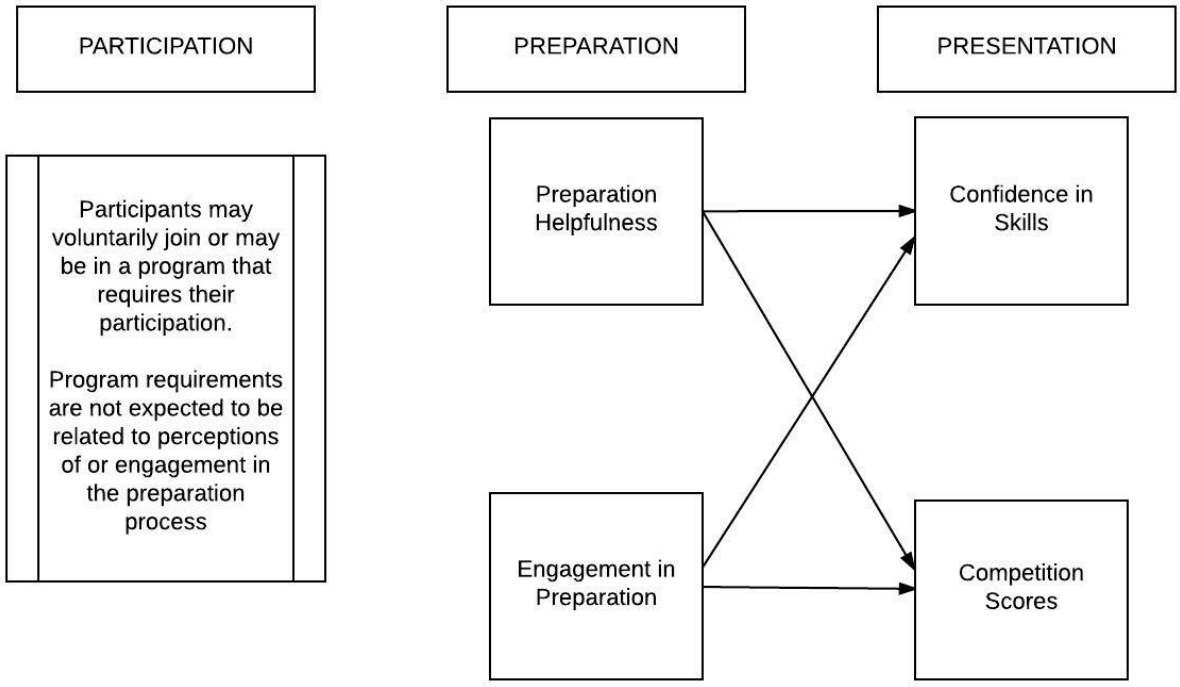

Figure 1. *Depiction of program theory.*

community (Boyle & Boice, 1998; Gardner & Barnes, 2007; Lovitts, 2001; Weidman, Twale, & Stein, 2001). Ultimately, graduate student development programs appear to receive less attention and evaluation than programs directed at development and retention of undergraduate students.

Self-Determination Theory

Self-determination theory comprises a set of theories explaining human motivation. The most applicable sub-theory to this evaluation is basic psychological needs theory (Deci & Ryan, 2001). The theory posits that when three basic psychological needs—relatedness, competence, and autonomy—are met, people become more motivated to accomplish difficult tasks. The authors applied it specifically in this study because autonomy—a feeling of control or belief that one's choices can effect intended change—was thought to potentially differ between students who chose to participate and those who were mandated to participate by their program. This study seeks to understand how program components of participation, preparation, and presentation fit together to support the development of graduate students.

Generating and Prioritizing Evaluation Questions

A major advantage of theory-driven evaluation approaches is the ability to gain deeper insight regarding how to evaluate a program by working closely with stakeholders to understand the entirety of the program (including its goals and values) and what the stakeholders want to learn from the evaluation (Donaldson, 2007; Fitzpatrick, Sanders, &Worthen, 2012). Theory-based evaluation uses "the construction of a plausible and sensible model of how a program is supposed to work" (Bickman, 1987, p. 5) to guide the evaluation. Thus, after an initial conversation with the stakeholder and a review of documents and literature, the evaluators returned to the stakeholder with a model of program components and expected processes (see Figure 1), as is typical of program theory-based evaluation (Donaldson, 2007).

As depicted in Figure 1 (and discussed previously as a part of the program theory), the components of the program to be evaluated include 1) the recruitment or requirement of students to participate in the program, 2) the training or other preparation that students undergo before competition, and 3) the competition itself. The student presentations, at any stage of the competition, show the skills of participants—the identified outcome of participating in the 3MP. Because the key stakeholder identified student engagement as a process through which students benefit from the program, the links of perceived choice, perceived helpfulness of preparing, and student engagement were also included in the program theory. Additionally, the program required some students to participate while others elected to be involved, which raised questions about students' motivation and the effect it might have on program outcomes—this is also included in the model. Finally, confidence was included as an additional check of the benefits of the program.

Once a program theory was agreed upon with the stakeholder, evaluators developed tools to assess the program components. Evaluators used both qualitative and quantitative methodological strategies to understand the program. Merging the qualitative and quantitative data resulted in a thick, rich description of the program that lent itself to use in a program-theory evaluation (Geertz 1974; Chen 1997). The 3MP program coordinator (the stakeholder) and the evaluation team came to consensus on three evaluation questions: 1) Who participates in the 3MP, and why? 2) How did students prepare to confidently communicate about their discipline? and 3) What was students' confidence and skill level after presentation?

Additional questions about program links – student perceptions of choice, perception of the training, and student engagement – were also explored to link the program components together in the evaluation. The linking questions include: 1) Do differences in program requirements relate to difference in perceived choice or student engagement? 2) What relationship, if any, exists between perceived choice, perceived helpfulness of the preparation process, and student engagement? 3) Do differences in confidence or competition scores exist between students

who did and did not attend the provided trainings? 4) What relationship, if any, exists between perceived helpfulness, student engagement, confidence levels, and competition scores?

After identifying the theory and determining evaluation questions, evaluators and stakeholders worked together to prioritize questions (Donaldson, 2007). One way to prioritize questions, as well as facilitate evaluation use, is to explore decision situations—points at which more information is needed to determine next steps. Evaluators and stakeholders can link each question to a decision situation to consider: Does the answer to this question lead me to make a better decision in the given situation? All evaluation questions, linking questions, and decision situation questions are listed in Table 1.

Table 1
Evaluation Questions

Primary Evaluation Questions	Linking Questions	Questions that Facilitate a Decision
1. Who participates in 3MP, and why?	1. Do differences in program requirements relate to differences in perceived choice or engagement?	1. Are changes needed to the recruitment process? 2. Should students be motivated to participate differently?
2. How did students prepare to confidently communicate their discipline's contribution to the public good?	2. What relationship, if any, exists among links and outcomes in the model?	3. Are changes needed to training or other preparation resources? 4. Do students need different/more engagement opportunities?
3. What was students' confidence level and competition score after presenting?	3. Do differences in confidence or competition scores exist between students who did and did not attend the provided trainings?	5. Is further evaluation needed to understand how the program could continue to support student outcomes?

Data Collection

IRB approval was obtained prior to the recruitment of participants. Data for this evaluation came from program judges and student participants. To answer all the evaluation questions, quantitative and qualitative data were collected concurrently. Quantitative survey methods were used to understand participation, perceived choice, training components, engagement, and student skills. In order to collect this data, an email survey was sent to students inviting them to participate in the evaluation. Email addresses were provided by the 3MP program coordinator. The surveys were accessible over a one-week period in the same semester in which a 3MP competition was being held. The competition judges' quantitative rubrics were also collected after the 3MP finals and were provided to the evaluation team by the 3MP program coordinator. Qualitative methods, including participant observation of training and a semi-structured interview with open-ended questions with the 3MP program coordinator were used to understand the students' training and, to a lesser extent, other preparation processes.

Instruments

Participation was assessed by the student demographics questions and the Perceived Choice Scale. The Perceived Choice Scale was adapted from the Self-Determination Scale (Sheldon, Ryan, & Reis, 1996). All other instruments were created by the evaluation team, based on program objectives described by the key stakeholder and a review of the previous competition's judges' rubric. Preparation was examined through a preparation helpfulness scale and a student engagement questionnaire. Additionally, a semi-structured trainer interview was

conducted as well as participant observation of the training seminar. To examine the presentation piece, students were asked to complete a confidence assessment, and judges provided competition scores. Table 2 provides a description of each data collection instrument/measure and the associated evaluation questions/model component.

Table 2
Evaluation Questions and Instruments

Measure	Purpose	Description of Items
Participation Questions: Which students participate? How autonomous are students' choices to participate?		
Student Demographic Questionnaire	Assessed the diversity of the student population participating in the 3MP and their reasons for participating	age, gender, ethnicity, disability status, program of study, first language, and student status
Perceived Choice Scale	Assessed the extent to which students' perceived their participation as an autonomous decision	5-item Likert-scale [range: 1 (Only A feels true) to 5 (Only B feels true)] *Example item:* *A. I participate in the 3MP because I choose to.* *B. I participate in the 3MP because I must.*
Training and Preparation Questions: What is the perceived helpfulness of the preparation process? How engaged are participants in preparation? What supports were provided for student development?		
Student Engagement in Preparation Questionnaire	Assessed the degree to which students prepared/engaged in scholastic activities during the 3MP process	List of 14 activities students may have engaged in during the 3MP process [Yes/No responses] *Example items:* *What did you do to prepare for the 3MP? (Prepared a script, studied training materials,* etc.)
Preparation Helpfulness Scale	Assessed the extent to which students perceived that the process of preparing for the 3MP helped them meet program objectives	9-item Likert scale [range: 1 (Strongly disagree) to 7 (Strongly agree)] *Ex. Preparing for the 3MP... Helped me consider the needs of a non-specialist audience.*
Semi-Structured Interview and participant observation of training	Gather information about the training to inform the context of Evaluation Question 2	Observation of in-person training seminar and face-to-face, semi-structured interview with training coordinator. Questions formatted in open-ended manner to elicit coordinator perspective on the training, its core messages, and how well the training helps to prepare students for competition objectives
Presentation Questions: How do students perceive their presentation skills after participating in the 3MP? How do others perceive students' presentation skills after participating in the 3MP?		
Competition Scores	Rated students on their 3MP performance, as noted by 3MP judges	Continuous scale from 20 to 80 based on a rubric of items. *Items within two components:*

		Comprehension/content & engagement/communication
Confidence Scale	Assessed whether students felt participation in the 3MP helped them develop skills that align with program objectives	11-item Likert scale [range: 1 = *not at all confident* to 7 = *exceedingly confident*] *Ex. How confident are you that you can… Explain how your discipline is beneficial to the public.*

Analysis

Data analysis plans were generated relative to the evaluation questions and utilized a convergent mixed-method analysis procedure (Creswell, 2014). Quantitative data analyses consisted of descriptive statistics (e.g., percentages, frequency counts, means, and standard deviations) and group comparisons (e.g., t-test, Mann-Whitney U). Correlations were used to assess consistency among responses (e.g., judges and students) and the relationship between links and the program outcome in the model. Multiple regression was not used due to sample size. Qualitative data were coded using a grounded theory methodology (Strauss & Corbin, 1994). Using the constant comparison method (Strauss & Corbin, 1990), important topics arising from the data were pinpointed and grouped together into relevant themes and categories. Data was coded and organized until saturation was met and no new topics arose.

Results

Surveys were initially emailed to 210 graduate students. Twenty-seven graduate students who participated in the 3MP during the fall 2016 semester provided data for the evaluation resulting in a 13% response rate. The 27 evaluation participants ranged in age from 21 to 47, with a mean age of 25 (*SD* = 3.06). The sample contained 48% males and 37% females, with 15% who chose not to provide their gender identification. In regards to race/ethnicity, 22% identified as White and 67% identified as Asian, and 11% chose not to provide their race/ethnicity. A total of 18 participants (66%) were international students, and 13 participants (48%) indicated that English was not their first language.

Participation Evaluation Questions: Which students participate? How Autonomous are Students' Choices to Participate?

Of the student respondents, 56% belonged to graduate programs that required students to participate in the 3MP. Despite this, only 22% said the mandate was a primary reason they participated. All respondents were asked why they participated in the program; 48% indicated it was due to an interest in skill improvement, 22% indicated it was a program mandate, 11% indicated it was for the cash reward, 4% indicated it was encouraged by their program, 4% indicated in was due to entertainment value, and 4% selected other (7% did not provide a reason for participating). There were no significant differences in responses by gender, race, or ethnicity in their reasons for participating.

Preparation Evaluation Questions: What is the Perceived Helpfulness of the Preparation Process? How Engaged are Participants in Preparation? What Supports were Provided for Student Development?

A total of 13 students indicated that they attended training, whereas 10 indicated they did not attend training. On average, students agreed that the training process helped to increase their presentation skills (M = 6.16, SD =.75 on a 7-point scale). Among the items on the Training Experience Survey, students indicated that the training was most helpful at developing their ability to "effectively manage their time" (M = 6.44) and that the training was

the least helpful concerning "improving their understanding of their topic" (M = 5.84). Another highly rated item was "pushed me to think of stories or evidence that illustrated my points" (M = 6.36).

Approximately 80 students attended the training program and only one participant from the two programs mandating attendance did not attend the training session. Additionally, there were two additional non-mandatory training sessions for non-thesis master's, specialist and graduate certificate students who volunteered to compete in the 3MP competition. Twenty-four students signed up to attend one of the two non-mandatory training sessions and less than 20 students actually attended the non-mandated training program.

Evaluation Question 3: How do Students Perceive their Presentation Skills After Participating in the 3MP? How do Others Perceive Students' Presentation Skills After Participating in the 3MP?

The dataset revealed a negatively skewed distribution in relation to confidence in presentation skills among 3MP participants. Most students reported that they were "pretty confident" to "exceedingly confident" that they had obtained the necessary skills to present their research to an audience. Means and standard deviations associated with each skill are presented in Table 3. Students felt particularly confident about communicating clearly and holding an audience's attention as evidenced by their high means (6.00 and 6.04 respectively) and small standard deviations (.96 and .90).

Table 3
Means and Standard Deviations Associated with Level of Confidence in Presentation Skill

Presentation Skill	*M*	*SD*
Give a good explanation of topic to someone you just met.	6.04	0.76
Explain your topic thoroughly without boring someone.	5.78	0.93
Communicate clearly about your discipline.	6.04	0.90
Communicate to a non-specialist audience effectively.	5.74	1.06
Avoid "jargon" when communicating your ideas.	5.78	1.12
Illustrate your points with stories or evidence.	5.96	1.02
Effectively manage you time when presenting.	5.81	1.11
Explain how your discipline is beneficial to the public.	6.07	1.04
Prepare a presentation in a concise format.	6.00	0.96
Prepare a visually effective presentation.	6.00	1.28
Hold your audience's attention with your topic.	6.00	0.76
Total	5.90	0.78

Note. N = 27; responses are based on a 7-point scale, where 1 = Not at all confident, 2 = a little bit confident, 3 = somewhat confident, 4 = moderately confident, 5 = confident, 6 = very confident, and 7 = exceedingly confident.

Linking Questions: Do Differences in Program Requirements Relate to Differences in Perceived Choice, or Engagement in Training/Preparation?

Two constructs that potentially linked program components were identified. First, the mandate of some programs to require students to participate was examined. As shown in Table 4, comparing students in programs that do mandate participation versus students in programs that do not showed no significant difference by program mandate in perceived choice, engagement in training, or hours spent preparing. There was a trending difference in engagement tasks,

with higher engagement among those not mandated to participate [$t(23) = 1.85$, $p = .08$]. Finally, the correlation between perceived choice and engagement tasks ($r = .34$, $p = .102$) was not statistically significant, but could be considered practically significant given the evaluation's small sample size.

Table 4
Comparisons Between Students Mandated to Participate and Those not Mandated to Participate in the 3MP on Perceived Choice, Engagement in Training, Hours Spent Preparing, and Engagement Tasks

	t	*df*	*p*
Perceived Choice	.94	24	.94
Engagement In Training	- .63	8	.54
Hours Spent Preparing	.28	20	.78
Engagement Tasks	1.85	23	.08

What Relationship Exists, if any, Among Links and Outcomes in the Model?

This linking question examined the connection between engagement / preparation and presentation skills. Engagement was measured in preparation tasks, behavior during training, and hours spent preparing. Skills were measured by the Confidence in Skills Questionnaire and the three rounds of judging across the semester. Correlations were analyzed between engagement and skill variables (see Table 5), and we found that the number of reported hours of engagement significantly correlated to confidence in skills, and the number of reported tasks completed in preparation (i.e,. engagement in preparation) had a significant, strong correlation to competition scores. While the other engagement scores did not correlate to the judges' scores, we believe that data might have been more informative if scores (which were provided to the evaluation team by the 3MP program coordinator) were disaggregated by the judging content areas.

Table 5
Correlations Among Key Variables

	Perceived Choice	Preparation Engagement	Engagement Hours	Preparation Helpfulness	Skills Confidence	Competition Scores
Perceived Choice	–					
Preparation Engagement	.34	–				
Engagement Hours	.31	.41	–			
Preparation Helpfulness	.18	.20	.42*	–		
Skills Confidence	.05	.22	.42	.63**	–	
Competition Score	.07	.55*	.15	.28	.29	–

Do Differences in Confidence or Competition Scores Exist Between Students Who Did and Did Not Attend the Provided Trainings?

To answer this linking question, an independent samples *t*-test was conducted to determine if those that attended training demonstrated more confidence in skills than those that did not attend training, and the results revealed there were no statistically significant differences between the groups [$t(21) = .36$, $p = .73$], suggesting that students

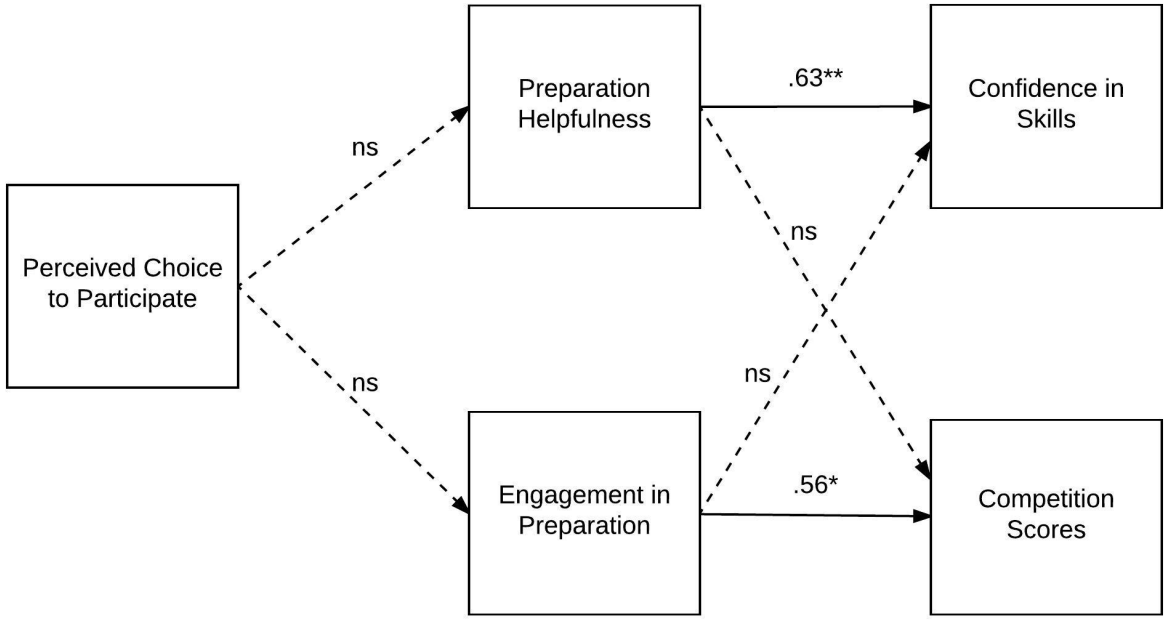

Figure 2. *Program evaluation: Correlations between the program components.*
Note. $*p < .05$; $**p < .01$

who attended the training did not obtain more confidence in their skills than those who did not attend the training.

Program Evaluation

Our final assessment of the program considers the program components first, then assesses how they work together (Donaldson, 2007), as outlined in the program theory. The program evaluation highlights that the sample of participants is not diverse, but their participation in the program is perceived to be due to their own choice (rather than required). The preparation process is seen as helpful and the training is convenient for those who chose to attend. There was a high level of engagement in the training and a largely varying level of engagement in preparation tasks as a whole. Confidence in the presentation skills were high. Overall, with a possible lack of diversity in the participants, the program components of recruitment, preparation, and skill development seem to be working as expected. The components of perceived choice and engagement are particularly important to link program components together, so will therefore be discussed as part of the program theory evaluation.

The program evaluation tested whether the decision to mandate participation in the 3MP would affect students' perceptions, experiences, or engagement in the process. Engagement in the process of preparation for the competition was also expected to facilitate the path from program support for student preparedness and the acquisition of skills. In the theory evaluation, no evidence supported the stakeholders' concern that the program mandate negatively affected students' perceptions, experiences, or engagement in the 3MP. Correlations (shown in Figure 2) between the program requirement and perception of choice and perception of helpfulness of the preparation process were both very low. However, student ratings of the helpfulness of the preparation process was significantly correlated with students' confidence in their skills. Further, engagement in the preparation process correlated with competition scores.

Strengths, Limitations, and Conclusions

The strengths and limitations of the program can be organized by its component parts—participation and recruitment, preparation and training, and skills and confidence. Strengths of participation and recruitment exist in the perceptions of students--students who participate generally felt autonomous in their decision to engage in the 3MP. Strengths of the training and preparation processes were many, with weaknesses being harder to find. Participants found the preparation processes very helpful, and their rating of how helpful the process was related to the confidence they had in their skills.

None of the aspects of training or preparation were rated low by the students and intended goals were met. Further, while most participants said the training was convenient and the preparation process was very helpful, open-ended comments collected from student surveys suggested that online training (asynchronous or streamed) or evening trainings be available or that the presence of online materials (e.g., slides and videos) were easier to access. A small number had difficulty finding the videos or were not aware of the online resources until the training. These anecdotes could indicate the need for additional opportunities to access training.

Once students decided to participate and prepared, their skills were tested in the various levels of competition. Within the competition, students' report of their engagement in preparation tasks was correlated with their scores. However, this relationship diminished after the first round. One interpretation is that the later rounds were more difficult for participants to compete in due to competing against a more diverse group – particularly more students whose first language is English and fewer students and judges who are accustomed to their interdepartmental jargon. Thus, we recommend that stakeholders consider the influence of language in the judging of the competition, potentially providing additional supports for bilingual students. An alternative explanation for this finding, however, is that the smaller sample size reduced the variability of participants, and therefore limited the range of scores from which correlations could be observed.

This brings us to note that this evaluation was limited by the scope and quantity of its sample, which could mean the evaluation captured the perspective of those most interested in the 3MP. An additional limitation is the aggregated judge's rubric data. We may have seen more correlations between engagement and skill if we could look at components of skill individually, but this was not possible with the existing data with which we worked. We address these limitations by acknowledging that, while these data represent an important portion of participants, additional participants and additional data may have provided a wider and more accurate account.

Finally, with regard to student outcomes, strengths and limitations of the skills developed for the competitions were the most difficult to evaluate, as self-reported ratings of confidence did not correlate to the judges' scores at the competitions. We believe this issue is more of a limitation of the evaluation itself than of the program. However, there is evidence that confidence was high amongst most students and that a relationship exists between how participants engaged in preparation tasks and how well participants did in the competition. These two pieces of evidence show that the program is working as expected – giving students an opportunity to practice an important professional skill. The theory evaluation, along with overall conclusions and strengths and limitations, provide stakeholders with the data and evidence they need to understand both how the program works and answer the decision-situation questions posed in the evaluation planning phase.

Lessons Learned

For future evaluations of this and similar programs, we recommend targeting a more diverse pool of participants by reaching out to a wider variety of programs and student types. Particularly in light of the program's aim to highlight scholarship's contribution to the public good, a diverse participation in the

program may facilitate the clarity and skill with which participants take their scholarship into their various communities. However, though the number of participants is relatively small, we feel it important to note that the program that was evaluated is one of the most visible programs provided by the Graduate College for graduate students, and the Graduate College used this evaluation data to: 1) determine that the program was worthwhile and will continue at our institution and 2) make changes for upcoming iterations for the program. In addition to this practical use of evaluation results, the findings of this evaluation also contribute to the knowledge base about programs for graduate students, which is an important part of the focus and use of evaluation (Owen, 2006) that should not be overlooked, no matter how seemingly small the sample. This is also important because little is known about graduate student involvement and outcomes (Okahana et al., 2016).

While the need for graduate college student development programs is a finally recognized vital component for graduate student graduation and satisfaction (Keeling, 2004), program evaluations are not always conducted on these programs to determine their success. To our knowledge, no program evaluations have focused on the 3MT or 3MP programs (at our institution or elsewhere). This evaluation provided evidence that the 3MP was a successful professional development program for graduate students. The theory-driven approach adopted for the present evaluation was beneficial given the 3MP program had not been evaluated before and the evaluators were working with a key stakeholder that was not familiar with the evaluation process. Therefore, it became pertinent throughout the evaluation to foster the key stakeholder's understanding of the evaluation process via encouraging inquiry, building rapport, and cultivating an environment where formative evaluation can be an ongoing process. This evaluation could serve as a framework for evaluation of similar and more broadly implemented 3MT competitions and graduate schools and universities.

Through the course of conducting this evaluation, we found that using a program theory evaluation approach helped the stakeholders conceptualize what they thought their program should do (student outcomes) and why they thought it should work. In situations such as the one in this study, where evaluation was not considered as a part of the program development and/or when program outcomes are not clearly stated from the start, theory-based evaluation approaches can be helpful. The central work of theory-based evaluation lies in developing a theory for why a program should achieve its desired outcomes through engagement with key stakeholders, evaluator expertise, and/or social science research to create linkages between program actions, goals, and outcomes (Chen, 1990, Weiss, 1997, p. 78). We found this approach to be absolutely essential in facilitating the evaluation process, and we recommend this approach continue to be used in similar situations in higher education.

Further, the use of the theory-based evaluation approach also helped facilitate evaluation capacity building for the department (i.e., the Graduate College) that oversaw the program. Fitzpatrick et al. (2011) described how evaluations can be used to empower by "(1) providing stakeholders with tools for assessing the planning, implementation, and self-evaluation of their program and (2) mainstreaming evaluation as part of the planning and management of the program/organization" (p. 209). This seemed to be case in this evaluation, because this process helped stakeholders think through all phases of the program, articulate its purpose, use information to make changes, and begin incorporating pathways for data collection and thus program evaluation into this and other programs they oversaw. Clearly, the perspective gained from this evaluation served as a guide for program improvement and capacity building. Capacity building has been described as "a context-dependent, intentional action system of guided processes and practices for bringing about and sustaining a state of affairs in which quality program evaluation and its appropriate uses are ordinary and ongoing practices within and/or between one or more organizations/programs/sites" (Stockdill, Baizerman, & Compton, 2002, p. 8).

In this evaluation, to encourage the key stakeholder to see the importance in

continuing ongoing evaluation, emphasis was placed on presenting the evaluation findings in an understandable and non-technical format. For example, path diagrams were presented to articulate the program-theory evaluation in a concise, easy-to-follow layout, which maximizes the utility of the evaluation for the key stakeholder. Additionally, by identifying the primary audience (i.e., a key stakeholder with limited knowledge of program evaluation), we were able to tailor the evaluation report to their information needs and, hence, increase the probability of the findings being used. In this case, we contend that both the evaluation approach (theory-based evaluation) and the transmission of information could be seen as intimately linked to capacity building, especially considering this was the first time the stakeholders had been involved in program evaluation. We recommend future evaluations on this program (and others similar to it in higher education) follow the same approach to help facilitate utility of findings and to build evaluation capacity.

Finally, the theory-based evaluation approach used in this evaluation allowed for a discussion about the components of a program that are critical to achieving the desired outcomes (Donaldson, 2007; Fitz-Gibbon & Morris, 1975). It also contributed to building evaluation capacity by cultivating an environment where evaluation was understood and valued. Thus, because this evaluation was frequently limited by its sample size, we recommend that evaluations of this program and others like it continue, and future evaluations should consider incorporating data collection as part of the normal "paperwork" to be completed by participating students. This would allow for a broader and more accurate understanding of the program's influence across participants and would allow for parts of the evaluation to be "embedded" into the program itself, thus continuing to ensure that information is collected to allow stakeholders to make evidence-based decisions regarding program improvements.

References

3MT (Producer). (2015, September 9). *Promoting yourself with 3MT* [video file]. Retrieved from https://vimeo.com/138709200

Astin, A. W. (1984). Student involvement: A developmental theory for higher education. *Journal of College Student Personnel*, *25*(4), 297-308. Retrieved from https://www.researchgate.net/profile/Alexander_Astin/publication/220017441_Student_Involvement_A_Development_Theory_for_Higher_Education/links/00b7d52d094bf5957e000000/Student-Involvement-A-Development-Theory-for-Higher-Education.pdf

Bickman, L. (1987). The functions of program theory. *New Directions for Program Evaluation, 33*, 5-18. doi: 10.1002/ev.1443

Boyle, P., & Boice, B. (1998). Systematic mentoring for new faculty teachers and graduate teaching assistants. *Innovative Higher Education*, *22*(3), 157-179. https://doi.org/10.1023/A:1025183225886

Brousselle, A., & Champagne, F. (2011). Program theory: Logic analysis. *Evaluation and Program Planning, 34*(1), 63-78. http://doi.org/10.1016/j.evalprogplan.2010.04.001

Chen, H. T. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage.

Chen, H. T. (1997). Applying mixed methods under the framework of theory-driven evaluations. *New directions for evaluation*, *1997*(74), 61-72.

Creswell, J. W. (2014). *A concise introduction to mixed methods research*. Sage Publications.

Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227-268.

Donaldson, S. I. (2007). *Program theory-driven evaluation science: Strategies and applications*. New York: Erlbaum.

Fitz-Gibbon, C.T., & Morris, L.L. (1975). Theory-based evaluation. *The Journal of Educational Evaluation, 5*(1). Retrieved from http://www.cem.org/attachments/publications/CEMWeb022%20Theory%20Based%20Education.pdf

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2012). *Program evaluation: Alternative approaches and practical*

guidelines (4th ed.). Upper Saddle River, NJ: Pearson.

Gardner, S. K., & Barnes, B. J. (2007). Graduate student involvement: Socialization for the professional role. *Journal of College Student Development*, *48*(4), 369-387. doi: 10.1353/csd.2007.0036

Hartnett, R. T. (1965). Involvement in extra-curricular activities as a factor in academic performance. *Journal of College Student Development*, *6*(5), 272-274. Retrieved from https://muse.jhu.edu/journal/238

Kuh, G. D., Schuh, J. H., Whitt, E., Andreas, R. E., Lyons, J. W., Strange, C. C., ... & MacKay, K. A. (1991). *Involving colleges*. San Francisco: Jossey-Bass.

Lovitts, B. E. (2001). *Leaving the ivory tower: The causes and consequences of departure from doctoral study*. Lanham, MD: Rowman & Littlefield.

Okahana, H., Feaster, K., & Allum, J. (2016). *Graduate enrollment and degrees: 2005 to 2015.* Washington, DC: Council of Graduate Schools. Retrieved from http://cgsnet.org/ckfinder/userfiles/files/Graduate%20Enrollment%20%20Degrees%20Fall%202015%20Final.pdf

Owen, J.M. (2006). *Program evaluation: Forms and approaches* (3rd ed.). New York: The Guilford Press.

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students* (Vol. 2). K. A. Feldman (Ed.). San Francisco, CA: Jossey-Bass.

Renn, K. A., & Reason, R. D. (2013). *College students in the United States: Characteristics, experiences, and outcomes*. San Francisco, CA: John Wiley & Sons.

Schuh, J.H. (2009). *Assessment methods for student affairs*. San Francisco, CA: Jossey-Bass.

Sheldon, K., Ryan, R., & Reis, H.T. (1996). What makes for a good day? Competence and autonomy in the day and in the person. *Personality and Social Psychology Bulletin*, 22, 1270-1279. https://doi.org/10.1177/014616729622 12007

Stockdill, S.H., Baizeman, M., & Compton, D.W. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. *New Directions for Evaluation, 93,* 7-26. doi: 10.1002/ev.39

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* Newbury Park, CA: Sage.

Strauss, A., & Corbin, J. (1994). Grounded theory methodology. In N.K. Denzin & Y.S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 217-285). Thousand Oaks, CA: Sage.

Terenzini, P. T., & Pascarella, E. T. (1991). Twenty years of research on college students: Lessons for future research. *Research in Higher Education*, *32*(1), 83-92. doi: https://doi.org/10.1007/BF00992835

Thomas, G., Clewell, B., & Pearson Jr, W. GRE Board. (1992). *The role and activities of American graduate schools in recruiting, enrolling, and retaining United States' Black and Hispanic students.* (Report No. 87-08) Princeton, NJ: Educational Testing Service, Graduate Record Examinations Board

Tinto, V. (1993). Building community. *Liberal Education*, *79*(4), 16-21. Retrieved from https://www.aacu.org/liberaleducation

University of Queensland. (2017) *Three Minute Thesis*. Retrieved from https://threeminutethesis.uq.edu.au/about

Weidman, J. C., Twale, D. J., & Stein, E. L. ERIC. (2001). *Socialization of Graduate and Professional Students in Higher Education: A Perilous Passage? ASHE-ERIC Higher Education Report*, (Volume 28, Number 3*).*, San Francisco, CA: Jossey-Bass, Publishers.Weiss, C. H. (1997). Theory-based evaluation: Past, present, and future. *New Directions for Evaluation*, *76*, 41-55. doi: 10.1002/ev.1086

Weiss, C.H. (2000). Which links in which theories shall we evaluate? *New Directions in Evaluation, 87*, 35-45. doi: 10.1002/ev.1180