An instrument for assessing the public communication of scientists

Rachel Collier Murdock
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd
Part of the Communication Commons, and the Rhetoric Commons

Recommended Citation
https://lib.dr.iastate.edu/etd/15586

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
An instrument for assessing the public communication of scientists

by

Rachel Collier Murdock

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Rhetoric and Professional Communication

Program of Study Committee:
Jean Goodwin, Co-major Professor
Stacy Tye-Williams, Co-major Professor
Gary Ockey
Luke LeFebvre
Michael Dahlstrom
Charles Kostelnick

The student author and the program of study committee are solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2017

Copyright © Rachel Collier Murdock, 2017. All rights reserved.
TABLE OF CONTENTS

LIST OF FIGURES .................................................................................................................. vi

LIST OF TABLES ...................................................................................................................... vii

ACKNOWLEDGMENTS ........................................................................................................... viii

ABSTRACT .............................................................................................................................. ix

CHAPTER 1.0 ASSESSMENT AND SCIENCE COMMUNICATION TRAINING .............. 1

1.1 Introduction ....................................................................................................................... 1

1.2 Communication Training is Important for Scientists ....................................................... 4
  1.2.1 Communication Experts .......................................................................................... 5
  1.2.2 Formative Evaluations ........................................................................................... 9
  1.2.3 Summative Evaluations .......................................................................................... 11
  1.2.4 Defining Goals of Science Communication ............................................................. 14

1.3 Assessment Could Help Shift Science Culture ................................................................. 16
  1.3.1 Extend Definition of Scholarly Work ..................................................................... 17
  1.3.2 Provide Legitimacy for Science Outreach Activity .................................................. 18
  1.3.3 Counter the Sagan Effect ....................................................................................... 19

1.4. Assessment of Science Communication Requires an Analytic Scale ......................... 20

1.5 Current Assessment Instruments for Scientists’ Communication ................................. 26

1.6 Method for Developing a Science Communication Assessment Instrument ............... 33
  1.6.1 Domain Analysis .................................................................................................... 35
  1.6.2 Domain Modeling ................................................................................................. 36
  1.6.3 Conceptual Assessment Framework ....................................................................... 37
  1.6.4 Assessment Implementation ................................................................................... 39
  1.6.5 Assessment Delivery .............................................................................................. 41

CHAPTER 2.0 DOMAIN ANALYSIS .................................................................................. 43

2.1 Method of Gathering Samples .......................................................................................... 43
  2.1.1 Science Societies .................................................................................................... 43
  2.1.2 Government Agencies ........................................................................................... 44

2.2 Method for Gathering Information .................................................................................. 45
  2.2.1 Method for Evaluating Science Societies ................................................................ 45
    2.2.1.1 Gather Sample of Data .................................................................................... 45
    2.2.1.2 Unitizing and Grouping Data ........................................................................ 49
    2.2.1.3 Creating Context for the Data ....................................................................... 50
    2.2.1.4 Relate Results to the Research Questions ...................................................... 50
  2.2.2 Method for Conducting the Literature Review ............................................................ 51

2.3 Results of Analysis of Samples from Science Societies ................................................ 52
  2.3.1 Results of the Review of Science Society and Government Entity Websites .................. 52
  2.3.2 Results of Literature Review .................................................................................. 58

CHAPTER 3.0 DOMAIN MODELING .................................................................................. 67

3.1 The Rubric Should Aid in Scientists’ Communication Training .................................... 68
3.1.1 Formative Assessment .................................................................69
3.1.2 Summative Assessment ...............................................................69
3.1.3 Quantitative Analysis .................................................................70
3.2 Arguments in Favor of Using the Library Talk as the Work Product that is the
Subject of Assessment ...........................................................................71
3.3 Arguments in Favor of the Knowledge, Skills and Abilities Included in the
Assessment Instrument .........................................................................73
3.3.1 Arguments for Excluding Certain Knowledge, Skills and Abilities from
the Assessment ....................................................................................74
3.3.2 Knowledge, Skills, and Abilities Included in the Assessment of
Scientists Speaking to Public Audiences .................................................76
   3.3.2.1 Scientists speaking to public audiences will be able to
   connect science information to the specific audience to which they
   are speaking .......................................................................................77
   3.3.2.2 Scientists speaking to public audiences will be able to
   communicate complex scientific ideas and make them clear to the
   public .................................................................................................79
   3.3.2.3 Scientists speaking to public audiences will be able to
   increase their audience’s understanding of science and science
   processes ............................................................................................81
   3.3.2.4 Scientists speaking to public audiences will be able to
   humanize scientists and help them seem trustworthy and
   knowledgeable ...................................................................................84
   3.3.2.5 Scientists speaking to public audiences will be able to
   engage audiences in interactions and conversations about science ....85
3.4 Arguments in Favor of Evidence Used for Supporting the Identified KSAs ........88
   3.4.1 Scientists speaking to public audiences will be able to explain the
   relevance and importance of science information to the specific audience to
   which they are speaking .........................................................................89
   3.4.2 Scientists speaking to public audiences will use concrete, direct
   language and analogies to communicate complex scientific ideas and make
   them clear to the public ........................................................................89
   3.4.3 Scientists speaking to public audiences will explain specific science
   processes and techniques to their audiences ........................................95
   3.4.4 Scientists speaking to public audiences will use techniques such as self-
   disclosure and immediacy to help scientists seem trustworthy and personable ....98
   3.4.5 Scientists speaking to public audiences will engage in discussion and
   interactions with their audiences .......................................................101

CHAPTER 4.0 THE CONCEPTUAL ASSESSMENT FRAMEWORK .........................106
4.1 The Student Model .............................................................................106
4.2 Evidence Model ................................................................................108
   4.2.1 Rubric Development ..................................................................110
   4.2.1.1 Development of Science Communication Items .....................111
   4.2.1.2 Public Speaking Competency Rubric (PSCR) Adaptation ........113
   4.2.1.3 General Rubric Layout and Design ......................................115
4.3 Task Model .......................................................................................116
4.4. Assembly Model ................................................................. 119
4.5 Presentation Model ............................................................. 120
4.6 Delivery System Model .......................................................... 120
4.7 Initial Revisions and First Tested Version of the Public Science
Communication Rubric ................................................................. 121
  4.7.1 The Raters ........................................................................ 122
  4.7.2 The Rated Presentations ...................................................... 123
  4.7.3 Initial Feedback .................................................................. 126
  4.7.4 Pilot Testing, Round 1 ......................................................... 132
  4.7.5 Revisions to Rubric and Codebook ....................................... 139
  4.7.6 Pilot Testing, Round 2 ......................................................... 145

CHAPTER 5.0 ASSESSMENT IMPLEMENTATION ............................................. 149
  5.1 Initial Norming Session with Coders ........................................ 151
  5.2 Revisions ............................................................................. 156
  5.3 First Round of Testing ........................................................... 157
  5.4 Revisions to Rubric and Code Book ........................................ 159
  5.5 Second Round of Testing ........................................................ 165

CHAPTER 6.0 ASSESSMENT DELIVERY, LIMITATIONS AND FUTURE
RESEARCH .................................................................................. 178
  6.1 Assessment Delivery .............................................................. 178
  6.2 Limitations .......................................................................... 179
  6.3 Future Research .................................................................... 182

7.0 DISCUSSION ........................................................................... 184
  7.1 The Need for Public Science Communication and its Assessment .... 184
  7.2 Defining the Desired Outcomes of Public Science Communication .... 185
    7.2.1 Additional Categories: Persuasion and Information to Make Decisions .... 187
    7.2.2 Application to Persuasive or Decision-Making Measures .............. 188
  7.3 Operationalizing the Desired Knowledge, Skills, and Abilities .......... 190
  7.4 Testing Rubric Validity and Working with Raters ......................... 192
    7.4.1 Challenges with Language Testing ........................................... 192
    7.4.2 Challenges with Science Communication Rating ........................ 193
  7.5 Choosing, Training, and Working with Raters ............................. 196
    7.5.1 Choosing Raters ................................................................. 196
    7.5.2 Rater Training ................................................................. 199
  7.6 Bringing Communication Expertise to Scientists’ Training ............... 203

8.0 CONCLUSION ......................................................................... 205

REFERENCES ............................................................................. 210

APPENDIX A: RESULTS OF THE DOMAIN ANALYSIS CONTENT ANALYSIS
STEP. BROAD GOALS FOR SCIENTISTS COMMUNICATING WITH PUBLIC
AUDIENCES .................................................................................. 227

APPENDIX B: RUBRIC FOR ASSESSING PUBLIC COMMUNICATION BY
SCIENTISTS, FIRST DRAFT ............................................................ 229
APPENDIX C: 2-7-17 DRAFT OF RUBRIC .................................................................231
APPENDIX D: 3-8-17 - RUBRIC FOR ASSESSING PUBLIC COMMUNICATION BY
SCIENTISTS .............................................................................................................233
APPENDIX E: 3-20-17 - RUBRIC FOR ASSESSING PUBLIC COMMUNICATION BY
SCIENTISTS SENT TO FINAL CODERS BEFORE FIRST MEETING –
SCIENCE COMMUNICATION ELEMENTS ONLY .............................................235
APPENDIX F: 3-24-17 ONLY SCIENCE PORTION OF THE RUBRIC (PSCR
PORTION DID NOT CHANGE FROM THIS POINT ON) INCLUDES
SCORING/NOTE TAKING SHEET USED FOR ROUND 1 OF FINAL TESTING ...237
APPENDIX G: 3-24-17 SCIENCE ONLY RUBRIC WITH SCORING/NOTE TAKING
PAGE USED FOR ROUND 2 OF FINAL TESTING .............................................240
APPENDIX H: FIRST CODE BOOK FOR APPS – THE ASSESSMENT OF PUBLIC
PRESENTATIONS BY SCIENTISTS .................................................................243
APPENDIX I: FINAL CODE BOOK FOR APPS – THE ASSESSMENT OF PUBLIC
PRESENTATIONS BY SCIENTISTS CODE BOOK, INCLUDING THE FULL
FINAL RUBRIC .......................................................................................................251
LIST OF FIGURES

Figure 4.1. First Rubric Draft - Illustration of the layout of the first draft of the rubric prior to the first round of pilot testing..........................................................116
Figure 5.1. Portion of the language construct from the rubric as revised 3/24/17..........................157
LIST OF TABLES

Table 1.1. Knowledge, Skills, and Abilities Found Most Often in the Science Society Texts .....58
Table 4.1. Raters for Pilot Study, their Qualifications and Participation ..........................122
Table 4.2. Initial Pilot Testing Round 1, Overall Agreement and Agreement by Categories ....138
Table 4.3. Overall Data, Pilot Study Rounds One and Two ..............................................146
Table 5.1. Raters for the Final Two Rounds of Coding ..................................................151
Table 5.2 Round 1 - Full Testing ....................................................................................158
Table 5.3. Results of Final Testing, Round Two ................................................................166
Table 5.4. All Rounds with Overall Scores Showing Simple and Adjacent Agreement ..........176
ACKNOWLEDGMENTS

I would like to thank my spouse, Joseph, for his unflagging support and encouragement and my children for their optimism and patience. I thank my parents, Robert and RaeAnn Collier, for instilling in me a love of learning and for always encouraging me to pursue my goals. I am grateful my father lived to see me complete my PhD. His passing in July 2017 was a great loss. In addition, I would like to thank my committee co-chairs, Jean Goodwin and Stacy Tye-Williams for their consistent confidence, positivity and support. I also wish to thank the members of my committee for their advice and cheerful guidance.

I express my gratitude to Adah Lesham and the fellows in the GK-12 SYMBI program at Iowa State University for the opportunity to learn from and work with such dedicated and brilliant young scientists. I also owe Lisa Schreiber a debt of gratitude for her congeniality, her permission to use parts of her Public Speaking Competency Rubric, and her interest in making the tools for developing excellent communication skills as freely and easily accessible to as many audiences as possible.
ABSTRACT

An instrument for valid, quantitative assessment of scientists’ public communication promises to promote improved science communication by giving formative feedback to scientists developing their communication skills and providing a mechanism for summative assessment of communication training programs for scientists. A quantitative instrument also fits with the scientific ethos, increasing the likelihood that the assessment will gain individual and institutional adoption. Unfortunately, past assessment instruments have fallen short in providing a methodologically sound, theory-based assessment instrument to use in assessing public science communication. This dissertation uses the Evidence Centered Design (ECD) method for language testing to develop and test the APPS—the Assessment for Public Presentations by Scientists—a filled-cell rubric and accompanying code book based on communication theory and practice that can be used to provide formative and summative assessments of scientists giving informative presentations to public, non-scientist audiences.

The APPS rubric was developed by employing an extensive domain analysis to establish the knowledge, skills, and abilities most desired for scientists who speak to public audiences, based on a methodical review of scientific organizations and a systematic review of science communication scholarship. This analysis found that scientists addressing public audiences should speak in language that is understandable,
concrete, and free from scientific jargon, translating important scientific information into language that public audiences can understand; should convey the relevance and importance of science to the everyday lives of audience members; should employ visuals that enhance the presentations; should explain scientific processes, techniques, and purposes; should engage in behaviors that increase the audience’s perceptions of scientists as trustworthy, human, and approachable; and should engage in interactive exchanges about science with public audiences. The APPS operationalizes these skills and abilities, using communication theory, in a detailed, user-friendly rubric and code book for assessing public communication by scientists. The rubric delineates theory-based techniques for demonstrating the desired skills, such as using explanatory metaphors, engaging in behaviors that increase immediacy, using first-person pronouns, telling personal stories, and engaging in back-and-forth conversation with the audience.

Four rounds of testing provided evidence that the final version of the APPS is a reliable and valid assessment, with constructs that measure what they are intended to measure and that are seen similarly by different raters when used in conjunction with rater training. Early rounds of testing showed the need to adjust wording and understanding of some of the constructs so that raters understood them similarly, and later testing showed marked improvement in those areas. Although the stringent interclass agreement measure Cohen’s kappa did not show strong agreement in most measures, the adjacent agreement (where raters choose scores that are within one point
of each other) was high for every category in the final testing. This shows that although raters did not often have exactly the same score for speakers in each construct, they nearly always understood the construct similarly.

The agreement ratings also accentuate the study’s finding that the raters’ backgrounds may affect their abilities to objectively score science speakers. Testing showed that science raters had difficulty separating themselves from their inherent science knowledge and had difficulty objectively rating communication skills. Therefore, this study finds that scientists can act as communication raters if they are trained by practicing rating science presentations as a group to norm scoring and by studying communication skills discussed in the code book. However, because of the possible difficulty separating themselves from their intrinsic science knowledge and their lack of experience in identifying excellent communication practices, the assessment of science speakers will nearly always be more accurate and the communication performance of scientists more enhanced when utilizing communication experts to help train and assess scientists in their science communication with public audiences.

Therefore, the APPS can be a valuable tool for improving the knowledge, skills, and abilities of scientists communicating with public audiences when used by communication training programs to provide prompt, specific feedback. Given the reliability limitations, the rubric should not be used for high-stakes purposes or for “proving” a speaker’s competence. However, when used in a science communication
training program with consistent raters, the APPS can provide valuable summative and formative assessment for science communicators.
CHAPTER 1.0 ASSESSMENT AND SCIENCE COMMUNICATION TRAINING

1.1 Introduction

Water and soil scientists in Idaho meet with city leaders to discuss ways to mitigate the effects of global warming on the area’s water supply. An animal scientist in New England enlists the help of citizens to gather data on roadkill they find as they drive the country for her population study. A group of community members meet at a local bar after hours to hear a scientist speak about the latest innovations in fluid science and how those innovations affect brewery technology. Science-oriented blogs, Twitter accounts, and Facebook pages abound. As these examples indicate, after decades of encouragement for scientists to communicate with various publics (Heagerty, 2015; Hermens, 1986; Royal Society, 1985; Safina, 2012), scientists and other stakeholders do generally agree and act upon the idea that the public communication of science is important and that such communication should be prioritized for the benefit of society at large, members of the public, and scientists themselves.

Despite this general agreement on its value and importance, there are challenges to public communication by scientists. While some challenges arise because of institutional resistance or an obstruction of resources (Bandelli & Konijn, 2012), many of the challenges arise because of attitudes scientists hold. For example, some scientists do not want to participate in communication or outreach activities because such activities take time away from their “real” work (Ecklund, James, & Lincoln, 2012; Safina, 2012),
some scientists don’t want to participate because they lack confidence (sometimes rightly so) in their public communication skills (Ecklund et al., 2012; Meredith, 2010), and some scientists fear ridicule and contempt from peers if they participate in public communication activities (Foote, Krogman, & Spence, 2009; Vergano, 2011). Meanwhile Kuehne, et al. (2014) contend that graduate students in science fields are not given adequate resources to learn science communication skills while in graduate school, so they are unprepared to contribute to science communication activities after graduation. Wynne (2006) posits that scientists hesitate to interact with the public because they believe there is a public mistrust of science and their efforts will be unappreciated.

Communication scholars and scientists have suggested a variety of ways to address these issues, in particular by increasing communication training for scientists (Heath et al., 2014; Kuehne et al., 2014; Neeley, 2013; Nisbet & Scheufele, 2009) and shifting the culture of science. This culture change would include adjustments to the incentive and promotion structures under which scientists operate so that public communication and outreach is recognized and rewarded rather than ridiculed and reviled (Ecklund et al., 2012; P. Jensen, Rouquier, Kreimer, & Croissant, 2008).

While there are legitimate arguments made in favor of both increasing the amount and quality of communication training for scientists and changing the culture of promotion and tenure in science, both of these approaches to increasing science communication face many roadblocks. Despite the fact that there is no simple solution to
the challenges, there are steps that can be taken in order to mitigate some of these challenges and improve the outcomes of both of these approaches. One such step is more frequent and more effective assessment of scientists’ public communication.

This dissertation outlines several ways in which increased assessment could enrich and improve science communication training and offers some possible areas where increased assessment could be part of a larger effort that might make progress in the conversations on shifting the culture of science promotion and tenure. It then discusses the best type of assessment for evaluating public communication by scientists, reviews the current instruments available for the purpose, and then discusses the method this project uses for developing the assessment. The method, evidence centered design, or ECD, is outlined and the steps taken for this study are described. The resulting filled-cell rubric is called APPS: The Assessment for Public Presentations by Scientists. APPS and its supporting codebook along with the reliability testing done with the rubric are offered as a means of assisting in the training of scientists to engage in public communication.
1.2 Communication Training is Important for Scientists

There are scholars from both science and communication fields who regularly propose that the key to increasing the amount and quality of the public communication of science is increasing and improving communication training for scientists. Nisbet & Scheufele (2009), for example, suggest a course for developing communication skills in young scientists, while Kuehne, et. al (2014) go further, proposing a five-prong approach to improving science graduate students’ communication and outreach skills. Such proposals come not only from individual scientists, but also from funding agencies. The NSF, in addition to encouraging outreach and communication activities with its “Broader Impacts” requirement on each funding application, also hosts a “Becoming the Messenger” training series specifically focused on teaching scientists to become “ninja” communicators who work strategically to get their messages out through media channels (Druschke & McGreavy, 2016). Additionally, the venerable AAAS sponsors a media fellows program that places science, engineering, and math fellows at media organizations and actively encourages more and better communication training for scientists (“AAAS Mass Media Science & Engineering Fellows Program,” 2017; Leshner, 2007). Even science-loving actor Alan Alda is in the game, working with Stony Brook University to use storytelling and improvisation methods to train scientists to be more engaging communicators (Filler, 2013). With such encouragement, the number of
communication training programs for science PhD candidates and more advanced scientists is increasing, at both academic and professional institutions.

While great strides are being taken in both the envisioning and implementing of such programs, there are yet improvements that need to be made to make the programs more valuable to the participants and to demonstrate their importance to those outside the programs. More and improved assessment of scientists can help increase the number and efficacy of these programs in four important ways. First, better assessments can bring communication experts into the communication training of scientists. Additionally, such assessments can provide impactful formative feedback for participants in these programs. Third, assessments can be constructed to provide summative feedback in keeping with the scientific ethos thereby improving the prestige of and funding for communication training programs. Lastly, quality assessments can focus attention on the most important aspects of science communication, which helps in developing quality science communication training programs. Each of these outcomes will result in better training for scientists and better results for communication programs. I will review each of these outcomes below.

1.2.1 Communication Experts

A well-developed communication assessment instrument will bring communication expertise into the communication training programs of scientists. Currently, communication training in science programs is often implemented and
coordinated by scientists rather than communication experts. While scientists certainly know a great deal about science, they often are not trained in and are not aware of communication research, including public speaking research. In their 2014 study, Neeley, et al. found that scientists often give lip service to the need for collaboration with outside experts in developing communication training for scientists but found that very little of this collaboration happened in actual practice. This is unfortunate, since even those scientists who seem to excel at outreach and public communication may not know exactly what it is they do with their communication that makes it successful nor understand exactly how to distill those useful methods and skills into patterns and instructions that they can share with other scientists.

On the other hand, communication specialists have researched and studied just those kinds of topics and are skilled in conveying this information to others. They are also knowledgeable about the best and most successful practices of public communicators generally, in keeping with the current research in the field. Many of them are also experts in assessing communication. Therefore, communication specialists can help science communication programs by bringing their expert knowledge of excellent communication practices with them to science communication situations. They can also put their communication analysis skills to work by discovering and bringing to the fore the most important communication practices uniquely important to public science communication and then operationalizing those practices so that they can be
accurately assessed. In fact, Neeley, et al. (2013) noted that one of the primary benefits seen when there is increased cooperation with communication experts in science communication training is better communication assessment. With their expertise in communication assessment, communication experts can apply sound principles of communication to science communication contexts, and they can provide valuable expertise in developing assessments that address the aspects of communication that are most important to scientists and that measure these aspects accurately.

Lastly, an assessment developed by communication experts can be used by properly trained, communication-focused scientists in addition to communication experts. While involving expert communicators in the communication training of scientists would be ideal, as Neeley, et al. (2013) noted, there are scientists who resist involving communication experts in training scientists.

Because of negative attitudes that exist among some scientists, such as the belief that science is undervalued or oversimplified by non-scientists (Kahan, 2015), the idea that members of the general public do not know much about science and scientific processes (J. D. Miller, 1983), and the assertion that non-scientists do not have a “place” in scientific discussions (Fabj & Sobnosky, 1995), it is possible that scientists may feel assessments will be more valid or legitimate if done by scientists themselves rather than by communication experts enlisted from other academic disciplines. Additionally, it may
be easier for some scientists to accept critique and suggestions from members of their own discourse community rather than those they consider to be outsiders.

Nevertheless, while scientists may prefer to have critique from within their own peer community, such critique and advice may not give them the help they need. According to one study of outreach and communication activities, over 40% of scientists’ outreach involves giving presentations to or interacting with lay audiences (Ecklund et al., 2012). With public presentations being such a vital component in scientists’ public communication and considering that public presentations to lay audiences are substantially different from the type of presentations that scientists give to their peers, it is important for scientists to have such presentations evaluated by those who have expertise in this area – communication specialists.

A quality, academically-sound assessment instrument developed by communication experts can meet the needs of those being trained and still meet the needs of those running training programs who may feel uncomfortable handing over assessment responsibilities to “outside” experts. In order for this to work effectively, the assessment instrument needs to be well-developed and include adequate documentation and training materials. Communication experts can train the program leaders or designated assessors to use the assessment instrument properly. Ideally, communication specialists would also conduct norming sessions to promote inter-coder reliability. At that point, if the program coordinators preferred to have ownership of the assessment of
their own participants, they could take over those responsibilities while still providing many of the benefits of expert communication training.

1.2.2 Formative evaluations

Not only will a quality communication assessment instrument bring communication expertise into the science communication training community, it will also allow for formative assessments of science communicators, which can help improve scientist’s overall communication skills. Formative assessments can help scientists recognize strengths and weaknesses, allowing them to become better, more confident communicators as they venture out into the world of public communication.

In order to help individual scientists improve their communication skills, formative feedback should be used effectively. Taras’s research (2003) confirms that evaluation, particularly supportive, prompt feedback, leads to improved public speaking outcomes, such as better message clarity, better audience understanding, and better self-evaluation by the speaker (in other words, speakers are better equipped to accurately analyze their own communication efforts when they get quality formative assessments). According to Stevens and Levi (2013), well-prepared assessment instruments allow assessors to provide this kind of feedback. Formative assessments can be particularly helpful to scientists, since such assessments are usually part of low-stakes exercises which allow communicators to experiment with new techniques with few negative consequences. A properly developed assessment instrument will allow assessors in
formative situations to give prompt, detailed feedback on any new communication efforts. This kind of feedback allows a communicator to make rapid changes to develop their communication, leading to more significant improvements long term.

In addition, regular, formative assessments have the potential to reduce the fear and increase the confidence scientists have about speaking to public audiences. Work such as that done by Hunter, Westwick and Haleta (2014) shows that regular, assessed public speaking experiences significantly reduce public speaking anxiety. A quality assessment tool makes formative assessment meaningful yet low-pressure enough to be valuable in reducing fear of speaking and will likely increase the number of practice presentations trainees are asked to give, since both student and instructor will be more likely to recognize a purpose in practice presentations with quality feedback.

Therefore, an assessment instrument developed specifically for science speakers can be an important element for use in completing formative assessments in public communication training programs for scientists. These formative assessments have the potential to assist scientists in learning public presentation skills, improving self-awareness about their public communication, improving self-evaluation about their speaking in public, and reducing anxiety about presenting to the public. In this way, a better formative assessment instrument can lead to improvements in public science communication training, and potentially assist in the improvement of public science communication.
1.2.3 Summative Evaluations

In addition to providing helpful formative assessments, a tested assessment instrument that can be used as a quantitative measure can also offer excellent summative evaluation (Stevens & Levi, 2013). Summative evaluations can be used for individual and programmatic evaluations, and could therefore possibly lead not only improvement on the part of individual communicators, but also to stronger, more valued, and more valuable overall science communication programs (Ewell, 1991).

Summative assessments can be used to show the improvement of individual communicators over time or can be used to show their progress in individual categories. Such assessments can also be used to show the overall work being done in the program and the benefits of the communication training as a whole. This kind of assessment, in combination with other types of assessments, is important in justifying a science communication program to programs, departments, and grant administrators. For example, summative evaluations can be used to demonstrate the value of a communication training program in supporting the NSA’s broader impacts requirements or similar statements on grant applications and annual reports. Lori Wingate, a research director at Western Michigan University, pointed out that principal investigators seeking funds from the NSA are required to show evidence of their broader impacts accomplishments. “Describing achievements and presenting evidence of the quality and impact of those achievements shows reviewers that the proposer is a
responsible steward of federal funds, can deliver on promises, and is building on prior successes” (2016, p. 2). Wingate notes that since the NSA requires evidence of achievements, researchers need to provide data (Wingate, 2015) to support claims of achievement in broader impacts. A quantitative rubric can assist in this process.

Summative evaluations are also important in supporting these training programs. Such training programs generally take place at a university, and universities require evidence of achievement. Often this evidence is required of university because other stakeholders, particularly governments which help fund universities, require such evidence (J. Burke, 2005; Carey & Schneider, 2010). Ewell (1991) said that since the beginning of the educational accountability movement, there has been agreement on the types of educational assessments that should be required -- namely, assessments should result in quantitative data that is easy to understand and that indicates institutional or programmatic performance. Given these expectations for programmatic assessments, a good assessment instrument can help provide some of the data that can be used to support claims of achievement and overall value and efficacy of a science communication training program.

Additionally, summative assessments can provide support for science students’ participation in communication training programs. Evidence suggests that such engagement can help scientists’ academic work (P. Jensen et al., 2008), but communication training programs are seldom required in Master’s or PhD programs for
science students, even in programs where communication training classes or programs are offered. Showing evidence of the benefits of participating in a training program might encourage more students to participate in communication training. Additionally, quantitative program assessment could be a piece of helping to make the case to administration officials that communication training for scientists is valuable and effective. It may also encourage administrators to allocate resources to administering and improving such programs. Providing quantitative evidence is often vital in gaining that support, particularly in programs directed at scientists.

The scientific ethos places value on such things as measurement, comparison, and data. As noted by Robert Merton decades ago (Merton, 1973), many people believe science to be governed by four norms: Universalism (truth claims are subjected to pre-established, impersonal criteria), Communism (an imperative for the communication of data and research findings), Disinterestedness (institutional checks are in place that prevent scientists from pursuing their career goals ruthlessly), and Organized Skepticism (all research is checked by rigorous, structured scrutiny of peers). While others have critiqued these norms and shown that they do not necessarily function in reality (Grundmann, 2012), this view of science does persist, both among scientists and among those who observe scientists. Therefore, these ideas about science and its processes should be acknowledged by communication experts working with scientists, and these experts should develop assessments that complement this understanding. By
the same token, assessment instruments that fit this scientific ethos are likely to be more readily accepted by scientists as being valid and useful than will other methods of assessment. Therefore, by using a reliable assessment instrument that utilizes a quantitative approach, science communication program administrators can provide data that resonates with scientists, supports broader impacts statements, illustrates the need for communication training, and encourages the allocation of resources to communication programs.

1.2.4 Defining Goals of Science Communication

The final advantage of a well-supported assessment tool for science communication is that it will assist in delineating the goals of the public communication of science. There have been several surveys of scientists asking for their reasons for engaging in outreach activities (Besley, Dudo, Yuan, & Ghannam, 2016; Davies, 2008; E. Jensen & Holliman, 2009) and essays enjoining scientists to engage in public communication for a variety of reasons the authors find important (Burns, O’Connor, & Stocklmayer, 2003; Eagleman, 2013; Safina, 2012). There are also treatises from organizations giving their own reasons for the necessity of scientists interacting with publics (House of Lords, 2000; Many experts, many audiences: Public engagement with science and informal science education, 2009; Neeley, 2013). However, all of these reasons and calls from different areas should to be brought together, analyzed and distilled into the most important goals for public science communication as proposed by all of these
stakeholders. Defining these goals and the knowledge, skills, and abilities scientists should demonstrate in order to meet those goals will help not only with the assessment of public science communication, but with other aspects of training as well.

There are differences of opinion as to what knowledge, skills, and activities are most important for science communicators and how they should be developed. In scientist Stewart Justman's opinion, for example, science literacy work needs to be done in the schools by teachers (Vergano, 2011) not in public venues by scientists – but others disagree, claiming that informing the public is an important job for scientists (Davies, 2008). A challenge to the process is scientists' distrust of the media. A National Science Foundation study found that scientists distrust the media and consider journalists to be generally uninformed and interested only in the sensational (Board, 2000). Additionally, scientists may have difficulty adjusting their technical language when talking to publics. As Martin-Sempere, Garzon-Garcia, and Rey-Rocha (2008) pointed out, scientists who try to communicate with the public must adapt their communication to audiences and communication styles with which they are unfamiliar, which requires extra effort and which introduces more chances of being unsuccessful. Even more, many scientists are hesitant to become public communicators because they feel that such interactions with the public may turn them from “pure scientists” into “advocates” for a particular point of view. While this may seem a trivial distinction to others, to a scientist immersed in the ethos of science, such concerns are anything but trivial. For scientists inculcated with the
Mertonian norms, the view of science as something “super” ordinary does persist, not least in the minds of scientists themselves (Lackey, 2007).

These tensions indicate an important need for reflection about what scientists are trying to accomplish with public science communication. Effective assessment instruments can distill all the encouragements, calls, and contradictions about the purposes for public science communication and help define well-supported knowledge, skills, and abilities that scientists should demonstrate when speaking to public.

1.3 Assessment Could Help Shift Science Culture

While the primary goal of an assessment instrument for the public communication of science is to improve the training of scientists, the use of the assessment in its summative role could have a small but positive influence in helping shift the science culture of promotion and tenure. While suggestions that the culture of promotion and tenure in science needs to be changed and encouragements to change it are plentiful, actually changing ideas about science communication can be, to borrow a metaphor, like turning around a huge ocean liner in a narrow channel while moving against the current. Changes are slow, and often resisted. Additional motivations and methods for making changes may be necessary, and improved assessment of public communication could play a small but significant role in helping shift the culture of promotion and tenure within the scientific community.
1.3.1 Extend Definition of Scholarly Work

While the promotion and tenure culture in science is well established, there are many potential benefits of changing that culture. Traditionally, scientists earn tenure by publishing in recognized, peer-reviewed, scientific journals in their specific fields. Being awarded research grants also has a positive influence on tenure decisions. However, scientists’ contributions to communication, science communication, or interdisciplinary journals are not regarded as valuable to their advancement, particularly for scientists who work in academic settings. Additionally, many common types of science communication work done by scientists are rarely acknowledged or given weight in tenure or other promotion decisions. Some of these types of communication include speaking to public audiences, developing educational programs to interest students in pursuing STEM careers, working with governmental bodies, consulting with farmers or outreach groups, and making presentations to special interest groups about advancements in the field that might have practical applications to lay persons.

In a personal essay in the American Physical Association’s APS News, scientist Carl Safina (2012) outlined other reasons scientists are hesitant to engage in public communication. He mentioned the time it takes from teaching, from home, and from other things that need doing. He also noted the lack of compensation (either financial or professional) for communicating science, the lack of vitae augmentation, the lack of progress toward tenure – and the possibility that outreach efforts will hurt tenure
chances. He also noted that communicating science to the public is still seen as unprofessional and is considered something best left to science writers. Changing the culture so that more value is placed on public communication and outreach activities could lead to more scientists engaging in those activities. Having a way to assess the quality of those outreach efforts may give them more credibility and encourage their use in promotion and tenure decisions.

1.3.2 Provide Legitimacy for Science Outreach Activity

A well-researched and strongly supported assessment instrument used in the assessment of science communication effort can help give the needed legitimacy to science communication and outreach in the eyes of scientists. As previously noted, despite consistent encouragement to engage in public outreach, many scientists ridicule and avoid such interactions. Developing a supported and validated assessment instrument can be used as a step forward in legitimizing and substantiating public science communication in the eyes of scientists, particularly when assessment instruments complement science culture. Assessment measures appropriate for assessing scientists will utilize quantitative scales using consistent, impersonal criteria and providing data that can be analyzed and compared by individual and groups. By providing this kind of data, assessments will be more welcomed by scientists, who tend to be most comfortable with quantitative, analytic methods of observation and evaluation.
1.3.3 Counter the Sagan Effect

The extended definition of outreach work and the increase in the legitimacy of such work can lead to another advantage for scientists and science culture: the mitigation of the so-called Sagan Effect. This term is used to describe the phenomenon of scientists being reluctant to engage in outreach work because such work may not only fail to help scientists earn recognition and reward, it may, in fact, earn them the disdain of colleagues (Ecklund et al., 2012; P. Jensen et al., 2008). The name comes from Carl Sagan, a great science communicator and popularizer of science, who was denied admission to the National Academy of Sciences in 1994 in a vote by his peers. The reason given for his rejection was that he was not a serious science professional since he spent so much time communicating with the public, despite his numerous academic articles. More than one scientist has proposed that, despite the strong encouragement that scientists currently get to participate in the public communication of science, the culture of disdain for science popularizers hasn’t changed a great deal since Sagan’s experiences. If the culture surrounding outreach and public science communication were to change so that such communication was seen as legitimate, valued, and rewarded, the chilling consequences of the Sagan Effect might be mitigated. Using a validated assessment instrument that fit the psyche of science could be one incremental step that will help scientific and academic institutions ease into the needed change of viewing science outreach as a positive rather than a negative activity.
1.4. Assessment of Science Communication Requires an Analytic Scale

In the previous section, I have argued that a well-researched, valid assessment instrument for scientists’ public communication activities would be make an important contribution to both science communication training and to the acceptance of science outreach and communication work as being valid and important. Nonetheless, an assessment method needs to be carefully chosen and developed in order to gain acceptance in the scientific community and thereby deliver the desired results. In this section, I will describe the type of assessment instrument best able to fit the criteria needed for a public science communication assessment instrument.

There are several ways oral communication can be assessed. One of these ways is by using a holistic scale (Fulcher, 2003), also called a global scale (Bachman, 1996). When using these scales, an assessor gives an overall rating - either a number grade, letter grade, or some other agreed-upon measure – of a student’s overall performance on a particular communication task. As Fulcher (2003) noted, holistic/global scoring may be guided by a rating scale, but the single score is meant to capture the overall quality of the communication. While such scores are simple to assign, quick to use, and easy to quantify, Bachman and Palmer (1996) point out some drawbacks to using holistic scoring. They say such scoring makes it difficult to know exactly what the score reflects, difficult to assign a level when the communicator may perform well on one aspect of the rating but not in another aspect, and difficult to weight components properly, since the
score tries to reflect such a broad range of skills (209). These characteristics make them less desirable for use by scientists, who prefer quantifiable data that can be analyzed in detail.

Another type of scoring mechanism is what Schreiber, Paul, and Shibley (2012) call a rating scale. By their definition, a rating scale has multiple criteria, or constructs, that it measures. However, the rating scale gives just one description of a skill and then lists a series of numbers or words to use in ranking how well the communicator did at demonstrating the skill. For example, the first criteria may be: “The communicator used an attention-getting device to get the attention of the audience” followed by “1,” “2,” “3,” “4” or “Excellent,” “Very Good,” “Good,” “Fair,” and “Poor.” Schreiber, Paul and Shibley suggest that while a rating scale is quick to grade and fairly easy to produce, it, like the holistic scale, has several drawbacks. For example, the difference between designations such as “Good” and “Fair” or “3” and “4” is unclear, leaving the possibility for a wide range of variation in scoring. Additionally, the rating scale leaves students with little feedback to use in improving their performance. This type of scale can provide quantifiable data, but it is limited in its value to the communicator, making it less effective in formative situations and less meaningful overall.

Bachman and Palmer, Fulcher, and Schreiber, Paul, and Shibley all suggest using an analytic scale (also called a filled-cell rubric or a descriptive rubric) to obtain the most effective and helpful rating of communicators’ work. Like a rating scale, an analytic scale
describes multiple constructs, or skills, that go into competently completing a task, such as making a public presentation about science. However, instead of giving one word or one number options for scoring each particular aspect of the task, the analytic scale offers “performance standards for each outcome or competency on the rubric, and consequently makes explicit the difference between advanced and proficient performance” (Schreiber et al., 2012, p. 212).

Analytic scales are excellent for assessing public science communication for several reasons. First, they can be easily quantified, giving scientists numerical data that can be compared and analyzed. Mertler (2001) points out that analytic scales allow assessors to assign each level of competence a numeric value. The scale can then be scored in the individual sections, or constructs, first and then those individual scores can be added together to get an overall, numeric score. This allows evaluators to see numbers, either by overall competence or broken down by individual competencies. This degree of detail also allows easy comparisons from semester to semester, over years, and between presenters. This level of precision makes analytic scales the “gold standard” of assessment, particularly useful for high-stakes purposes such as accreditation, funding, and program continuance (Schreiber et al., 2012, p. 212), all important considerations in science circles.

An analytic scale also makes it possible to more exactly define how scientists should demonstrate that they are meeting the goals of public science communication.
Constructing the scale requires thought and analysis about the learning objectives for the task or assignment being assessed and descriptions of how those objectives, in this case, the goals of public science communication, are operationalized. Developing this analytic scale can encourage specificity in describing the performance outcomes, particularly if principles of criterion-based testing (Bachman & Palmer, 1996) or evidence centered design (R. J. Mislevy, Almond, & Lukas, 2003) are followed. When using either of these methods, a rubric developer would carefully analyze the “domain” or construct that they are trying to assess and methodically identify the abilities scientists would demonstrate to show mastery of the identified constructs. In this case, identifying the construct would mean carefully analyzing the desired abilities a scientist would demonstrate when engaging in public science communication and then specifying what kinds of words and actions demonstrate a mastery of that ability, something important in creating an assessment instrument that improves public science communication training. This careful work helps ensure that the assessors know exactly what they hope to see from the scientists and exactly how goals are met, and it helps ensure that the wording of the analytic scale developed to assess the scientists’ skills and efforts accurately reflects masterful demonstrations of those abilities.

Analytic scales are also favored by many assessment experts because they can be used effectively for both summative and formative purposes. Instructors in a training program will have many duties to fulfill, and, as important as assessment is in
improving performance, instructors giving qualitative or holistic feedback could find it overwhelming or oppressive to give such feedback in formative settings. However, as Stevens and Levi (2013) point out, analytic scales save time in grading, because they give explicit feedback to the communicator, which means the assessor needs to spend less time writing or recording comments. Therefore, assessors are able to complete evaluations more quickly and return them to learners promptly. This means the scientists who are working on their communication will be able to more easily remember the thought processes they used when creating and presenting a talk or seminar and will be better able to incorporate feedback into their next presentation.

Analytic scales also provide more detailed feedback than other types of assessments such as holistic scores or rating scales, making the feedback more valuable and helpful to those using it, in this case, scientists creating public presentations. When using analytic scales, an assessor can circle or highlight specific words or phrases in the skill description found in the boxes on the scale, giving the speaker exact feedback about the areas in which they need to improve. Mertler (2001) said that the degree of feedback gained by both the person being assessed and person doing the assessing when using an analytic scale is significant, making their use a substantial improvement over other scoring methods. What’s more, not only are scientists able to improve their presentations, but science communication trainers are able to improve their coursework based on the weaknesses they identify in the science presentations. Therefore, a well-
developed analytic scale will fit the need for an assessment instrument that provides
excellence in both formative and summative contexts.

Of course, any analytic scale is only effective and useful if the assessors using the
scale are trained in its use. As pointed out by both Bachman and Palmer (1996) and
Schreiber, Paul & Shibley, (2012), coders should be trained to use the scale properly and
any differences in understanding the constructs being measured should be carefully
discussed and resolved. No scale will be effective at measuring communication well if it
is not administered carefully and consistently. Any analytic scale developed for
assessing scientists will need to include instruction for raters and a code book which will
help assessors identify examples of both excellent and less competent communication,
and communicators (in this case, scientists) will need to be aware of the constructs
outlined in the rubric in order for it to be useful for formative assessment.

Therefore, an effective assessment instrument for assessing the oral public
communication of scientists will be a filled-cell, or analytic, rubric that is carefully
researched so that it assesses the primary knowledge, skills, and abilities for public
speaking and public science communication as expressed by scientists, science
communicators, communication experts, and science institutions. It will require training
to use properly, but will provide a fairly simple, relatively quick way for trained
assessors to provide both specific, formative feedback for incremental improvement as
well as quantitative, summative feedback for evaluation, analysis and comparison. It
will also be tested for validity and reliability through vetted linguistic and communication techniques. The Assessment for Public Presentations by Scientists, or APPPS rubric, reflects these requirements for an effective instrument for assessing public science communication.

1.5 Current Assessment Instruments for Scientists’ Communication

It is first important to determine whether or not such an instrument currently exists. In this section, I examine the four assessment instruments created specifically for assessing public communication by scientists and one analytic rubric developed to assess public speaking presentations generally. Each of the first four was developed to assess a specific type of public communication by scientists, while the public speaking rubric was intended to have broad application to many situations. Following is a brief review of each and an explanation of why none is entirely adequate for the purpose of assessing public oral presentations by scientists.

The first assessment instrument is from Baram-Tsabari and Lewenstein (2012), and is an instrument for assessing the writing skills of scientists as they communicate with the public. The authors point out that there is a particular “language of science” that scientists learn as they are trained in their field, but this language is not the language needed as they communicate with the public. The authors focused on counting specific instances of desired written communication to create a quantitative method of assessing writing for public audiences. For example, one measure the authors used was
the numbers of “jargon” and “non-jargon” words used in written summaries of students’ research.

Aspects of the authors’ domain analysis, or their gathering of important concepts that public science communication should address, will be useful for developing an analytic scale for public presentations by scientists. However, the method of analysis the authors used was intended for written work, and the evaluation is somewhat complicated, with results calculated using mathematical formulas. These evaluation methods would be difficult to use with oral communication due to the complexity of the calculations and specificity of the criteria (i.e., identifying particular words, counting the number of occurrences). Additionally, if using the formulas suggested in the article, it would be difficult to provide rapid formative feedback to communicators, and it could be difficult to explain the analysis metrics to the communicators being assessed.

Another innovative method proposed by Sharon and Baram-Tsabari (2013) uses computational linguistics to create a “yardstick” for measuring the use of scientific jargon in public oral presentations by scientists. Jargon is an area of concern in public science presentations, and this work is valuable background. However, the study was intended to “lay the groundwork” for evaluating jargon and its use by scientists in oral presentations, not provide an assessment tool for analyzing oral presentations. Additionally, giving a good science communication presentation involves much more than simply avoiding the use of jargon.
A 2008 paper by and Gonsalves does important work in defining some of the issues in public science communication and creates a detailed analytic rubric. The authors developed the rubric to assist in a Graduate to Kindergarten through 12th grade (GK-12) grant from the NSF. These grants put science graduate students in secondary school classrooms as science specialists. The rubric provides a useful focus on self-evaluation and peer-evaluation as well as a rubric’s use in training. Many ideas in this paper are valuable for those training scientists to interact with the public. However, this paper focuses largely on pedagogical and learning concerns, particularly for K12 students, and is therefore not as broadly applicable as necessary for the purpose of the current study. The descriptions within the rubric are somewhat lengthy as well, which is excellent for detail but makes prompt assessment and feedback more challenging.

The final rubric for science presentations is from the Florida Institute of Technology (Tankersley, Bourexis, & Kaser, n.d.) and has similarities to the Sevian and Gonsalves rubric as it was also developed as part of a GK-12 program with an emphasis on assessing scientists who are teaching K12 students. The developers called it the InStep rubric.

While a useful instrument, the Tankersley rubric is not easily accessible, since it was not published in a journal or posted on a public forum. Instead, it must be specifically requested through an email to the developers. The latest update to the website was made in 2010. Though I attempted several times, I was not able to get a
copy of the most recent version of the rubric until after my testing was complete, so I could not analyze it fully in preparing my rubric. The scale had been updated since I first acquired a copy. The scale being difficult to get from the source makes it less attractive as an assessment option for new users. The fact that the rubric is now being used in the commercial “Boot Camps for Scientists” may be a factor.

As stated, the InStep rubric was developed to assess scientists speaking in educational settings, and was then revised several times to be used in more general contexts. While there are useful constructs being measured by the InStep rubric, its theoretical foundations and specific measures fall short, particularly in the context of communication scholarship. In the 3.0 version from July 2010, the authors list eleven reference sources as the foundation for the rubric. A review of the sources used to develop the rubric reveals little reference to sources that integrate communication theory. One source is a book written by an electrical engineer, another is a well-known manual written by a public speaking coach with a BA in communication studies. That book is focused on speaking tips for business leaders. The reference list cites other similar texts about speaking and education. These works are certainly valuable, but do not reflect a purposeful study of the knowledge, skills, and abilities most important for scientists to exhibit when speaking to the public nor do they reflect an academic, theoretical foundation in communication, both of which could strengthen the rubric. Additionally, the most recent version, 4.0, of the rubric is eleven pages long (decreasing
the chances that it will be read carefully, particularly as a speaker is being assessed) with an additional one-page scoring sheet to be used as the speaker is being assessed. The scale uses measures such as “Proficient,” “High Developing,” and “Developing,” rather than numerical measures recommended by the most recent rubric research. As discussed in the prior section, numerical ratings allow comparisons over time, summaries of skills, and easy-to-see measures. Additionally, descriptive words such as “high developing” tend to have such varying connotations for each rater that they are less valuable in assessing individuals and programs and have negative implications for interrater reliability.

That said, the InStep assessment tool has admirable qualities, particularly in their acknowledgement and emphasis on the fact that assessment alone will not lead to steady communication improvement. Instead, scientists need to be trained, given opportunities to speak to public audiences regularly, and be assessed with specific, rapid feedback. As will be shown, several of the constructs have some similarity to the constructs in the Assessment for Public Presentation by Scientists, or APPS, including the relevance and language constructs. Two of the other constructs touch on similar ideas as the APPS, namely the InStep constructs of “technology” and “questions,” which show some overlap with the “visuals” and “engagement” constructs in the APPS rubric. The construct criteria do not appear to have clear theoretical support in the measure descriptions, and several lack specific descriptions of the exact skills and abilities that
will demonstrate mastery of the constructs. The InStep rubric does not reveal any
reliability testing results for the rubric or any information about the selection or
development of the constructs for either its science-based constructs or its public
speaking-focused constructs.

Now that I have reviewed the four rubrics developed specifically for science
communication evaluation, I look at a general public speaking rubric. Schreiber, Paul
and Shibley (2012) developed a well-researched, general use public speaking rubric,
which they call the PSCR, or Public Speaking Competence Rubric. This analytic scale
provides excellent support for general public speaking skills such as organization and
delivery. This universal rubric, however, does not address those constructs specific to
public science communication, a vital element in any rubric presented as an assessment
vehicle for scientists. Even if a public speaking expert were convinced that the PSCR
covered every possible speaking situation, scientists would likely not agree. Scientists
believe that their work is unique and important, and may balk at using a general rubric.
While scientists may not explicitly say these exact words, they do indicate a propensity
to consider their own judgement to be superior to others and their own profession to be
more valued than others. For example, professor Matthew Hornsey said, “People don’t
act like scientists, weighing up evidence in an even-handed way” (Healy, 2017),
implying that laypersons don’t have the same good judgement that scientists have.
Oceanographer Carl Safina (2012) said, “…Here’s the problem. Virtually no one outside
of science understands why and how any of this matters. Inside of science, hardly anyone gives it a thought, or realizes, the exceptional value that scientific thinking, not just scientific findings, would have in wider society...wield the knowledge, the value, or just the informed perspective that you have,” again implying that scientists are a special breed using better thinking and more intelligence than the average citizen.

In addition to, and perhaps connected to, this idea that scientists are special and think in different and better ways than the general public is the fact that scientists are constantly told that they need to engage in this communication with publics, but are less often told why or what they should accomplish, leaving them striving for many different outcomes when they do engage with the public (Dudo & Besley, 2016). An effective analytic scale for scientists will address the outcomes important to scientists and help provide normative guidelines for the outcomes that are most important to the science community at large. For example, a general public speaking rubric may specify that the content should address “…a topic appropriate for the audience and occasion” and “provide supporting material (including electronic and nonelectronic presentational aids) appropriate for the audience and occasion” (Schreiber et al., 2012). These general guidelines provide no normative direction helpful to scientists, leaving the scientist and the assessor having to manufacture for themselves what is appropriate. There is a great deal of literature that indicates what objectives should be met when scientists communicate with the public, as the following domain analysis will show. An analytic
scale that acknowledges the value scientists see in their way of thinking and the work they do and that defines and delineates the normative objectives for public science interaction has more chance of being accepted by scientists than a general rubric that does not acknowledge the unique place of science in society and the unique goals science communicators have. Therefore, as will be seen in forthcoming chapters, the APPS rubric combines the generally applicable delivery and organization elements of the PSCR with elements unique to public science communication.

1.6 Method for Developing a Science Communication Assessment Instrument

This dissertation creates a science communication rubric, the Assessment for Public Presentations by Scientists, or APPS, that fills the requirements of an assessment instrument for the public communication of science. Namely, it is an analytic scale, or filled-cell rubric, that is well-researched and tested. The analytic scale is detailed enough to be valuable to both the communicator and the assessor without being so detailed as to be prescriptive. Now that the type of assessment instrument best suited for public science communication assessment is identified, it is important to determine how to develop such a rubric. In this section, I will outline the method that was identified as best for developing such a rubric.

Rubrics can be developed in an unsystematic way where the assessor or a group of interested people write down several ideas about what knowledge and skills they think a person should demonstrate while completing a task. The assessor(s) then
attempt(s) use that rubric to assess those attempting to meet those objectives. As assessors see how the communicators actually perform the task, they adjust the rubric to reflect the practices of the speakers. This method allows an assessor to put a rubric together quickly, but gives questionable results that are not useful for comparison purposes. In addition, early communicators are penalized by the assessor’s incomplete understanding of the desired results.

Linguistics offers two more methodical ways of producing a rubric: the task-based method and the evidence-centered method. The first centers on the task to be done – what the communicator should be able to accomplish - in this case, the public presentation of science work by a scientist. The second method, ECD, is a construct-centered approach to test design, rather than a task-based approach. According to Mislevy and Risconscente (2005), “A construct-centered approach (to assessment design) would begin by asking what complex of knowledge, skills or other attributes should be assessed…” (p. 3). After deciding on the skills to measure, an ECD assessment designer asks what behaviors or abilities show that a person has the knowledge and skills desired, and lastly, what kinds tasks or performances a person should be able to complete that allow him or her to demonstrate those behaviors and abilities (R. Mislevy & Riconscente, 2005). The evidence centered design, or ECD, method provides the most complete method that provides the most reliable and valid testing instrument because it requires a more detailed and careful development process leading to a more reliable
result. To more fully define the concepts that will allow a person to design and implement a full test, the ECD approach also includes a structured framework that lets assessment designers use common vocabulary and create specific and explicit assessment designs that operationalize the desired constructs. Additionally, ECD is the method used by the most widely known and respected language testing service, and it is known outside language testing circles. This will give the analytic scale developed by this project more credibility outside communication circles.

This section briefly outlines what an EDC framework will look like in regard to the assessment of scientists speaking to the public in what this dissertation refers to as “library talks,” or informal presentations to intelligent but uninformed members of the general public. These presentations often take place in a library, city building or other public gathering place. This section briefly outlines each part of the ECD framework and explains how it will be undertaken for this study.

1.6.1 Domain Analysis

The purpose of the domain analysis is to discover what an assessment should measure. For example, if I wanted to assess the quality of cheesecakes, I would speak to the top pastry chefs and cheesecake lovers to determine what qualities create an ideal cheesecake. Is it texture? Is it sweetness? Is it firmness? Is it the crust? I don’t know until I speak to many experts and aggregate their responses to come up with the most important cheesecake qualities. For the domain analysis of this science presentation
assessment project, I needed to find out what knowledge, skills, and abilities, called KSAs in ECD literature, are most important for scientists to exhibit when they speak to the public. Instead of characteristics or techniques for baking, I was looking for characteristics of the best presentations by scientists to public audiences. To find out what those are, I consulted many experts, both scientists and science communication experts, to understand their perspectives on what KSAs, or demonstrations of knowledge, skills, and abilities, are most valued and most needed when scientists speak to the public.

In conducting the domain analysis, I used two analyses of data. First, I researched websites for stated goals of public science communication and/or desired speaking abilities from scientists, science communicators, and scientific organizations. Next, I searched academic articles to find out what skills scientists who engage in science communication are told they need to develop. I also evaluated the exhortations given by scientists and science communication scholars encouraging scientists to develop and demonstrate particular abilities when communicating with public audiences.

1.6.2 Domain Modeling

The purpose of the domain modeling step is to examine all of the knowledge, skills, and abilities identified through the domain analysis and determine which of them should go into the assessment. For example, if I discover that crust, texture, sweetness, and the amount of time the baker spends whipping the cream cheese are the important
qualities of the best cheesecakes, I decide which of those qualities to assess. In that case, I may decide that it is impossible to know how long the cream cheese was whipped by the chef when I am assessing the final product, so I would argue that assessing crust, texture, and sweetness would give me adequate measures to make a judgement. Therefore, the domain modeling expresses the argument for a particular assessment based on the results of the domain analysis. The domain analysis identifies the specific abilities needed for situations that are unique to public science communication and the domain modeling determines which of those abilities should be part of the assessment.

While there are those who argue that all public communication situations are essentially the same and can be assessed using common tools, the domain analysis shows specific objectives that multiple stakeholders have for public communication by scientists, such as “attracting young people to STEM careers,” and “explaining the impact of science in the lives of audience members.” While the rubric does not address every objective discovered in the domain analysis, it does identify a number of science communication specific, science-focused skills that necessitate a rubric specifically aimed at science communicators; the skills that are most crucial to the effective public communication of science.

1.6.3 Conceptual Assessment Framework

The next part of the ECD approach to creating a language assessment tool is the conceptual assessment framework, or CAF. This section includes the framework or
“blueprint” for explaining each element of the assessment. In the hypothetical cheesecake example, this is where I would discuss, for example, what makes a wonderful crust, and then describe those tasty-crust qualities in terms of assessment criteria. I would offer support for why those criteria are important, and then describe the circumstances in which the bakers would need to create their cheesecakes to be assessed. Would they be able to bake a cheesecake in their own kitchen and bring me a sample? Would I travel to their kitchens? Would I bring the bakers to a central location and have them create a cheesecake in my kitchen? These kinds of details are covered in the CAF. In other words, the CAF gives the nuts and bolts for implementing the assessment and illustrates how it will work.

The elements of the CAF are:

- **The Student Model**: How much of what the assessment hopes to measure does a student (or communicator) have? This section expresses what the assessment will measure. In the case of my project, the student model is, generally, can the speaker give an effective presentation about a science topic to a lay audience?
- **Evidence Model**: The evidence model tells us how the speaker will give us the evidence that they can do what is being asked of them. In this case, the evidence is the scientist speakers giving a brief presentation about an aspect of science that they work with, and hopefully using good science communication techniques while doing so.
- **Task Model**: This section describes exactly what information and instruction the communicators are given. In this section, the actual analytic scale is developed and tested. For this science communication assessment, APPS, it will be shown that preliminary testing helped finalize the rubric in the task model stage. The scale was pilot tested using previously recorded public presentations by PhD candidates from a variety of science and engineering fields who presented their research to public audiences.
• Assembly model: This section asks how much data is needed for a fair analysis, or how long the speakers will present and what needs to be included in order to assess speakers fairly. In this case, the presentation needs to be long enough for the scientists to demonstrate all the knowledge, skills, and abilities requested, determined to be at least 10 minutes.

• Presentation Model: This section explains how a task appears to the person being assessed. In some assessment situations, the presentation model might be true/false or essay test on a piece of paper or on the computer. In the case of the APPS, communicators are not being given a multiple-choice test or a prompt for an impromptu speech, but instead are presented with the task well in advance so they have time to prepare a presentation.

• Delivery System Model: This section expresses the methods used for assessment and the circumstances under which the rubric could be implemented. In the case of the science communication rubric, I discuss not only the situation under which the recorded speeches were given, but other circumstances under which this rubric might be used.

1.6.4 Assessment Implementation.

At this stage, all the operational elements specified in the CAF were finalized and prepared. If I were testing my hypothetical cheesecake ratings, I would ask a few chefs to bake a cheesecake and then I would test my rating system for excellent cheesecake by having at least two different raters use my assessment materials to rate the cheesecakes and see how the assessment materials worked in actual practice.

For my science communication rubric, then, this layer of the process included final testing, attempting to mimic the circumstances under which the rubric would be used. Final testing used four raters; two scientists with advanced degrees in science and two communication specialists with experience teaching public speaking and assessing speakers. The raters assessed some of the pre-recorded science presentations and rated
the presentations using the rubric and code book, giving suggestions for revisions and improvements. As part of the testing, I analyzed the reliability of the ratings using two reliability methods. The first method is the percent agreement commonly used in many fields, including language testing, which measures how often raters were in exact agreement with one another. The second measure is percent adjacent agreement, as is commonly used in language testing, although less often cited in other fields. Adjacency agreement measures how often raters were within one point of each other in their ratings. For example, if one rater awarded a score of 3 and another awarded a score of 2, they are considered to be in agreement. For comparison, I also include the stringent measure of Cohen’s kappa, which attempts to account for chance agreement in raters.

While in some fields, codes are tweaked until raters are channeled into rating instances exactly the same way, in language testing, this approach would result in a prescriptive guideline allowing little room for authenticity, as defined by Bachman and Palmer (1996). These authors say that a test’s usefulness is measured through its qualities of reliability, construct validity, authenticity (how realistic the language situation is), interactivity (how much it involves the scientist’s individual characteristics in completing the task), impact (ways the test might affect the society, the community in which the communication takes place and individual communicator), and practicality (how readily available the resources are to conduct the test). Therefore, the purpose of rubric testing in this case is not to create a strict code into which speakers are categorized
so that the results can be analyzed in order to draw broad conclusions applicable to a wide population. Instead, the testing attempts to create a useful language test by increasing test reliability so that scoring results would be consistent across different test situations with different characteristics (i.e., scientists as raters or communication specialists as raters) and construct validity (that we can interpret a score as an indicator of the abilities we want to measure). This is why, in this project, more attention is paid to simple and adjacent agreement than to Cohen’s kappa, although the kappa numbers are reported in testing results.

1.6.5 Assessment Delivery

At the assessment delivery stage, the assessment is applied and performances are evaluated. In this case, the study suggests situations in which this rubric and codebook can be utilized to greatest effect. (R. Mislevy & Haertel, 2006).

By using an ECD approach to create an analytic scale that provides quantitative data, this project produces an assessment tool for public science communication that has the potential to be broadly applicable and yet provides meaningful, actionable results. This method provides a tested, validated rubric that can be trusted by scientists and communicators alike to provide meaningful summative and formative assessments of public science communication by scientists.
In this chapter, I have stated the argument for the creation of a filled-cell rubric for the purpose of assessing scientists speaking to public audiences. I outlined the primary characteristics that an effective rubric for this purpose will have: a foundation in communication theory and practice, the applicability of the rubric for prompt, supportive formative and summative evaluations, a quantitative scale for accuracy and numerical data, and a rubric that addresses the most important knowledge, skills, and abilities that scientists should show based on a careful examination of the literature and public calls from scientists and science communication experts. I have established that the current rubrics or systems for evaluating scientists’ public communication do not include all the elements needed for an effective assessment instrument. Lastly, I have shown that a rubric for assessing public science communication should be developed using a systematic, academically sound methodology. I chose evidence centered design as the most effective method for developing the APPS rubric. In the next chapters, I will describe the development of the APPS through each step in the process of ECD, beginning with the domain analysis.
CHAPTER 2.0 DOMAIN ANALYSIS

According to Mislevy and Haertel (2006), as the first stage in assessment design, domain analysis marshals the information that will provide the grounding for assessment designs. It allows those developing the test to understand what knowledge structures are in use in a domain, what is valued knowledge and work, task features, common representational forms, and performance outcomes. These kinds of information represent what is valued in this domain by teachers, researchers, and domain experts in the field.

2.1 Method of Gathering Samples

In order to “gather substantive information about the domain of interest (scientists communicating with public audiences) that (would) have direct implications for the assessment” (R. Mislevy & Haertel, 2006, p. 4) of scientists when communicating with the public, I determined to look at what scientists are saying and being told about communicating with publics first, by scientific societies and authorities and second, by academic literature.

2.1.1 Science Societies.

I looked at two types of entities that might give guidance or advice to scientists. First, I looked at scientific societies, since these are places where scientists speak to other scientists and have conversations about their profession and their specialties. I speculate that the reasons given by peers, colleagues and superiors in the field for public science
communication would carry the most weight to scientists because scientists will recognize that their peers understand the scientific world and understand the concerns a scientist might have about communicating with the public generally. I limited the analysis to societies headquartered in the U.S., since international emphasis on science communication is slightly different than the emphasis within the U.S.

2.1.2 Government Agencies.

Second, I looked at government agencies that engage in science activities and fund science research. First, I looked at the National Science Foundation (NSF) and the National Institutes of Health (NIH). These government agencies grant funds to scientists to carry out their research; therefore, their requests carry much the same weight with scientists as the scientists’ employers’ requests do. While scientists may not approve of or agree with all NSF or NIH policies, they will likely comply with the requirements these institutions impose rather than forfeit grants from the agencies. Additionally, the NSF makes some grants specifically for projects that engage the public. Since scientists want research funds, they are likely to pay attention to what the NSF values institutionally. Finally, I looked at NASA because the organization represents the ultimate in science research to many in the United States and because it has a history of science outreach.
2.2 Method for Gathering Information

2.2.1 Method for Evaluating Science Societies.

I used the method of qualitative content analysis to gather and analyze texts that described the KSAs scientists most want to see when they are their peers communicate with public audiences. Krippendorff (2013), in his overview of qualitative content analysis, gives a guide for using the methodology in a variety of settings. The process involved the following:

1. Gathering a sample of data from a medium or a group of participants.
2. Unitizing or grouping data based on words, assertions, or proposals by using excerpts or quotes from the samples.
3. Creating a context for the data by using the researcher or coders’ own understandings about the contexts where the data were gathered.
4. Relating the findings to a research question. This was accomplished by distilling the sample texts into descriptions of the knowledge, skills and abilities, or KSAs, that could be supported by communication research and measured in a rubric.

2.2.1.1 Gather Sample of Data

To complete the first step, I needed to first identify the population of scientific societies from which I would choose the texts. I started with the list of 252 affiliates of the American Association for the Advancement of Science (AAAS) in addition to the AAAS itself. The AAAS is the largest general scientific society in the world, and therefore the largest in the United States. Its criteria for an organization to be considered for affiliate status include a size requirement (200 members) and a longevity requirement (at least 5 years of existence). This vetting process by the AAAS ensures that any organization
examined will be of significant import. The standing of the AAAs in science circles in the U.S. also make it likely that any scientific society of any standing and importance will apply for affiliate status.

To choose the population for the study, I went through the list of 252 affiliate organizations on the list. If an organization was headquartered outside of the United States, I eliminated it from the sample. Similarly, if a society focused on medical specialties, engineering, or social sciences, I eliminated it from the sample. I included societies from specialties that are considered “life” and “earth” sciences such as geology, biology, agronomy, neuroscience, physics, animal science, astronomy, and microbiology. I eliminated regional and special interest societies (such as those for members of a specific race or gender), since the goals of such societies may vary from those of a strictly science-based society. This process resulted in 30 organizations to search. Added to the federal agencies, the total was 33 organizations.

To gather texts that described the organizations’ attitudes toward and encouragements for of science communication that might have implications for assessing that communication, I first looked at the “About” page to identify any science communication-related statements. If there was no “About” page or the about page contained no text (having only photographs or visual elements), I went to the “Mission Statement” or “Purpose” page to look for the overall mission information. After examining the “About” page, I conducted three searches within the website:
1. Why communicate science
2. Public science communication
3. Public outreach.

I chose these search terms out of many possible terms, since in preliminary searches these terms seemed to return richer results than the following possibilities:
“why engage” “engage” “why public engagement” “why science communication”
“outreach” “engagement” “science communication.”

After completing the searches, I identified the total number of results for each search terms and entered the number into spreadsheet. Since searches are sorted by relevance, in order to get the most relevant returns and be best able to understand each organization’s attitude toward public science communication, I took the most relevant returns from each site, up to the top 100 for each search term.

After conducting the searches, I established a population of 6,383 search results. To determine an appropriate sample texts to examine, I calculated the size sample I would need to achieve a confidence level of 95% with a 5% margin of error. The calculation indicated that 363 results would give a representative sample of the population (Smith, 2013).

To gather the 363 samples, I utilized the random number feature in Excel to gather 363 random numbers between 1 and 6,383. I ordered the entire 6,383 population, and then located each of the individual samples as indicated by the random numbers,
saving each web page as a .pdf file, recording the URL, and assigning each page a unique identifier which indicated the organization from which the hit came and the search term that I used to collect entries from the page. I continued, collecting all the results in a spreadsheet. If Excel returned duplicate random numbers, I requested a new number (replacement). The sample texts included a wide range of artifacts, including, but not limited to, newsletters, conference programs, meeting abstracts, webpages, and white papers.

When gathering the 363 samples, if a hit was the exact duplicate of a previous result, I skipped that result and instead took the next result. If a result addressed something other than public communication of science (such as internal communication or intra-society communication) I skipped the result and moved to the next result. Additionally, if the results addressed activity other than scientists engaging with the public (for example, if the association sent a PR team or an executive to talk to legislators), I did not count that as public science communication for this study. As I examined results, if the result was inadequate, for example, simply an announcement of an event or a page of short announcements, I moved to the next sample so that I would get robust returns and understand clearly the goals science societies have for their members when they engage with the public. In this way, I gathered 363 sample texts that described the knowledge, skills, and abilities (KSAs) that are considered important by science societies.
As I evaluated the individual sources, I first looked for “should” statements – normative reasons given to scientists for communicating and engaging with the public. I then gathered other comments or implications about communicating science to the public. For example, when an article highlighted the fact that a researcher took time to create lessons for middle school teachers, I inferred that the fact that the society highlighted this kind of activity indicated that the organization was interested in scientists engaging in public communication to inform publics about science and to increase understanding of science in the highlighted way, in this case, in a formal educational setting.

2.2.1.2 Unitizing and Grouping Data

Several iterative steps went into unitizing and grouping data. After gathering the 363 samples, I entered the examples into a spreadsheet and identified all the evidences I could see of scientist/public interaction in these sample texts. I hired a coder to examine each sample independently while I examined each example as well. We had a brief training session in which I showed the coder some of the samples. We discussed our understanding of science communication and our experience with examining the concept. We then went over one quote and each determined what label or category of knowledge, skill, or ability we would place that example into. When we both felt comfortable that we each understood what kinds of directives and implications we were
looking for, we separated to complete the full analysis. For this first phase, we did not try to agree upon phrasing or categories in advance.

We assigned each sample text one or more skills or abilities that we determined were stated or implied within that text and indicated either what knowledge, skills, and abilities, or KSAs, a scientist should demonstrate when speaking to public audiences or something that the scientist should accomplish when speaking to a public audience. We each examined the texts at least twice, creating categories and re-assigning texts to different groups as we refined our understanding of the set. We met again to discuss the findings.

2.2.1.3 Creating Context for the Data

Applying our knowledge and understanding of science communication and our experience with both public presentations and science communication, the coder and I regrouped individual categories into broader categories. We reexamined, regrouped, and adjusted several times in the process of applying context, as explained in the upcoming sections.

2.2.1.4 Relate Results to the Research Questions

The last step in this part of the domain analysis was to look specifically at how the results related to the research questions. We made revisions and adjustments to the categories as needed to be sure we could relate the categories to the objectives, or
research question, of this project. In this case, the process of specifically applying results to the objectives of the study resulted in some revisions to categories and wording.

2.2.2 Method for Conducting the Literature Review

To further define the important goals for public science communication by scientists, I also completed a literature review of articles by scientists and by communication scholars who gave reasons for their public communication. To gather the sample, I searched “Communication and Mass Media Complete” and EBSCO Host’s “Communication Abstracts,” and “Google Scholar” for articles appearing between 1980 and 2016. I searched using the following 18 search terms:

1. Why communicate science
2. Why science communication
3. Why science outreach
4. Why science engagement
5. Science communication goals
6. Purpose of science communication
7. Purpose of science engagement
8. Purpose of science outreach
9. Goals of science communication
10. Goals of public science communication
11. Goals of public science outreach
12. Goals of science outreach
13. Goals of public science engagement
14. Goals of science engagement
15. Why public science engagement
16. Why public science communication
17. Why public science outreach,
18. Why public communication of science.
From the search returns, I gathered papers that addressed public science communication, gathering samples that had both descriptive and normative statements about public science communication. When I encountered ten results in a row that were irrelevant to the topic, I ended examination for that particular search term. As I added samples to the literature list, I eliminated duplicates.

2.3 Results of Analysis of Samples from Science Societies

Following the methods previously mentioned, the coder and I looked for mentions of both normative and descriptive examples of scientists engaging with publics.

2.3.1 Results of the Review of Science Society and Government Entity Websites.

Evidences of knowledge, skills, and abilities in the sample texts ranged from clear admonitions to engage in communication with the public to accolades for those who did engage in such outreach to descriptions of activities that included such interactions to job ads that included outreach as part of the job description. An example of admonitions or encouragements to interact with the public included this example from the Society of Neuroscience, which said, “Get involved in public outreach, from Brain Awareness and science teaching to advocacy” (“Join the Conversation,” 2014). One coder originally coded this sample into two categories, one they called “communicate science information and benefits to the public” and another they labeled “influence policy and
policy debate,” while the other rater designated it as belonging to a policy debate category.

An example of a description of a scientist/public interaction is found on the AAAS website, where one of the sample texts reads, “Print materials and workshops were developed to help promote understanding of both scientific facts and each other’s concerns for the benefit of both scientific advancement and the public at large” ("The Perceptions Project," 2016), which one coder placed in a “communicate science and its benefits” category and the other in a “giving scientific information to help people make decisions” category as well as a category labeled “tell the benefits of science in everyday life.”

An announcement from the Society for Integrative and Comparative Biology is indicative of another type of artifact that I discovered in my analysis – an advertisement for a workshop or class that scientists can attend to learn how to better communicate science to public audiences: "You know that your work tries to uncover basic truths about the way the world works. But explaining the importance of studying esoteric subjects like Gila Monster venom or the architecture of corals to ordinary people can be a struggle. This year’s workshop on Public Outreach, brought to you by the SICB Public Affairs Committee and the Symposium on the Morphological Diversity of Intromittent Organs, will help members of the SICB community learn about and discuss effective ways to make the value of basic and seemingly arcane research subjects clear to a broad
audience” (Final Program, 2016). One coder put this into a “communicate science and its benefits” grouping and the other into a category labeled “tell the benefits of science in everyday life.”

This job description from the Weed Science Society of America is an example of a job advertisement that includes outreach requirements: “The Director is also responsible for developing and implementing an applied research and educational outreach program” (Director, Desert Research & Extension Center and Vegetable Crop Cooperative Extension Specialist with The University of California, Division of Agriculture and Natural Resources in Holtville, CA,” 2014), which both coders placed in categories emphasizing education and outreach.

A few of the samples did not contain any references that discussed interactions between scientists and the public, while some artifacts contained numerous examples. I recorded each instance that I found within each artifact.

After the first iteration of unitizing and grouping the sample texts, one coder had 16 categories and the other coder had 20 categories. Several of these categories were quite specific, like “Develop web-based tools to communicate research results to the general public,” “Promote understanding and awareness in the public of the impacts of climate change in the community,” and “Exhibit a tone that is free from judgement and self-aggrandizement.” These categories were clearly too detailed to use in a rubric.
Therefore, the next iteration in the unitizing and grouping step of the analysis was to group these specific statements into broader categories that identified over-arching themes in the knowledge, skills, and abilities identified in the sample texts. We both then took the individual knowledge, skills, and abilities categories we had identified and developed broader categories (see details below) into which they fit. At this stage, we each had 14-15 categories (See Appendix A for full listing).

We then worked together on the next grouping iteration, discussing our findings, looking for patterns, and determining which of the 14-15 categories we had each identified were similar to each other.

We next attempted to create context for our findings based on our knowledge and experience with science communication, public speaking, and public science presentations. To do so, we worked with the categories – analyzing, discussing, and synthesizing – until we had agreed in broad terms on what we thought the basic knowledge, skills, and abilities that science societies and scientists wanted to see demonstrated through public science communication. At this point, we organized most of the findings into eight broader categories. They were:

1. Communicate science information and benefits to the public
2. Increase trust in and emphasize the human side of science and scientists
3. Audience awareness and adaptation
4. Interpret/translate science for the public
5. Increase the publics’ interest in science
6. Engage in educational outreach opportunities
7. Influence policy and policy debate
8. Provide scientific information to help people make good decisions

A handful of texts indicated knowledge, skills, or abilities that scientists should demonstrate when speaking to public audiences, and then encouraged scientists to develop particular public speaking and communication skills in order to be able to demonstrate those desired outcomes. These were grouped according to the KSAs identified at first and the public speaking aspects were not categorized. A small number of texts (fewer than five) that said the reason a scientist would communicate with public audiences was to improve her/his own public communication skills, and these texts were excluded.

Finally, we applied the findings of the analysis to the research questions, or in this case, objectives, of the study. To do so, we developed a spreadsheet indicating the coding categories we had determined and placed the coded sample texts into the appropriate category, creating a frequency chart and then comparing those results and frequencies to the objectives of the study.

Upon examining the data, there was a need to parse out further categories, as it become clear that we had grouped some of the texts into categories that were too broad so that some of the categories were identifying at least two different KSAs and some sample texts were not adequately described by the eight category headings. Therefore, the categories were again expanded to the following final categories:
1. Increase public interest in science and science topics
2. Build public understanding of science through sharing information
3. Clearly explain complex scientific concepts, processes, and ideas
4. Engage in dialogue with the public; learn from them as well
5. Share the wonder/joy/excitement of science
6. Build trust in and the reputation of scientists
7. Explain the relevance of science to the lives of the public
8. Adjust the message for the specific audience
9. Provide scientific information to make good decision/make suggestions
10. Influence policy; inform policymakers
11. Advocate for more science funding
12. Encourage young people to enter STEM fields

Making suggestions to the public as to how they should think or what they should do was originally its own category but was collapsed into number nine as the coders felt code nine encompassed both concepts, although making suggestions might be considered more of a persuasive goal, it was decided that if a scientist is going to provide scientific information to allow people to make good decisions then it was likely that a scientist would suggest that the public use the information to make those good decisions.

As shown in Table 1.1, the most frequently stated ability that scientists are encouraged to have is to increase the understanding of science/research topics and processes, and the next most commonly stated ability is to engage in a back and forth interaction with audience members.
Table 1.1. Knowledge, Skills, and Abilities Found Most Often in the Science Society Texts

<table>
<thead>
<tr>
<th>Ranking</th>
<th>Category of Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most Common</td>
<td>Increase audience understanding of science/research topics and processes</td>
</tr>
<tr>
<td>2nd Most Common</td>
<td>Engage with the audience in back and forth communication</td>
</tr>
<tr>
<td>3rd Most Common</td>
<td>Interpret and clearly explain science concepts</td>
</tr>
<tr>
<td>4th Most Common</td>
<td>Communicate the relevance and importance of science</td>
</tr>
<tr>
<td>5th Most Common</td>
<td>Provide accurate scientific information so people can make good choices</td>
</tr>
</tbody>
</table>

A preliminary study used a similar but smaller sample gathered by using the DMOZ Open Directory and searching “top science organizations.” The results were reviewed by only one rater (Murdock, 2013). The knowledge, skills, and abilities that the preliminary study identified as being desirable for scientists when they are interacting with the public were similar to those found by the more recent full study. The motivations identified by the preliminary study were: 1) informing the public about science, 2) providing information that will help people make good decisions, 3) increasing the public’s understanding of science, 4) fulfilling the moral/ethical obligation scientists have to communicate their results, 5) obtaining or maintaining funding for research, 6) communicating the joy and wonder of science, and 7) encouraging youth to enter STEM fields as a career, confirming that similar themes run through these types of texts.

2.3.2 Results of Literature Review.

To analyze the motivations for public interaction between scientists and publics as found in the academic literature, I reviewed each article identified through the
academic searches done with “Communication and Mass Media Complete” and “Google Scholar” as well as through EBSCO Host’s “Communication Abstracts.” After eliminating duplicate results and results that addressed topics other than public science communication, the sample included 62 papers. While not a comprehensive sample, it was a robust sample. Additionally, the analysis showed that similar themes were emerging from the papers and few unique results were identified after the first 30 samples. There was also a great deal of overlap between the results of the analysis of the website samples (both the preliminary and the full studies) and the results of the literature review. I will summarize some of the findings here.

In a 1985 report, the Royal Society (in England) outlined some of the goals it saw for increased communication by scientists with publics. Two of those goals included the broad-reaching goal of promoting national prosperity and increasing economic progress. It was theorized that one of the ways increased public understanding of science could reach that goal was by raising the quality of public and private decision making – as people understood more, the thought was, then they would make better policy choices at the national, local and individual level. The second goal was enriching the individual lives by “increasing pleasure and insights from science” (Royal Society, 1985). According the report, each of these goals were to be accomplished through reaching the underlying broad goal: Get the public to understand basic science better.
The assumption was that as they better understand science, members of the public would better recognize the value of science and therefore be more inclined to support scientific innovations and opinions, as noted in the discussion of PUS. Presumably people would be more willing to support scientific research if they had positive attitudes about the value of the research (Stilgoe & Wilsdon, 2009).

A later attempt at defining the goals for science communication came from Burns, O’Brien, & Stocklmayer (2003), who identified the “vowel” goals for science communication:

- **A – Awareness.** Increasing the awareness of the uninformed public about a science concept can increase scientific literacy or encourage people to engage in further learning.
- **E – Enjoyment.** A significant end goal of science communication may be a positive feeling and attitude about science at a superficial level or a deeper level of personal satisfaction coming from discovery and involvement in science.
- **I – Interest.** This may range from a voluntary engagement with science activities to a desire to become a scientist.
- **O – Opinion.** A goal for science communication is for members of the public to “change, reflect on, form, or reform” their attitudes on science and society.
- **U – Understanding.** Another goal is that the public better understand science and how it works.

This heuristic reiterates the Royal Society goal of the public’s comprehending science content, processes, and social factors. The concept is that such understanding leads to increased science literacy and engagement (Burns et al., 2003).

A recent book from the National Academy of Science identified five goals for scientists who communicate with the public, goals that are quite similar to those
identified in both website review and the literature review. The goals identified in the book are: 1) Sharing the findings and excitement of science (this goal really includes two divergent goals in one statement), 2) Increase appreciation for science as a useful way of understanding and navigating the modern world (in other words, showing that science is relevant to daily life), 3) Increase knowledge and understanding of the science related to a specific issue, 4) Influence people’s opinions, behaviors, and policy preferences, and 5) Engage with diverse groups so that their (those groups’) views about science related to important social issues can be considered in seeking solutions to societal problems that affect everyone (National Academies of Sciences, 2016). These categories are similar to the most-often-mentioned categories revealed through the domain analysis.

Another study of scientists who engage in science communication activities found that the scientists expressed seven different goals or reasons for their involvement in public engagement, several of which mirror the findings from scientific society texts, including the moral obligation to engage with the public, the desire to share the joy and wonder of science, and the desire to recruit more young people into STEM careers. Additionally, this study found that scientists engage with the public because they enjoy it, they want to counter negative stereotypes of scientists, they want to create a greater awareness of science and they want a public accountability for science (E. Jensen & Holliman, 2009).
Critics and researchers noted that a majority of these findings focus on the scientists – what scientists gain themselves in terms of enjoyment, funding, compliance from the public in accepting the scientists’ goals and understanding of the world, more science students and thus affirmation of their career choice, more influence in policy making decisions on a personal and public level, and greater acceptance in their social spheres. This emphasis on the scientists in science communication is seen as problematic by many critics, since it largely ignores one important half of the equation in the public communication of science – namely the public – and discounts opportunity for the public to communicate with scientists and have its goals met. Additionally, these goals reflect the “deficit model” in understanding the relationship between scientists and publics. This model discounts the knowledge, interest and intelligence of the public, focusing entirely on the specialized knowledge of the scientists. The deficit model is linked to what Holliman and Jensen call “first order” or “top down” interactions. The goal of this type of communication is often that the public understands and accepts the scientist’s perspective. This model of science communication, while still practiced, is panned by critics (Irwin & Wynne (Eds), 1996).

Critics propose a more interactive approach to science communication. The concept of PES (or PEST – Public Engagement with Science and Technology) is a goal articulated by other critics and practitioners round the world, as noted by several researchers (Armstrong, Payne, Deas, & Catchpole, 2013; Irwin, 2009; Russell, 2010).
Broadly stated, the goal of this kind of public communication of science is a two-way communication between scientist and public, where the scientist gains from the public’s perspective and the public gains from the scientists’ perspective through interactions, rather than from the public passively listening to the scientist.

Recent scholarly research increasingly points to the goals of PES or PEST as being the most helpful and productive goals for science communication. For example, the Center for Advancement of Informal Science Education suggested in a 2009 report that Public Engagement with Science (PES) is a worthy goal for many informal science communication encounters. The PES(T) model suggests that while publics need to learn science to participate in modern society, in science communication situations the “focus should be on the valuable perspectives and knowledge publics bring from their lives that enhance the discussions of science and issues of science-related societal issues” (Many experts, many audiences: Public engagement with science and informal science education, 2009). Many of the samples in the literature review address the perception of science communication scholars that public/scientist interactions need to be more about engagement and less about knowledge transfer. For example, the entire May 2016 issue of the journal Public Understanding of Science, which featured several articles examining the question, “In science communication, why does the idea of deficit always return?” and offered several possible answers to that question. Generally speaking, scholars encourage PES interactions rather than deficit interactions.
In recent years, John Besley and his colleagues like Anthony Dudo have done intense study of scientists and their interactions with public audiences, both from the perspective of the scientists as well as the perspective of communication experts who train or consult with scientists in their outreach efforts. For example, their 2016 study looked at scientists’ intermediate objectives as well as their long-term goals for public engagement and looked at how communication experts who train scientists understand scientists’ goals and objectives and help them reach these objectives. The authors noted that scientists often have long-term goals for interactions. Communication trainers said scientists see public engagement as integral to the process of science, they want better communication skills, they want more funding, they want to promote the value of science, they want to influence policy decisions, they felt a duty to those who fund their research, they want to be role models, and, universally, they want to increase the public’s knowledge of science (Besley et al., 2016).

Overall, a literature review showed about 20 distinct kinds of knowledge, skills, and abilities that would be beneficial for scientists’ work in communicating with public audiences. The most commonly cited were 1) increasing the public’s knowledge of science and 2) facilitating engagement between scientists and public audiences or close variations on those themes. The majority of the other skills and abilities or objectives for public science communication mentioned in the literature largely correspond with those found in the webpage analyses. These results include:
1. Sharing the excitement/wonder/joy of science
2. Correcting misinformation about science
3. Participating in a public good
4. Fulfilling a moral or professional obligation
5. Adjusting scientific messages to specific audiences
6. Speaking with clarity
7. Clearly communicating complex concepts and findings
8. Improving the public perception of science
9. Engaging in ethical communication
10. Communicating the importance and value of science
11. Giving good information so that people can make good decisions
12. Encouraging young people to pursue STEM careers
13. Raising public awareness of science
14. Becoming more effective communicators, including improving specific communication skills
15. Personal enjoyment – fulfillment from sharing with the public.

There were also a handful of unique motivations found in the literature review that were not found in the website examinations, including helping people think like scientists, teaching the public to be problem solvers, and being ethical communicators.

So, in this chapter I used two different methods for gathering data about what scientists and science communication experts want scientists to do when they communicate with public audiences: an examination of websites published by science societies and science-related government entities, and an examination of the academic literature as found on library search engines. In both of these sources, I found common themes.

The two most commonly mentioned goals from both sources for the public communication of science were: 1) increasing the public’s knowledge of science and 2)
facilitating engagement between scientists and public audiences or close variations on these themes. The five themes mentioned most often in the search were 1) increasing the audience’s understanding of science research and processes, 2) engaging with the public in different ways, 3) interpreting and explaining science, 4) communicating the relevance and importance of science, and 5) providing accurate scientific information so that people can make good choices. In the following chapter, I will explain how the knowledge, skills, and abilities emphasized in the domain analysis were operationalized for the rubric.
CHAPTER 3.0 DOMAIN MODELING

In this section of the ECD, I construct the arguments in favor of assessing the public presentations of scientists and for using ECD to do so. As Mislevy and Haertel (2006) point out, there are two approaches to the modeling when using ECD. This project takes a sociocultural approach, looking at science and science communication as it functions “within a community of practice,” and addressing the protocols and forms of communication as used in that setting. In Mislevy and Risconscente (2005), they describe the domain modeling layer of the ECD as taking the form of a narrative. It includes “…coherent descriptions of proficiencies of interest, ways of getting observations that evidence those proficiencies, and ways of arranging situations in which students can provide evidence of their proficiencies” (p. 10). In this context, I present an argument for the specific test being used to assess a scientist speaking in the public, an argument that is supported by the evidence gathered in the domain analysis. In this case, there are two foci of that argument: One, that the test and the accompanying rubric are, in fact, needed in the current science and science communication context, and two, that the test and the rubric are good representations of competency in the area of public communication by scientists. Therefore, this chapter uses the evidence gathered in the domain analysis to create an argument supporting the choice of the task to be assessed. It then identifies the knowledge, skills, and abilities that should be included in the rubric used to assess the task and gives support for the inclusion of those particular KSA’s. Lastly, this chapter
provides supports for the evidences that the speakers will display to show that they do possess the knowledge, skills, and abilities that give evidence of their competence.

3.1 The Rubric Should Aid in Scientists’ Communication Training

Much of my argument demonstrating that the test is, in fact, needed is presented in the introduction to this study. There I presented the argument that a filled-cell rubric to assess the public communication of scientists is necessary and will be an asset to both scientists who speak to public audiences and communication experts who work with these scientists. In this section, then, I outline and offer support for the goals that this specific filled-cell rubric, the APPS, is assessing. I will identify the goals I believe the rubric needs to meet and then show how this specific APPS rubric will help provide better training for scientists who communicate in the public and make the argument that such a rubric will be better accepted by scientists than other forms of assessment.

Training scientists to communicate is an increasingly important aspect of preparing both PhD candidates in science fields and established scientists working to increase their public outreach activities. Encouragement for this kind of training comes from communication scholars, science organizations, and governmental agencies, in addition to scientists themselves (Cantley, 2016; Leshner, 2007; Nisbet & Scheufele, 2009), and the numbers of science communication training programs is increasing. A filled-cell rubric will be a help to both the scientists trained in these programs and those doing the training in those programs.
3.1.1 Formative Assessment.

A high-quality, theory-based, filled-cell rubric is effective for use as a formative assessment. In fact, the assessment’s use in formative contexts, giving feedback as the scientists learn communication skills, is the most powerful benefit that this Assessment for Public Presentations by Scientists offers. It can encourage scientists by giving prompt, detailed feedback that can help increase confidence and decrease anxiety (Hunter et al., 2014). Studies show that having frequent opportunities to speak to public audiences or practice oral presentations as if they were speaking to public audiences will help scientists learn communication skills and increase their level of comfort with speaking in public (Finn, Sawyer, & Schrodt, 2009; Gray & McMaughton, 2000). Such frequent practice will also help assessors learn to quickly recognize the knowledge, skills, and abilities that reflect effective public communication by scientists.

This APPS rubric is not meant for use in high stakes contexts, such as determining entry into an academic program or deciding funding for a scholarship. Instead, it is meant to be used primarily as a teaching and evaluation tool for the benefit of scientists learning to speak to public audiences and assessors participating in the training of such scientists.

3.1.2 Summative Assessment.

Not only can a quality, filled-cell rubric be used for formative assessment, but it can also be used for summative assessment. Scientists can be evaluated at the beginning
and end of a training program to see if they have improved or to identify areas where their speaking could use further attention. A completed copy of the APPS rubric showing a scientist’s progress can be given as a “take-away” to a scientist who completes a training course, in which case it is used as a tool for continued learning and a way to direct or guide future public interactions.

3.1.3 Quantitative Analysis.

As they work within a culture that values programmatic and individual quantitative assessment (Merton, 1973), it is beneficial that scientists can assign numerical values to each level of a filled-cell rubric and then use those numerical values to create reports. Numerical reports allow scientists to demonstrate the value and efficacy of training programs and improvement or lack thereof in participants in those programs. Assessors can aggregate scores and show trends, individual improvements, areas where scientists need to improve their skills, and group trends over time. Although this Assessment for Public Presentations by Scientists (APPS) would not be appropriate to use as the entire basis for support of a training program, its quantitative nature does allow it to be used and valued as an important element of programmatic or individual assessment reports.
3.2 Arguments in Favor of Using the Library Talk as the Work Product that is the Subject of Assessment

In this portion of the argument, I argue for assessing informative presentations by scientists, which I call “library talks.” The stated purpose of this assessment is to evaluate informative public communication by scientists in the context of a “library talk” situation. Originally deriving from the “book talk” done by a librarian about a particular book, the library talk evolved as a way for the librarian to talk about aspects of the library outside of books. These talks were seen as public relations tools for librarians (Kent, 1984).

Anecdotal evidence shows that librarians and groups such as “Friends of the Library” in a variety of locations sponsor library talks addressing a broad array of subjects, such as a new book in the library’s collection, a historical event in a particular area, the collecting of historical items during revolutions, the history of crime in an area, the use of solar energy, and the importance of computer security ("Computer security topic of library talk in Bernardsville," 2013; Dixon, 2014; Mick, 2012; Stimpfle, 2013). According to descriptions of such presentations in the popular press as well as descriptions by a library scientist (Kent, 1984), the library talk generally features an author or expert of some kind making a presentation to an audience composed of interested members of the public. Libraries extend the invitation for audience members to attend such talks through public announcements at the library, posts on websites, or
messages through email lists, with audience members self-selecting to attend. These talks may be on any subject and may feature scientists, particularly if they are held at a university library or a library associated with a research facility or science museum. However, there are also similar science-oriented talks held at government buildings and other community sites. This type of talk may be sponsored by a citizen’s group, a city government, an academic association, or an interest group rather than a library.

For purposes of this study, talks that are similar to library talks in purpose and context are called library talks, even if they are not held in or sponsored by a library. In fact, there are a variety of formats used for the science library talk. One is a science café, where a scientist joins a group of interested members of the public in a café, pub, or restaurant, gives brief opening remarks, and then engages in a conversation with members of the public about a particular science topic or a range of science topics. There are also talks at science centers or zoos that often include a good deal of demonstration as well as speaking, citizen science program talks, where citizens who are involved in gathering data for a research project talk with the scientist operating the project, talks given as part of eco-tourism trips, and presentations at science festivals or competitions (Burns et al., 2003).

Library talks provide an accepted, recognized way for scientists to meet with members of the public and have interactions with them. They can also provide an opportunity for a scientist to initiate an interaction with the public, as a scientist can
approach a library (or similar organization) and ask for a time and venue to make a presentation. Such proposals are accepted and expected in a library environment. Because of its potential to reach so many members of the public at a location where the public already feels comfortable, the library talk is a form of informal science communication that is rich with possibilities for increased interaction between scientists and publics.

3.3 Arguments in Favor of the Knowledge, Skills and Abilities Included in the Assessment Instrument

In this section, I construct an argument in favor of using specific knowledge, skills, and abilities in the assessment instrument, based on the domain analysis research. First, I delineate the items identified in the domain analysis but not included in the assessment instrument and the reasons for not including these items. I then outline the knowledge, skills, and abilities (KSAs) scientists should be able to demonstrate in public presentations, based on the research done in the domain analysis. I develop claims that the proficiencies of interest when scientists speak to the public fall into six primary areas. I claim that these areas, out of all the possible categories identified in the domain analysis, are the ones that should be used in an assessment of the public communication of scientists. I then specify what particular types of evidence will be included in the assessment as a way of gauging whether or not the scientist is demonstrating the desired KSAs.
3.3.1 Arguments for Excluding Certain Knowledge, Skills and Abilities from the Assessment.

By focusing on the library talk, this proposed assessment deliberately does not address other kinds of public interactions that scientists may have. For example, one purpose or possible outcome of public communication by scientists that appeared repeatedly in the domain analysis was that of scientists influencing public policy. Such communication is deliberately persuasive in nature rather than informative. This proposed assessment, however, is not appropriate for assessing persuasive speaking by scientists, which may also include lobbying government leaders or pushing for increased or continued funding for science. While persuasion, advocacy, and lobbying are all types of public communication in which scientists might participate, that kind of communication is more frequently used in special circumstances or requested by special interest groups connected with science. While the goals of these groups may be laudable, the audience is often limited to those with the power to make funding or resource decisions and the presentations have more narrow purposes.

When a scientist is invited or asks to speak to a more general, public audience, the purposes are often more informative than persuasive. The purpose of a library talk, for example, is nearly always informative. Because the informative presentation is more broadly applicable to a wider range of audiences than is persuasive speaking and because, as I will show below, informing public is the focus of the majority of the stated
outcomes desired when scientists speak to the public, this assessment focuses on those broad-focused, informative presentations rather than more narrowly-focused, persuasive communications.

Besides advocating for public policy changes, there are a few other skills and abilities that are related to informative communication and are indicated in the domain analysis that are not included in this assessment instrument for a variety of reasons. First, the assessment does not address formal science in classroom situations, instead focusing on informal science communication where members of the public voluntarily seek opportunities to listen to and speak with scientists. There are other assessment instruments that are specifically designed to assess educational situations. This assessment also does not assess whether or not scientists give public audiences valuable information for making decisions. While there are communication measures that can indicate that information being shared by speakers may be beneficial in helping the listener make decisions, this purpose is more complicated to assess and measure than other knowledge, skills, and abilities. What’s more, this outcome is mentioned less often than other important indicators of effective public science communication. The rubric must be as easy as possible to learn to use and must be of a length that makes assessment manageable for both experienced and inexperienced assessors. For these reasons, “giving good information to help people to make decisions” was not included in the rubric as a measure of a scientist’s ability to speak effectively to public audiences.
Finally, this assessment does not address internal motivations or personal rewards scientists may get from interactions with members of the public. Some of the findings of the domain analysis indicated that scientists have intrinsic motivations for public science communication, such as “I communicate with the public because I enjoy it,” or “I feel a moral obligation to communicate the results of my research to the public that pays for the research.” While these outcomes may be important to some scientists, as they are internal measures, they cannot be assessed by an outside observer. Therefore, I mention these items as being important motivations for some scientists to engage in public outreach, but do not include those measures in the rubric itself.

3.3.2 Knowledge, Skills, and Abilities Included in the Assessment of Scientists Speaking to Public Audiences.

While all of the abilities, goals, and outcomes identified in the domain analysis are interesting and many are important in the public communication of science, I focus the assessment instrument on those outcomes that the domain analysis noted as being most commonly cited by scientists, scholars, and communication experts, those that are mentioned most often by individuals who train science communicators, and those which, based on communication theory and research, are most effectively operationalized for a rubric. Using these criteria originally yielded five categories, or constructs, for inclusion in the assessment. The five are summarized in the list below, and then the following paragraphs explain each criterion. Some of these definitions and
explanations were refined during the testing process, so the following paragraphs generally reflect the initial understanding of the measure and later paragraphs reflect the updated understanding.

Criteria for the Science Communication Rubric: Scientists will be able to...

1. Connect science to the audience so that it is relevant and significant to audience members
2. Communicate complex science ideas clearly so they can be understood
3. Help audience members understand science and the processes of “doing” science
4. Portray scientists as “human,” good, and trustworthy individuals
5. Interact with the public through conversation and dialogue.

3.3.2.1 Scientists speaking to public audiences will be able to connect science information to the specific audience to which they are speaking.

The ability to relate to their specific audience is a commonly assessed ability in public speaking. However, scientists have a particularly difficult job when it comes to this skill. Broadly, a general public speaking assessment looks for the speaker to create what is called “identification,” or some connection between the speaker and the audience (K. Burke, 1969). This helps the audience member feel that the speaker is similar to them and understands their needs, interests, and concerns.

The difficulty of this task is heightened when the speaker is a scientist, however. From the outset, the audience senses a separation between themselves and the scientist. The audience knows that that the scientist has years of schooling and research experience, making the scientist very much unlike the audience member in many ways.
Therefore, the scientist as speaker needs to show the audience that he or she is like the audience member in at least some ways, and that the speaker 1) understands the audience’s level of expertise on the subject, 2) can help the audience understand this science topic as being relevant and important to the audience members’ lives, and 3) can help the audience see science as interesting, appealing and exciting.

The perception of difference is wider between the public and scientists than it is with experts from other fields. Americans hold scientist in high esteem, with 79 percent saying science has a positive effect on society and 70 percent saying that government investments in engineering and technology and basic science usually pay off in the long run (Funk & Rainey, 2015). Another study showed that over 90 percent of Americans believe scientists are “helping to solve challenging problems” and are “dedicated people who work for the good of humanity” (Gannon, 2014). However, this respect for scientists and these beliefs can lead members of the public to a feeling of separation from science, of standing in awe of science and scientists and of being intimidated by science (Jacobs, 2011). Additionally, the same Pew Research study that notes how respected science and scientists are in the public perception additionally notes that there are significant differences in the public’s and scientists’ opinions on science topics, such as whether or not it is safe to eat genetically modified foods (Yes: 37%, public, 88% scientists) and build nuclear power plants (Yes 45% public, 65% scientists) (Funk & Rainey, 2015). In order to bring those opinions closer together, scientists need to understand the public and why
they feel they way they do, and then find areas of common ground so that they can reason together and create the sense that the topics scientists study are relevant to the lives of the public. Therefore, because of the perceived and real distance between scientists and lay publics, the scientist must make a particular effort to close the gap between expert and lay person and clearly focus the message to the specific audience to which she or he is speaking. The focus of this aspect of the rubric is not to simply create a connection in the way that every public speaker should try to do, but instead, to create a connection between the science that the scientist does and the members of the public. The speaker is trying to bridge the gap between her or his view of the importance of the research (s)he is doing and the view of the audience member who may not see much value in counting the number of ants who live under a particular kind of plant, for example. Therefore, the primary focus of this construct is connecting the importance of the science topic to the lay audience.

**3.3.2.2 Scientists speaking to public audiences will be able to communicate complex scientific ideas and make them clear to the public.**

Science can be complex and difficult for non-experts to understand. In fact, as earth scientist and science communicator Roger Aines points out, many magnificently intelligent people may know little about science (Aines, 2016). Additionally, even scientists often do not understand the work of other scientists outside their area of expertise. Each specialized area of science has its own jargon and “paradigm,” as
Thomas Kuhn suggested, and scientists learn their own language as well as laws and theories that provide an explanation of the world that is understood by those initiated into that paradigm (Kuhn, 1970). Aines said, “It may be funny to joke about ‘drinking from a fire hose’ when a presentation is incredibly dense, and we scientists enjoy the challenge of absorbing information at a high rate. But that metaphor is entirely too apt when applied to a non-scientist. You can’t drink from a fire hose – almost all the water escapes you, even if you get some.” He pointed out that scientists who fail to explain ideas clearly for an audience that is uninitiated into the paradigm will not be heard. (Aines, 2016).

In the domain analysis, I discovered that scientists and science organizations frequently say that explaining complex ideas clearly should be an outcome of public communication by scientists. On one website, the American Geophysical Union talked about their how their Sharing Science Program emphasizes eliminating jargon from the vocabulary of scientists speaking to the public, noting that words that have a specific meaning to scientists may have another meaning to members of the public (Hanlon, 2016). Another of its webpages holds up as an example of clear, helpful science communication some articles that are “clearly written” with helpful background information and colorful, well-chosen graphics (Union, n.d.). By applying communication theory, science metaphor research, and PowerPoint research, it is
possible to assist scientists in communicating complex science clearly and in a way that is accessible and understandable to members of the public.

3.3.2.3 Scientists speaking to public audiences will be able to increase their audience’s understanding of science and science processes.

While the previous criterion primarily addresses language use, this criterion looks at the ability of the scientist to help a member of the public understand more about the processes of “doing” science. More than just simplifying language or explaining her or his own research clearly, it is important that the scientist add to the public’s understanding of the methods, practices and processes of scientific experiments. In fact, the most often mentioned desired outcome or skill that the domain analysis revealed was that scientists share information about science and the processes of science with public audiences.

There are many critiques of the basic sharing of science information. Some scholars say that simply sharing information, which, as mentioned previously, is called a “deficit” interaction, is not an effective way of engaging the public’s interest in science or even increasing the public’s understanding (Nisbet & Scheufele, 2009; B. Wynne, 2006). However, as evidenced in the domain analysis, large numbers of scientists remain convinced that more and better explanations of science to the public will help scientists accomplish their other goals (Davies, 2008). And the fact remains that scientists do know more about science processes and outcomes than the non-scientist, general public.
Therefore, even interactions with scientists that emphasize the “public engagement with science” that is more popular among science communication scholars (E. Jensen & Holliman, 2009) need to start with the public getting some information they didn’t already have. As William Bodmer, who was instrumental in promoting science communication during its years of rapid growth in the 1980s, argues, engagement (with science) can’t come without some kind of understanding of the science being discussed (2010).

Throughout the domain analysis, scientists and scholars argued for better understanding of science and science processes. For instance, when a scientist hesitates to claim that something is a “fact,” some members of the public believe an issue still under debate, when in fact, the culture of science is to always acknowledge there are unknowns, even when a question is generally settled (Ceccarelli, 2011). Many scientists believe scientists in general have a responsibility to make this kind of information clear to publics (Fischhoff, 2007), and believe that being open about the processes and values that underlie science will increase public trust in scientists (Irwin, 2009). Some scientists also believe that scientists should be actively involved in correcting misinterpretations of science and in being honest about what scientific and technical advances can and cannot do (Lackey, 2007). Other scholars claim that “stripping away” the air of mystique that surrounds science and helping members of the public see it as a difficult, messy process
will help the public respect scientists more as people and also understand better the way science changes and develops (Fabj & Sobnosky, 1995).

In the domain analysis, I found comments such as this one from the American Water Resource Association “We found that our focus groups were relatively uneducated about science in general and about environmental issues in particular. Once their awareness of those issues was raised, however, their interest also appeared to increase” (Halverson & Burton-Radzely, 1999) and this from the National Speleological Society: “Instead, let it do what it is best equipped to do: bring science to the non-scientific public” (Palmer, 1996). Others argue that the public needs a better understanding of both statistics and risk (Bodmer, 2010) and that the public should better understand the place of uncertainty in science.

For these reasons, communicating information about science processes, choices made by scientists, uncertainties of scientific findings, results of studies, and descriptions of the scientific processes (and how the actual process of “doing” science may differ from the “scientific process” taught in K-12 situations), among other topics, are important tasks in which scientist should engage. An effective rubric for the assessment of public science communication should evaluate how well scientists share these things with their audiences.
3.3.2.4 Scientists speaking to public audiences will be able to humanize scientists and help them seem trustworthy and knowledgeable.

While Colquhoun (2009) suggests that scientists blog rather than speak orally to public audiences, he believes, as do many of those whose work was reviewed for the domain analysis, that having scientists communicate directly with the public rather than through intermediaries, such as paid corporate or university communication professionals, will lead to increased trust in and a sense of the humanity of scientists. While his general positivity hearkens to the myth that “all communication is good communication,” there is, in fact, research that shows that the right kinds of communication can promote trust, connection, and empathy between communicators.

To further illustrate the need for scientists to build trust with publics, a study in 2005 found that teenagers viewed scientists as valuable, but “not like them,” and “not normal and attractive men and women.” When asked to sketch a scientist, most drew a person with a headful of crazy, white hair, lab coat, and thick glasses, as shown in a 2006 study (Rosenberg). Since the 1960’s, the public’s perceptions of scientists have been that scientists are difficult to comprehend and odd (Rosenberg, 2006).

Therefore, although scientists enjoy respect, they do not necessarily enjoy trust and credibility as persons who have a good understanding of the world the “rest of us” live in. This construct attempts to identify ways scientists can increase public’s trust in
scientists and their perceptions that scientists are “normal” people who are personable and have the well-being of others in mind.

3.3.2.5 Scientists speaking to public audiences will be able to engage audiences in interactions and conversations about science.

The scholarly conversation surrounding the public communication of science has been dominated by the concept of public engagement with science as opposed to the public understanding of science for a number of years. The public understanding of science (PUS) model of sharing science information is known as the deficit model: the public has a deficit of knowledge; the scientists fill that deficit. This model discounts the knowledge, interest and intelligence of the public, focusing entirely on the specialized knowledge of the scientists. The deficit model is linked to what Holliman and Jensen call “first order” or “top down” interactions (2009). The goal of this type of communication is often that the public understand and accept the scientist’s perspective. This model of science communication, while still practiced, is panned by critics (Irwin & Wynne (Eds), 1996).

 Critics propose a more interactive approach to science communication. The concept of PES (or PEST – Public Engagement with Science and Technology) is a goal articulated by critics and practitioners round the world, as noted by several researchers (Armstrong et al., 2013; Irwin, 2009; Russell, 2010). Broadly stated, the goal of this kind of public communication of science is a two-way communication between scientist and
public, where the scientist gains from the public’s perspective and the public gains from scientists’ perspectives through interactions, as opposed to the public passively listening to the scientist. Recent research increasingly points to the goals of PES or PEST as being the most helpful and productive goals for science communication. For example, the Center for Advancement of Informal Science Education suggested in a 2009 report that public engagement with science is a worthy goal for many informal science communication encounters. The PES(T) model suggests that while publics need to learn science to participate in modern society, in science communication situations the “focus should be on the valuable perspectives and knowledge publics bring from their lives that enhance the discussions of science and issues of science-related societal issues” (Many experts, many audiences: Public engagement with science and informal science education, 2009). Meanwhile, the worthy goals that scientists have for PUS science communication may also be reached through PES(T) science communication.

Different researchers have categorized interactions between scientists and the public differently. Jensen and Holliman (2009) offer a model similar to those proposed by other researchers, dividing the interactions into First, Second and Third order interactions. First Order interactions follow the PUS or deficit model, with a member of the public interacting with the scientists in a way that privileges the scientists and maintains his or her position of power. This might be a standard question/answer interaction where the member of the public asks a question and the scientist answers the
question from the position of expertise, or it might involve a scientist inviting a member of the audience to participate in an activity or an experiment by holding, pouring, touching or throwing something. A second-order model envisions a discussion, a two-way interaction between scientists and the public where the two parties have more of a symmetrical relationship, an interaction that would be considered PES(T). This kind of engagement requires more accountability on the part of the scientists and the public, and operates on more of a consensus basis than first order interactions, with scientists not necessarily being granted privileged status automatically, although such privileging may occur during the interaction (E. Jensen & Holliman, 2009). On the other hand, third order interactions involve scientists and publics engaged in a deliberation and debate, together setting the agenda for discussion. When interactions are third order, there is not only input from the public and interaction with the scientist, but there is also disagreement and critiques from the public that are accepted and processed as valid by scientists. Jensen and Holliman say that rather than seeing them as problematic or threatening, scientists who engage in third order interactions find disagreement and critical discussion on the social implications of science as being “societal resources to be valued” p. 38). An example of second and third order communication took place in the U.K., for example, when the government organized deliberative forums to discuss genetically modified foods and their place in the food chain and organized sessions to deliberate about nanotechnology and how it would be used in society. The participants
for these groups were carefully chosen and invited to attend the discussions. Participants included members of the public from a range of demographic backgrounds as well as scientists and other specialists (Irwin, 2009).

Despite the emphasis from researchers, scholars, and some government officials on second and sometimes third order communication, there is still a preponderance of deficit model communication taking place during science communication interactions, including many informal science interactions (Davies, 2008; Russell, 2010). However, the domain analysis showed that not just researchers but also scientists and science organizations value these interactions when scientists communicate with the public. Therefore, this rubric includes a construct that measures the ability of scientists to interact with members of the public as an aspect of public science communication that should be assessed. However, due to the nature of the data available for analysis, this outcome cannot be evaluated in this study, and therefore is not included in the results section.

3.4 Arguments in Favor of Evidence Used for Supporting the Identified KSAs

Here I outline the observable variables that will be assessed in order to show a scientist’s knowledge, skills, and/or abilities in each of the five areas to be assessed for the communicating science portion of the rubric.
3.4.1 Scientists speaking to public audiences will be able to explain the relevance and importance of science information to the specific audience to which they are speaking.

To show KSAs in this area, speakers will include elements that help the audience members see the relevance of the science topic to their own lives. To do so, the speaker will emphasize values, goals, and/or experiences that the scientist speaker and the audience member may have in common, as suggested by Burke (1969). The speaker should call attention to the areas (s)he has in common with audience members, not just as a person, but as a scientist (Larson, 2012). Scientists will describe how and why the science they pursue is relevant to the audience to whom they are speaking, and encourage agreement about the joy and excitement of science through relating that excitement to the values of the listeners. As theorized by communication scholars, the scientist will avoid areas of disagreement, particularly at the beginning of the presentation, and instead focus on how the scientist’s point of view is “consistent with what they (the audience members) believe” (Lucas, 2012).

3.4.2 Scientists speaking to public audiences will use concrete, direct language and analogies to communicate complex scientific ideas and make them clear to the public.

Certainly, scientists speaking to public audiences need to use clear, understandable language that is free from complex vocabulary or “jargon.” However, they can use additional techniques to help their language be understandable to their audiences. For example, studies such as those completed by Knudsen (2003), Boyd (1993) and Cat (2001) show that metaphors, comparisons, and analogies are good ways to
explain complex scientific principles to lay publics. Metaphor in the linguistic tradition focuses on the Theory of Cognitive Linguistics, which has developed over the past approximately three decades. This theory sees metaphor as something that “permeates daily conventional language” (White, 2003, p. 132) and infuses all conversation and language. Lakoff and Johnson (1980), for example, note that the metaphor “argument is war” is so deeply embedded in our modern culture that we habitually use war metaphors to describe arguments and what we do when arguing (such as “winning” a point, “attacking” a weak point, or “retreating” from an unsuccessful argument) and we have difficulty thinking of argument in any other terms (for example, a negotiation or a collaboration).

Cat (2001) notes that there has been some historical bias against metaphor in science communication, but scientists now are often encouraged to use metaphor to explain difficult concepts. In one example of this work, Boyd (1993) proposes that scientific metaphors should be divided into two different types of metaphors. One type is generative or theory-constructive metaphors that are generally used in scientific discourse within the science community and which cannot be paraphrased because there is no other way to talk about a particular phenomenon (such as “the genetic code.”). There are also pedagogical or exegetical metaphors such as “messenger RNA” used to explain or illustrate a “scientific phenomenon for which a perfectly adequate, alternative original expression exists” (p. 485).
While Knudsen disagrees with Boyd’s distinct delineation between the two types of scientific metaphor, she does say that the development of metaphor in scientific situations differs from the application of and development of metaphors in everyday situations. In fact, Knudsen’s own work (2003) expands the research into science metaphors to examine how metaphor is used in communication to public, non-expert audiences by examining metaphors in science journal articles and comparing them to articles in a science magazine for lay readers. She shows that the metaphors in science move back and forth between theory-building and pedagogical so that the strict delineations Boyd suggests are impossible to make. Nevertheless, she says, metaphors can have strong explanatory power.

Typically, the metaphors scientists consciously use to explain their work to lay audiences are pedagogical metaphors. Scientists use them to compare a scientific concept or process unknown to the audience with something that is known to the audience. However, scientists may also use cognitive metaphors – those metaphors that are unconscious and permeate regular conversation and speech. For a scientist, these subconscious metaphors are different than those of a lay person, since the scientists’ cognitive structure and experience differs from that of a lay person. For example, a neuroscientist might describe the brain as a vessel or as an independent agent (“his brain turned on him”) where such descriptions may not make sense to a lay person who has not studied the brain in the same way or with the same assumptions (Knudsen, 2003).
Studies show that the use of deliberate, pedagogical or teaching metaphors is more effective for clarifying complex ideas than are the cognitive or subconsciously-used theory-building metaphors, and show that the use of one consistent metaphor is more effective than the use of multiple metaphors. For example, I assessed audience reaction to some of the presentations made by scientists which were assessed for this study. In one of these presentations, a scientist speaking about polymers called them “chains” “strings of beads” “building blocks” “networks” and “systems” – five different metaphors for the same item in a single presentation – and also personified the polymers by saying the molecules in a polymer “liked” or “didn’t like” one another. Audience feedback forms showed that audience members found that the speaker who used these varied metaphors was unclear. Conversely, a speaker who used one metaphor consistently throughout a presentation was given high scores in clarity, and, on feedback forms, audience members mentioned the single metaphor as helping with clarity. Therefore, the rubric instructs assessors to listen for comparisons, and instructs assessors to pay attention to deliberate analogies or metaphors that are clearly meant for teaching.

Another way scientists can add clarity to presentations is by using helpful visuals, and this assessment looks for visuals that add clarity to a presentation. PowerPoint presentations are now de rigueur for science presentations, and scientists can use PowerPoint effectively for lay audiences as well. However, scientists will need to use different techniques than they normally use for scientific presentations in order to create
PowerPoint slides that are effective for public audiences. Research shows that scientific presentations, particularly scientific presentations to the public, can be confusing to the audience. The speaker needs to make explicit assertions and explain things clearly, both orally and visually, if they are to help audiences understand complex ideas (Alley, Schreiber, Ramsdell, & Muffo, 2006). So, scientists using PowerPoint should consider not only scientific principles but also design principles when creating presentations, particularly when creating presentations meant to increase understanding in public audiences. Research by Tufte (2003), Doumont (2004), Alley & Neeley (2005), Mackiewicz (2007, 2008), and Durso, Vlad, Burnett & Stearman (2011) does not always agree on details, but does agree on broad, overall suggestions about clarity in scientific visuals.

First, visuals accompanying oral communication should differ from those for written communication. Slides should not serve as speaker notes. Ideally, slides addressing technical topics should contain a short sentence or two (no more) and a visual element that contains a “visual argument,” or support for the text (Alley & Neeley, 2005; Alley et al., 2006; Gross & Harmon, 2009). One study specifically found that visuals that are highly integrated with the text got more attention and were remembered longer than those that were decorative or least integrated with the text (Slykhuis, Wiebe, & Annetta, 2005). There should be no other text on a slide (including lists or bullet points) unless necessary to support the visual. The “headline” sentences should be written in active voice, using a positive, rather than negative, tone. As a whole, the slides should support an overall message, with at least one article suggesting
that presentations use narrative organization. While I do not suggest that every scientific presentation to a public audience should use a narrative pattern of organization, it is ideal for the slides to have a unified feel with an overarching theme to clarify and support the oral presentation.

Any charts or graphs should be simplified and the detailed labels and tick marks removed. Scatter plots are confusing to audiences and should be avoided, as should stacked bar charts and three-dimensional charts. Simple bar charts and line graphs are easily understood by lay audiences (Tufte, 2003).

Fonts can be serif or sans serif, but need to have a professional appearance. Gil Sans and Souvenir Lt. are top fonts for clarity and professionalism, as are Tahoma, Arial and Verdana. Fonts should be no smaller than 22 points for text and 16 or 18 points for references, legends and labels (Durso et al., 2011; Mackiewicz, 2008). Color and animation (outside of color photos) should be used sparingly and only for effect. There should be a good contrast between the background and the text on the slide, but speakers should avoid the red/green color combination, 3-D graphics, and excessive shapes and colors on a single slide or throughout the presentations (Durso et al., 2011).

These recommendations may contradict advice scientists are given by fellow scientists, but are supported by communication and technical communication research. Scientists speaking to public audiences should keep in mind that the visuals for a public
audience need to be constructed differently than visuals for a scientific audience, particularly one made up of fellow experts in the field. Standard advice for good PowerPoint construction also applies, such as using a consistent color and design theme throughout the presentation, using limited numbers of fonts and colors (speakers should use the same two or three fonts and colors throughout the presentation) and making the visuals simple, legible, and interesting.

3.4.3 Scientists speaking to public audiences will explain specific science processes and techniques to their audiences.

Many scientists believe that the public should have a better understanding of how science works and understand better how scientists do their jobs and what assumptions guide their thinking. A better understanding of how science works will, they believe, lead to more appreciation for science and scientists and more respect for scientific findings and advice from scientists. To give the public a better understanding of how science works, scientists should give basic explanations of science methods, assumptions, and processes. The goal of these explanations is not to give the audience every detail of planning and decision making – that much information would be overwhelming and unnecessary. Instead, scientist speakers should give explanations appropriate to the audience about how a specific principle operates or how the scientist gathers data in their work. As revealed in the domain analysis, some of the primary themes scientists believe public audiences should better understand are the creative,
difficult, incremental way the work of science is actually done versus the neat “scientific method” non-scientists learn in school, the team nature of scientific discovery, the methods scientists use for gathering data, and the role of uncertainty and risk in science (Hilgartner, 1990; C. R. Miller, 2003).

For example, a scientist talking about the theory upon which they base their work could explain that a theory is not just a guess, but instead is a logical attempt to explain observed phenomenon and predict outcomes. Theories may be confirmed, and if they are, they become important tools for scientists. A scientist could explain briefly how the theory (s)he uses developed over time and perhaps how an older theory was superseded by the current more accurate or comprehensive theory (Siegfried, 2014).

A recurring theme coming from scientists is that the public should understand that scientific discovery or movement does not happen in a bubble – the image of the isolated scientist working isolated in the lab into the night is a largely inaccurate one. Therefore, scientists speaking to public audiences should acknowledge the contributions of other scientists whose work they used to build their own research, they should point out the help of the lab techs and field techs, recognize the statisticians who help with data analysis, and acknowledge the other scientists who have expertise in various aspects of the project. Simply mentioning the other members of a group that assisted the scientist when gathering data, showing a photo of several scientists working together on
an experiment, or discussing how scientists talk together about how to solve thorny problems indicates to a public audience that science is a team effort.

Scientists can also tell an audience about how they gather data. For example, the Archeological Institute of America (AIA) suggests that scientists show and/or explain their methods of gathering data through bringing actual data gathering instruments to a presentation or showing photos and videos of the process. Watching or seeing photos of data gathering is engaging to audiences and helps them understand how science work is actually done (Maskas, 2014). Understanding more about how scientists gather and record data makes the science seem less mysterious and more concrete to the audience.

Similarly, brief explanations of how scientists make decisions can help public audiences understand science as an incremental process. Talking about choices such as deciding what questions to explore, what data to gather, and what materials to use to gather them helps audiences be aware that first, there are many different options available to a scientist making these decisions and that many different decisions need to be made wisely at each step, and second, that each of the decisions have consequences for how the science is carried out and what information the project will gather.

Lastly, helping audiences have a realistic view of uncertainty and risk is considered important to the public communication of science. Scientists can explain how certain a particular concept is or how much there is to learn. Scientists by nature tend to
hedge on saying anything is “known” or “certain,” since additional information could always come to light, but scholars recommend that when talking to a public audience, scientists be clear about which approaches and understandings enjoy nearly universal agreement and which are still in the initial stages of explanation. Similarly, scientists often want to present all possibilities that could occur, but publics may not understand some of those possibilities as being remote or unlikely. Scientists should use layperson’s terms to make the actual levels of risk more apparent.

So, scientists speaking to public audiences should talk about the decision-making processes they go through and some of the choices they make to help break down some of the mystery of science, talk about the actual processes of setting up an experiment or study and gathering data, talk about the teams with which they work, and be more direct in talking about what things are more certain and less certain, more of a risk and less so.

3.4.4 Scientists speaking to public audiences will use techniques such as self-disclosure and immediacy to help scientists seem trustworthy and personable.

A significant body of communication research shows that public audiences are much more willing to listen to and accept the arguments of people that they deem as credible or trustworthy. In addition, the public is more trusting of people who seem more “real,” “human,” and similar to themselves. This trustworthiness boils down to the communication concept of credibility – judgments made by the perceiver (or recipient of
a message) that the communicator is believable. The elements of credibility as outlined in communication research are competence, trustworthiness, dynamism or charisma, and composure (Benoit & Strathman, 2004). Dynamism and composure are encompassed in the delivery construct in the public speaking portion of the rubric. However, specifically building trust and connection to an audience is of special concern to scientists, and numerous studies show that there are specific things a speaker can do to increase these elements of credibility.

One effective method of building trust and increasing personal appeal is to engage in self-disclosure. Self-disclosure is the voluntary revealing of any personal information to someone who would not otherwise know this information. Such revelations lead to connection, increased trust, and the perception of closeness in a relationship (Wheeless, 1978; Wheeless & Grotz, 1977). Revelations that show a speaker to be vulnerable or fallible or that seem counter to the speaker’s self-interest are particularly effective in building a closeness between speaker and listener (Pratkanis & Aronson, 2001; Reis & Patrick, 1996) Self-disclosure may involve a scientist talking about his or her family, but for this construct of building trust between a scientist and the audience, the self-disclosure should involve the scientist talking about his/her science work. For example, the scientist might tell a story about an event that happened in the lab. One scientist talked about how surprised he and his partners were at how sticky a substance that they developed was. This disclosure made him seem more human and
“real” to the audience while simultaneously expressing the excitement of discovery aspect of science and making him seem fallible. All of these things – self disclosure, excitement, and fallibility – can increase the connection between the speaker and the audience. Too much or inappropriate self-disclosure, however, can have a negative effect, so speakers need to be aware of situation and audience and disclose in ways that are appropriate for the context.

Another ability scientists can use that increases their “humanity” and creates a sense of connection between them and their audience is to make use of inclusive pronouns rather than first person pronouns. If a speaker uses terms such as we, our, and us, they are perceived as more inclusive, more connected, and more trustworthy. Using first person pronouns such as I, me, and my gives the perception of arrogance, of individualism, and of distance. (Dreyer, Dreyer, & Davis, 1987; Fitzsimons & Kay, 2004).

Additionally, scientists can use concrete, specific, and unambiguous language in their presentations rather than using abstract terms. When a speaker uses concrete words that the audience can understand quickly and easily, the audience perceives the speaker as truthful. The audience is also able to create clear mental images more easily when listening to concrete language. Psychologically, when something is easier to imagine, then more truthful it seems to be (Hansen & Wanke, 2010).
Communication supplies another measure that can increase perceptions of trustworthiness: the principle of immediacy (Baringer & McCrosky, 2000). A principle often invoked in instructional communication when seeking to increase the connection between student and teacher, immediacy identifies several behaviors that tend to increase the sense of psychological availability, warmth, and closeness between people. In the context of the public communication of science, the measures that are particularly applicable include the use of appropriate humor, moving physically closer to the audience, smiling at the audience, and looking directly at the audience (Richmond, McCroskey, & Johnson, 2003). While some of these behaviors may also contribute to the “nonverbal communication” elements of the public speaking score, when considering the goals of science communicators, these immediacy measures also contribute to the trustworthiness and approachability of the speakers as scientists, and are important to assess as an element of science communication.

Therefore, using communication skills and abilities such as appropriate self-disclosure, inclusive pronouns, and concrete language can accomplish the goal of scientists of seeming more human, approachable, and trustworthy.

3.4.5 Scientists speaking to public audiences will engage in discussion and interactions with their audiences.

Numerous scientists and science scholars advocate scientists engaging with the public in ways that involve conversations, questions, and mutual learning – both the
public learning from the scientists and the scientists learning from members of the public. Wynne (1992) proposes that rather than considering themselves the “experts,” scientists should learn from the lay people who are involved with science-related situations, trusting their social networks, relationships, and identities as much as the scientists trust their science. Other researchers encourage scientists to find ways to share the traditional authority of science with the public by inviting them to make decisions about what is discussed. Davies (2008) notes that scientists may believe that audiences are not interested in discussion and engagement, and believe that many members of the public like the “whiz-bang” aspects of science without the boring facts. However, he argues that scientists need to think less about one-way methods of engagement and more about multi-way, context-dependent debate where scientists engage with publics. Therefore, scientists engaging with publics should deliberately encourage the audience to speak, to question the speaker, and to engage in dialogue.

Dudo and Besley (2016) point out that scientists participating in public communication are most driven to engage with the public to defend science from misinformation and educate the public about science rather than to build trust or establish relationships with publics. Therefore, scientists speaking to the public should quell any defensive instincts if audience members question them and attempt to give useful information while listening to the audience and acknowledging their experiences and understandings. Scientists should avoid the urge to sound authoritarian, as though
they are the final word on any question, and instead use techniques such as asking the audience what they understand about a topic or what experiences they have had before and then share the scientific perspective on the topic.

Nisbet and Scheufele (2009) suggest that speakers get to know as much about their audiences as possible prior to the event so that the scientist can engage with the audience based on systematic, empirical understanding of the audience’s values, knowledge, and attitudes, among other things.

Scientists are encouraged to have the audience answer questions, through raise of hands or voice responses, during a presentation as well as after. Speakers can ask for brief personal experiences from audience members, have an audience member participate in a brief demonstration, or have an audience member touch and describe an artifact for the rest of the audience. These are just a few of the suggested ways a speaker might engage an audience during a presentation. The speaker should, of course, engage in a question and answer period after the presentation as well. Although situations can arise that are challenging to handle, such as an audience member going on too long about a personal experience or an audience that lapses into a prolonged silence rather than choosing to ask questions, the opportunity for the audience to interact with the speaker is one of the primary objectives of science communication opportunities (Many experts, many audiences: Public engagement with science and informal science education, 2009), and practice and training will help speakers manage audience interactions.
To reiterate, the engagement aspect of public communication by scientists is not specifically assessed in this study because the data available to us for testing did not include the question and answer periods at the end of the presentations, so this component of the rubric should be tested and modified as it is used in the future. However, since this element is so important to public science communication, it needs to be included in this rubric.

So, in this chapter I have established the characteristics that a rubric should have in order to help with the communication training of scientists, including that the rubric should work well for both formative and summative assessments of a scientist’s public communication and that a rubric should offer quantitative data. I presented the argument in favor of using the “library talk” variety of presentation that should be assessed by this rubric, and I then argued in favor of omitting some knowledge, skills, and abilities from the rubric, while including the following abilities:

1. Explaining the relevance of the science topic to the audience.
2. Using concrete, direct language and analogies or comparisons, when appropriate, to help audience members understand complex science concepts.
3. Explaining processes, methods, and assumptions about science.
4. Using communication techniques such as self-disclosure and immediacy to build trust and personal connections with the audience.
5. Engaging in discussion and interaction with members of the audience.

Now that I have established the specific knowledge, skills, and abilities to be measured by the rubric, I will show how those will be operationalized for the rubric.
This takes place in the CAF, or conceptual assessment framework, in the following chapter.
CHAPTER 4.0 THE CONCEPTUAL ASSESSMENT FRAMEWORK

The conceptual assessment framework is the third layer of the ECD assessment method. This is the section in which the justification for and argument in favor of the assessment moves from a narrative approach found in the domain modeling layer into the practical matters of implementing a test and creating the materials that will be used. The variables that may exist in the test, the task(s) to be performed, and the scoring mechanisms to be used are all finalized. (R. Mislevy & Riconscente, 2005). In this chapter, then, I will describe the student model, or the five skills that scientists will be expected to display. I will then describe the evidence model, or how the scientists will be assessed – what evidence they will show indicating that that they have the necessary skills as identified in the student model. As a part of the description of the task and the rubric to assess the task, I describe the development of the rubric itself. I then describe the task model, or the description of the exact task scientists will complete, and the process of the pilot testing of the rubric.

4.1 The Student Model

The student model outlines the focal knowledge, skills, and abilities, or the primary knowledge, skills and abilities targeted by this test and rubric along with other knowledge, skills, and abilities that may be required by this design pattern (R. Mislevy & Riconscente, 2005). The KSAs that I identify as a result of the domain analysis and domain modeling are outlined below.
Through this test and application of the rubric, I wish to measure the ability of scientists to give an effective, informative presentation on a scientific topic to a lay, public audience. I also wish to measure the ability of the scientists to seamlessly include six (as will be show in the pilot testing narrative, the five constructs became six) recommended sub constructs of communicating science to the public, including the following:

1. Scientists speaking to public audiences will be able to explain the relevance and importance of the science information to the specific audience to which they are speaking.
2. Scientists speaking to public audiences will be able to use language to express complex ideas clearly and in a manner that is adapted to the understanding of the audience to which the scientist is speaking.
3. Scientists speaking to public audiences will use visual aids, particularly electronic visual aids, in a way that enhances their presentations by following applicable guidelines for speakers with technical expertise creating visual aids for non-expert audiences.
4. Scientists speaking to public audiences will explain one or more science concepts and processes such as steps in an experiment, data gathering, uncertainties in science, and/or the team nature of science exploration.
5. Scientists speaking to public audiences will present themselves, and scientists in general, as trustworthy, friendly, approachable, and knowledgeable.
6. Scientists speaking to public audiences will engage audiences in interactions and conversations about science.

The means of implementing these KSAs and the means for assessing them, the before-mentioned filled-cell rubric, will be outlined in the evidence model, which follows this student model. The method of developing the rubric is included in the task evidence model section.
4.2 Evidence Model

The evidence model outlines what a participant, or “student,” will do to show evidence of his or her knowledge, skills, and abilities that are to be measured through this test. Mislevy & Haertel say that the evidence model acts as a bridge between the student model variables and the task model. In their explanation, they note the two components of the evidence model, those being the evaluation component and the measurement model. In other words, the evidence model outlines what you are going to assess and how you are going to assess it (2005). As part of the evaluation component, the assessors look at the qualities of what is being produced. Observable variables are identified, as are the evaluation procedures, such as “answer keys, scoring rubrics with examples, and automated scoring procedures” (R. Mislevy & Haertel, 2006).

Out of all the possible evidences that could be presented, the evidence I have chosen to evaluate for this study is a library talk of approximately 10 minutes. Other assessors who make use of the rubric could choose to increase the recommended length of the presentation to 15, 20, or even 30 minutes. However, the data samples to which we had access for this study are approximately 10 minutes long, and 10 minutes is a common time limit for brief, informative presentations. A shorter presentation would not allow the speaker enough time to demonstrate competence in these areas.

The evaluative aspect of the evidence model is an analytic scale (rubric) that will assess speakers on how well they demonstrate the observable knowledge, skills, and
abilities recommended for a science speaker as called for by this study. The observable variables were developed using communication and science communication research. The evaluation procedure in this case is a filled-cell rubric that allows the assessor to assign a value to each of the observable variables.

To aid in this evaluation, a code book accompanies the rubric. (See Appendix H for the first version of the code book). This code book gives detailed explanations of each of the elements in the rubric so that all people using the assessment instrument, regardless of their experience or background in public speaking and/or science, have a collective understanding of the desired knowledge, skills, and abilities that the speakers should demonstrate, and can feasibly rate the speakers based on the criteria outlined in the filled-cell rubric and the explanations offered in the code book. The code book gives specific examples of evidences that an assessor can look for when using the rubric. For example, the following sample section from the first version of the code book focuses on metaphor and assessing a scientist’s use of metaphor when using unambiguous language to present science topics to lay audiences. The code book, including this section on metaphor, underwent several revisions during the process of testing and revising. This is the first iteration of the metaphor section.
**Metaphor:** One technique is metaphor and other comparisons. Research shows that pedagogical metaphors that connect to the audience members’ experiences are effective at explaining complex ideas as are theory-building metaphors. A pedagogical metaphor is one that is meant to teach about a concept or phenomenon, something that describes existing knowledge and for which other possible ways of explaining exist. A theory-building metaphor is one that is used as a new way to explain a scientific phenomenon which cannot be paraphrased because there is not another way to talk about the phenomenon (Boyd, 1993). These distinctions are so fine as to be impossible to make, and it is not necessary to distinguish between them for purposes of assessing speakers. Primarily, scientists should avoid metaphors and comparisons that are based on knowledge inherent to being a scientist and have no explanatory power to people who do not have scientific expertise.

a. **Example 1 – pedagogical metaphor** – messenger RNA, proteins as building blocks

b. **Example 2 – theory-building metaphor** – genetic code

c. **Example 3 – cognitive metaphor** – polymer chain (used as if all present would understand, with no explanation of how polymers are chains or why they might be thought of as chains).

An expert speaker will use a few metaphors or comparisons throughout the presentation rather than offering up multiple metaphors.

The best metaphors/comparisons compare the new object or concept to something that is familiar to audience members and gives an accurate picture of the scientific concept.

4.2.1 Rubric Development.

After determining the KSAs that were not only the most important for science speakers to possess but also that could reasonably be assessed with a filled-cell rubric, it was necessary to develop the format for the rubric and specific items to be assessed in each row of the rubric. In this section, I discuss developing the assessment criteria for each of the KSAs covered by the rubric, and support the choice of those criteria through
communication research. Originally, the science communication portion of the rubric included five elements, as delineated below. (See Appendix B).

4.2.1.1 Development of Science Communication Items.

For the competency “Scientists speaking to public audiences will connect the relevance of science information to the specific audience to which they are speaking,” I emphasized connecting science concepts to audience through common goals, values, and/or experiences, as suggested by communication research. The next KSA I included was making a case for the relevancy and importance of science to the lives of the audience, and lastly, indicating excitement about and the joy of science for reasons relevant to the audience. At the various levels, I attempted to use words showing levels of competency, such as “clearly and consistently” for level four, “inconsistent” for level three, “some connection,” “not explicitly relevant” and “not explicitly connected” for level two, “and little connection” for level one.

For the competency “Scientists speaking to public audiences communicate complex scientific ideas and make them clear to the public,” I focused on metaphors and comparisons, understandable language that considers audience knowledge, and simple, clear visual aids that follow superior design principles. At level four, I used the qualifiers “interesting,” “deliberate,” “build,” and “excellent.” At level three, the qualifiers were, “deliberate, but not entirely appropriate,” “generally,” and “helpful.” For level two, the
qualifying words were “some,” “not relevant,” “not fully developed,” “complex or unclear,” and “ignore design principles.”

For the next section, “Scientists speaking to public audiences explain to their audience science processes and methods,” the scoring scheme focused on defining scientific terms, descriptions of scientific processes, explanations of why particular steps are needed, giving reasons for choices, the team nature of science, and explanation of uncertainties. The qualifying terms at level four were “clearly” and “emphasizes,” while at level three, they were “usually,” “may,” “without explaining,” “mention,” “give reasons for choices, but not all.” At level two, the qualifiers included “mostly,” “may attempt,” “without complete success,” and “fail to completely…”

For the KSA “Scientists speaking to public audiences will use techniques such as self-disclosure and immediacy to help scientists seem trustworthy and knowledgeable,” I focused on self-disclosure, personal stories and/or ideas, inclusive terms such as we, our, and us, concrete language, and positive portrayals of scientists. For the qualifiers, I used “frequent,” “appropriate,” and “uses” at level four. At level three, I used the terms “some,” “mostly,” “largely,” and “seem.” At level two, the qualifiers were “some form,” “sometimes,” “some,” and “may or may not.” At level one, the qualifying terms were “little,” and “dispassionately,” while for the zero level, the rubric includes qualifiers such as “little to no,” “few or no,” and “no.”
Lastly, for competency that “The scientists speaking to public audiences will engage audiences in interactions and conversations about science,” I focused the assessment on the scientist creating opportunities for interaction with the audience, taking opportunities to ask questions, the speaker acting as learner, and engaging in active dialogue. The qualifying terms at level four included “multiple opportunities” and at level three, “opportunities,” “does not give them much time,” “little dialogue.” At level two, the qualifiers were “ask but not respond,” “few, if any,” “no indication,” and “no dialogue.” For level one, I used the qualifiers, “not asked questions,” “limited,” and “no dialogue” while at level zero, I used the qualifiers, “authoritative air,” “does not give power,” “finality,” and “unbending.” These terms and qualifiers will be tested in subsequent tests of the rubric.

4.2.1.2 Public Speaking Competency Rubric (PSCR) Adaptation.

To fully assess a scientist’s public speaking competence, it is important to evaluate the basic public speaking issues, such as organization and eye contact. The issue of what competencies are core to basic public speaking skills has been examined thoroughly since about the 1970s, when the National Communication Association (NCA) launched a detailed effort to gather the core competencies that a public speaker should portray. Since such research is beyond the scope of this study, but needed for an effective rubric to assess the communication competence of scientists, I turned to others who have previously developed a filled-cell rubric to assess those competencies.
Therefore, with permission of the authors, the first part of the rubric includes measures taken from the Public Speaking Competency Rubric (PSCR), developed by Schreiber, Paul & Shibley (2012). The first section of the APPS (items 1-6) includes issues of organization and delivery common to nearly any public speaking opportunity. The 11 criteria included in the PSCR for informative presentations are listed below:

The Student (or speaker):
1. Selects a topic appropriate to the audience and occasion
2. Formulates an introduction that orients audience to topic and Speaker
3. Uses an effective organizational pattern
4. Locates, synthesizes, and employs compelling supporting materials
5. Develops a conclusion that reinforces the thesis and provides psychological closure
6. Demonstrates a careful choice of words
7. Effectively uses vocal expression and paralanguage to engage the Audience
8. Demonstrates nonverbal behavior that supports the verbal message
9. Successfully adapts the presentation to the audience
10. Skillfully makes use of visual aids
11. Constructs an effectual persuasive message with credible evidence and sound reasoning

In choosing which of the 11 elements to include in the public speaking portion of the rubric, I chose those items that address the most general public speaking criteria, that are important to nearly any presentation, and that are well-suited for general assessment. I did not include those that addressed in a general way the criteria that would be addressed more specifically by the science communication portions of the rubric. Using these criteria, I included items one, two, three and five, which address topic choice and organization in a presentation. I also included the two items specific to
delivery, item seven (focusing on the use of voice, or vocalics) and item eight (focused on body movement and eye contact), with some alterations. Item eight did not correlate as expected in the original tests (Schreiber et al., 2012), and the authors surmised that it might be due to combining too many elements of non-verbal communication into one item. Additionally, item seven uses the term “paralanguage” to refer to vocal delivery factors, when in fact the term can refer to any communication outside the spoken word, including gestures and facial expressions, factors addressed in item eight. Therefore, it was important to make the division between the two non-verbal elements clearer. Further testing should be done on these elements to see if they function as hoped with the slightly altered wording. Outside of those two minor changes, the wording on the first six elements of the rubric is taken by permission directly from Schreiber, et al. (2012).

The second portion of the rubric, items 7-12, addresses scientist-specific knowledge, skills, and abilities that were identified through the current study.

4.2.1.2 General Rubric Layout and Design.

Originally, I created two rubrics; one with rating levels going from zero to four and another using rating levels one to five. I also included a descriptive word with each level, such as “Advanced,” “Proficient,” or “Basic.” Shown in Figure 4.1 is a section of the rubric in its original layout and format with some of the wording for the first category of the science portion of the rubric showing. This illustration shows the “zero to
four” scoring format, as this is the format settled on through discussion with raters and advisors, and shows the rubric to be in portrait layout, as it was originally conceived. For the full first draft of the rubric, see Appendix B.

<table>
<thead>
<tr>
<th>Connect science information to a specific audience in a specific situation</th>
<th>Advanced 4</th>
<th>Proficient 3</th>
<th>Adequate 2</th>
<th>Basic 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clearly and consistently connects science concepts to ideas familiar to audience, creates common ground by strongly emphasizing common goals, values and/or experiences, makes a strong case for the relevance and importance of the science to the audience’s lives, indicates the excitement and joy of science for reasons relevant to the audience</td>
<td>Specifically connects science concepts to ideas familiar to the audience, inconsistent mentions of common goals or values, discusses the importance and relevance of the topic to the audience, expresses excitement and joy in working with this science topic</td>
<td>Makes some connection between the audience and the science, some mention of common goals or values, may mention the importance of the topic but does not make it explicitly relevant to the audience, excitement and joy of science topic not explicitly connected to audience</td>
<td>Makes little connection between the audience and the topic, few if any mention of common goals or values, does not make the topic relevant to the audience, excitement and joy not expressed and/or not connected to the audience</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.1. First Rubric Draft - Illustration of the layout of the first draft of the rubric prior to the first round of pilot testing.

The length of the rubric was an area of importance as well, since an analysis of other rubrics noted the length and complexity of those documents as being a deterrent to their use. I formatted the rubric in portrait view, with one page containing the PSCR elements and one page containing the science communication elements.

4.3 Task Model

The task model includes the aspects of assessment situations that are likely to evoke, or cause the participant to produce, the desired evidence. It also includes aspects of the assessment situation that can be varied in order to shift difficulty or focus of the
test (R. Mislevy & Riconscente, 2005). For this particular assessment, the task model is envisioned as follows:

You are asked to prepare a presentation of approximately (10 to 40) minutes explaining your research or some aspect of your research to an “intelligent but uninformed” public audience of between (10 and 40) people. You can assume that your audience members are interested in knowing more about science but do not have any special knowledge about science or scientific procedures.

Your presentation should include the elements of any good public presentation, such as an introduction that grabs the audience’s attention and introduces the main ideas you will be talking about, a well-organized body section that emphasizes clear main points about your topic, and a strong conclusion that reviews your main points and gives the audience members something to think about going forward.

The presentation should be delivered naturally, with little reliance on notes. You should attempt to speak to the audience in a conversational manner, using language that they can understand, while maintaining the integrity of the science concepts you are presenting. Please allow the audience to ask questions at the end of your presentation, and be prepared to have a conversation with audience members about the topic.
As you speak, we ask that you particularly try to communicate to the audience the importance and significance of the topic while using language the audience understands. We encourage the use of analogies, comparisons and metaphors that compare the scientific topic to something with which audience members are more familiar, helping them understand how your science applies to their lives. We also encourage you to integrate explanations of some of the processes involved in carrying out your science. These processes may include collaboration, experimentation, data gathering, and managing uncertainty. If your science could have positive impacts on your audience’s lives, it is ideal for you to point those positive impacts out.

As you speak, you should consider this an opportunity for the audience to get to know a scientist on a personal level. Sharing your excitement for and love of science and sharing personal reactions and experiences can help build a connection to the audience.

Please include visuals in the form of PowerPoint slides, a Prezi, or other standard format visuals. You may use non-digital visual aids; however, if you are going to show the audience an actual object, be sure it is large enough for the audience to see well. Keep in mind the non-technical nature of your audience’s understanding as you prepare your visual aids and the spoken portion of your presentation. Minimize the use of text in your visual aids, and use large, simple graphic elements.

The presentation will be scored on a rubric. This rubric might be shared with you either before or after the scoring is completed. In most cases, the assessment of your communication is a formative process that is meant to help you become a more effective public communicator of science. In some cases, the assessment may become part of a personal report or programmatic report to demonstrate the broader impacts of a research project or to show community involvement for a particular grant or program of which you are a member. It may also be used in measuring the efficacy of a training program.

We encourage you to continue to seek out public communication opportunities with your science!
4.4. Assembly Model

This section asks how much data is needed for a fair analysis, or how long the speakers need to present and what needs to be included in order for a science library talk to be assessed fairly. In this case, the data collected is the presentation content, so the presentation needs to be long enough for the scientists to demonstrate all the knowledge, skills, and abilities requested. Scientists who are being trained in science communication or scientists who are engaging in public outreach are invited to give a (10-40) minute presentation about an aspect of their research. Based on experience and the literature review, 10 minutes was set as the minimum time needed to demonstrate the desired KSAs. The actual time allotted to an individual will differ according to the purpose of the assignment.

If a scientist is being assessed as part of a training course or as a way of showing what (s)he can do in a public presentation, then the presentation might be closer to the 10-minute mark in the interest of assessing several presenters in a short period of time. However, if a scientist will be assessed in a more natural setting, such as giving an actual library talk, the length of the presentation will likely be longer, ranging from about 20 to 40 minutes. Therefore, the specific time guidelines will be altered according to the situation. The scientists who were assessed in order to test the rubric for this study were given the assignment of preparing a presentation of approximately 10 minutes.
4.5 Presentation Model

In this case, the presentation model (or how the test will look or be administered to the scientist) is outlined in the Task Model. Many language assessments consist of a multiple choice or true/false test given on paper or an essay prompt given on a computer or on paper. Speaking assessments are often impromptu, with the person being tested receiving the prompt and then having only a few minutes or no time at all to prepare. However, this assessment is different from any of those assessments. Instead of a paper test or an impromptu speaking assignment, those being assessed with this rubric will instead be given the task well in advance and will have time to prepare for the assessment. They will then stand in front of the assessor or a small group of audience members that includes the assessor and present her or his presentation.

4.6 Delivery System Model

This section outlines the methods used for assessment and the circumstances under which the rubric could be implemented. The scientists who give these presentations will likely be members of a class that trains scientists in public communication or members of a grant team that are being encouraged to participate in the public communication of science. In either case, the scientists being assessed with the rubric will likely be working to improve their communication skills. The optimal context for the implementation of this rubric is in the context of a training course or training support of some kind where the scientist has access to the rubric and codebook, is
receiving instruction in communication skills and the purposes of science communication, and is receiving formative assessments on a regular basis.

Other speakers who could feasibly be assessed with this scale are researchers or other members of a lab who prioritize public communication of science or who are members of a grant team who would like to show that they are effectively communicating their findings to the public. It is conceivable that the speakers might be asked to speak to audiences of different ages (children, teens, adults) or divergent backgrounds (high school or college students enrolled in a science class, perhaps, rather than a public audience with no science background at all). As mentioned, assessors can assign presentations of varying length, such as 15, 25, or 40 minutes. It is not recommended that the presentation portion (the portion that includes mostly a scientist speaking, not the interactive, engagement portion) of the event last longer than 40 minutes. There is a high likelihood that a lay audience will lose interest in even the most fascinating science topic after that length of time.

4.7 Initial Revisions and First Tested Version of the Public Science Communication Rubric.

After drafting of the initial rubric and code book, I recruited raters to assist in reviewing the documents and assessing speeches in order to test the rubric for reliability and validity. The testing was meant to highlight any areas that needed further clarification and modification and see if the rubric seemed to be testing the abilities and
skills that it was designed to test. It was also meant to see if different raters were seeing and understanding the constructs similarly.

4.7.1 The Raters

The raters were MA or MFA students in the English department at Iowa State University. They volunteered to assist with the work out of interest in the project and to earn a stipend of $100. (See Table 4.1)

Table 4.1. Raters for Pilot Study, their Qualifications and Participation

<table>
<thead>
<tr>
<th>Rater</th>
<th>Qualifications</th>
<th>Participation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rater 1</strong></td>
<td>MA Student 4 Semesters teaching Public Speaking Interest in Science Communication</td>
<td>Completed both rounds of testing</td>
</tr>
<tr>
<td><strong>Rater 2</strong></td>
<td>First year MFA student Teaches composition – no public speaking Interest in sustainability, but no background in science communication</td>
<td>Completed first round only</td>
</tr>
<tr>
<td><strong>Rater 3</strong></td>
<td>First year PhD student Teaches composition – no public speaking Interest in Rhetoric of Science</td>
<td>Completed first round only</td>
</tr>
<tr>
<td><strong>Rater 4</strong></td>
<td>First year PhD student -Rhetoric 4+ semesters teaching Public Speaking No science expertise or interest</td>
<td>Completed both rounds of testing</td>
</tr>
<tr>
<td><strong>Rater 5</strong></td>
<td>4 Semesters teaching Public Speaking Works for sustainability lab MFA Student</td>
<td>Completed both rounds of testing</td>
</tr>
</tbody>
</table>

One of the raters was the same MA student who assisted with the work of the domain analysis, a student in a rhetoric program with some experience (about 4 semesters) teaching public speaking sections, and who has an interest in science communication. The second rater was an MFA student in the last semester of the
program. This student had also taught public speaking for about 4 semesters and was teaching composition classes at the time of the testing. The MFA student worked as a communication specialist for a lab that specializes in sustainability. The third rater was a first-year PhD student with more than 4 semesters experience teaching public speaking. These three raters participated in the initial evaluation of the rubric and the two pilot rounds of testing.

The next rater is a first-year PhD student with an interest in science rhetoric. This rater is the teacher of record for first year composition classes, and has never taught public speaking. The final rater is a first-year MFA student with some interest in sustainability, but no experience with science communication or public speaking. These last two raters participated in the initial rubric evaluation and the first pilot round of coding, but were unable to complete the second round of pilot testing.

4.7.2 The Rated Presentations

The 79 presentations by scientists that were used for rater assessment in this study were provided by the SYMBI program, an NSF-sponsored GK-12 grant at Iowa State University. The primary focus of the SYMBI program was placing graduate students in STEM (science, technology, engineering, and math) fields into K-12 classrooms to provide science support and expose students and teachers to “real” scientists (Ufnar, Kuner, & Shepherd, 2012). The SYMBI program included an additional component for public communication. Program administrators asked fellows to give a
presentation about their doctoral research to an “intelligent but uninformed audience,” usually consisting primarily of undergraduate students. Occasionally staff or professors attended a presentation, but largely the audience members were undergraduate students. The presentations were professionally recorded by university videographers for later assessment and were posted to a private, password-protected video storage site in compliance with IRB requirements.

The fellows were given the assignment to prepare a presentation of approximately 10 minutes that would be given to a “lay” audience. They gave one presentation at the beginning of their year as a fellow and one at the end of their year as a fellow. If the fellow was going to continue for a second year in the program, (s)he would give two presentations on subsequent days; one serving as the ending of their first year as a fellow and one serving as the beginning of their second year as a fellow.

Fellows were instructed to prepare visual aids appropriate for a non-expert audience. In the early years of the grant, the fellows were simply told to prepare a presentation with visual aids and were not given any indication of how the presentations would be assessed. After the first approximately two years, the fellows asked to see the assessment criteria. The rubric was a “by-the-seat-of-the-pants” rubric that was developed without using a methodological approach; however, having access to the rubric seemed to help fellows understand some basics of presentation skills and help them prepare for the presentations.
In the first two years of the grant, fellows did not receive training specific to public science communication, but in subsequent years, they did receive some training. Training varied from year to year. On one occasion, the fellows participated in a short version of the “Portal to the Public” training that is offered by some science centers and science organizations to help scientists with their informal public science communication (Selvakumar & Storksdieck, 2013). Other times, the fellows had different kinds of training, such as talks by science communicators who were considered “expert” by the SYMBI organizers.

Fellows came from many different fields, including agronomy, biology, genetics, and various areas of engineering. Some were passionate about public communication and took opportunities to engage with public audiences outside the SYMBI program, while others were less interested in public science communication. Fellows demonstrated a range of skills, from expert to basic, in preparing and delivering their messages.

Since these fellows had such diversity in research topics, interest in public communication, and communication skill level, and they were given a task similar to the one that is envisioned for the APPS rubric (although less detailed), these presentations made excellent models for testing the APPS.
4.7.3 Initial Feedback

Prior to beginning the coding process, I contacted the raters to give them copies of both versions of the APPS (starting with a zero and starting with a one in the rating scale), a copy of the codebook, and a sample presentation by a scientist for them to assess using the sample rubric. I asked that they review both documents for usability, assess the codebook for how helpful or not it was in understanding the rubric, give feedback on using the zero to four scale or the one to five scale, and give feedback on what could be improved in the APPS and codebook after using them to assess a presentation. Raters had several weeks over winter break to complete this part of the assignment.

The initial feedback from raters included the recommendation to use the zero to four scale to rate the speakers, citing the need to have a way to indicate to a speaker that the desired element did not seem to be present at all. While there has been discussion and many decades of debate on this issue (Kidd, Parry-Giles, Beebe, & Mello, 2016; Morreale & Backlund, 2007; Stevens & Levi, 2013), the suggestion to have a scale beginning at zero mirrors advice from other public speaking instructors (Schreiber et al., 2012), from the majority of the reviewers, and the majority of the raters on this project, both the pilot raters and the later raters. Therefore, the APPS used a zero base from this point forward.

Raters did find the code book to be useful, and suggested some minor edits. One also suggested that the long length of the descriptions, both on the rubric and in the code
book, made it difficult to absorb the information quickly and to use the rubric effectively. While I knew that using the rubric would become quicker and more manageable with use, I also knew that it was important that the layout and design of the rubric and code book contribute their ease of use. While I thought it important to use the same wording in the code book and on the rubric, I looked at ways to adjust the rubric and codebook based on this feedback.

After receiving the initial feedback from the raters, I then sent the rubric to an advisory committee consisting of three experts in communication studies and rhetoric, one expert in language testing, one expert in rhetoric and composition, particularly visual rhetoric, and one expert in science communication, with a background in both science and science communication.

This advisory committee had several suggestions for the rubric, although less advice regarding the code book. One concern was repeated by most of the experts and one of the raters: the categories did not seem to be unique enough. In particular, the reviewers had a difficult time differentiating between the measure that asks a scientist to communicate complex ideas and make them understandable to the general public and the measure that asks the scientist to increase the audience’s knowledge and understanding of science and scientific processes.
For example, one reviewer said that since he did not think we are measuring audience reaction (which the rubric does not measure), then what was the contrast between the process of science and the ideas of science. He asked if it was the difference between the process and the results of a scientific experiment or research project. This was not the distinction I was trying to make, which would be too fine a distinction in any case. Instead, the distinction was between the clarity of the language and the description of science processes. I attempted to address this concern in my next revision, but it continued to be a concern for several more rounds of testing until I made more dramatic changes down the road that made the distinctions clearer.

Another comment made by at least one reviewer was that the distinctions between the qualifiers weren’t clear enough to make it obvious which category a speaker would fall into. For example, the reviewer pointed out that “basic” and “not present” seemed too close to be helpful. This particular reviewer felt the distinctions in the PSCR elements were also unclear, but I did not revise that part of the rubric. While I did take note of this critique, I did not make significant improvements in this area until later in the testing process when the raters also expressed concerns about that aspect of the rubric.

Another reviewer suggested removing the qualitative descriptor words (“Advanced” for level four, “Proficient” for level three, etc.) from the rubric. The reviewer, someone with extensive experience assessing and teaching public speaking,
said that the descriptor words are likely to bias the assessments, since every rater will have a different view of what “Advanced” or “Proficient,” etc. means, and could get caught up in those descriptor words rather than the descriptions of the proficiencies laid out in the rubric. I made note of this objection, although the descriptor words are part of the PSCR. The numerical score is the important score for scientists in any case, and the loss of the descriptor words would cause little disruption. However, since the PSCR did use descriptors and many other rubric samples that I examined also used the descriptor words, I initially retained them for the first round of pilot testing. However, I determined to listen and watch closely for any problems with the descriptors. I found that the descriptors did indeed cause confusion among raters at our first coding meeting, with raters assigning a score based on their connotative understanding of the descriptor words. Therefore, I dropped the descriptor words from the rubric before the first round of pilot testing.

This reviewer, who had extensive public speaking experience but little to no science communication experience, wanted to know how Schreiber, et al.’s dimensions that I did not use in the rubric differed from the science communication specific measures I created based on the domain analysis. For example, the reviewer asked how “Increase the audience’s knowledge and understanding of science and scientific processes” is any different from Schreiber, et al.’s dimensions, “Demonstrates a careful choice of words,” and/or “Locates, synthesizes and employs compelling supporting
materials.” However, those measures from the PSCR are too general for use by scientists seeking to accomplish specific goals with their public communication.

Schreiber et al.’s wording is necessarily vague, and as the article (2012) states, the rubric was used in public speaking classrooms, not in science-specific contexts. A general public speaking rubric measures general public speaking knowledge, skills, and abilities. In a broad-reaching public speaking classroom, that is adequate and likely desirable. However, in a scientist-specific training course, one seeking to nurture public engagement practices in scientists that will meet the needs and desires of science stakeholders and science communication scholars, the broad measures are inadequate for several reasons. Scientists, seeing a high score for the measure, “Locates, synthesizes, and employs compelling supporting materials” are unlikely to imagine that the speaker who earned that high score is accomplished at helping members of the public understand the processes of science better. Instead, the scientist would possibly assume that the speaker used sources well in the presentation. Using sources well is a lower priority for science speakers speaking to public audiences. A scientist, by weight of her/his degree and research credentials, has a great deal of credibility with the general public (Funk & Rainey, 2015). Unlike a classroom speaker, the credibility of science speakers does not rest on citing multiple, credible sources. And unlike general classroom speakers, scientists speaking to public audiences do not need to convince the audience that they are qualified to speak about their research.
Additionally, scientists are more likely to use a rubric developed specifically for their needs and wants. The stated purposes of this Assessment for Public Presentations by Scientists are to bring communication expertise into the training of science communicators and to discover and assess desirable public communication knowledge, skills, and abilities specific to scientists. The domain analysis demonstrated repeatedly that scientists do have desired outcomes for their public presentations that differ from the general goals of a classroom public speaking presentation. In order to be useful to the target audience, the rubric must reflect those desired outcomes, and must reflect them in language that is appealing to scientists. Like every other discipline, scientists think of themselves as being unique from other disciplines, as doing work that is exclusive to science, and as being important in a distinct way. Therefore, the science communication-specific measures are important to this rubric and its purpose, and several of the PSCR measures do not meet that purpose.

Another concern raised by the reviewers was the use of the PSCR elements. Although they are not being tested, validated, or modified in this study, it is important to bring the basic elements of delivery and organization into the assessment if the APPS is to be used effectively to improve the public speaking of scientists. The reviewers wanted to be sure the creators of the PSCR approved of the use of certain elements for the science communication rubric. On Feb. 6, 2017, Lisa Schreiber, the lead author of the PSCR, generously gave permission for elements of the PSCR to be used in the APPS, if
the PSCR paper is cited properly in the APPS (Personal communication, Feb. 6, 2017). Accordingly, I inserted a reference to the Schreiber, et al. (2012) paper into the rubric directly under the first six elements, which are adapted from the PSCR. I also inserted a similar reference into the code book.

4.7.4 Pilot Testing, Round 1

The purposes of testing the rubric are twofold. First, testing, particularly pilot testing, offers opportunity for identifying areas in the rubric where language needs to be clarified or meaning is ambiguous. Testing helps identify constructs that raters may have trouble scoring or trouble differentiating between two adjacent scores and allows refining of those definitions to assist in later coding. Secondly, rubric testing helps identify the reliability and validity of a rubric. If a rubric is valid, it is accurate in measuring what it claims to be measuring, and if a rubric is reliable, it will measure the construct consistently across raters. In language testing, as described earlier, we are looking for a variety of factors in a good test, including reliability, construct validity, authenticity, interactivity, impact, and practicality. This test was designed to be authentic and interactive, have impact, and be practical. The testing helps ensure reliability and construct validity, two tightly correlated factors (Bachman & Palmer, 1996).

For the APPS testing, the agreement testing was also designed to determine if raters without particular knowledge of public speaking could be trained to accurately
rate public science communication. The first pilot round of testing had raters with public speaking experience and raters without, while the final rounds of testing had raters with public speaking experience and raters with science expertise and no public speaking experience.

After gathering and analyzing feedback from the reviewers and the raters, I set up an initial rater meeting to go over the rubric and participate in a norming session where we all watched and assessed the same speech and compared our assessment results. One of the raters, rater 5, despite agreeing to the time via a Doodle poll and receiving multiple email reminders about the meeting, forgot to attend the meeting. I met with that rater separately the next day and attempted to recreate the coding meeting as closely as possible. In analyzing the results, rater five was in agreement with other raters a majority of the time and was not an outlier at any time, so the individual training was deemed to be effective.

At the meeting, we discussed the challenges the raters had with assessing the presentation that was sent to them as a preliminary practice presentation. We discussed many comments and concerns they expressed, one of which was the context in which the presentations were given. Raters did not feel I had adequately explained to them the rhetorical situation for the presentations being assessed, and therefore felt they were unable to adequately assess the item that asked if the science topic was adequately connected to the audience. I explained the situation to them at that time (PhD scientists}
speaking to audience made up of non-scientist members of the university community, usually undergraduates who were attending the presentations to earn extra credit for their communication classes.) In addition to discussing the rhetorical situation with the raters during our meeting, I inserted a paragraph describing the situation into the codebook and highlighted it. While this description of the rhetorical situation was removed when the codebook was made generally available, it was useful for testing purposes. In addition to the explanation of situation, I added a note explaining that the final element (engagement) would not be assessed during testing because there was not adequate evidence available to evaluate this competency. In most cases, the filming of the presentations was cut off just before the question and answer period at the end of the presentation.

After the evaluating the presentation in order to normalize our scoring at the meeting, we discussed ways to make the language construct appear to be more clearly about language and the processes element to be more clearly about explaining how science works. The raters suggested that one of the problems with the language competency was that it was trying to cover too many items. They suggested creating a separate category for the use of visual aids. Although using clear visual aids can help increase the clarity of a science presentation, including the visual elements with a competency focused on oral communication was too complex and made assessing that item difficult. That suggestion was followed, since visual aids are an expectation of a
science presentation of nearly any kind, and it is important that the science speakers understand what visuals work or do not work, particularly in a science context. When presenting to other scientists, science speakers will use technical vocabulary, detailed charts and graphs, and a great deal of text in their visual aids to make their points (Alley et al., 2006; Doumont, 2004). Therefore, I did agree to create a separate construct for visuals, despite the fact that they add to clarity in a presentation, because visuals need to be assessed as an important part of a scientific presentation, and because of the important contribution communication research can make to using visuals in science presentations.

To help with the layout and design issues identified in the preliminary comments, I suggested that using a one or two-word headline or title for each competency would help with usability. Therefore, instead of having long sentences describing the construct, the rubric had one hierarchical element that stood out as defining that competency. Additionally, it was determined that it would be easier to assess the presentations and talk about the assessments if each of the competencies were numbered rather than identified only with a title. Therefore, the public speaking competencies became one through six, and the science competencies became seven through twelve, with the addition of number nine, visuals. At this point in the revision process, there was little change to the verbiage within the cells, with the exception of separating the visuals as a
competency out of the language competency (See Appendix C for the rubric as revised at that stage in the review process.).

When the revisions to the rubric were complete, I sent the revised rubric and links to ten presentations to the raters. They asked for two weeks to finish coding the ten presentations, so we set a date to send in their assessments by February 20th.

As noted in prior chapters, to calculate interrater reliability for this study, I used three measures: Simple percent agreement, a measure favored by language testing experts that is appropriate for rubrics, can accommodate multiple raters, and is less stringent than some other measures of reliability; simple adjacency agreement, which shows how often raters are within one point in their ratings rather than complete agreement; and Cohen’s Kappa, a stringent measure (Neuendorf, 2017) that accounts for possible chance agreement, is widely accepted in communication studies, and accommodates multiple raters.

In language testing, adjacent agreement is commonly used to determine if raters are scoring a speaker similarly, since it indicates a general agreement or lack thereof, which shows that raters are or are not seeing a construct similarly. If the adjacency agreement is high, then the constructs are seen as adequately measuring the skills and abilities that the test was designed to measure and doing so reliably among raters. However, if raters are scoring speakers differently (i.e., one rater gives a score of 2 and
another gives a score of 4), that indicates a discrepancy in understanding that needs to be addressed either through rater training or modification of the rubric.

Some other measures of inter-rater reliability only accommodate two raters, often functioning on a binary scoring system: either a quality is present or it is not present. These ratings were clearly inappropriate for this study. Scoring a rubric, with its multiple constructs and multiple levels within each construct, is much more complex and has a greater chance for differences. Therefore, I used both simple pairwise agreement, simple adjacency agreement, and Cohen’s kappa, a form of agreement that accommodates various raters and accounts for chance agreement. There is not a clear-cut guideline for interpreting any of these results, but ideally simple agreement would be somewhere around or over 50% and adjacent agreement would be over 80%. For Cohen’s kappa, at least .2 is suggested for fair agreement and .4 for moderate agreement.

Results from the first round of pilot testing showed that the raters did had fairly strong adjacency agreement overall, which was encouraging. Simple agreement and Cohen’s were not as strong, showing that there was work to do in clarifying the details of some of the measures. Overall, the agreement for the full data set was a simple agreement of 34.6%, an adjacency agreement of 83.2%, and a Cohen’s kappa measure of .067. Looking at the individual categories indicated areas for improvement and areas of strength, as shown in Table 4.2.
Table 4.2. Initial Pilot Testing Round 1, Overall Agreement and Agreement by Categories

<table>
<thead>
<tr>
<th>Measure</th>
<th>Simple Agreement</th>
<th>Adjacent Agreement</th>
<th>Cohen’s Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Data</td>
<td>34.6%</td>
<td>83.2%</td>
<td>.067</td>
</tr>
<tr>
<td>Relevance</td>
<td>36%</td>
<td>90%</td>
<td>.085</td>
</tr>
<tr>
<td>Language</td>
<td>30%</td>
<td>83%</td>
<td>-.025</td>
</tr>
<tr>
<td>Visuals</td>
<td>42%</td>
<td>94%</td>
<td>.11</td>
</tr>
<tr>
<td>Processes</td>
<td>34%</td>
<td>79%</td>
<td>.056</td>
</tr>
<tr>
<td>Trustworthy</td>
<td>31%</td>
<td>70%</td>
<td>.11</td>
</tr>
</tbody>
</table>

As previously mentioned, there is no set agreement level that every simple agreement, adjacent agreement, or Cohen’s situation must reach. Instead, those guidelines depend on the context. In the case of pilot testing this rubric, it was important that raters were using the rubric in largely the same way and assessing largely the same constructs (Osborne, 2008). Therefore, the results from the first round of pilot testing showed that, although there are some areas with more agreement than others, there was still a good deal of work to do to establish interrater agreement with the rubric. It was apparent that either the measures needed to be presented differently or worded so that they measured one construct more clearly, and/or raters needed more training and conversation in order to arrive at agreement, as Table 4.2 shows.

An examination of the scores and agreement revealed issues to discuss with the raters, and also highlighted areas of continuity. Overall, exact agreement was fairly low, but an 83% adjacency agreement indicated that raters were seeing concepts similarly.
The constructs of relevance and visuals were particularly strong, and language – again, while lower in exact agreement – was quite strong in adjacency agreement, indicating a general agreement on what the construct was measuring and when a person was displaying the KSAs measured by the construct.

However, the constructs of illustrating science processes and showing scientists to be trustworthy and human showed lower agreement overall, in both exact agreement and in adjacent agreement. This indicated discrepancies in those constructs and the understanding of how to apply those constructs. It was interesting to note that the raters with the highest levels of disagreement with each other on these constructs were the two that had general communication background (composition) rather than public speaking-specific experience. Another note was that the two raters who showed the least amount of confidence in rating meetings – expressing uncertainty frequently, apologizing for their ratings on frequent occasions – had the lowest agreement on the trustworthy/human construct. The trustworthy/human construct had the most variation in agreements out of all the constructs.

4.7.5 Revisions to Rubric and Codebook.

It was important to look at all of the constructs, but particularly those with the least agreement, and attempt to adjust the rubric and/or the understanding of the raters as it applied to the rubric to improve the agreement in those categories. I was most concerned about clarifying first, the trustworthy/human construct, second, the process
construct, and then lastly the language construct, and was less concerned about the visuals and relevance constructs.

I scheduled a second meeting with the coders on March 8. After the meeting was scheduled, one coder contacted me to say that she would be unable to complete a second round of coding, due to time constraints. Another coder did agree to attend the meeting and complete the second round of coding, but then did not attend the coding meeting and did not respond to my attempts to reach out and request feedback and additional coding for round two. The coders who dropped out were both of the raters with no public speaking experience. Therefore, there were three coders for the second round of pilot testing: the MFA student, the first-year PhD student in rhetoric and professional communication, and the MA student who helped throughout the domain analysis process. Each of these raters had experience teaching public speaking.

At the meeting, we discussed the areas of disagreement and each rater explained the reasons for their ratings. We talked in more depth about the situation, or context, in which the presentations they were scoring took place, and we discussed the fact that the rating scale should be thought of as a continuum, rather than as a measure that includes binary decisions. For example, one rater said something to the effect that, “If the speaker doesn’t say the word “we” at least three times, they get a one in the trustworthiness category.” We discussed the need to take all of the expressions of each competency into account and then assign the score based on the speaker’s overall performance in that
area, not count particular evidences of skill or look or one item as the measure that indicated competency or lack thereof in a category.

The most important change to come out of this discussion was the clarification of the “language” and the “process” constructs. In the original rubric wording, the skill of defining science terms was included in the “processes” construct, with the idea that explaining a term would help explain the science process with which it was connected. However, in practice, such definitions simply helped the language be clear so that the audience could understand what was being said. Additionally, defining terms is an activity that has to do with language, which is the focus of construct eight, understandable language. After scoring several presentations with the rubric, it was apparent that moving the skill of defining science terms to the “language” construct made sense and would help coders understand the constructs more clearly.

Additionally, the group discussed the use of the “concrete language” element as a part of the “trustworthy” construct. There is general agreement in the communication literature that a person who uses concrete language is seen as more trustworthy, but the raters noted that one reason that people who use concrete language are seen as more trustworthy is that they speak more plainly and clearly and seem to be more “like” their listeners. Since those attributes also contribute to a scientist speaking clearly, the coders felt that moving the “concrete language” element to the “language” construct would make that construct even clearer for future coders. The fact that concrete language also
does increase trust can be included in the code book, and coders can take that fact into account as they score a speaker on the “trustworthy” construct, but since it seemed to be better understood as a language issue, it made more sense combined in the language construct.

In this discussion, it also became clear that the raters’ understanding of the first construct, “relevance and importance,” was unclear, despite the raters’ general agreement in scoring that construct. Although the construct was titled “relevance and importance,” the wording within the cells talked first about connecting the topic to the audience. Audience connection, however, is not the most critical issue with this construct. I had attempted to make that clear through the title of the construct, but this second discussion made it apparent that it was not clear in the wording of the rubric measures, and it was necessary to adjust some of the wording in the cells to reflect more exactly what the construct should measure. Since the construct is first and most importantly meant to measure how well the scientist shows the relevance and importance of the science topic about which (s)he is speaking to the audience, I reworded those boxes to reflect that priority.

Another item that came up in the discussion was the fact that not every science presentation needs analogies or metaphors to make the points clear. Sometimes using concrete language, keeping a presentation free from science jargon and using audience-appropriate vocabulary can be adequate language accommodations. This is particularly
true when the science concepts being addressed are less complex, more familiar to the audience, and/or easier to understand. I changed the wording in the rubric to reflect the idea that scientists should use comparisons when they help clarify complicated ideas, but they are not absolutely required to earn a high score on the rubric. If they are used, they should be purposeful and increase understanding.

After discussion with the raters, I also changed the wording in the “process” construct cells. I wanted to express the idea that the speaker does not need to explain in great detail every decision and process they used in their project. Instead, speakers should explain some processes, data gathering procedures, information about working in teams, and/or information about navigating uncertainties. In an attempt to convey this message, I used the word “may” frequently. Again, the desired result was that a coder would understand that some of these measures would be present in the presentation, but not necessarily all of them.

Additionally, after using the rubric and examining the communication theory behind the inclusion of the elements, the raters and I agreed that sharing the “joy and excitement” of science, which was, at the time, included in the “relevance and importance” construct, contributed more directly to the “trustworthy” aspect of the rubric than the “relevance” aspect. The raters noted that being enthusiastic, excited, and passionate about their research fits into the other measures, such as immediacy, in that construct. The change helped to clarify the “relevance” measure and helped to
strengthen the “trustworthy” measure as well. In order to add and take away these elements, the wording in the “trustworthy” construct was largely redone.

The group also discussed more adjustments to the layout of the rubric. The numbering was seen as being useful, but the layout of the numbers did not contribute to an attractive document design. It was suggested that the numbers be put in a separate column to the side of the title and descriptor of the construct. To make this change work for the APPS layout, I switched from a portrait layout to a landscape layout in order to keep the length of the APPS to one page, front and back.

During the first pilot round, the raters were asked to use the rubric as a user would in actual practice, so they were asked to assess speakers using the PSCR measures at the first of the rubric along with the science communication constructs, even though the PSCR data would be collected for future testing and would not be analyzed for this project. At our meeting following the coding of the first round of presentations, the raters did want to spend time discussing in some depth their desire to accurately reflect the speaker competency in those basic public speaking areas covered by the PSCR, particularly those raters who had experience teaching public speaking. This led me to believe that the thought and concern being put into scoring the PSCR half of the rubric could be distracting raters from the task of primary importance for this study: the science communication measures. Therefore, in future rounds of coding, the raters were
instructed to look only at the science communication measures, not the entire rubric. The March 8 version of the rubric that reflects this change is found in Appendix D.

4.7.6 Pilot Testing, Round 2.

After completing the revisions based on our discussion, I sent the newly-revised rubric out electronically to the three coders who were able to complete the second round of coding. Using a random number generator, I assigned five more presentations to the coders. Due to travel and spring break plans, they asked to have until March 20th to return the scored rubrics. It is possible that having spring break during the coding period had a negative effect on the scoring, since two of the coders completed their coding on or after March 20th, so it is likely that the scoring was completed in a rush and that the coders had forgotten some of the discussion we had in conjunction with scoring the rubric.

However, despite the disruption of spring break, the intercoder reliability scores did improve significantly in round two over round one of the pilot study, both the overall scores and most of the individual construct scores, except in the “trustworthy” category, as shown in Table 4.3.
Table 4.3. Overall Data, Pilot Study Rounds One and Two

<table>
<thead>
<tr>
<th>Construct</th>
<th>Pilot 1 Simple</th>
<th>Pilot 1 Adjacent</th>
<th>Pilot 1 Cohen’s</th>
<th>Pilot 2 Simple</th>
<th>Pilot 2 Adjacent</th>
<th>Pilot 2 Cohen’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>34.6%</td>
<td>83.2%</td>
<td>.067</td>
<td>34.67%</td>
<td>80%</td>
<td>.07</td>
</tr>
<tr>
<td>Relevance</td>
<td>36%</td>
<td>90%</td>
<td>.085</td>
<td>46%</td>
<td>86.6%</td>
<td>.302</td>
</tr>
<tr>
<td>Language</td>
<td>30%</td>
<td>83%</td>
<td>-.025</td>
<td>53.3%</td>
<td>93%</td>
<td>.151</td>
</tr>
<tr>
<td>Visuals</td>
<td>42%</td>
<td>94%</td>
<td>.11</td>
<td>26.67%</td>
<td>66%</td>
<td>.083</td>
</tr>
<tr>
<td>Processes</td>
<td>34%</td>
<td>79%</td>
<td>.056</td>
<td>20%</td>
<td>86%</td>
<td>-.123</td>
</tr>
<tr>
<td>Trustworthy</td>
<td>31%</td>
<td>70%</td>
<td>.11</td>
<td>27%</td>
<td>86%</td>
<td>.037</td>
</tr>
</tbody>
</table>

These results indicated strong overall improvement, including strong improvement in the relevance and language categories, two areas of concern in the first pilot round of testing. These results, particularly in the two categories that were the focus of attention in the rating meetings, indicated that the adjustments made to the rubric and the discussions among the raters were helpful in defining those measures. The results in the visuals construct was puzzling and not easily explained. The raters disagreed more on the relative levels of clarity in the visuals and how much that presence of or lack of clarity in the visuals affected or did not affect the speakers’ presentations. Additionally, the results in the processes and trustworthy categories indicate that there was lower exact agreement in those areas. However, the adjacent agreements were higher, indicating that the raters were overall seeing the agreement more clearly.

One item of interest in examining the scoring was that the rater who was late in returning the scores, Rater Three in this round of testing, the MFA student who works in
a sustainability lab as a communication specialist and who has public speaking training, had the most overall disagreement with other raters in the visuals category, but the rater who showed a lack of confidence, Rater Two in this round, had the lowest levels of agreement overall, despite the fact that Rater Two expressed in meetings that (s)he attempted to guess what the other raters would choose and base his/her scores on that guess rather than making a best judgement based solely on personal observations of the speaker. It appeared that the attempts to guess at others’ ratings backfired and resulted in lower, rather than higher, agreement.

These first and second pilot rounds of coding helped expose some weaknesses in the rubric that were addressed, making the rubric and accompanying code book more effective and consistent for coders to use. The agreement was much stronger in several of the constructs, which was encouraging. It was now time to have a new set of raters work with the revised rubric and a revised code book so that raters with no pre-conceived ideas about the rubric or the process could start with a blank slate in evaluating the revised rubric. Therefore, pilot testing ended and the project moved forward to the official testing mode in the Assessment Implementation step of the ECD process.

To summarize, in this CAF chapter, I described the five knowledge, skills and abilities addressed in the rubric. I then explained through the evidence model that a public “library talk” of at least 10 minutes will be the evidence collected to demonstrate a scientist’s speaking ability. I discussed the use of the PSCR and its place in the rubric
development and discussed the development of the first draft of the rubric that implemented the five KSAs necessary for the science portion of the rubric. I then discussed how the five KSAs turned into six with the division of the visuals construct from the language construct. I also discussed the two rounds of pilot testing and the results of those rounds of testing. The agreement in these rounds shows room for improvement, but when looking at the data, it shows that the raters do seem to be understanding the constructs in basically the same way, and that the agreement is generally trending in the same direction for each construct, with visuals and processes being the lowest measures, but still showing reasonably strong adjacency agreement, except with the visuals in round two. However, the disagreement in visuals was nearly all from rater three who had to rush through the ratings, so it was important to revisit that construct with new raters before adjusting the construct. So, while the initial testing numbers could have been stronger, there was good evidence that the constructs were being understood in similar ways and that the rubric is testing the desired constructs.
CHAPTER 5.0 ASSESSMENT IMPLEMENTATION

After completing the preliminary review and pilot testing of the Assessing Public Presentations by Scientists (APPS) to catch major inconsistencies and challenges with the rubric, it was time to enter the assessment implementation phase of the project. In ECD, the assessment implementation phase is where the operational elements are finalized. As explained by Mislevy & Haertel (2006), the assessment implementation layer of the ECD concerns “constructing and preparing all of the operational elements specified in the CAF.” This includes producing test forms, finalizing rubrics, providing examples, and finalizing the task instructions. Most design decisions are finalized in this layer (R. Mislevy & Riconscente, 2005), with each decision being based on the principled, coordinated rationale of the assessment argument (R. Mislevy & Haertel, 2006). In the case of the APPS rubric, each decision is based on the stated purposes of the assessment, including providing formative and summative assessment for scientists who are working on improving their presentation skills for lay audiences.

For this layer of the project, four raters tested the rubric by applying it to a number of different presentations by scientists. The presentations were chosen from the list of 79 available presentations by using a random number generator. The randomization was done with replacement so that no presentation was assessed more than once by the same group of raters. The four raters who participated in the assessment implementation layer of the project were recruited from a large, land grant
university in the Southeast. The raters were paid an hourly wage for their participation in the study. Coding meetings and norming sessions were held via the Zoom virtual meeting application, professional version. This application allows all participants in a meeting to see each other and allows any member of the group to share her/his computer screen with everyone else in the group. In this way, all coders could view and assess presentations simultaneously.

For the purposes of the study, it was important to have both communication experts and science experts as raters. Since previously cited research shows that scientists are reluctant to invite communication specialists to collaborate with them on communication training, one of the purposes of the study was to see if learning about the applicable communication theories and techniques through the code book and rubric would help scientists be able to effectively assess scientists speaking to public audiences even without formal communication training.

One coder was a second-year doctoral student in a communication, rhetoric, and digital media program. This coder has over three semester’s experience teaching public speaking, and has no science training. The next coder is a PhD student in the same program. This coder has taught public speaking more than three semesters, and has also taught extensively in science communication with an emphasis on environmental communication. The third coder has completed a master’s degree in fisheries, wildlife, and conservation biology. This coder had no experience teaching public speaking, and is
trained as a research biologist/ecologist. The fourth coder is in the fifth year of a PhD program in statistics, and has worked with multiple scientists on research projects. The coder has no experience teaching public speaking and identifies as a trained scientist.

Table 5.1 shows the participants and their qualifications.

<table>
<thead>
<tr>
<th>Rater</th>
<th>Qualifications</th>
<th>Participation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comm Specialist 1</td>
<td>2nd Year Doctoral Student 6 Semesters teaching Public Speaking No science background or interest</td>
<td>Both Rounds – participated in final meeting via audio and written responses</td>
</tr>
<tr>
<td>Comm Specialist 2</td>
<td>PhD Student – Completed Final Defense 6+ Semesters teaching Public Speaking Environmental Communication emphasis</td>
<td>Both Rounds – Missed final meeting, sent comments</td>
</tr>
<tr>
<td>Scientist 1</td>
<td>Master’s Degree in Wildlife/Fisheries No Public Speaking</td>
<td>Both Rounds</td>
</tr>
<tr>
<td>Scientist 2</td>
<td>Trained and active as a research scientist Fifth year PhD student No public speaking background Worked w/ scientists on research projects</td>
<td>Both Rounds – 2nd Round Outlier</td>
</tr>
</tbody>
</table>

5.1 Initial Norming Session with Coders

Prior to the initial meeting with the raters, I sent them the newly-revised rubric (the science communication page only – See Appendix E), the code book, and a link to one of the science presentations. I asked that they review both the codebook and the rubric for clarity and ease of use, and make note of any suggested revisions. I then asked that, after reading through the codebook and rubric, they perform a practice assessment by viewing the presentation to which I had sent them a link and use the rubric to assess the speaker. At our first meeting, the first item on the agenda was to gather comments
about the code book and the second was to gather any suggestions raters had for revision and increasing clarity in the rubric. The initial discussion quickly focused on the rubric as much as the code book, including suggestions about design and wording in both documents.

Raters had many comments about the design elements of both the rubric and code book. All of the raters said the code book was useful, but that they had to go back to the code book several times to read the descriptions and decide which score to give in a particular category. Two of the raters suggested that the visual design of the code book be changed, making better use of bolding, bullets, numbers, or similar hierarchical devices to make it easier to spot the different criteria that applied to each construct. Another coder suggested simplifying the language in the code book or highlighting single words or short phrases to help with clarity. Similarly, they all said the rubric was hard to read quickly as they were assessing the presentation. There was so much type in each box that it was hard to see quickly what KSAs they should be watching for. Coders also said they would like to see bold, bullets, or numbers used in the rubric so they could easily see what KSAs applied to the construct as they were scoring a presentation. Two of the coders said they liked the examples in the code book, but did not like having to jump back to the code book to be reminded of the examples as they were scoring, and the other two coders agreed.
Raters also had suggestions for the wording in the rubric and the code book. One of the specific suggestions was to use the same words in the code book descriptions as are used in the rubric descriptions. The raters thought that using more of the identical words would help make the meaning clearer and the assessment easier. Additionally, two of the raters said they had trouble choosing between scores, such as a two and a three, and thought that more specificity and clarity in the qualifier words (such as “mostly” “often” and “rarely”) would help. They also pointed out that the actual criteria differed in some of the rows. For example, in the visuals construct, I used the word “graphics” in the “four” box, but used “charts and graphs” in box “three.” Additionally, in the trustworthy construct, in box “two” it read, “lacks self-disclosure,” but in box “four,” it doesn’t mention self-disclosure as being an important measure. The coders also pointed out that sometimes the qualifier words seemed very similar in two adjoining cells. For example, in one of the measures, both the “zero” cell and the “one” cell said that the speaker showed “no” sign of the quality being measured.

We also spent a significant portion of our meeting time discussing the layout of the rubric. A universal complaint from the coders was that there was no space to write notes or justifications for their scores on the rubric. Two of the coders specifically said they wanted to write down the words and phrases a speaker used that supported each construct. The other two coders agreed that they wanted room for notes. Another coder asked for bullets in each rubric box so that the measures were easier and faster to
identify. I expressed concern about using bullets, since that would certainly make the rubric boxes too large to fit on a single page and possibly too large for two pages. One of the coders had developed an alternative layout that gave room for notes, but did not use wording from the rubric. We discussed the need to test the filled-cell rubric as described in prior research, and brainstormed how to meet their needs as assessors while still maintaining the format and benefits of the filled-cell rubric. I expressed my desire to have the rubric be a practical, useful tool for assessors, not simply theoretical. I suggested making the changes we had discussed to the actual rubric wording and design and then adding a note-taking/scoring sheet that would be printed on the reverse of the rubric. This sheet would have a space for note-taking, would contain one or two examples from the code book, and would highlight some of the qualifier words for each measure. I agreed to look at a way to incorporate such a note-taking page while maintaining the integrity of the rubric.

We had a rather extended discussion about the use of the word “may.” Some coders felt that the word indicated that the speaker “should” do each thing that was preceded by the word “may,” while I expressed my intent that the word should indicate that it was not necessary for the speaker to demonstrate each item listed in the rubric, but instead was expected to include at least some of those items or discuss one or two of them in more depth. At the end of the discussion, we agreed that my intent was not made entirely clear through the use of the word “may” and that it would be important to
try to word the rubric in such a way that the intent was clearer. In other words, it was important that coders understand that no one criterion was non-negotiable, but that some combination of the criteria in each construct should be present. The score a rater gives would reflect how expertly that criterion was met. For example, in the “processes” construct, a speaker need not tell about methods and decision making and data gathering and uncertainty and working with a team. Instead, he or she would talk in depth about at least one of those concepts or in some depth about two or three of them. The rater would assess how well the speaker did at conveying some explanations of processes in science and award a score according to the skill displayed by the speaker.

Other agenda items included reviewing the scores each gave on the sample presentation and discussing reasons for the scores. We discussed the meaning of relevance and clarified the meaning of that construct. We discussed that if a speaker made a brief mention of some reason that her/his research is important at the beginning of the presentation, it would mean that speaker scored higher than a zero in the relevance category, but would not necessarily score a four. Instead, a person would score a four if (s)he gave several strong reasons for the importance and relevance of his/her research or if (s)he mentioned the reasons several times throughout the presentation. We then scored another presentation based on our discussion of the rubric and compared scores.
5.2 Revisions

Following the first coders meeting, I made extensive revisions to the rubric. As I suspected, trying to use bullet points made the boxes too large, and I quickly ended the attempt. However, I did go through each cell and attempt to make the measures being assessed consistent through each scoring level. Additionally, I made the key words in each cell bold to help them stand out more. Throughout this process, it was necessary to make the wording more concise and direct in order to fit in the cells and remain a one-page document. This process made the rubric more precise and more descriptive.

The next project was creating the note-taking/scoring sheet. I wanted it to be useful, complementary to the rubric, and consistent with the measures as written in the rubric. I began with the number and title of each construct just as it is portrayed on the rubric. I then pulled the key words from the description and put those under the title, rather than using the full sentence, as was found in the rubric. I then created a space about 5” wide for use while taking notes. On the right, I created boxes for scores from four to zero, but only included the qualifier words in the boxes on the rubric. For an example of the layout for one construct, see Figure 5.1 below.
I emailed the revised rubric, including the note taking/scoring sheet, to the coders along with links to the five presentations that were to be evaluated in the first round of coding. In the email, I included instructions for using the scoring sheet and a reminder of the deadline for returning the scored rubrics. See Appendix F for this version of the rubric.

5.3 First Round of Testing

Each of the coders returned the completed rubrics before our next scheduled meeting, which was held one week after our first meeting. Results showed solid levels of agreement in some areas, and evidence of confusion in other areas, as shown in Table 5.2.
Table 5.2 Round 1 - Full Testing

<table>
<thead>
<tr>
<th>Measure</th>
<th>Simple Agreement</th>
<th>Adjacent Agreement</th>
<th>Cohen’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Data</td>
<td>36%</td>
<td>83%</td>
<td>.14</td>
</tr>
<tr>
<td>Relevance</td>
<td>22%</td>
<td>90%</td>
<td>.059</td>
</tr>
<tr>
<td>Language</td>
<td>32%</td>
<td>96.6%</td>
<td>.169</td>
</tr>
<tr>
<td>Visuals</td>
<td>28%</td>
<td>86%</td>
<td>.118</td>
</tr>
<tr>
<td>Processes</td>
<td>33.3%</td>
<td>76.6%</td>
<td>.082</td>
</tr>
<tr>
<td>Trustworthy</td>
<td>30%</td>
<td>56.6%</td>
<td>.069</td>
</tr>
</tbody>
</table>

Each of the first four measures had strong adjacent agreement, with language and relevance showing high adjacent agreement, indicating that these measures are consistently seen in the same way and do measure what they are intended to measure.

The visuals category also had strong adjacent agreement, which was a positive result since the visuals category had an extremely low level of agreement in the second round of pilot testing. This testing seemed to confirm that the problem with the visuals rating was related to individual raters and not the construct. Since this was the first round of testing for these coders, it was encouraging that they were able to reach strong levels of agreement in three of the five constructs on the first round of testing. Nonetheless, there was need for improvement in both the process construct and the trustworthy construct.

Discussing the trustworthy/human and process measures became a primary agenda item for the next raters’ meeting, particularly the trustworthy/human construct. I wanted to see if I needed to do more training with the raters, since the first group of raters initially
had difficulties with those two constructs, or if we needed to do more adjusting to the constructs themselves and the wording of the rubric.

5.4 Revisions to Rubric and Code Book

After analyzing the results of the first round of testing, I held a second raters meeting. The goals for the meeting were to see what went well with the first round of coding and how we could clarify the APPS and code book to increase agreement across the board, particularly in the processes and trustworthy/human categories. The raters did say that they like the note-taking/scoring sheet to help the scoring process and found it useful to have, particularly as they were trying to score a speaker at the same time as they were watching a presentation. They did suggest adding a section to the code book about the overall philosophy and about how to use the rubric and the note-taking page.

Next, I turned to the areas where the raters lacked agreement. The first area we examined was the category of trustworthiness. Some of the raters were confused about including “trustworthiness” in the same category as “personable.” They felt that a measure of trustworthiness should be found in another category equating to “knowledgeable in the field,” or “credible to speak about this topic.”

We discussed the fact that “trustworthy” in this case did not mean “credible,” as in “they know what they are talking about and they’re qualified to have an opinion on this topic.” We discussed the literature that says that people in general respect scientists,
but that they have a very different opinion about many topics related to science than many scientists do (Funk & Rainey, 2015). Therefore, the question in not whether the scientist is credible as a scientist, but whether an audience member would trust a scientist as they would a friend or trusted advisor, and whether the audience members would listen to their opinions about science topics. One of the raters suggested thinking about the questions, “Would I want to sit down and have coffee with this person?” or “If I had thousands of dollars to fund research in this area, would I give it to this scientist?” as a way to understand this measure. We discussed adding verbiage to the code book to make the definition of trustworthiness clearer. The coders also discussed the idea of feeling inspired by the speaker, inspired by the research or inspired by the scientist’s passion and suggested adding that wording to the rubric.

Some of the raters were again counting the number of times the speaker said, “we,” “us,” or “our,” and giving the speaker a high score in trustworthiness if they used those words frequently and a low score if they used them never or rarely. We discussed the fact that the trustworthiness measure is a combination of the inclusive language, enthusiasm about the topic (which can be expressed in a variety of ways), and positivity about science and scientists. We discussed again that no one factor should be considered the deciding factor in the score the speaker earns. One of the speakers we rated as part of the meeting had an approachable, personable demeanor which included some humor related to the science topic. The coders suggested that being able to joke about the
science was another way to build a rapport and trust with the audience. The communication concept of immediacy measures feelings of warmth and closeness, and humor is one of the measures that can contribute to those feelings of closeness. Therefore, a scientist who is able to add science humor to a presentation about science could increase the audience’s trust in her/him as a scientist. Following this discussion, I added some immediacy measures, including humor, to the rubric and codebook, but with the emphasis that raters should be measuring the speaker’s humor when it comes to science, not just general humor, and measuring whether the person seems trustworthy as a scientist, not simply on a personal level (i.e., as a mother, child, worker, etc.) If a speaker does use these immediacy measures, it should increase trust in the speaker as a scientist.

Another area where it was important to increase agreement was in the measure of increasing the audience’s understanding of scientific processes. I found that the assessors with science backgrounds wanted the speakers to give more explanation of processes, more detail in the explanations, and more background on why the decisions were made. For example, during one presentation, one of the coders with a science background gave the speaker a score of two, while the coders with communication background gave the same speaker a four rating. The science coder said that the explanation from the speaker was “probably right for the audience” but the rater felt that the rubric required that the speaker give more detail. We discussed the idea that if explanations were appropriate
for the audience, then they were appropriate for this situation and should be awarded a higher score. In fact, the rubric is specifically designed for use in the context of a public audience, and therefore if the speaker’s presentation was right for a public audience, it was right for the rubric.

The coder with the science background (Scientist 1) expressed that she had too many questions after hearing the brief explanations given by the speaker, and said that there was no way the experiment or methods could be replicated based on the information given in the presentation. We discussed the fact that these presentations did not have the same purpose as a scientific article, which would give full details so an experiment could be duplicated. Instead, the explanations should be general, giving an overall overview of science processes and concepts as applied to this specific research project rather than an in-depth explanation of this research project. Since the rater had this concern on more than one of the presentations, it was important to try to mitigate those concerns. We discussed ways that we could add information to increase this understanding of the “processes” construct into the language of the code book and the rubric as much as possible so that another scientist with a similar bias who was trying to use the rubric would be better able to understand the measure. The other coders did not express this same confusion about the measure at this time, but it was important to know that this scientist did have this difficulty, since one of the primary intended audiences for the rubric is scientists who are training other scientists to be good public
communicators. The areas in which this scientist was most demanding of the speakers were areas in which the scientist had personal expertise, while the scientist was more generous and less demanding of speakers who spoke on topics that fell outsider his/her areas of expertise. It was useful to recognize that scientists acting as evaluators of other scientists’ public presentations may have difficulty separating their inherent knowledge and the demands of their particular field from their work as a rater.

Despite relatively high scores in the measure of “relevance to listener,” the coders wanted to discuss the construct, since they did have questions regarding scoring the measure. First, they expressed the concern that it was difficult to know if the audience was concerned about a particular topic such as topsoil or food prices or fossil fuels. They indicated that it would be necessary to simply make a guess, and that they could not know for sure. We discussed the fact that we did, in fact, need to make a “guess,” or a reasonable judgement as to whether the speaker made effective attempts to relate the topic’s importance to the audience in a way that could reasonably be assumed to have some relevance to many members of that general audience (i.e., college students, older adults, families, teens). Although the rater clearly won’t know the attitudes and interests of each member of the audience, there are some general ideas and values that are often held by particular groups of people. If the speaker talked about how the topic related to some or one of those values/ideas, then it could be said that the speaker was making the topic at least somewhat relevant to the audience, and the speaker would score higher
than a zero. The actual score between one and four would be determined by how clearly, how well, and how often this relevance was indicated. One of the coders suggested asking the questions, “If I were a member of this audience’s demographic, would I feel that this scientist is trying to improve my life? Would I feel that (s)he is trying to improve the world I live in?” I agreed that I would add that question to the code book and would add wording to the “relevance” construct to further clarify the ideas we discussed.

After this meeting, I revised the code book and rubric according to the adjustments we discussed. I changed some of the language in the rubric as we determined, and I added more clarification and explanation to the code book. I added the questions that the coders suggested that a person using the rubric might ask themselves when trying to assess particular measures. Additionally, I made changes to the wording and layout of the individual sections of the code book in an attempt to make them more user-friendly and easier to use. As an example, the revised version of the “metaphor” section of the language construct is found below:

**Make thoughtful use of metaphors/analogies:** One technique that increases clarity is the use of metaphors and other comparisons. Research identifies several different types of metaphors and comparisons; however, the important thing for this assessment is that, when comparisons, metaphors, or analogies would help explain a complex point, the speaker uses them, and does so deliberately, as a way to help the audience understand.
Essentially, when scientists use comparisons, they should use deliberate comparisons that are developed to increase audience understanding and help with the clarity of their speaking. They should avoid metaphors and comparisons that are based on knowledge inherent to being a scientist and are not helpful to people who do not have scientific expertise (called cognitive metaphors).

**Example 1 – Better (pedagogical, or teaching, metaphors) – messenger RNA**: proteins are building blocks for the cell

**Example 2 – Better (theory-building metaphor that explains a process)** – The cell contains what we call a genetic code, a blueprint for building a new cell...

**Example 3 – Worse (cognitive metaphor)** – so the polymer chains make substances that are plastic-like or these organisms all come from the same family, so… (metaphors are used instinctively, not deliberately, and are not used to clarify a concept or practice – mixing metaphors)

An expert speaker will use a few metaphors/comparisons throughout the presentation rather than offering up multiple and/or contrasting metaphors in the same presentation. The best metaphors/comparisons compare the new object or concept to something that is familiar to audience members, and they give an accurate picture of the scientific concept.

After completing the revisions, I then I emailed each rater a copy of the revised APPS rubric and the revised code book (see Appendix G) along with a list of ten more randomly selected presentations for them to assess. The raters did not think they could complete ten assessments in a week, and asked for two weeks to complete assessments on ten presentations by scientists.

### 5.5 Second Round of Testing

After receiving the scored rubrics for the last ten presentations, I analyzed the results. Although I anticipated strong agreement, based on the last round of testing and
the improvements made to the rubric and code book since the last round of testing, the results were not as strong as I had hoped overall, but there were several encouraging findings. Table 5.3 shows the results from the second round of final testing.

Table 5.3. Results of Final Testing, Round Two

<table>
<thead>
<tr>
<th>Measure</th>
<th>Simple Agreement</th>
<th>Adjacent Agreement</th>
<th>Cohen’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Data</td>
<td>39%</td>
<td>82.6%</td>
<td>.114</td>
</tr>
<tr>
<td>Relevance</td>
<td>33.3%</td>
<td>65%</td>
<td>.068</td>
</tr>
<tr>
<td>Language</td>
<td>42%</td>
<td>90%</td>
<td>.128</td>
</tr>
<tr>
<td>Visuals</td>
<td>40%</td>
<td>83.3%</td>
<td>.148</td>
</tr>
<tr>
<td>Processes</td>
<td>33%</td>
<td>90%</td>
<td>0</td>
</tr>
<tr>
<td>Trustworthy</td>
<td>45%</td>
<td>83.3%</td>
<td>.196</td>
</tr>
</tbody>
</table>

As is shown in the table, overall simple agreement measures increased in most cases, with the exception of a .3 drop in the processes construct, which was a positive finding. Adjacent agreement stayed about the same for the overall, language, and visuals constructs, although relevance was an anomaly, as will be discussed. Most encouraging was the strong increases in adjacent agreement, in the processes and trustworthy/human constructs as well as strong increases in Cohen’s and simple agreement in the trustworthy construct. This indicated that raters were seeing that construct similarly, and discussion showed that was largely true.

I was interested in discussing with the coders some of the possible reasons for discrepancies. It was encouraging that the changes made to the rubric and codebook,
which all of us viewed as significant improvements, did seem to help the coding. However, when examining the details of agreement, there were some definite outliers, and I was interested in hearing explanations.

I did have at least one idea of a factor that may have contributed to the lack of agreement: time. I asked raters to send me their evaluations at least a day prior to our next coder meeting, if at all possible. One of the raters finished the coding the next morning, since a new job was affecting the ability of the rater to work on this project after that date. The end to funding due to the job change caused that particular rater to need to rush through the job and be unable to watch presentations repeatedly while scoring, although (s)he said (s)he was able to complete the ratings within the hours (s)he was allowed for the purpose. One other rater submitted completed rubrics two days before the meeting, while the last two submitted rubrics the night before the meeting and early on the morning of the meeting, respectively. Therefore, one of the raters had to hurry through the job, while three of the raters finished (and perhaps started) the scoring after a gap of two weeks from our last discussion. Although I was unable to test the idea that these time-related issues might have led to raters paying less attention to the project or to forgetting some of the coding issues we had discussed, it seems possible that the time lapse between discussing the constructs and coding the presentations added to the various other concerns (new jobs, conference presentations, grading, etc.) that the raters said interfered with their abilities to complete the coding more promptly or under less
time pressure may have had an effect on the rater’s abilities to code. Other factors that
might have had an effect on the raters were the fact that they had ten presentations to
score rather than five as they had in the first round. Scoring ten rather than five
presentations gives more opportunity for disagreement and could lead to rater fatigue,
especially if the raters tried to complete all scoring at one sitting. Additionally, I felt the
scoring might have been affected by the fact that there were some presentations included
in the sample that I considered challenging to score.

At the raters meeting, Comm Specialist Two, the rater with public speaking
expertise (and science communication expertise) was not able to attend the meeting
because of the previously mentioned limits on employment hours, so I relied on the
notes they offered for input. The second rater with public speaking experience (no
science communication experience – Comm Specialist One) attended via the Zoom app,
as did the two scientists, but a weak Wi-Fi signal at the Comm Specialist One’s hotel
meant that we could not see the video feed from the hotel and that comments were
contributed via the text function in the Zoom program rather than orally.

The raters did have a few more specific ideas for improving the rubric and code
book. For example, Scientist Two did suggest that the word “knowledge” be removed
from the heading of the “Trustworthy” construct, since we had discussed the fact that
the construct is not measuring how much science knowledge the speaker has. Although I
had changed the wording in the rubric cells, I failed to change it in the top-level title box.
Additionally, raters once again noted that the differentiation between a rating of two and a rating of three was difficult for them to make. They suggested that there should be more distinct qualifier words to help them make the judgement in all measures, but particularly in the visuals measure. They also asked that the descriptions in boxes three and four in each measure be sure to say something positive that they should look for that was in direct contrast to the things noted as negatives at the lower ratings. One last suggestion was that I use a more descriptive word than “clear” in the visuals section. They noted that the word, particularly when applied to graphics or images, can have several meanings such as the following: crisp and sharp in terms of a resolution, easy to see, large, or comprehensible.

At the meeting, I let the raters know about our lower levels of agreement in some categories and then discussed those areas with the most variation, or the widest range, in their responses. While discussing these specific situations, some interesting issues arose that shed additional light on some of the discrepancies and why they arose. For example, both scientist raters expressed the idea that, based on our adjustments to the rubric, they felt that a speaker needed to be bubbly and outgoing in order to be trustworthy. One of the scientists said that to her, the more extroverted speakers were more convincing, although based on the rubric, she had to give a good score to those who were not as bubbly or extroverted but who conveyed a sense of excitement and commitment to their work in other ways. The second scientist said that it was impossible to tell the level of
excitement or interest from the speaker’s behavior or words, and that she found the construct difficult to measure. We did discuss the fact that extroverts can have an advantage in some ways, but that audience members and raters do observe other verbal and nonverbal cues that give them an impression of trustworthiness and relatability. Comm Specialist One agreed, saying that speakers who convey their knowledge through humor and/or self-disclosure can seem equally credible as an extroverted speaker. Speakers do not need to be exuberant or outwardly social to give the sense that they genuinely care about their topic and that they are good people to have working on the issue.

Another interesting discussion came when talking about the discrepancies in the scores the raters gave a person who spoke about new battery technology that he was developing. Some raters thought the speaker had done an adequate job of stating the relevance of the topic, since he had shown a picture of an exploding laptop battery and talked about how his innovations would help avoid such problems. Scientist One, however, scored the speaker low on the relevance measure, since he used “something that will never really happen” and “was far-fetched” to try to show the relevance of his topic. Other raters immediately mentioned actual instances of exploding or burning batteries in computers and cell phones, such as the recent Galaxy 7 problems and Dell computer battery melt downs. It was interesting that Scientist One’s lack of context in that case caused her to give the speaker a lower rating than others who had a context for
exploding batteries and thought that the speaker did an adequate job. After the discussion, all agreed that the speaker could have referred to the specific incidents by name (possibly saying something like, “...similar to the Dell laptop battery problems in 2006 and 2017”) rather than assuming that everyone would know about lithium battery explosions and fires as actual incidents.

Another discussion that focused on the construct of visuals shed additional light on some of the discrepancies in the reliability measures. Speaker 50 was given high marks (3 or 4) by all raters in all categories but one, with the exception of the scores in the visuals measure. In visuals, two raters gave the speaker a score of two and two raters gave him a score of three. This led to some discussion of differentiating more clearly between ratings of two and three, as I mention previously in this section. However, even more revealing was the discussion from those who rated the speaker’s visuals as three. All raters said that the speaker did well at using actual objects as visual aids, but that when it came to the PowerPoint slides, some slides were cluttered with too many elements. One of the raters who gave the speaker a three on visuals was Comm Specialist One, who said that the speaker had more great visuals that were really helpful than he did cluttered visuals, and so the rater felt the well-done visuals were more helpful than the cluttered visuals were harmful. This led her to weigh out the positives and negatives of the visuals in the presentation and award a score of three. The second person who gave a score of three was Scientist Two. The reason this rater gave for giving
the speaker a three rather than a two was that she (the rater) understood what the
speaker was saying in the presentation (the speaker did have excellent speaking skills)
and therefore the cluttered visuals weren’t a problem because she didn’t need them to
understand the presentation. In other words, she disregarded the elements of the rubric
that were specific to the construct being measured because the speaker did well in
another construct that, to her, was related.

This tendency to rely more on inherent knowledge or feelings than on the rubric
was more pronounced when the raters were scoring speaker number 49. When I saw
that this presentation would be one of the presentations evaluated by the raters, I was
immediately interested in seeing the results of their scoring. This speaker had a great
presence, was outgoing and passionate about the science work, and connected well with
audiences. However, in the specific presentation randomly selected for viewing
numbered 49, this speaker chose to minimize the amount of speaking she did about the
actual science work that was important to her research and instead focus on portraying
the wonder, awe, and joy of science and working in science. When the scores were
complete, the two public speaking experts had given the speaker high scores for
language, visuals, and trustworthiness, but low scores (one) for relevance and explaining
the processes of science. However, the two scientists were taken in by the engaging, awe-
inspiring attention-getter, which occupied the first half of the scientist’s speaking time,
and awarded the speaker fours in relevance. The discrepancies in scoring this speaker alone accounted for a significant drop in the overall relevance agreement.

When I discussed presentation #49 with the raters and asked the two scientists what the speaker had said that specifically made a connection between the relevance of her science work and the lay audience hearing the presentation, they both realized that they didn’t actually note anything that was said. They both rated the speaker high on relevance when what they were really scoring was the trustworthy construct. In fact, when I told the Scientist Two that the speaker’s research was in turtles, not swarm theory (the science concept highlighted in the speaker’s engaging, attention-getting video), the second scientist had no memory of turtles being a part of the presentation, much less the focus of the speaker’s study. Scientist One did score the speaker lower (a two) in the processes construct, as did both communication experts, but Scientist Two again scored the speaker high (three) in processes, even though Scientist Two was not clear on the type of work the speaker did and even less clear on how she did it. Scientist Two was so taken with the speaker’s basic public speaking and trustworthy construct skills that it seemed she was unable to objectively score the speaker on the other science communication constructs.

The last area of discussion that was enlightening was a discussion on the decrease in adjacent agreement in the relevance construct (although simple agreement increased by 11.3%); a category in which the raters had shown high agreement in the first round of
scoring. We discussed specific instances in the relevance category in which Scientist Two had given a score of four, the highest score possible, and the other raters had given a score of one or two. In all instances, the coders who had given the lower scores said that they felt the speaker had only alluded to a possible relevance with their science without making the relevance clear. Scientist Two, who gave the highest possible scores, said, in all cases, something similar to this: “Well, I have background in (science concept) so I know that (science concept) is related to (another science concept) so when the speaker mentioned (this concept), I guess I just put it together in my mind. I mean, I didn’t need the speaker to draw me a path or anything; I know how those things connect, so I thought the speaker was pretty clear.” In this case, the rater specifically did not take the point of view of a lay audience member, but instead rated the speaker as a fellow scientist and relied on the inherent scientific knowledge possessed as a scientist rather than rating the presentations as the rubric specifies; as appropriate for public, lay audiences who do not have the science expertise.

While scoring and interrater reliability agreement in the final round may have been affected by the speed with which the ratings were done and the amount of time that had passed between the coding meeting and the actual rating being done, it was nonetheless clear from the discussions in the coding meetings and the scores given that both scientists, and particularly Scientist Two, had difficulty divorcing themselves from their inherent expertise in order to rate a speaker for a public audience. Scientist raters
had difficulty putting aside their personal scientific knowledge and their personal
expectations of scientific talk and behavior in order to accurately score science
presentations for public audiences. While their ability to take the perspective of a lay
person was better when assessing some speakers than others (particularly speakers who
spoke about topics that were outside the rater’s areas of scientific expertise), the results
of these coding sessions do seem to reaffirm the advisability of involving communication
experts in the assessment of public presentations by scientists.

After the final coding session and after having more time to examine the data
more carefully, I made two observations. First, many of the disagreements, or variations
in coding, seemed to come from Scientist Two. In fact, removing scientist two from the
ratings increases the overall adjacent agreement to 86%. Second, despite some errors in
rating and some complete disregard for instructions, the simple agreement ratings did
improve between the first and second rounds of official testing and the process and the
trustworthy/human adjacent ratings increased significantly, showing that those
constructs were being viewed more consistently by all raters. As illustrated by the strong
adjacent agreement, often the range of disagreement was only one point rather than the
three and four points of difference that were more common in the pilot testing stages.
Overall, as shown in Table 5.4, testing shows that, despite problems with the raters being
influenced by their intrinsic expertise, the APPS and the individual constructs in the
APPS do measure quite consistently the knowledge, skills, and abilities that are most
important for scientists to demonstrate when speaking to public audiences, and raters generally agree on how those constructs are demonstrated.

Table 5.4. All Rounds with Overall Scores Showing Simple and Adjacent Agreement

<table>
<thead>
<tr>
<th>Construct</th>
<th>Simple</th>
<th>Adjacent</th>
<th>Simple</th>
<th>Adjacent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>34.6%</td>
<td>83.2%</td>
<td>34.67%</td>
<td>80%</td>
</tr>
<tr>
<td>Relevance</td>
<td>36%</td>
<td>90%</td>
<td>46%</td>
<td>86.6%</td>
</tr>
<tr>
<td>Language</td>
<td>30%</td>
<td>83%</td>
<td>53.3%</td>
<td>93%</td>
</tr>
<tr>
<td>Visuals</td>
<td>42%</td>
<td>94%</td>
<td>26.67%</td>
<td>66%</td>
</tr>
<tr>
<td>Processes</td>
<td>34%</td>
<td>79%</td>
<td>20%</td>
<td>86%</td>
</tr>
<tr>
<td>Trust/Human</td>
<td>31%</td>
<td>70%</td>
<td>27%</td>
<td>86%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Construct</th>
<th>Simple</th>
<th>Adjacent</th>
<th>Simple</th>
<th>Adjacent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>36%</td>
<td>83%</td>
<td>39%</td>
<td>82.6%</td>
</tr>
<tr>
<td>Relevance</td>
<td>22%</td>
<td>90%</td>
<td>33.3%</td>
<td>65%</td>
</tr>
<tr>
<td>Language</td>
<td>32%</td>
<td>96.6%</td>
<td>42%</td>
<td>90%</td>
</tr>
<tr>
<td>Visuals</td>
<td>28%</td>
<td>86%</td>
<td>40%</td>
<td>83.3%</td>
</tr>
<tr>
<td>Processes</td>
<td>33.3%</td>
<td>76.6%</td>
<td>33%</td>
<td>90%</td>
</tr>
<tr>
<td>Trust/Human</td>
<td>30%</td>
<td>56.6%</td>
<td>45%</td>
<td>83.3%</td>
</tr>
</tbody>
</table>

In this chapter, I discussed the assessment implementation and the final testing of the rubric. The first round of final testing showed agreement that was quite strong from the outset, and indicated that with more work on the process and trustworthy constructs, the agreement could be strong across the board.

However, the second round of final testing did not have agreement that was quite as strong, although the trustworthy category did improve significantly. There were ten
presentations to assess and several of them were challenging to evaluate, requiring self-awareness and discipline on the part of the rater. Additionally, it seemed that several of the raters were rushed to complete the work on time. However, the primary factor in the discrepancy in some of the results seems to be the difficulty on the part of the scientist raters to remove their preconceived knowledge and assumptions from their ratings of the speakers. One rater, Scientist Number Two, seemed to be a particular outlier. When the ratings for Scientist Two were removed, the agreement for round two of the final testing in most of the categories and overall increased. Finally, the strong adjacent agreement in the final round of testing indicates that, despite the difficulties arising from some inherent biases, raters’ understandings of the constructs are largely similar and that the adjustments to the rubric and code book did help with overall understanding and assessment of constructs. The results also indicate that the selection and training of raters is of utmost importance when using the rubric to give formative and summative assessments of science speakers.
CHAPTER 6.0 ASSESSMENT DELIVERY, LIMITATIONS AND FUTURE RESEARCH

6.1 Assessment Delivery

The assessment delivery layer of the ECD is where “students interact with the tasks, their performances are evaluated, and feedback and reports are produced” (R. Mislevy & Haertel, 2006, p. 20). This particular study does not take the step of delivering the task to scientists, training them, and then assessing their presentations. However, the current code book and rubric are useful tools for use in assessment delivery.

Analysis of the overall agreement scores resulting from testing the APPS rubric, even when including the scientist coders who had difficulty separating their innate knowledge from their assessment tasks, shows that there is general agreement between the raters on the constructs and that the constructs overall are reliable measures of the KSAs most desired in science speakers.

Analysis suggests that with coders who understand the APPS and code book and who are trained to use it properly, the APPS can be counted on to measure those constructs. With trained communication raters, and with consistent practice and training in the use of the rubric and the meaning of the constructs, the rubric should give relevant results in both formative and summative settings.

When implementing this assessment as part of a training program, it is likely that the same assessors will be assessing the speakers over time and as they are trained.
While a training program for scientists may have the same people do the communication training as do the assessment or they may choose to have different people handle each aspect of training, consistency in raters will likely help the measures be more effective as the raters come to an understanding of how they as a team are interpreting the construct requirements.

To aid in this process, coders who are assessing presentations in the context of assessment delivery should be training normatively; that is, during norming sessions, they should be told explicitly when something is or is not appropriately scored and why. The APPS has been tested and revised enough that it is testing the desired constructs, and therefore it will be most helpful to consistent and productive scoring if the raters are giving similar feedback to scientists speaking to public audiences.

### 6.2 Limitations

Despite the extensive research and testing that accompanied the development of this Assessment of Public Presentations by Scientists (APPS) rubric and codebook, there are limitations to the study. First, the rubric could be more effectively tested if it were used by communication specialists who are training scientists in public communication skills. It would also be advantageous if the scientists being assessed had the benefit of being trained in public communication skills and of seeing and reviewing the rubric and codebook during their training. Testing in this study utilized pre-recorded presentations
given by scientists who did not have the benefit of being systematically trained in communication skills and who did not have access to the APPS rubric or the codebook.

Another limitation of the study is that I used random sampling to determine which presentations to use to both train raters and to test the rubric. For easier norming and better training, it could have helped to choose samples that were clearly worse than average and clearly better than average for the norming sessions. This would have given the raters baselines from which they could judge and could have provided them guidance as they tried to apply the rubric. Raters could have been asked to norm their ratings using easy-to-assess presentations (clearly good or clearly bad) and then immediately assess more challenging presentations. These practices could have helped raters develop more agreement more quickly.

Another significant limitation related to the first is that the recordings raters used to assess scientists’ presentations stopped at the end of the formal presentations, prior to the question and answer period. Therefore, we were unable to fully test one crucial element of the rubric, construct 12, engaging with the audience. Some speakers do engage with the audience in various ways during their presentation (such as by asking questions and waiting for responses), but many speakers waited until the end of a presentation to solicit questions. Therefore, the interaction element has not been fully tested. This is a primary limitation to this study, particularly since the domain analysis
showed interaction with the public as being a highly desired outcome of public presentations by scientists.

A limitation worth noting from the data-gathering portion of the domain analysis is that in gathering data from science organizations, I focused on public communication by scientists only, not public communication by communication experts or others in an organization. If, for example, a science organization sent its public relations or communication team to talk to legislators or a school group, I did not document the report of that interaction. If, however, the organization arranged for bench scientists to have a Q & A session with legislators in an attempt to explain the importance of basic research or the challenges inherent in a policy decision, I did document that interaction. It is possible that the interactions by science communication professionals warrant their own study.

Another area of note is in the visuals category of the rubric itself. The visuals construct focuses on PowerPoint, although it does acknowledge and allow for the fact that some speakers could choose to use physical objects rather than PowerPoint slides or as a supplement to PowerPoint slides. Although scientists are generally devoted users of PowerPoint, the reliance of the rubric on a speaker using PowerPoint does limit the person rating the presentations, since, although there are other means for visual communication, this rubric doesn’t make significant allowance for those other means.
Finally, anyone using this rubric to help train scientists to make presentations to public audiences should recognize that this study does not analyze or report on the use of basic public speaking skills by scientists. Those skills are vitally important to scientists speaking to public audiences, and must be included in any training program. Those important skills are included in the rubric, thanks to the inclusion of several elements of the PSCR from Schreiber, et al. (2013), with more detailed descriptions of the constructs found in the code book for training purposes. However, it is strongly recommended that a person with a grounding in communication research and practice be employed to at least assist with the instruction of basic public speaking skills in a training program for scientists. Such an expert will be able to help with such things as reducing public speaking anxiety, creating powerful introductions, using helpful organization, and other valuable public speaking skills.

6.3 Future Research

While there are several possible studies that could extend from this effort to create an assessment instrument for public science communication, future research should include at least two different studies. First, future study should include analysis of the actual assessment implementation; the testing of the rubric in a natural setting, assessing scientists who have been trained in public communication skills and who have been taught the desirable knowledge, skills, and abilities for scientists who communicate with public audiences. Additionally, scientists in such a study should have access to the APPS
and codebook so that they are familiar with the constructs they are trying to achieve. If the APPS and codebook could be utilized under these conditions, ideally with guidance from a communication expert who assists the scientists in their training, they will provide the most realistic and productive results. Future study could assess the results of testing under those circumstances to see if the rubric tests the measures it is meant to test and is a reliable measure of those constructs in that situation.

A second area for future study is testing the engagement construct of the rubric. This ability to engage with public audiences is valued by virtually every stakeholder who speaks about science communication, but due to limitations in the available data, raters were not able to code for this measure. It is likely that the evidences of knowledge, skills, and abilities in this area as currently written in the rubric will need to be modified and the language adjusted in order to most effectively measure the construct. This work could take place simultaneously with the testing of the rubric in a natural setting, or it could be done as a separate study.
7.0 DISCUSSION

The process of constructing the Assessment of Public Presentations by Scientists revealed a number of findings that add to the body of knowledge regarding science communication and particularly public presentations by scientists.

7.1 The Need for Public Science Communication and its Assessment

The political climate in 2017 has brought to the forefront one of the primary reasons more public communication by scientists is critical. Recent events, such as the defunding of science programs, exiting the Paris agreement, and more vocal skepticism regarding climate change, illustrate what this study shows in Chapter 1: that although the public in general hold science and scientists in high esteem, individual members of the public often do not accept scientists’ findings and recommendations as useful or important to follow. The gap between scientists believe and what members of the general public believe about generally accepted scientific findings is sometimes wide, and scientists believe they can help close that gap as they spend more time interacting with public groups and building trust with public communities.

However, this study shows that many scientists do not engage in this outreach work. Although there are a variety of reasons for the reluctance, one of the primary reasons is that scientists do not feel qualified or competent to engage in public communication. If interactions with the public are carried out without focusing on the specific skills and outcomes important for public communication, then public science
communication will not have its desired effect. In fact, in the worst-case scenario, poor public science communication could have deleterious effects on public perceptions of scientists. As this study notes, one of the most frequently noted ways to improve and increase public science communication is improving and increasing the communication training that scientists receive. As outlined in the prior chapters, effective assessment is crucial to better communication training for scientists. Assessment makes expectations clear, allows meaningful feedback, creates a clear direction for growth and improvement, and builds scientists’ comfort and confidence in public speaking situations. Current rubrics are either not suited for assessing oral presentations or do not adequately integrate communication theory and are therefore less effective.

7.2 Defining the Desired Outcomes of Public Science Communication

As proposed by Wiggins (1989) in arguing for better educational assessments, before deciding to assess individuals from a group of people, it is necessary to know what it is important that they be good at doing. It is critical to define what performance demonstrates ability and only then be concerned about how to assess those performances. This is also true when assessing scientists. Therefore, this study did a thorough examination of the academic literature about public science communication, both by scientists and by communication scholars who study public communication by scientists to see what scholars say are the skills and abilities scientists should display when engaging in public communication. It also examines the communication available
on the websites of science organization to see what these organizations tell their members they ought to do when engaging in public science communication.

This thorough examination of the literature combined with an iterative process of qualitative content analysis showed about 20 to 25 different abilities or skills scientists are encouraged to demonstrate when they engage in public science communication, and showed that all of these abilities can be distilled into about 12 categories that are consistently repeated as important by stakeholders from across the science and communication spectrum. No single science presentation is likely to demonstrate 25 different skills and abilities, particularly since some are highly specific, but the domain analysis showed that many of the most often and most emphatically mentioned topics could be operationalized for assessment. The results showed that the five most commonly and emphatically encouraged skills and abilities for scientists to demonstrate in public communication situations are:

1. Increase audience understanding of science/research topics and processes
2. Engage with the audience in back-and-forth communication
3. Interpret and clearly explain science information
4. Communicate the relevance and importance of science to each person
5. Provide accurate information so people can make good choices

Other important skills and abilities scientists should demonstrate that were included in the rubric were sharing the wonder, joy and excitement of science along with improving the reputation of scientists with the public by building the public’s trust in scientists. A common theme through all the categories was a concern about audience,
which is an important rhetorical and communicative principle. The concept of audience appropriateness is found throughout the rubric.

7.2.1 Additional Categories: Persuasion and Information to Make Decisions

There were at least two other important categories of knowledge, skills, and abilities important for scientists to demonstrate when speaking to public audiences that were identified in the domain analysis. One was showing persuasive skills, including advocating for more science funding, influencing policy decisions, changing incorrect ideas about science, and encouraging young people to go into STEM fields. The final category was giving people useful, accurate information and specific suggestions to help them make better decisions.

For several reasons, these two groups of skills were not included in the APPS. Primarily, the persuasive elements were not included in order to keep the scope of the rubric narrow and focused on the more common informative speaking elements. The category of giving helpful information and suggestions, although measurable using communication theories and practices, would have required more complex scoring and nuanced understandings of communication techniques than could be easily taught to non-expert assessors. For the purposes of this rubric, and the intended audience of scientists using a communication theory based rubric to assess other scientists, it was determined that simplicity would be advantageous.
7.2.2 Application to Persuasive or Decision-Making Measures

It is important to note that the rubric could be modified in the future to include these categories. Many of the elements of this rubric could be applied to a rubric measuring persuasive speaking. Construct seven, the relevance and importance of science concepts, would be crucial to a persuasive argument about science, because it would be vital for an audience to understand a scientific concept as being important to their own lives and to the well-being of the nation in order to be persuaded to do anything about it. Construct eight, using unambiguous language and language techniques appropriate for the audience, is also important in any situation involving a scientist speaking to a public audience. If the message can’t be understood, it won’t be effective, regardless of the purpose of the message. Similarly, good visual design is important for any science presentation to the public. Scientists speaking to fellow scientists often use complicated graphics that would be in comprehensible to public audiences, and therefore not persuasive. Creating trust, immediacy, and goodwill with the audience is possibly even more important in persuasive speaking than in informative speaking, and would be imperative when a scientist is trying to persuade someone to action.

The explanation of processes would be less important in a persuasive situation, and other elements would need to be added to a rubric assessing persuasive public science communication. For example, in a persuasive presentation, scientists are more
credible if they call on other credible sources, including other scientists, even though such bolstering of credibility is not necessary in an informative presentation when a scientist is accepted as an expert. When (s)he is trying to persuade an audience to take action, however, it is important to show a consensus of scientific thought. Additionally, a persuasive rubric would need to account for persuasive communicative techniques, including specific calls to action and building believable, audience-appropriate arguments. Still, as noted, many elements of the current rubric would be applicable to a persuasive rubric.

The skill of giving public audiences useful information so that they can make wise decisions is another construct that was not specifically addressed in the rubric. Nonetheless, it is likely that scientists who are building their audience’s understanding of the relevance and importance of science topics and using descriptive, audience-appropriate language to explain methods and procedures in science will be conveying information that can be used by audience members. Additionally, those who train scientists to speak to public audiences can explain to scientists that giving audiences specific, science-based information that audience members can use to make good personal decisions is an excellent way to show the importance and relevance of their work, besides being a desirable ability for scientists to demonstrate when speaking to the public.
7.3 Operationalizing the Desired Knowledge, Skills, and Abilities

As has been noted by some scholars, there is great benefit to scientists inviting communication expertise into their communication training programs. As has also been noted, such invitations are not often extended. On those unusual occasions that the invitations are extended, such expertise is sometimes not accepted or applied. This study once again asserts that communication theory, research, and practice can benefit science communication training by identifying specific communication theories and practices that can help scientists learn and demonstrate the skills and abilities they desire to show when speaking to public audiences. Unlike other studies, this study specifically applies these communication concepts to a rubric that can be used to measure scientists’ public speaking competency.

Through investigation of communication literature and testing, this study develops the APPS rubric that identifies and assesses specific measures that demonstrate basic public speaking skills and science-specific public speaking skills. This study also develops a code book that briefly explains the theory behind the measures and explains what behaviors and skills scientists should demonstrate to show mastery of these measures. This study borrows items from the Public Speaking Competency Rubric (Schreiber et al., 2012) for the first six constructs on the rubric, which are measures of basic public speaking skills. This study does not test these items further, but does offer
explanation and clarification in the code book so that the constructs can be taught and measured to scientists in a communication training program.

The study does codify the most desired skills and abilities for scientists speaking to public audiences into six measurable areas: 1) showing the relevance and importance of the science topic, 2) using language techniques, such as comparisons, concrete words, and jargon-free definitions, to make science concepts clear to public audiences, 3) using good visual design techniques to create uncluttered, clean, well-designed visual aids to enhance the clarity of their presentations, 4) explaining some of the processes of science such as data gathering and measurement, methods for experimentation, steps in the process, and the team nature of science, 5) using communication techniques that increase the audience’s trust and liking for scientists, particularly the scientist who is speaking to them, and 6) engaging with the audience by having discussions, asking questions, and sharing in learning.

In this way, this study takes the vital step of bringing communication research and theory to the training and assessing of public science communicators. While other scholars have suggested assessing public science communicators, those scholars have not applied measurable communication skills to a comprehensive code book and rubric, in this case the APPS rubric, to give research-based, theoretically sound, specific constructs and actions that scientists can take to demonstrate that they have the skills needed to engage in meaningful, helpful public science communication. This rubric and
codebook help scientists prepare to be excellent science communicators by offering researched, tested, measurable communication techniques that will help scientists identify and display the knowledge, skills, and abilities necessary to meet the well-defined goals of scientists engaging in public communication.

7.4 Testing Rubric Validity and Working with Raters

This rubric was tested with four rounds of inter-rater reliability testing. The testing was effective at creating a stronger, more usable rubric with descriptive definitions of the constructs and can be used by both communication experts and science experts. While some of the constructs were shown to be have stronger intercoder reliability that others, one of the more important results that came from the testing is the understanding, well-known to those in the language testing field, that language testing is challenging.

7.4.1 Challenges with Language Testing

Assessing spoken or written language is difficult for many reasons. One of the primary challenges in language testing is developing a rubric that describes the skills and abilities that a person must display in great enough detail to be useful but not so great detail that the assessment becomes formulaic and does not acknowledge the unique individuality of language use. When completing coding or rating for a communication study, researchers often attempt to develop a more and more definitive code sheet or rating rubric so that coders will have little room for variation in their
scores. Often in communication studies, if an element is present in any form in a sample text, for example, the item is coded as a yes. If it is absent, the code is no.

However, language testing rejects such binaries out of hand. Filled-cell rubrics, the best of the available rubrics for this study, should have measures that are distinct and descriptive but never prescriptive. For example, no language testing rubric should look for a certain number of descriptive words or a certain number of first-person pronouns during a presentation. Instead, raters are looking at overall competence based on the specific constructs and descriptions noted in the assessment rubric.

Additionally, language is context-specific, which causes some difficulty for raters when they are assessing communicators in a staged or inauthentic context. What’s more, all raters bring biases and particular understandings to an assessment situation. Even those who consider themselves to be the most objective of raters bring their own experiences, knowledge, and preconceptions to their evaluation. Even more, agreeing on whether the speaker has performed the specific behaviors or said the specific words that indicate mastery of a construct at a particular level is difficult.

7.4.2 Challenges with Science Communication Rating

This study finds that assessing science communication is a particularly difficult, especially when the raters are scientists with little communication training. As stated previously, one of the goals of this study was to see if scientists, who generally have little
to no training in communication skills and what is required to demonstrate those skills, could use the codebook and rubric to become trained well enough to be objective and accurate assessors of scientists speaking to public audiences. Since scientists tend to eschew help from communication experts, it would be beneficial if they could take the materials, which are rooted in communication theory, and use them effectively for assessment. While this could be effective, this study suggests that such scientists would need to be chosen carefully and trained well.

Because of their implicit knowledge and training, some scientists have difficulty separating their own expertise from that of the speaker. An expert in a particular field may be particularly strict or demanding of a person from that same field, expecting more detail or more explanation from that person. In contrast, the rater who is an expert in the same field as the speaker may draw on their implicit knowledge of the science and “fill in the blanks” for the speaker. As a result, they may give the speaker higher scores than they should get in a particular category. Even raters without a science background could be affected by this tendency as they may be more generous or more demanding of a speaker who is speaking about a topic about which they have a little knowledge as a reaction to several presentations on topics about which they know little. However, scientists assessing science speakers may be more prone to this problem.

Scientists may also lack the practice of stepping back from communication techniques and analyzing them objectively. While raters with communication
backgrounds will be versed in communication concepts and will often have experience identifying and evaluating the demonstration of those concepts, raters who are scientists do not usually have practice in these skills and are often not used to noticing and evaluating techniques such as immediacy, audience connection, and even skillful organization. Without the experience of noticing and objectively analyzing rhetorical devices, scientists may score a speaker inaccurately because they are not aware that they are allowing their judgement to be clouded by rhetorical devices rather than objectively rating a speaker’s skill at using those rhetorical devices.

There were several examples of these tendencies in this study during rubric testing. Both of the scientists who were raters in the final rounds of testing made comments that showed fairly clearly that they were allowing their prior knowledge to cloud the way they assessed the speaker, and yet neither seemed to be aware that their comments indicated a poor implementation of scoring. For example, one of the scientists explicitly said that because of prior science experience, (s)he had, it was unnecessary for the speaker to make things explicit. However, the very measure that was supposed to be assessed was whether or not the speaker was making things explicit and clear for the audience of lay persons who would, by definition, not have the same science background as the speaker. Additionally, both scientists were so taken with the charisma and extroversion of two of the speakers that they gave the speakers undeservedly high scores in some areas. The communication experts, while giving those same speakers high
scores for trustworthiness and using language well, were able to maintain a critical eye and recognize that the speakers did not fulfill expectations in other areas of the rubric. Therefore, it is important to recognize that scientists may not naturally take to rating communication objectively.

7.5 Choosing, Training, and Working with Raters

Because of these and additional challenges, great care needs to be taken when selecting and training raters and when working with them to “norm,” or see the rating scale as you would like them to.

7.5.1 Choosing Raters

When choosing raters, it is important to choose carefully.

First, as mentioned before, if a training program allows, choose communication experts as assessors. Communication experts will often be easier to train quickly and will be better able to give effective feedback to scientists. Not surprisingly, scientists have different expertise than communication scholars, and their expertise is not necessarily helpful in the evaluation of communicative behaviors and skills. It is easier for communication scholars, even if they are not trained as assessors of public speaking, to understand the desired outcomes and objectively assess those outcomes. These professionals have dual advantages of not only being a representative of the lay
audience the scientists are trying to reach, but also of having the communication expertise to assess the knowledge, skills, and abilities of the scientists’ presentations.

Next, if a training program requires that scientists be used as raters to assess science speakers, the person(s) chosen should have a communication orientation through prior experience engaging in communication or through training in some field of communication. The training or experience does not necessarily have to be in public speaking, but could be in public relations, blogging, or K-12 outreach. This communication experience creates a rhetorical mindset of focusing on audience, message, and context and can help a scientist be better suited for assessing communication.

Related to this training, the raters should recognize communication as an important academic field and respect the theory and research done in the field. Scientists who consider humanities or social science research to be “lesser than” their field or “unimportant” or “common sense” may be less likely to take communication training seriously and be less likely to try to implement a rubric carefully. If the rater appreciates the theoretical and scholarly validity of such concepts as self-disclosure, immediacy, and identification, they are more likely to take the rubric seriously and consider their ratings carefully.
In addition, the person(s) chosen should be excited about the public communication of science and find it important and interesting, not a chore or a waste of time. Again, it will take effort for a scientist to adopt the mindset of a communication assessor, and it is imperative that they be excited about and interested in the opportunity so that they are willing to expend the mental effort and time that it will take to learn to do the work in a way that is helpful to scientists trying to improve their public communication skills. If the task is assigned against their will or demanded as a requirement of funding, etc., they are less likely to invest the effort into becoming an effective rater.

Lastly, raters should have strong self-confidence. It appears that raters who are confident in their abilities are more likely to internalize and apply training and less likely to attempt to adjust their scores to the what other raters are scoring or to what they think the rater trainer wants. Self-confident but not arrogant raters were able to learn from the training and take instruction without feeling that they needed to defend their choices and still had the ability to rate objectively.

In the course of rubric testing, raters lacking in self-confidence frequently asked what score others had given a speaker in a particular category and then adjusted their ratings, moving their scores up or down as they were scoring a speaker because thought that they tended to be “too strict” or “too generous” in scoring and so they tried to compensate for those natural tendencies as they were doing their individual scoring.
This led to these raters trying to score according to what they thought other raters were going to give the speakers rather than giving the speakers the scores they felt the speakers deserved. Rather than leading to scores that were more in line with the other raters, however, these attempts to guess what scores others would give the speaker often resulted in scores that were not in agreement with the other raters.

7.5.2 Rater Training

Regardless of their background or expertise, raters need to be carefully trained to act as raters using the APPS rubric. Rater need attention in at least three categories: understanding their implicit expertise and the biases that imposes, norming the ratings so that everyone using the rubric in a specific training program is using basically the same way, and consistency over time.

Some specific instruction raters need to understand is that they do bring bias with them to the rating situation. Scientists who are working as raters should be specifically instructed in the ways their implicit expertise may cause them to be too generous or too stingy when it comes to scoring speakers. Each rater, regardless of background, should be asked to identify areas of personal expertise and experience and then specifically note ways that expertise may cause implicit bias. Identifying these tendencies and biases in advance can help raters be more self-aware and can make those tendencies easier to point out if/when they become evident during training and norming sessions.
Directly instruct raters that they will need to step out of their knowledge bubble and apply the same levels of critical assessment to the entire broad range of presentations that they will be asked to assess and will need to deliberately avoid treating one group differently because they as raters are more familiar with the topic or like the speaker better or are more engaged with the speaker’s vocal style.

Raters should be required to read the code book carefully so that they understand the communication theory behind the assessment instrument and so that they recognize the importance of applying the principles as evenly and consistently as possible rather than “going with their gut” or “filling in the blanks” for the speaker. Raters with a science background may benefit from being introduced to the research articles that explain the underlying communication theories, if reading those articles will increase their understanding of the constructs and/or increase their willingness to follow the criteria as described in the rubric.

Rater training should begin with discussion where the supervisor or person who is operating the training program for scientists explains each rubric construct and how the construct will be visualized and operationalized for their program. Most of the training time should be spent viewing presentations and scoring them together. It is recommended that raters view sample presentations that are clearly well done and then presentations that are clearly poorly done. Raters should rate the presentations independently, then compare scores and discuss differences. Where there are
discrepancies, the training supervisor should give normative instruction as to the quality of the performance and what specific criteria are met or not met to warrant the score (Bukta, 2014). By having this normative instruction and repeated opportunities to rate presentations together, raters will more quickly understand the boundaries and guidelines of the process.

After norming scores on clearly good or clearly poor presentations, the training coordinator can introduce presentations that are more difficult to rate and have the raters practice several times with those presentations. This repeated practice will help the raters come to a collective understanding of the rating standards and understand the vision of the training coordinator for how each specific construct will be understood for their program. It is important to be clear about the goals of your rating. Your raters should know what you are looking for, so it’s important to be explicit during training.

For best results in a training program, the same raters should be used consistently over time, and the raters should be monitored consistently and checked for reliability in ratings on a frequent basis (Bukta, 2014). Studies show that scoring shifts over time, but that training does help with inter-rater agreement so that the assessment instrument is being used consistently over time. It is also important to check that raters are consistent with their own ratings (intra-rater reliability). Training effects do fade over time, so regular training is strongly recommended.
These challenges with raters and reliability are common to any language assessment, but can be particularly challenging for scientists who are not accustomed to rating language performance. Particular care should be taken to train these raters, and particular attention paid to frequent checks on their ratings. The more opportunities scientists have to rate communication on a consistent basis, the more aware they will become of communication skills and abilities and how to assess them, if they are trained consistently. Even experienced raters need rater training on a consistent basis, and inexperienced raters need such training even more (Alderson, Clapham, & Wall, 1995).

Helping scientists be self-aware, understand their implicit biases, and see the need for objectivity are useful topics to address in rater training for scientists, along with basic communication theory and practice and the specifics of the APPS codebook and rubric.

Finally, it is useful as a trainer of raters to recognize one’s own implicit knowledge and biases. If the trainer does have communication expertise, they need to recognize that scientists do not have the same background and implicit communication knowledge. Scientists training as raters and even experienced raters may not understand the connection between immediacy and trustworthiness, for example, and may not grasp the idea of identification. For a communication scholar, it may seem difficult to accept that a person could earn a PhD and never encounter Burke, but it is perfectly reasonable that it could happen. Being explicit with the raters and the speakers as you
guide the training and explaining even the basic concepts and theories behind the assessment will lead to greater understanding and more effective assessment.

7.6 Bringing Communication Expertise to Scientists’ Training

One of the purposes of this study was to see if communication theory and practices could be incorporated into a code book and rubric that could be read and applied by scientists without communication training and have those scientists act as assessors and produce reliable results. The results of developing and testing the APPS with both communication and science raters shows that the codebook and the APPS rubric can be useful to scientists who are asked to train or assess other scientists in communication skills. However, it also shows that it would be beneficial to have communication experts involved at least in the training of the scientists who are speaking to public audiences and any scientists who are assessing those public presentations. If possible, having communication experts involved in the assessment of the public science presentations would be beneficial.

While scientists can read the materials and gain knowledge about communication skills, they will not bring the same depth of understanding as someone who has studied and practiced with those concepts and theories for possibly decades. Just as it would be a misplacement of resources to put a communication expert in charge of analyzing gene expression, it would be a misplacement of resources to put a scientist, even one that enjoys and values public science communication, in charge of
communication training. If a training program does require that a scientist take the role of running a communication training program and/or serving as assessor of scientists’ public communication, it would be ideal for that person to have some background and training in communication and/or language assessment as well as science expertise. Similarly, any communication expert who is running a training program for scientists should have training in science and science communication in order to be most effective. Having a partnership between a science expert and a communication expert in a training and assessing situation would be ideal.

Drawing upon the wealth of research, theory, and practical knowledge in the fields of communication, public speaking, and language assessment can help science programs in creating and operating a training program for scientists who are learning to engage in public communication. Using the APPS rubric as part of a training program will help incorporate important communication principles into the training, allow for rapid and effective summative assessment, provide a means for quantitative summative and programmatic assessment, demonstrate the validity and effectiveness of communication training, provide numerical evidence of a program’s effectiveness, and incorporate the best practices for assessment into a training program.
8.0 CONCLUSION

This study demonstrates the urgent need for more and better communication training for scientists so they can learn to speak more effectively to public audiences that lack scientific expertise. It points out the encouragement from scientists, politicians, scholars, activists, and others for this increased communication to take place to meet a variety of goals. This study argues that more effective assessment is key to scientist’s communication training.

This study develops an assessment instrument, the Assessment for Public Presentations by Scientists, or APPS, that meets the needs for an assessment instrument for public presentations by scientists. First, it brings communication expertise to science communication assessment. Many science communication training programs are organized and operated by scientists, not communication experts. By introducing this assessment instrument developed by incorporating communication research, theory, and practice, scientists are adding to their expertise the complex and detailed understanding communication experts have of best communication practices as applied to the specific goals of scientists speaking to public audiences. Since this instrument identifies and assesses those skills most desired for public science communication and operationalizes them using tested communication practices, it provides an excellent foundation for assessment and training of scientists.
Studies have demonstrated that communicators improve their skills most rapidly when they are presented with supportive, prompt feedback. Such feedback also helps communicators more accurately assess their own communication. The APPS instrument allows communication raters to provide this type of prompt feedback specifically directed at the skills and abilities necessary for public science communication, leading to better presentations that meet the goals of the speakers and the audiences. In addition to leading to better presentations, the type of formative assessment this instrument encourages means that scientists are giving frequent, low-stakes presentations so that they can be evaluated often. These repeated opportunities to speak to an audience means less anxiety and increased self-awareness in their presentations.

The APPS rubric is also ideal for use in low-stakes, summative contexts such as providing a snapshot of a speaker’s improvement over the course of a training program or as one element out of several used to show administrators the value of a training program or to demonstrate the need of a training program.

The rubric provides numerical measures so that results can be reported and analyzed statistically, speaking the language of scientists and administrators. Scientists report results in numbers and may pay more attention and give more respect to quantitative results reported with as much objectivity as is possible in a communication assessment situation. Administrators are under constant pressure to justify programs, report results, and quantify progress. The APPS helps administrators with those
responsibilities and desires in addition to meeting its primary goal of helping scientists be better communicators.

Additionally, the ECD methodology used to develop the rubric is a methodologically sound, respected approach to assessment design. Using the ECD method to develop the rubric and code book not only means a better rubric that fills the needs of scientists and communication trainers, but it also means that the APPS was developed using a methodology that is well known and well respected in assessment circles as well as in administrative discourse communities. This respect for the methodology can be transferred to the rubric itself and could increase the acceptance of the measure for use in assessing scientists’ public presentations.

One of the primary benefits of this study is that it provides a thorough domain analysis consisting of an extensive literature review of both scholarly and popular documents to discover what specific knowledge, skills, and abilities scientists should demonstrate when they make presentations to public audiences. This study then operationalized the most important and most often mentioned skills using communication theory and practices so that the rubric is more useful than other existing rubrics as a means of assessing scientists’ public communication and as a tool for training scientists in communication skills. The six skills and abilities this rubric assesses are listed here:
1. Demonstrating the relevance and importance of science to public audiences.
2. Use language in a way that makes science concepts understandable and clear to the public.
3. Explain some of the processes, procedures, risks, and methods of doing science so that audiences can understand them.
4. Skillfully use visual aids to enhance the scientific message.
5. Increase the public reputation of scientists by portraying them as trustworthy, “human,” good, and “normal” people.
6. Interact with public audiences through conversation and dialogue.

Although this rubric is designed specifically for informative presentations, many of the skills and abilities it assesses are also important in persuasive speaking contexts and when giving the audience members valuable information they can use to make decisions, two other skills and abilities identified in the domain analysis as being desirable for scientists to demonstrate when speaking to public audiences. With some additions and alterations, the rubric can be adjusted to assess those situations as well as in the case of the informative “library talk” it was developed to assess.

Another important finding of this study is the illustration of the challenges inherent in assessing science communication. Language assessment is challenging in any context, given the wide variations in expression that are influenced by culture, gender, education, family experience, and more. However, assessing science communication is shown to be uniquely challenging. It is important to take great care in choosing and training raters who will assess science speakers and take great care in choosing the person(s) who will train the scientists. A person who has an extensive communication background and an understanding of scientific methodologies and tendencies is an ideal
candidate to act as training coordinator. The discussion section of this study has several guidelines that would ideally be followed in choosing and training raters.

Therefore, the APPS rubric and code book together comprise a comprehensive tool that can be used to effectively measure the public communication of scientists. Through careful domain analysis and domain modeling, the APPS codifies the most important knowledge, skills, and abilities for scientists to exhibit when they speak to the public; something that other rubrics have not done. This alone makes it a valuable tool for scientists who communicate with public audiences and those who train those scientists. The code book and APPS rubric also include the application of communication theory and practice as they apply to those important knowledge, skills, and abilities that scientists should demonstrate, giving scientists the benefit of decades of communication research to effectively meet the needs of their audiences and thereby fulfill their own goals in speaking to public audiences.
REFERENCES


10.1177/1064804611416583


doi:10.1371/journal.pone.0036240


Retrieved from *Final Program.* (2016). The Society for Integrative and Comparative Biology
McLean, VA. Retrieved from

http://www.sicb.org/meetings/2016/SICB%202016%20Final%20Program.pdf


doi:10.1080/08941920802653257


Retrieved from


doi:10.1093/biosci/biu051


http://www.publications.parliament.uk/pa/ld199900/ldselect/ldsctech/38/3802.html


Join the Conversation. (2014). Retrieved from

http://jneurosci.org/content/jneuro/35/9/local/advertising.pdf#search=%22Why communicate science%22


doi: [http://dx.doi.org/10.1080/01463370309370170](http://dx.doi.org/10.1080/01463370309370170)


Siegfried, T. (2014). Top 10 things everyone should know about science.


doi:10.1159/000092659

APPENDIX A: RESULTS OF THE DOMAIN ANALYSIS CONTENT ANALYSIS
STEP. BROAD GOALS FOR SCIENTISTS COMMUNICATING WITH PUBLIC AUDIENCES

When communicating with public audiences, the scientist will:

<table>
<thead>
<tr>
<th>Coder 1</th>
<th>Coder 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledge and address the ethical and legal implications of scientific research and discovery</td>
<td>Provide quality instruction materials for science teachers through all levels of education (informing)</td>
</tr>
<tr>
<td>Exhibit a tone that is free from judgment or self-aggrandizement (communication skills)</td>
<td></td>
</tr>
</tbody>
</table>

Make suggestions to publics

Provide good scientific information so people can make good decisions

Build the trust in and respect for scientists

Improve the communication skills of scientists

Clearly explain complex science concepts

Build knowledge of science research and processes

Engage with, have dialogue with the public

Establish bi-directional channels of communication with the public

Share the wonder/excitement of science

Promote understanding and excitement for the beauty of natural processes and phenomena

Connect messages with the specific audience

Exhibit sensitivity to different audiences and frame issues in ways that are directly relevant to them

Inform policy makers/influence policy

Influence policy debate by providing research-based information and advice to policymakers
| Advocate for science funding  
| Increase interest in science  | Promote interest and investment in science and scientific research among the general public |
| Encourage young people to enter a STEM career | Motivate young people, including those from underrepresented groups, to pursue a career in science |
| Change incorrect views of scientists and science | Dispel misconceptions about science, correct pseudoscientific knowledge |
| Acknowledge fears and perceptions about science, show how there have been positive results from scientific research as well as some negative | Communicate both the benefits and challenges that have been raised with science  
Promote understanding by the general public of the advantages and disadvantages of science |
| Encourage lab and campus visits by young people – increases knowledge, positivity about science, desire for STEM careers | Promote student visits to campuses, labs, on-site camps, and other places where scientists develop their work |
## APPENDIX B: RUBRIC FOR ASSESSING PUBLIC COMMUNICATION BY SCIENTISTS, FIRST DRAFT

(Note: All rubrics in the Appendices use reduced font size to comply with publisher requirements. Contact author for full-size rubrics.)

<table>
<thead>
<tr>
<th>Performance Standard: The speaker demonstrates ability to …</th>
<th>Advanced 4</th>
<th>Proficient 3</th>
<th>Adequate 2</th>
<th>Basic 1</th>
<th>Not Present 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select a topic appropriate to the audience and occasion</td>
<td>Topic engages audience, topic is worthwhile, and of interest and of an appropriate complexity for the audience and situation.</td>
<td>Topic is appropriate to the audience and situation and is of some interest to many members of the audience.</td>
<td>Topic is somewhat complex or not entirely relevant to many of the audience members. It is adequate for the situation.</td>
<td>Topic is too trivial, too complex, or inappropriate for audience; topic not suitable for the situation.</td>
<td>The topic is entirely too complex, is not stated clearly, or is inappropriate for the audience and situation.</td>
</tr>
<tr>
<td>Formulate an introduction that orients audience to topic and speaker</td>
<td>Excellent attention getter, firmly established credibility, sound orientation; topic, clear thesis, preview of main points cogent and memorable.</td>
<td>Good attention getter, generally establishes credibility, provides some orientation to topic, discernible thesis, previews main points</td>
<td>Attention getter is mundane; somewhat develops credibility; Awkwardly composed thesis; provides little direction for audience</td>
<td>Irrelevant opening; little attempt to build credibility; abrupt jump into body of speech; thesis and main points can be deduced but are not explicitly stated.</td>
<td>No evidence of opening technique; no credibility statement; little to no background on topic; thesis/statement of topic is unclear; no or unclear preview of points.</td>
</tr>
<tr>
<td>Use an effective organizational pattern</td>
<td>Very well organized; main points clear, mutually exclusive and directly related to thesis; effective transitions and signposts</td>
<td>Organizational pattern is evident, main points are apparent; transitions present between main points; some use of signposts</td>
<td>Organizational pattern somewhat evident; main points are present but not mutually exclusive; transitions are present but are minimally effective</td>
<td>Speech does not flow well; speech was not logically organized; main points were not clear, transitions present but not well formed</td>
<td>Organizational pattern not clear or nonexistent, few to no transitions; information sounds as if it is being randomly presented</td>
</tr>
<tr>
<td>Develop a conclusion that reinforces the thesis and provides psychological closure</td>
<td>Provides a clear and memorable summary of points, refers back to thesis/big picture, ends with strong clincher or call to action</td>
<td>Appropriate summary of points, some reference back to thesis, clear clincher or call to action</td>
<td>Provides some summary of points, no clear reference back to thesis, closing technique can be strengthened</td>
<td>Conclusion lacks clarity, trails off; ends in a tone at odds with the rest of the speech or brings in new information.</td>
<td>Little in the way of a conclusion; speech ends abruptly and without closure</td>
</tr>
<tr>
<td>Effectively use vocal expression to engage the audience</td>
<td>Excellent use of vocal variation, intensity and pacing; vocal expression natural and enthusiastic; avoids vocal fillers</td>
<td>Good vocal variation and pace, vocal expression suited to the situation, few if any vocal fillers</td>
<td>Demonstrates some vocal variation; enunciates clearly and speaks audibly; generally avoids fillers</td>
<td>Sometimes uses a voice too soft or indistinct for listeners to comfortably hear; little vocal variety, often uses fillers</td>
<td>Speaks much too loudly or softly, enunciation is lacking, speaks in monotone, poor pacing, distracts listeners with vocal fillers</td>
</tr>
<tr>
<td>Use eye contact, facial expressions, and body movement to support the verbal message</td>
<td>Posture, gestures, facial expression and eye contact well-developed, natural, and display elevated levels of poise and confidence</td>
<td>Postures, gestures and facial expressions are suitable for speech, speaker appears confident</td>
<td>Some reliance on notes, but has adequate eye contact, generally avoids distracting mannerisms</td>
<td>Speaker relies heavily on notes; nonverbal expression is stiff and unnatural</td>
<td>Usually looks down and avoids eye contact; nervous gestures and nonverbal behaviors distract from or contradict the message</td>
</tr>
<tr>
<td></td>
<td>Advanced 4</td>
<td>Proficient 3</td>
<td>Adequate 2</td>
<td>Basic 1</td>
<td>Not Present 0</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------</td>
<td>--------------</td>
<td>------------</td>
<td>---------</td>
<td>---------------</td>
</tr>
<tr>
<td><strong>Connect science information to a specific audience in a specific situation</strong></td>
<td>Clearly and consistently connects science concepts to ideas familiar to audience, creates common ground by strongly emphasizing common goals, values and/or experiences, makes a strong case for the relevance and importance of the science to the audience’s lives, indicates the excitement and joy of science for reasons relevant to the audience.</td>
<td>Specifically connects science concepts to ideas familiar to the audience, inconsistent mentions of common goals or values, discusses the importance and relevance of the topic to the audience, expresses excitement and joy in working with this science topic.</td>
<td>Makes some connection between the audience and the science, some mention of common goals or values, may mention the importance of the topic but does not make it explicitly relevant to the audience, excitement and joy of science topic not explicitly connected to audience.</td>
<td>Makes little connection between the audience and the topic, few if any mention of common goals or values, does not make the topic relevant to the audience, excitement and joy not expressed and/or not connected to the audience.</td>
<td>Connections between the audience and the topic are minimal or nonexistent, little or no mention of common goals or values, topic’s relevance to the audience is unclear, little or no excitement or joy expressed regarding this science topic.</td>
</tr>
<tr>
<td><strong>Clearly communicate complex ideas and make them understandable to the general public</strong></td>
<td>Develops interesting, audience-focused, teaching metaphors or other comparisons to make topic clear to audience, may make deliberate use of metaphors or analogies that help build science theories, clear, vivid language takes into account the audience’s knowledge, excellent visual aids use simplicity, clarity, and good design principles.</td>
<td>Comparisons/metaphors are deliberate but may not be entirely appropriate for the audience, theory-building metaphors or analogies may not appear deliberate, language is generally clear and vivid, visual aids generally follow design principles and are helpful in clarifying points.</td>
<td>Uses some metaphors or other comparisons in a way that enhances understanding but may not always be appropriate for the audience, language is usually clear, visual aids usually adequately incorporate good design principles.</td>
<td>Uses few metaphors or other comparisons, comparisons not relevant to audience or not fully developed, language is sometimes too complex or unclear, visual aids often ignore good design principles, may be unclear.</td>
<td>Uses few metaphors or other comparisons, comparisons not relevant to audience or not fully developed, language is sometimes too complex or unclear, visual aids often ignore good design principles, may add to audience confusion.</td>
</tr>
<tr>
<td><strong>Increase the audience’s knowledge and understanding of science and scientific processes</strong></td>
<td>Clearly defines any science terms used, includes descriptions of scientific processes, explains why particular steps are needed, explains reasons for choices, emphasizes the team nature of science, talks about uncertainties and how they are navigated.</td>
<td>Usually defines any unknown science terms, may describe science processes, may describe steps without explaining clearly why they are needed, may mention teams or uncertainties, or reason for choices but not all adequately.</td>
<td>Uses mostly clear language and largely defines terms used, may attempt to describe processes without complete success, may fail to completely describe the nature of teams or uncertainties.</td>
<td>Often uses scientific terms without clearly explaining or defining them, often fails to describe processes, makes only brief mention and little to no description of teams or uncertainties.</td>
<td>Uses many scientific terms without clearly defining them, fails to describe processes or give reasons for choices, little to no mention or description of teams, reasoning, or uncertainties in scientific processes.</td>
</tr>
<tr>
<td><strong>Humanize scientists and help them seem trustworthy and knowledgeable</strong></td>
<td>Engages in frequent, appropriate self-disclosure and uses personal stories and ideas, uses inclusive terms such as we, us, and our, uses concrete language, portrays scientists positively.</td>
<td>Engages in some self-disclosure and personal stories, mostly uses inclusive pronouns rather than first person singular, largely uses concrete language, scientists seem credible.</td>
<td>Self-disclosure and personal reflection is present in some form, sometimes uses inclusive pronouns, uses some concrete language, may or may not promote scientists.</td>
<td>Uses little self-disclosure or personal reflection, speaks largely in terms of “we,” speaks of scientists dispassionately, focuses on objectivity.</td>
<td>Little to no self-disclosure, few or no personal stories, little to no inclusive language, no portrayals of scientists in a positive, credible light.</td>
</tr>
<tr>
<td><strong>Engage in dialogue and interactions about science with public audiences.</strong></td>
<td>Audience members have multiple opportunities to ask and respond to questions, audience makes comments, speaker is also a learner, there is active dialogue between audience and speaker during the presentation.</td>
<td>Audience members have opportunities to ask questions and make comments, speaker listens to audience but does not give them much time, even during Q&amp;A, little dialogue with audience.</td>
<td>Audience members have opportunity to ask questions but not respond to questions, audience makes few if any comments, speaker gives no indication of learning, no dialogue with audience.</td>
<td>Audience members are not asked questions, interactions between audience and speaker limited to a single Q&amp;A period at the close of the presentation, no dialogue w/ audience members.</td>
<td>Speaker consistently assumes an authoritative air, does not give speaking power to audience, if audience questions are allowed, speaker responds with finality and unbending tone.</td>
</tr>
</tbody>
</table>
## APPENDIX C. 2-7-17 DRAFT OF RUBRIC

<table>
<thead>
<tr>
<th>Performance Standard:</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Topic:</strong> Appropriate for speaker and audience</td>
<td>Topic engages audience, topic is worthwhile, and of interest and of an appropriate complexity for the audience and situation.</td>
<td>Topic is appropriate to the audience and situation and is of some interest to many members of the audience.</td>
<td>Topic is somewhat complex or not entirely relevant to many of the audience members. It is a adequate for the situation.</td>
<td>Topic is too trivial, too complex, or inappropriate for audience; topic not suitable for the situation.</td>
<td>The topic is entirely too complex, is not stated clearly, or is inappropriate for the audience and situation.</td>
</tr>
<tr>
<td><strong>2. Introduction:</strong> Orients audience to topic and speaker</td>
<td>Excellent attention getter, firmly established credibility, sound orientation to topic, clear thesis, preview of main points cogent and memorable</td>
<td>Good attention getter, generally establishes credibility, provides some orientation to topic, discernible thesis, previews main points</td>
<td>Attention getter is mundane; somewhat develops credibility; Awkwardly composed thesis; provides little direction for audience</td>
<td>Irrelevant opening; little attempt to build credibility; abrupt jump into body of speech; thesis and main points can be deduced but are not explicitly stated.</td>
<td>No evidence of opening technique; no credibility statement; little to no background on topic; thesis/statement of topic is unclear; no or unclear preview of points</td>
</tr>
<tr>
<td><strong>3. Organization:</strong> Effective for topic and audience</td>
<td>Very well organized; main points clear, mutually exclusive and directly related to thesis; effective transitions and signposts</td>
<td>Organizational pattern is evident, main points are apparent; transitions present between main points; some use of signposts</td>
<td>Organizational pattern somewhat evident; main points are present but not mutually exclusive; transitions are present but are minimally effective</td>
<td>Speech does not flow well; speech was not logically organized; main points were not clear, transitions present but not well formed</td>
<td>Organizational pattern not clear or nonexistent, few to no transitions; information sounds as if it is being randomly presented</td>
</tr>
<tr>
<td><strong>4. Conclusion:</strong> Restates main points, provides closure</td>
<td>Provides a clear and memorable summary of points, refers back to thesis/big picture, ends with strong clincher or call to action</td>
<td>Appropriate summary of points, some reference back to thesis, clear clincher or call to action</td>
<td>Provides some summary of points, no clear reference back to thesis, closing technique can be strengthened</td>
<td>Conclusion lacks clarity, trails off; ends in a tone at odds with the rest of the speech or brings in new information.</td>
<td>Little in the way of a conclusion; speech ends abruptly and without closure</td>
</tr>
<tr>
<td><strong>5. Voice:</strong> Vocal expression engages the audience</td>
<td>Excellent use of vocal variation, intensity and pacing; vocal expression natural and enthusiastic; avoids vocal fillers</td>
<td>Good vocal variation and pace, vocal expression suited to the situation, few if any vocal fillers</td>
<td>Demonstrates some vocal variation; enunciates clearly and speaks audibly; generally avoids fillers</td>
<td>Sometimes uses a voice too soft or articulation too indistinct for listeners to comfortably hear; little vocal variety, often uses fillers</td>
<td>Speaks much too loudly or softly, enunciation is lacking, speaks in monotone, poor pacing, distracts listeners with vocal fillers</td>
</tr>
<tr>
<td><strong>6. Non-Verbal:</strong> Eye contact, facial expressions, and body movement to support the verbal message</td>
<td>Posture, gestures, facial expression and eye contact well-developed, natural, and display high levels of poise and confidence</td>
<td>Postures, gestures and facial expressions are suitable for speech, speaker appears confident</td>
<td>Some reliance on notes, but has adequate eye contact, generally avoids distracting mannerisms</td>
<td>Speaker relies heavily on notes; nonverbal expression is stiff and unnatural</td>
<td>Usually looks down and avoids eye contact; nervous gestures and nonverbal behaviors distract from or contradict the message</td>
</tr>
</tbody>
</table>

*These elements are adopted with permission from the PSCR (Schreiber, et al, 2012).*
<table>
<thead>
<tr>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>7. Relevance and Importance:</strong> Connects science to lives, shows importance, excitement, and relevance of science</td>
<td>Clearly and consistently connects science concepts to ideas familiar to audience, creates common ground by strongly emphasizing common goals, values and/or experiences, makes a strong case for the relevance and importance of the science to the audience’s lives, indicates the excitement and joy of science for reasons relevant to the audience</td>
<td>Specifically connects science concepts to ideas familiar to the audience, inconsistent mentions of common goals or values, discussed the importance and relevance of the topic to the audience, expresses excitement and joy in working with this science topic</td>
<td>Makes some connection between science concepts and the audience, some mention of common goals/values, may mention importance of topic but may not make it explicitly relevant to the audience, excitement and joy of science topic not explicitly connected to audience</td>
<td>Makes little connection between the audience and the topic, few if any mention of common goals or values, does not make the topic relevant to the audience, excitement and joy not expressed and/or not connected to the audience</td>
</tr>
<tr>
<td><strong>8. Language:</strong> Uses language to make complex clear, metaphor and comparisons</td>
<td>Develops interesting, deliberate, audience-focused, teaching metaphors or other comparisons to make topic clear to audience; may make deliberate use of comparisons that help build science theories; uses clear, vivid language that considers audience’s current knowledge, particularly about science</td>
<td>Comparisons/metasporh s are deliberate and are generally helpful for the audience based on their understanding, theory-building metaphors or analogies may not appear entirely deliberate, language is generally clear and vivid, reflects general audience understanding, including about science</td>
<td>Uses some metaphors or other comparisons in a way that enhances understanding but may not always be appropriate for the audience, may use analogies or metaphors that don’t make sense to the audience, may occasionally use language that is complex or unclear</td>
<td>Uses few metaphors or other comparisons, comparisons may be irrelevant to audience or not fully developed, language is sometimes too complex or unclear, frequently uses scientific terms that are outside the audience understanding</td>
</tr>
<tr>
<td><strong>9. Visuals:</strong> Excellent visual aids, use simplicity, clarity, and good design principles. Visuals focus on pictures or other graphic depictions, easy to understand, small amounts of text, non-electronic visuals are clearly visible</td>
<td>Visual aids generally follow design principles and are helpful in clarifying points, but may have some aspects that are not as clear, slides may be a bit cluttered</td>
<td>Visual aids usually adequately incorporate good design principles, but some may use too many elements, seem a bit cluttered, have too much text or some complex diagrams</td>
<td>Visual aids often ignore good design principles, may often be unclear, or may use too much text, charts/graphs may be too plentiful and/or too complex, graphics may be complicated</td>
<td>Visual aids often ignore good design principles, may cause audience confusion. May use too many colors, fonts, charts, or graphics, text-heavy slides, confusing graphics</td>
</tr>
<tr>
<td><strong>10. Explain Science Processes:</strong> Clearly defines any science terms used, includes descriptions of scientific processes, may clearly explain why particular steps are needed, explain reasons for choices, emphasize the team nature of science, and/or talk about uncertainties and how they are navigated</td>
<td>Usually defines any unknown science terms, may describe science processes, will usually clearly describe steps and why they are needed, and/or make good efforts to mention teams, uncertainties, choices, and/or other scientific processes</td>
<td>Uses mostly clear language and largely defines terms used, may attempt to describe processes without complete success, may fail to completely describe the nature of teams, uncertainties, choices and other processes</td>
<td>Often uses scientific terms without explaining or defining them, often fails to describe processes, makes only brief mention and little to no description of teams, uncertainties, choices, or any other scientific processes</td>
<td>Does not explain terms, fails to describe processes or give reasons for choices, little to no mention or description of teams, reasoning, uncertainties, or other scientific processes</td>
</tr>
<tr>
<td><strong>11. Human and Trustworthy:</strong> Help scientists seem approachable, knowledgeable</td>
<td>Engages in frequent, appropriate self-disclosure and/or uses personal stories and ideas, uses inclusive terms such as we, us, and our, uses concrete language, portrays scientists positively</td>
<td>Engages in regular self-disclosure and/or personal stories, mostly uses inclusive pronouns rather than first person singular, largely uses concrete language, scientists seem credible</td>
<td>Self-disclosure and/or personal reflection is present in some form, sometimes uses inclusive pronouns, uses some concrete language, may or may not promote scientists</td>
<td>Uses little self-disclosure or personal reflection, speaks largely in terms of “I” rather than “we,” speaks of scientists dispassionately, focuses on objectivity</td>
</tr>
<tr>
<td><strong>12. Engagement:</strong> Have dialogue and interactions about science with public audiences.</td>
<td>Audience members have multiple opportunities to ask and respond to questions, audience makes comments, speaker is also a learner, there is active dialogue between audience and speaker during the presentation</td>
<td>Audience members have opportunities to ask questions and make comments, speaker listens to audience but does not give them much time, even during Q&amp;A, little dialogue with audience</td>
<td>Audience members have opportunity to ask but not respond to questions, audience makes few if any comments, speaker gives no indication of learning, no dialogue with audience</td>
<td>Audience members are not asked questions, interactions between audience and speaker limited to a single Q&amp;A period at the close of the presentation, no dialogue w/ audience</td>
</tr>
</tbody>
</table>
### APPENDIX D: 3-8-17 - RUBRIC FOR ASSESSING PUBLIC COMMUNICATION BY SCIENTISTS

<table>
<thead>
<tr>
<th>Performance Standard:</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Topic:</strong></td>
<td>Topic engages audience, topic is worthwhile, and of interest and of an appropriate complexity for the audience and situation.</td>
<td>Topic is appropriate to the audience and situation and is of some interest to many members of the audience.</td>
<td>Topic is somewhat complex or not entirely relevant to many of the audience members. It is adequate for the situation.</td>
<td>Topic is too trivial, too complex, or inappropriate for audience; topic not suitable for the situation.</td>
<td>The topic is entirely too complex, is not stated clearly, or is inappropriate for the audience and situation.</td>
</tr>
<tr>
<td><strong>2. Introduction:</strong></td>
<td>Excellent attention getter, firmly established credibility, sound orientation to topic, clear thesis, preview of main points cogent and memorable</td>
<td>Good attention getter, generally establishes credibility, must be oriented to topic, discernible thesis, previews main points</td>
<td>Attention getter is mundane; somewhat develops credibility; Awkwardly composed thesis; provides little direction for audience</td>
<td>Irrelevant opening; little attempt to build credibility; abrupt jump into body of speech; thesis and main points can be deduced but are not explicitly stated.</td>
<td>No evidence of opening technique; no credibility statement; little to no background on topic; thesis/statement of topic is unclear; no or unclear preview of points</td>
</tr>
<tr>
<td><strong>3. Organization:</strong></td>
<td>Very well organized; main points clear, mutually exclusive and directly related to thesis; effective transitions and signposts</td>
<td>Organizational pattern is evident, main points are apparent; transitions present between main points; some use of signposts</td>
<td>Organizational pattern somewhat evident; main points are present but not mutually exclusive; transitions are present but are minimally effective</td>
<td>Speech does not flow well; speech was not logically organized; main points were not clear, transitions present but not well formed</td>
<td>Organizational pattern not clear or nonexistent, few to no transitions; information sounds as if it is being randomly presented</td>
</tr>
<tr>
<td><strong>4. Conclusion:</strong></td>
<td>Provides a clear and memorable summary of points, refers back to thesis/big picture, ends with strong clincher or call to action</td>
<td>Appropriate summary of points, some reference back to thesis, clear clincher or call to action</td>
<td>Provides some summary of points, no clear reference back to thesis, closing technique can be strengthened</td>
<td>Conclusion lacks clarity, trails off; ends in a tone at odds with the rest of the speech or brings in new information.</td>
<td>Little in the way of a conclusion; speech ends abruptly and without closure</td>
</tr>
<tr>
<td><strong>5. Voice:</strong></td>
<td>Excellent use of vocal variation, intensity and pacing; vocal expression natural and enthusiastic; avoids vocal fillers</td>
<td>Good vocal variation and pace, vocal expression suited to the situation, few if any vocal fillers</td>
<td>Demonstrates some vocal variation; enunciates clearly and speaks audibly; generally avoids fillers</td>
<td>Sometimes uses a voice too soft or articulation too indistinct for listeners to comfortably hear; little vocal variety, often uses fillers</td>
<td>Speaks much too loudly or softly, enunciation is lacking, speaks in monotone, poor pacing, distracts listeners with vocal fillers</td>
</tr>
<tr>
<td><strong>6. Non-Verbal:</strong></td>
<td>Posture, gestures, facial expression and eye contact well-developed, natural, and display high levels of poise and confidence</td>
<td>Postures, gestures and facial expressions are suitable for speech, speaker appears confident</td>
<td>Some reliance on notes, but adequate eye contact, generally avoids distracting mannerisms</td>
<td>Speaker relies heavily on notes; nonverbal expression is stiff and unnatural</td>
<td>Usually looks down and avoids eye contact; nervous gestures and nonverbal behaviors distract from or contradict the message</td>
</tr>
</tbody>
</table>

*These elements are adopted with permission from the PSCR (Schreiber, et al, 2012).*
<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>7. Relevance &amp; Importance:</strong> Connects science to lives, shows importance and relevance of science</td>
<td>Makes a compelling case for the relevance and importance of the science topic to the audience’s lives, creates common ground by strongly emphasizing common goals, values and/or experiences, clearly and consistently connects science concepts to ideas familiar to the audience.</td>
<td>Clearly discusses the importance and relevance of this science topic to the audience, makes an effort at creating common ground with audience through shared values and goals, inconsistently connects science concepts to ideas familiar to the audience.</td>
<td>May mention importance of topic but not make it explicitly relevant to the audience, may mention of common goals/values, may make some connection between science concepts and the audience.</td>
<td>Only brief mention of topic’s relevance to the audience, few, if any, mentions of common goals/values, makes little connection between the audience and the topic.</td>
</tr>
<tr>
<td><strong>8. Language:</strong> Uses language to make the complex clear, definitions, metaphors, and comparisons</td>
<td>Uses clear, concrete, understandable language that considers audience’s current knowledge, particularly about science; defines any complex science terms used; when appropriate, uses comparisons and analogies to make complicated ideas clearer; metaphors used are interesting, deliberate, audience-focused, teaching metaphors that clarify ideas</td>
<td>Uses generally clear, concrete and understandable language that is mostly appropriate to audience’s science knowledge. Any comparisons/analogies seem deliberate and are generally helpful for the audience based on their understanding, analogies are appropriate for audience and thoughtfully expressed.</td>
<td>May sometimes use complex or unclear language, may fail to define all science terms. Some analogies may not always be appropriate for or make sense to the audience, may not use comparisons or examples when they would help explain and clarify the topic.</td>
<td>Uses several complex terms without defining them; language is often inappropriate for audience’s level of understanding. Any comparisons may not be less relevant to audience or not fully developed, may show lack of awareness of or lack of adaptation to audience</td>
</tr>
<tr>
<td><strong>9. Visuals:</strong> Clear, simple visual aids using good science communication design techniques</td>
<td>Excellent visual aids that use simplicity, clarity, and good design principles. Visuals focus on pictures or other graphic depictions, easy to understand, use minimal text, only use clear graphs and charts, any non-electronic visuals are clearly visible</td>
<td>Visual aids generally follow design principles and are helpful in clarifying points, may have some aspects that are not entirely clear. Slides may be slightly cluttered, but are generally helpful and thoughtfully designed.</td>
<td>Visual aids usually adequately incorporate good design principles, but some may use too many elements, seem a bit cluttered, have too much text or some complex diagrams</td>
<td>Visual aids often ignore good design principles, may often be unclear, or may use too much text; charts/graphs may be too plentiful and/or too complex; graphics may be complicated</td>
</tr>
<tr>
<td><strong>10. Explain Science Processes:</strong> Explains procedures, methods, and/or uncertainties.</td>
<td>Emphasizes a number of scientific processes. May explain in some detail particular steps in an experiment or in gathering data, may explain why particular choices were made, may emphasize the team nature of science, and/or discuss navigating uncertainties and/or risk in science.</td>
<td>Describes several scientific processes, will usually clearly describe steps and why they are needed, and/or make good efforts to meet particular uncertainties, choices, and/or other scientific processes such as data gathering, may not in appropriate depth.</td>
<td>May attempt to describe processes without complete success, may attempt but fail to adequately describe the nature of teams, uncertainties, choices and other processes at the audience’s level of understanding.</td>
<td>Describes scientific processes only occasionally, makes only brief mention with little to no description of teams, uncertainties, choices, or any other processes, or uses too much complicated detail.</td>
</tr>
<tr>
<td><strong>11. Trustworthy &amp; Personable:</strong> Scientists seem approachable, excited, knowledgeable</td>
<td>Expresses an excitement for and/or passion about science, especially the topic of the presentation. May explain in frequent, appropriate self-disclosure and/or uses personal stories and ideas, uses inclusive terms such as “we,” and our, portrays scientists positively</td>
<td>Demonstrates an excitement about and/or passion for science. Engages in regular, descriptive and personal stories, mostly uses inclusive pronouns rather than first person singular, speaks positively about scientists and their work</td>
<td>Excitement and enthusiasm for science is indicated at some level. Sometimes, engages in occasional personal reflection and uses inclusive pronouns, may or may not be positive about scientists</td>
<td>Expresses little emotion or enthusiasm for science work. Uses little self-disclosure or personal reflection, speaks largely in terms of “I” rather than “we,” speaks of scientists dispassionately, focuses on objectivity</td>
</tr>
<tr>
<td><strong>12. Engagement:</strong> Have dialogue and interactions about science with public audiences.</td>
<td>Audience members have multiple opportunities to ask and respond to questions, audience makes comments, speaker also a learner, there is active dialogue between audience and speaker during and after the presentation.</td>
<td>Audience members have opportunities to ask questions and make comments, speaker listens to audience but does not give them much time, even during Q&amp;A, little dialogue with audience.</td>
<td>Audience members have opportunity to ask but not respond to questions, audience makes few if any comments, speaker gives no indication of learning, no dialogue with audience members.</td>
<td>No questions asked of audience during presentation, audience and speaker interaction limited to a Q&amp;A period at the close of the presentation, no dialogue w/ audience</td>
</tr>
</tbody>
</table>

Topic’s relevance to the audience is unclear, little or no mention of common goals or values, connections between the audience and the topic are minimal or nonexistent.
APPENDIX E: 3-20-17 - RUBRIC FOR ASSESSING PUBLIC COMMUNICATION BY SCIENTISTS SENT TO FINAL CODERS BEFORE FIRST MEETING – SCIENCE COMMUNICATION ELEMENTS ONLY
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Makes a compelling case for the relevance and importance of the science topic to the audience’s lives, creates common ground by strongly emphasizing common goals, values and/or experiences, clearly and consistently connects science concepts to ideas familiar to audience</td>
<td>Uses clear, concrete, understandable language that considers audience’s current knowledge, particularly about science; defines any complex science terms used; when appropriate, uses comparisons and analogies to make complicated ideas clearer; any analogies that are used are interesting, deliberate, audience-focused, teaching analogies that clarify ideas</td>
<td>Excellent visual aids that use simplicity, clarity, and good design principles. Visuals focus on pictures or other graphic depictions, are easy to understand, use minimal text, only use clear graphs and charts, any non-electronic visuals are clearly visible</td>
<td>Emphasizes a number of scientific processes. May explain in some detail particular steps in an experiment or in gathering data, may explain why particular choices were made. May emphasize the team nature of science, and/or discuss navigating uncertainties and/or risk in science.</td>
<td>Expresses an excitement for science and other processes, or uses too much self-disclosure and/or personal reflection, is often inappropriate for audience’s level of understanding. Any comparisons may not be less relevant to audience or not fully developed, may show lack of awareness of or lack of adaptation to audience</td>
<td>Audience members have multiple opportunities to ask and respond to questions, audience makes comments, speaker is also a learner, there is active dialogue between audience and speaker during and after the presentation</td>
</tr>
<tr>
<td>3</td>
<td>Clearly discusses the importance and relevance of this science topic to the audience, tries to create common ground with the audience through shared values and goals, inconsistently connects science concepts to ideas familiar to the audience</td>
<td>Uses generally clear, concrete and understandable language that is mostly appropriate to audience’s science knowledge. Any comparisons/analogies seem deliberate and are generally helpful for the audience based on their understanding, analogies are appropriate for audience and thoughtfully expressed</td>
<td>Visual aids generally follow design principles and are helpful in clarifying points, but may have some aspects that are not entirely clear. Slides may be slightly cluttered, but are generally helpful and thoughtfully designed.</td>
<td>Describes several scientific processes, will usually clearly describe steps and why they are needed, and/or make good efforts to mention teams, uncertainties, choices, and/or other scientific processes such as data gathering, maybe not in appropriate depth.</td>
<td>Demonstrates an excitement about science is indicated at some level. Self-disclosure and/or personal reflection is present in some form, sometimes uses inclusive pronouns, may or may not be positive about scientists</td>
<td>Audience members have opportunities to ask questions and make comments, speaker listens to audience but does not give them much time, even during Q&amp;A, little dialogue with audience</td>
</tr>
<tr>
<td>2</td>
<td>May mention importance of topic and/or risk in science. Emphasizes both the relevance of science to the audience, may mention common goals/values, may make some connection between science concepts and the audience</td>
<td>May sometimes use complex or unclear language, may fail to define all science terms. Some analogies may not always be appropriate for or make sense to the audience, may not use comparisons or examples when they would help explain and clarify the topic</td>
<td>Visual aids usually adequately incorporate good design principles, but some slides may use too many elements, seem a bit cluttered, have too much text, or include some complex diagrams</td>
<td>May attempt to describe processes without complete success, may attempt but fail to adequately describe the nature of teams, uncertainties, choices and other processes at the audience’s level of understanding</td>
<td>Expresses little emotion or enthusiasm for science work. Uses little self-disclosure and/or personal reflection, speaks in terms of “I” rather than “we,” speaks of scientists dispassionately, focuses on objectivity</td>
<td>Audience members have opportunity to ask but not respond to questions, audience makes few if any comments, speaker gives no indication of learning, no dialogue with audience members</td>
</tr>
<tr>
<td>1</td>
<td>Only brief mention of topic’s relevance to the audience, few, if any, mentions of common goals or values, makes little connection between the audience and the topic</td>
<td>Uses several complex terms without defining them; language is often inappropriate for audience’s level of understanding. Any comparisons may not be less relevant to audience or not fully developed, may show lack of awareness of or lack of adaptation to audience</td>
<td>Visual aids often ignore good design principles, may often be unclear or may use too much text; charts/graphs may be too plentiful and/or too complex; graphics may be complicated</td>
<td>Describes scientific processes only occasionally, makes only brief mention with little to no description of teams, uncertainties, choices, or any other processes, or uses too much complicated detail.</td>
<td>Fails to clearly describe any scientific processes. Little to no description of, data gathering, reasoning, uncertainties, choices, teams, or other scientific processes or uses too much irrelevant detail.</td>
<td>No questions asked or interactions had with audience during presentation, audience and speaker interaction limited to a Q&amp;A period at the close of the presentation, no dialogue with audience</td>
</tr>
<tr>
<td>0</td>
<td>Topic’s relevance to the audience is unclear, little or no mention of common goals or values, connections between the audience and the topic are minimal or nonexistent</td>
<td>Language is often too complex or unclear, frequently uses scientific terms that are outside the audience understanding. Uses few or no metaphors or other comparisons, or comparisons are not relevant to audience or not fully developed</td>
<td>Visual aids often ignore good design principles, may cause audience confusion. May use too many colors, fonts, charts, or graphics, text-heavy slides, confusing graphics</td>
<td>Fails to clearly describe any scientific processes. Little to no description of data gathering, reasoning, uncertainties, choices, teams, or other scientific processes or uses too much irrelevant detail.</td>
<td>Demonstrates no passion or excitement for science. Little to no self-disclosure and/or personal stories, may use mostly “I” language, uses few or no descriptions of scientists in a positive, credible light</td>
<td>No questions asked or interactions had with audience during presentation, audience and speaker interaction limited to a Q&amp;A period at the close of the presentation, no dialogue with audience</td>
</tr>
</tbody>
</table>

236
APPENDIX F: 3-24-17 ONLY SCIENCE PORTION OF THE RUBRIC (PSCR PORTION DID NOT CHANGE FROM THIS POINT ON) INCLUDES SCORING/NOTE TAKING SHEET USED FOR ROUND 1 OF FINAL TESTING
<p>| Relevance &amp; Imporance: | 4 | Connects science to lives, shows importance and relevance of science. Makes a compelling case for the relevance and importance of the science topic to the audience’s lives. Creates common ground by strongly emphasizing common goals, values, and/or experiences, clearly and consistently connects concepts to ideas familiar to audience. Clearly discusses the importance and relevance of this science topic to the audience, tries to create common ground with the audience through common values and goals, usually connects science concepts to ideas familiar to the audience. May mention importance of topic but not make it explicitly relevant to the audience, inconsistent connection to common goals/values, may make some connection between science concepts and the audience. Only brief mention of topic’s importance and/or relevance to the audience, few, if any, mentions of common goals or values, makes little connection between the audience and the topic. Topic’s relevance and importance to the audience is unclear, little or no mention of common goals or values, connections between the audience and the topic are minimal or nonexistent. |
| Language: | 3 | Uses clear, concrete, understandable language that considers audience’s current knowledge, particularly about science; defines any complex science terms used; when appropriate, uses comparisons and analogies to make complicated ideas clearer; any analogies that are used are interesting, deliberate, audience-focused, teaching analogies that clarify ideas. Uses generally clear, concrete, and understandable language that is mostly appropriate to audience’s science terms usually defined; any comparisons or analogies seem deliberate and are mostly helpful to the audience at their understanding, analogies are appropriate for audience and thoughtfully expressed that clarify ideas. May sometimes use complex or unclear language, may fail to define all science terms. Some analogies may not always be appropriate for or make sense to the audience, may not use comparisons or examples when they would help explain and clarify the topic. Uses several complex terms without defining them, language is often inappropriate for audience’s level of understanding. Any comparisons may not be less relevant to audience or not fully developed, may show lack of awareness of or lack of adaptation to audience. Language is often too complex or unclear, frequently uses scientific terms that are outside the audience understanding. Uses few or no metaphors or other comparisons, or comparisons are not relevant to audience or not fully developed. |
| Visuals: | 2 | Nearly all visual aids use simplicity, clarity, and good design principles. Visuals focus on images or other graphic depictions, are easy to understand, use minimal text, only use clear graphs and charts, non-electronic objects are clearly visible. Visual aids are usually clear, simple, and helpful, but may use several images or graphics per slide or smaller than ideal font size. Slides may use more than minimal text, most charts/graphs are clear, any objects visible. Visual aids somewhat clear, but some slides may use too many visuals, be cluttered with graphics or too much text, may include some complex charts/graphs, physical objects hard to see. Visual aids lacking in clarity and simplicity, images or graphics are unclear, use too much text, may have too many charts/graphs, may be too complex, non-electronic objects may not be visible. Visual aids are confusing, cluttered, and unclear. Most have too many elements. Slides may be text-heavy; poor design of charts, graphs, images. |
| Explain Science Processes: | 1 | Explains a number of scientific processes. May explain in some detail certain steps in an experiment or in gathering data, explains to the audience appropriate choices in some depth. May emphasize the team nature of science, and/or discuss navigating uncertainties and/or risk in science. Describes several scientific processes, will often clearly describe steps and why they are needed, and/or make good efforts to mention teams, uncertainties, choices, and/or other scientific processes such as data gathering, maybe not in great depth. May attempt to describe processes without complete success, may discuss data gathering, uncertainties, teams, and/or choices, but explanations are incomplete or too complex for audience understanding. Describes scientific processes only occasionally, makes only brief mention with little to no description of teams, uncertainties, choices, or any other processes, or uses too much complicated detail. Fails to clearly describe any scientific processes. Little to no description of, data gathering, text may be small. Slides may be text-heavy; poor design of charts, graphs, images. |
| Trustworthy &amp; Personable | 0 | Scientists seem approachable, excited, knowledgeable. Expresses an excitement for and/or passion about science, especially the topic of the presentation, engages in frequent, appropriate self-disclosure and/or uses personal stories and ideas, frequently uses inclusive terms such as we, us, and our, portrays scientists positively. Demonstrates an excitement about and/or passion for science. Engages in regular self-disclosure and/or personal stories, mostly uses inclusive pronouns rather than first person singular (I, me), speaks positively about scientists and their work. Excitement and passion for science is indicated at some level. Self-disclosure and/or personal stories are present in some form, sometimes uses inclusive pronouns, may or may not be positive about scientists and science work. Infrequent use of self-disclosure and/or personal stories, uses “I” and “me” language, uses few or no descriptions of scientists in a positive light. Demonstrates no passion or excitement for science. Little to no self-disclosure and/or personal stories, may use mostly “I” language, uses few or no descriptions of scientists in a positive light. |
| Engagement: | | Audience members have multiple opportunities to ask and respond to questions, audience makes comments, speaker is also a learner, there is active dialogue between audience and speaker during and after the presentation. Audience members have opportunities to ask questions and make comments, speaker listens to audience but devotes less than ideal time to questions and interactions, has a Q&amp;A session, less dialogue with audience. Audience members have opportunity to ask but not respond to questions, audience makes few if any comments, speaker gives no indication of learning from audience, little dialogue with audience member. No questions asked or interactions had with audience during presentation, audience and speaker interaction limited to a Q&amp;A period, no dialogue w/ audience. Speaker consistently assumes an authoritative air, does not ask questions, and speaker interaction limited to a Q&amp;A period, no dialogue or interaction with audience, may not invite questions. |</p>
<table>
<thead>
<tr>
<th>Examples and Notes</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Relevance &amp; Importance</td>
<td>Example: “Because of our research, we know there are more birds of this species left that we originally thought, and we are better able to preserve the species because we know what habitats they prefer and what their migration routes are.”</td>
<td>Compelling</td>
<td>Strongly</td>
<td>Clearly</td>
<td>Mention</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Strongly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consistently</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consistent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Connection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Importance Relevance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Common values</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Language</td>
<td>Example: “The fish all died in the experiment” “Proteins are like building blocks for the cell”</td>
<td>Clear</td>
<td>Defines</td>
<td>Generally</td>
<td>Sometimes</td>
</tr>
<tr>
<td>Clear, concrete</td>
<td></td>
<td>Comparison</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Definitions Metaphors</td>
<td></td>
<td>s when</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparisons to known world</td>
<td></td>
<td>appropriate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audience focus</td>
<td></td>
<td>Deliberate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Importance</td>
<td></td>
<td>Teaching</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevance</td>
<td></td>
<td>Audience</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Visuals</td>
<td>Example: Large image, little text, clear, legible text, simple charts and graphs, easy to follow layout</td>
<td>Clear</td>
<td>Large</td>
<td>Usually</td>
<td>Somewhat</td>
</tr>
<tr>
<td>Clear, large text</td>
<td></td>
<td>image</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Pictures</td>
<td></td>
<td>Limited</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple Visuals Basic Charts</td>
<td></td>
<td>text</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brief Text</td>
<td></td>
<td>Simple</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncluttered</td>
<td></td>
<td>Visible</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Science Processes Explanations</td>
<td>Example: “We measure the eggs in mid-July, since by that time the mothers have left the nests, and we aren’t disturbing any of the nesting patterns.”</td>
<td>Explains</td>
<td>A number</td>
<td>Describes</td>
<td>Attempt</td>
</tr>
<tr>
<td>Data gathering</td>
<td></td>
<td>of</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedures/Steps</td>
<td></td>
<td>Some detail</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methods Uncertainties</td>
<td></td>
<td>Emphasize</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Discuss</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>In some</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>depth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Trustworthy &amp; Personable Excited</td>
<td>Example: “I get to swim with turtles on my spring break, and I can’t think of anything I’d rather be doing.”</td>
<td>Expresses</td>
<td>Engages</td>
<td>Demonstrates</td>
<td>Indicated</td>
</tr>
<tr>
<td>Passionate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledgeable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-disclosure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inclusive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Engagement Dialogue, Interactions Questions, Shared Learning</td>
<td>Example (during the presentation): “How many of you have ever used building bricks like Legos®? Let’s see your hands.”</td>
<td>Multiple</td>
<td>Active</td>
<td>Opportunity</td>
<td>Not respond</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX G. 3-24-17 SCIENCE ONLY RUBRIC WITH SCORING/NOTE TAKING PAGE USED FOR ROUND 2 OF FINAL TESTING
<table>
<thead>
<tr>
<th>Question</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Relevance &amp; Importance: Connects science to lives, shows importance and relevance of science</td>
<td>Makes a compelling case for the relevance and importance of the science topic to the audience’s lives.</td>
<td>Clearly discusses the importance and relevance of this science topic to the audience, tries to create common ground with the audience through common values and goals, usually connects science concepts to ideas familiar to the audience.</td>
<td>May mention importance of topic and show some relevance to the audience.</td>
<td>Only brief mention of topic’s importance and/or relevance to the audience.</td>
<td>Topic’s relevance and importance to the audience is unclear, little or no mention of common goals or values.</td>
</tr>
<tr>
<td>8. Language: Uses language to make the complex clear, definitions, metaphors, and comparisons</td>
<td>Uses clear, concrete, understandable language that considers audience’s current knowledge, particularly about science; defines any complex science terms used; when appropriate, uses comparisons and analogies to make complicated ideas clear; any analogies that are used are interesting, deliberate, audience-focused, teaching analogies that clarify ideas.</td>
<td>Uses generally clear, concrete and understandable language that is mostly appropriate to audience’s, science terms usually defined, any comparisons or analogies seem deliberate and are mostly helpful to the audience at their understanding, analogies are appropriate for audience and thoughtfully expressed.</td>
<td>May sometimes use complex or unclear language, may fail to define all science terms. Some analogies may not always be appropriate for or make sense to the audience, may not use comparisons or examples when they would help explain and clarify the topic.</td>
<td>Uses several complex terms without defining them; language is often inappropriate for audience’s level of understanding. Any comparisons may not be less relevant to audience or not fully developed, may show lack of awareness of or lack of adaptation to audience.</td>
<td>Language is often too complex or unclear, frequently uses scientific terms that are outside the audience understanding. Uses few or no metaphors or other comparisons, or comparisons are not relevant to audience or not fully developed.</td>
</tr>
<tr>
<td>9. Visuals: Uses clear, concrete, understandable language that considers audience’s current knowledge, particularly about science; defines any complex science terms used; when appropriate, uses comparisons and analogies to make complicated ideas clear; any analogies that are used are interesting, deliberate, audience-focused, teaching analogies that clarify ideas.</td>
<td>Nearly all visual aids use simplicity, clarity, and good design principles. Visuals focus on images or other graphic depictions, are easy to understand, use minimal text, only use clear graphs and charts, non-electronic objects are clearly visible.</td>
<td>Visual aids are usually clear, simple, and helpful, but may use several images or graphics per slide or smaller than ideal font size. Slides may use more than minimal text, most charts/graphs are clear, any objects visible</td>
<td>Visual aids somewhat clear, but some slides may use too many visuals, be cluttered with graphics or too much text, may include some complex charts/graphs, physical objects hard to see.</td>
<td>Visual aids lacking in clarity and simplicity, images or graphics are unclear, use too much text; may have too many charts/graphs, may be too complex, non-electronic objects may not be visible.</td>
<td>Visual aids are confusing, cluttered, and unclear. Most have too many elements, text may be small. Slides may be text-heavy, poor design of charts, graphs, images.</td>
</tr>
<tr>
<td>10. Explain Science Processes: Explains procedures, data, choices, and/or uncertainties.</td>
<td>Explains a number of scientific processes. May explain in some detail certain steps in an experiment or in gathering data, may explain particular choices in some depth. May emphasize the team nature of science, and/or discuss navigating uncertainties and/or risk in science.</td>
<td>Describes several scientific processes, will often clearly describe and explain why they are needed, and/or make good efforts to mention teams, uncertainties, choices, and/or other scientific processes such as data gathering, maybe not in great depth.</td>
<td>May attempt to describe processes without complete success, may discuss data gathering, uncertainties, teams, and/or choices, but explanations are incomplete or too complex for too complex for audience understanding.</td>
<td>Describes scientific processes only occasionally, makes only brief mention with little to no description of teams, uncertainties, choices, or any other processes, or uses too much complicated detail.</td>
<td>Fails to clearly describe any scientific processes. Little to no description of data gathering, uncertainties, choices, teams, or other scientific processes, or uses too much irrelevant detail.</td>
</tr>
<tr>
<td>11. Trustworthy &amp; Personal: Expresses an excitement for and/or passion about science, especially the topic of the presentation, engages in frequent, appropriate self-disclosure and/or personal stories, mostly uses inclusive pronouns rather than first person singular (I, me), speaks positively about scientists and their work.</td>
<td>Expresses an excitement for and/or passion about science. Engages in regular self-disclosure and/or personal stories, mostly uses inclusive pronouns rather than first person singular (I, me), speaks positively about scientists and their work.</td>
<td>Excitement and passion for science is indicated at some level. Self-disclosure and/or personal stories are present in some form, sometimes uses inclusive pronouns, may or may be positive about scientists.</td>
<td>Expresses little excitement or passion for science work. Infrequent use of self-disclosure and/or personal stories, uses &quot;I&quot; and &quot;me&quot; than inclusive pronouns, scientists portrayed as distant, not positive.</td>
<td>Demonstrates no passion or excitement for science. Little to no self-disclosure and/or personal stories, may use mostly &quot;I&quot; language, uses few or no descriptions of scientists in a positive light.</td>
<td></td>
</tr>
<tr>
<td>12. Engagement: Have dialogue and interactions about science with public audiences.</td>
<td>Audience members have multiple opportunities to ask and respond to questions, audience makes comments, speaker listens to audience but devotes less than ideal time to questions and interactions, has a Q&amp;A session, less dialogue with audience.</td>
<td>Audience members have opportunity to ask but not respond to questions, audience makes few if any comments, speaker gives no indication of learning from audience, little dialogue with audience.</td>
<td>No questions asked or interactions had with audience during presentation, audience and speaker interaction limited to a Q&amp;A period at the close of the presentation, no dialogue w/ audience.</td>
<td>Speaker consistently assumes an authoritative air, does not ask questions, gives no indication of learning, no dialogue or interaction with audience, may not invite any questions.</td>
<td>Speaker consistently assumes an authoritative air, does not ask questions, gives no indication of learning, no dialogue or interaction with audience, may not invite any questions.</td>
</tr>
<tr>
<td></td>
<td>Examples and Notes</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>---</td>
<td>-------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>
| 7. | **Relevance & Importance**  
Connection  
Importance  
Relevance  
Common values  
Goals | Example: “Because of our research, we know there are more birds of this species left that we originally thought, and we are better able to preserve the species because we know what habitats they prefer and what their migration routes are.” | Compelling  
Strongly  
Clear  
Consistent | Clearly  
Tries  
Usually | Mention  
Not explicit  
Some  
Inconsistent | Brief  
Mention  
Little  
connection | Unclear  
No mention  
Minimal  
Nonexistent |
| 8. | **Language**  
Clear, concrete  
Definitions  
Metaphors  
Comparisons to known world  
Audience focus | Example: “The fish all died in the experiment” “Proteins are like building blocks for the cell” | Clear  
 Defines  
Comparisons when appropriate  
Deliberate  
Teaching  
Audience | Generally  
Understandable  
Mostly  
Seem  
Deliberate  
Appropriate | Sometimes  
Not always defined  
Inappropriate to audience  
Absence of comparison | Complex  
Inappropriate  
Not relevant  
Lack of audience awareness | Unclear  
Complex  
Not defined  
No comparison | Not relevant  
No audience focus |
| 9. | **Visuals**  
Clear, large text  
Large Pictures  
Simple Visuals  
Basic Charts  
Brief Text  
Uncrowded | Example: Large image, little text, clear, legible text, simple charts and graphs, easy to follow layout | Clear  
Large  
image  
Limited  
text  
Simple  
Visible  
Uncrowded | Usually  
Mostly  
clear  
Smaller than ideal  
Mostly  
minimal text | Somewhat  
Cluttered  
Too much text  
Complex | Lack clarity  
Unclear  
Too many  
Not visible | Confusing  
Cluttered  
Unclear  
Small  
Poor design |
| 10. | **Science Processes**  
Explanations  
Data gathering  
Procedures/Steps  
Methods  
Uncertainties | Example: “We measure the eggs in mid-July, since by that time the mothers have left the nests, and we aren’t disturbing any of the nesting patterns.” | Explains  
A number  
Some detail  
Emphasize  
Discuss  
In some depth | Describes  
Several  
Good  
efforts  
Mention  
Not as much depth | Attempt  
Discuss  
Incomplete  
Too complex | Occasional  
Brief  
Mention  
Little  
Too much detail | Fails  
Little to no  
Too much  
irrelevant detail |
| 11. | **Trustworthy & Personable**  
Excited  
Passionate  
Knowledgeable  
Self-disclosure  
Inclusive | Example: “I get to swim with turtles on my spring break, and I can’t think of anything I’d rather be doing.” | Expresses  
Engages  
Frequently  
Positively | Demonstrates  
Regular  
Mostly | Indicated  
Some level  
Present  
Sometimes  
May or may not | Little  
Infrequent  
Distant | No  
Little to no  
Few |
| 12. | **Engagement**  
Dialogue  
Interactions  
Questions  
Shared Learning | Example (during the presentation): “How many of you have ever used building bricks like Legos®? Let’s see your hands.” | Multiple  
Active  
Opportunities  
Listens | Not respond  
Few  
No  
Little | No  
Limited  
No | Authoritative  
Yes  
No  
Not |
APPENDIX H. FIRST CODE BOOK FOR APPS – THE ASSESSMENT OF PUBLIC PRESENTATIONS BY SCIENTISTS

Code Book

Purpose:
This rubric can be used to assess informal, public presentations by scientists. The presentations that can appropriately be assessed with this rubric include what we call “library talks”: presentations where scientists are invited to share their research, knowledge, expertise, and opinions with a group of intelligent but unaware, non-scientist members of the general public.

Limitations
This rubric is specifically not appropriate for situations in which the scientist has a deliberate persuasive agenda, such as when the scientist is seeking funding, attempting to influence policy decisions, or trying to win favor for a particular science museum or organization.

Sections:
This rubric is divided into two basic sections. The first section is borrowed from the Public Speaking Competency Rubric developed by Schreiber, Paul, & Shibley (2012), and addresses issues of delivery and speech organization. The second section is unique to this assessment, and addresses important skills and abilities that, according to our research, scientists should demonstrate when speaking to public audiences.

Definitions and Descriptions:
This code book describes the terms used in the rubric and gives examples of speakers who expertly fulfilled the expectations of each element. If, after reading this code book, you have more questions about the measure or what constitutes competence and excellence in that area, please contact the rubric developers.
Performance Standards:

1. **Select a topic appropriate to the audience and the occasion:**
   All presentations will address a science topic, so given that caveat, is the topic:
   a. **Engaging:** Is the topic framed, or presented, in a way that is interesting/appealing to an audience of this demographic mix such as age, education level, etc.? If the audience is mixed, is it presented in a way that appeals to a wide range of the audience members?
   b. **Complexity:** is it presented in a way that makes it complex enough to be interesting but not so complex that the audience is unable to grasp the concepts?
   c. **Situation:** Is the topic right for the situation or context? In other words, does it take into account where and when the presentation is being given, the audience, the time of year, the current world/area conditions, etc.?

2. **Formulate an introduction that orients audience to topic and speaker:** (The attention getter should be first and the preview last, but other elements can be in and order)
   a. **Attention getter:** The speaker should start off with something that grabs the audience’s attention such as a quote, a story, a song, a video clip, a startling statistic – something that pulls the listener in. Unless the speaker has not been introduced, he/she should not begin by stating their name and/or lab affiliation. That makes for a dull introduction. If the speaker was not introduced by someone else, then stating their name is fine, but it should be followed immediately by an attention getting device.
   b. **Establish Credibility:** The speaker lets the audience members know why she/he is worth listening to on this topic. The simple fact that they are a scientist working in a lab is not adequate for firmly establishing credibility. They should express why they’re an expert on this particular topic. Ideally, they’ll say something about how long they’ve studied the issue, why they became interested in the topic or how it personally affects them. The best expressions of credibility include a personal connection to the topic as well as authoritative credibility such as a degree in the field or extensive study of the topic.
   c. **Sound orientation to the topic:** The speaker gives the audience an idea of why this topic is of interest to them. The connections between the topic and the audience should be continually made throughout the presentation, but during the introduction, the audience members should understand that this topic does affect them and their lives in some way and should understand that it is of interest to them.
      1. **Example 1:** “You may not realize it, but there are polymers all around you every day of your life. Your water bottle is a polymer, the pan you use to cook your eggs probably has a polymer on it, the toothpaste you use contains polymer. Polymers are all around you.”
   d. **Clear thesis:** The speaker clearly states the main idea of the presentation at some point in the introduction. In an ideal thesis statement, the speaker tells the
audience generally what the presentation is going to be about and clearly defines any complex terms so that the audience clearly understands what they’re discussing.

e. **Preview of main points:** The speaker should let the audience members know the basic direction the speech is going to take by ending the introduction with a clear and direct statement of the main points that are coming up in the speech. The preview will forecast the organizational pattern of the speech as well. The speaker should not give out details in the preview, but should clearly summarize each of the upcoming main points.

   i. **Example 1:** “So today I’ll step you through the process we take in getting a fuel from leaves – which gives us a product that is called biofuel. We’ll talk about the plants we work with and how we choose them, then we’ll find out how we learn to squeeze oil from the plants, then how we make large amounts of oil, and then how we can turn that into fuel for your cars.”

3. **Use an effective orientation pattern:** Often presentations are organized by time (this comes first, this comes second…) or by topic (here’s one important thing you should know, here’s another important thing you should know…) There are other ways to organize a presentation, such as problem and solution (here’s something that was really causing difficulty for a group of people; here’s what we did or what we are doing to fix it).

   Speakers can use any pattern of organization, but it should be clear, make sense to the audience, and be clearly stated. If speakers seem to be rambling around different ideas related to a topic, then they are not demonstrating good organization.

   a. **Main points clear:** Ideally a speaker will state each main point as they get to it in a way that makes it clear for the audience that they are moving on to the next section of the talk. To make listening to the speaker as easy as possible, the speaker could structure each main point similarly.

      i. **Example 1:** “The first step in making biofuel is choosing the plants or plant products we’re going to use…. (supporting information comes here). The next step in making biofuel is squeezing the oil out of the plants we’ve chosen…. (supporting info goes here). The next step in making biofuel is turning that oil into large amounts of oil… (supporting info goes here). The last step in making biofuel is processing that oil so it can be used in cars… (supporting material).

   b. **Mutually exclusive:** Each section of a presentation should address a single issue or process. The speaker can certainly refer to something they’ve already discussed, but should avoid the wandering from point to point. When in the section about producing large amounts of oil, the speaker should not jump to telling a story about how the oil is processed for cars and then jump back to producing large amounts of oil, since it confuses listeners.
c. **Related to the thesis:** Main points should be on point as previewed in the introduction. There are always many interesting things to talk about, but a speaker needs to choose the handful of points they are going to make (usually from two to five main points in a 10 to 20-minute presentation) and stick with those ideas. Particularly with science topics where an audience may have little background and understanding, bringing in too much information just confuses the listeners.

d. **Effective transitions and signposts:** Transitions and signposts lead the listener from one point to the next in a speech. Transitions may show a causal relationship between main points, a complementary relationship, or a contrasting relationship. They may indicate a chronological relationship as well. A speaker should orally speak transitions and signposts, not rely on a PowerPoint slide to make the move or simply jump to the next point. Signposts can be as simple as one word.

i. **Examples of signposts:** next, then, first, now, move on to, turn to, elaborate on, most importantly, to begin

ii. **Examples of transitions:**

   1. **Internal:** “Now that we’ve talked about choosing the plants, let’s look at what we do with the plants once we get them”

   2. **Internal summary:** “So now we know that we have to look at the plant’s molecular structure, its rate of growth, and the amount of energy it stores when we decide if it’s a good plant to use for biofuel. Now let’s look at actually using the plant.”

   3. **Internal preview:** “Now that we’ve discussed choosing the plant, we’ll discuss getting oil through breaking down the proteins, extracting the oil, and collecting the oil.”

4. **Develop a conclusion that reinforces the thesis and provides psychological closure:** Each presentation needs a conclusion. Speakers should avoid comments such as “that’s it” or “that’s all I got” or anything similar. A good conclusion restates the main points or summarizes them generally, connects to the thesis, and ends with a strong take-away/clincher/call to action. There should be a transition of some kind between the last main point and the conclusion, even if it is as simple as “So, in conclusion...” or “To wrap up....” or “So let’s review what we’ve learned...”

   a. **Example 1:** “So to summarize, there are many complicated steps in creating biofuel, including finding the best plants for the job, getting the oil out of those plants and making big supplies of it, and then turning that oil into something our cars can use. The process takes time, but is a smart way to use the plant waste that we are always going to have. Your car and the farmer’s tractor really can run on Plant Power!”

5. **Effectively use vocal expression and paralanguage to engage the audience:** Effective speakers are expressive in with their voice. They avoid speaking in monotone or having a lift or drop in their voice at the end of every sentence. They exhibit what is called vocal
variety in their speaking. They vary their tone, pitch, rate of speaking, and intensity and they avoid vocal fillers (words such as “uh,” “um,” “like,” “you know,” and “well” that are often said unconsciously, indicate nervousness or lack of forethought, and serve as a distraction to the listener). Ideally, the speaker sounds like she/he is having an animated conversation with a good friend, telling the friend a story. They should sound enthusiastic about the topic and natural in their speaking.

6. **Use eye contact, facial expressions, and body movement to support the verbal message:** A speaker’s face and body movements should enhance their speaking. Ideally, the speaker will show the following types of body movement:
   a. **Posture:** Erect, natural and balanced (equal weight on each foot). Avoids slouching over a podium or leaning to one side or the other.
   b. **Eye Contact:** Frequent, if not constant, eye contact with all audience members. Looks around at the whole audience, not just one or two people. Glances at notes or PowerPoint only occasionally, not constantly. Body is facing forward, toward the audience, not angled toward a screen or visual aid.
   c. **Facial Expressions:** Varies facial expressions as appropriate to what is being said. Avoids keeping the same expression throughout the talk. Smiles, raises eyebrows, shows surprise, or uses whatever expressions are appropriate for the talk.
   d. **Movement and Gestures:** Avoids standing in one spot the entire presentation. Instead, moves naturally around the speaking area. Moves hands and arms in natural gestures, indicates important items on visual aids, uses gestures to illustrate points.
   e. **Confidence and poise:** Is competently incorporating most of the above behaviors. Seems at ease and happy to be speaking.

7. **Connect science information to a specific audience in a specific situation:** The speaker should give clear connections between the science and individual lives of the audience members. Speakers should show awareness of audience member’s interests, concerns, and desires by the way they make these connections. If they are speaking largely to college students, they might connect their topic to making money, social lives, jobs students may get in the future, and so on. If they are talking largely to families with children still at home, they might mention safety, consumption, making life easier, budgeting, education, or transportation. If they are talking to older adults, they might emphasize security, how something might help grandchildren, how an innovation could save them money, or how the topic relates to a hobby they may enjoy. An excellent speaker will make these connections frequently throughout the presentation.

Ideally the speaker deliberately gives examples of ways they have values and interests in common with the audience, such as a desire to make life easier for those with disabilities or those who are concerned about aging parents or those concerned about the high cost of living.
Speakers will ideally make explicit connections between the science they are discussing and the audience member’s daily lives. Speakers will express what they find interesting and exciting about science and what the audience might find interesting/exciting about the science.

8. **Clearly communicate complex ideas and make them understandable to the general public:** Scientists work with complex ideas and processes that are difficult for non-experts to understand. Ideally, scientists giving a public presentation will develop clear ways to help non-experts understand the complex ideas. There are several communication techniques that have been shown to help communicate clearly.

**Metaphor:** One technique is metaphor and other comparisons. Research shows that pedagogical metaphors that connect to the audience members’ experiences are effective at explaining complex ideas as are theory-building metaphors. A pedagogical metaphor is one that is meant to teach about a concept or phenomenon, something that describes existing knowledge and for which other possible ways of explaining exist. A theory-building metaphor is one that is used as a new way to explain a scientific phenomenon which cannot be paraphrased because there is not another way to talk about the phenomenon (Boyd, 1993). These distinctions are so fine as to be impossible to make, and it is not necessary to distinguish between them for purposes of assessing speakers. Primarily, scientists should avoid metaphors and comparisons that are based on knowledge inherent to being a scientist and have no explanatory power to people who do not have scientific expertise.

- **Example 1** – **pedagogical metaphor** – messenger RNA, proteins as building blocks
- **Example 2** – **theory-building metaphor** – genetic code
- **Example 3** – **cognitive metaphor** – polymer chain (used as if all present would understand, with no explanation of how polymers are chains or why they might be thought of as chains).

An expert speaker will use a few metaphors or comparisons throughout the presentation rather than offering up multiple metaphors.

The best metaphors/comparisons compare the new object or concept to something that is familiar to audience members and gives an accurate picture of the scientific concept.

**Clear, vivid language:** Makes use of descriptive words, precise (not generic) nouns and verbs, and simple, plain language. It avoids scientific jargon or “inside” terminology in favor of crisp, to-the-point declarative sentences.

- **Example 4** – **(jargon)** The biota exhibited a one hundred percent mortality response.
  **(plain, crisp English)** All of the fish died.
- **Example 5** – **(precise language)** Generic noun: house; More precise nouns: cabin, shack, cottage, mansion, basic ranch, mansion. Generic verb: leave behind; Stronger verb: abandon
9. **Increase the audience’s knowledge and understanding of science and scientific processes:** Ideally, science speakers increase the audience’s knowledge of science and scientific processes through several means. One is that the scientist defines any scientific terms they use in clear English. If they are going to err, they should err on the side of defining a term for the public. The definitions may be brief, but they should be clear. Another thing scientists can do is explain processes, steps, and reasons for these steps and processes. These explanations should also be clear, concise, and in language familiar to the audience.

Specific aspects of science and science processes that are often emphasized and are great to emphasize in public presentations are the team/cooperative nature of science, uncertainty and how the uncertainty is managed. It is particularly good when scientists note how they build on the work of others, work in teams, write research papers together, and incorporate other’s ideas into their experiments. It is also important that scientists speak confidently and knowledgably about uncertainty – explaining how scientists think about uncertainty, how they adjust for it, and how it affects results of scientific tests.

10. **Humanize scientists and help them seem trustworthy and knowledgeable:** The best way to build trust and trust is through self-disclosure. Self-disclosure is simply the revealing of personal information to others. This can be information about things that happen in the lab or outside their science work, but the comments should be person in nature. Ideally there are instances of self-disclosure that include vulnerability – indications that the scientist has challenges or weaknesses. This tends to create a sense within the audience that scientists are like them and are trustworthy. Statements of self-disclosure:

   a. **Example 6** – “We couldn’t figure out why this stuff was so sticky; I showed my lab partner what was happening and he couldn’t figure it out either.”

   b. **Example 7** – “My daughter is afraid of birds, so I don’t bring her to the lab very often…”

   c. **Example 8** – “I think one of the perks of my job is crawling around in the river with turtles”

Additionally, scientists can tell stories about themselves and their work, and portray themselves and other scientists in a positive light by talking about themselves and their colleagues and their work in a way that meshes with the values of their audience.

**Concrete language:** Scientists can also increase the trust the audience has in them by using concrete language. Studies show that a speaker using concrete rather than abstract terms gives listeners a sense of confidence in the speaker and trust in what they say. Often there is a continuum between abstract and concrete, and a scientist’s language will fall somewhere on the continuum. In general, it is better to be closer to the concrete than to the abstract.

**Inclusive Language:** Speakers build relationships of trust with their listeners when they use inclusive language. Inclusive language frames the speaker as part of the group, and can
even include the audience in that group. For example, scientists who use “we,” “our,” and “us” rather than “I” or “you” will do a better job at connecting with and building a relationship of trust with the audience.

d. **Example 9:** Abstract language: “To excel at school, you need to work hard.”
Concrete language: “To excel at school, you need to attend class, read before attending class, review and revise all homework, and take notes that you review each week.”

e. **Example 10:** Non-inclusive language: “I identified the bacteria by isolating it in the lab, and then I tested each of the samples.” Inclusive language: “We worked on the project from May through November.”

11. **Engage in dialogue and interactions about science with public audiences.**
Since we do not have adequate data to properly test this aspect of speaking, we will skip it this time.
APPENDIX I: FINAL CODE BOOK FOR APPS – THE ASSESSMENT OF PUBLIC PRESENTATIONS BY SCIENTISTS CODE BOOK, INCLUDING THE FULL FINAL RUBRIC
Assessing Public Presentations
By Scientists (APPS)

A Code Book and Rubric to assist scientists in making effective presentations to public audiences.

Are you a scientist who wants to learn how to better engage with the public? Are you a communication specialist or a mentor of scientists who wants to help scientists better engage with the public? If so, this code book and rubric can help! It draws on rigorous communication theory and practice as well as a thorough examination of the knowledge, skills, and abilities science societies, scientists, and government entities want scientists to exhibit when they engage with the public. Using these documents for training and assessment can lead to better skills and improved, interactive, and interesting presentations.
Code Book for the Assessment of Public Presentations by Scientists (APPS)

**Purpose:**
This code book is to be used with the Assessment of Public Presentations by Scientists (APPS) rubric. The APPS rubric is designed to assess informal, informative, public presentations by scientists. The presentations that can appropriately be assessed with this rubric include what can be called “library talks”: presentations where scientists are invited or offer to share their research, knowledge, expertise, and opinions with a group of intelligent but uninformed, non-scientist members of the general public. The rubric may also be used to assess other similar informative presentations by scientists.

**Limitations:**
This rubric is intended for use as a formative or summative assessment when evaluating scientists giving informative presentations to public groups. This rubric is specifically not appropriate for situations in which the scientist has a deliberately persuasive agenda, such as when the scientist is seeking funding or attempting to influence policy decisions, or is trying to win favor for a particular science museum or organization. This rubric and code book should be used in conjunction with a training program where scientists learn public communication skills, have access to the rubric and codebook and receive instruction on how to develop the skills outlined in those documents, have the opportunity to practice those skills regularly, and get rapid, helpful feedback through the rubric and instructor comments.

**Sections:**
This rubric is divided into two basic sections. The first section, used by permission of the authors, is adapted from the Public Speaking Competency Rubric (PSCR) developed by Schreiber, Paul, & Shibley (2012), and assesses a speaker’s delivery and speech organization. The second section is unique to this assessment, and addresses important skills and abilities that, according to our research, scientists should demonstrate when speaking to public audiences.

**Definitions and Descriptions:**
This code book describes the terms used in the rubric and gives examples of speakers who expertly fulfilled the expectations of each element. If, after reading this code book, you have additional questions about any of the measures or what constitutes competence and excellence in a particular area, please contact the rubric developers.

**Scoring**
A speaker’s performance can be thought of as falling on a continuum. There is not one specific
measure for success or failure in any category. For example, there is not a number of lines of
text or number of images on a slide that drops a speaker from a high score to a low score in
visuals. Rather, in each category the speaker should be assessed as to how well (s)he is fulfilling
the ideal knowledge, skills, and abilities outlined in that category based on a range from
expertly demonstrating all those abilities (4) to not expressing those qualities at all (0). The
wording in the rubric cells should help assessors place a speaker’s performance on that
continuum.
Assessed Performance Standards:

Public Speaking Elements –

1. Topic

“The speaker will select a topic appropriate to the audience and the occasion”

All presentations will address a science topic, so given that caveat, is the topic:

a. Engaging: Is the topic framed, or presented, in a way that is interesting/appealing to an audience of this demographic mix such as age, education level, etc.? If the audience is mixed, is it presented in a way that appeals to a wide range of the audience members?

b. Complexity: Is it a topic that is complex enough to be interesting but not so complex that the audience is unable to grasp the concepts?

c. Occasion/Situation: Is the topic right for the situation or context? In other words, does it take into account where and when the presentation is being given, the audience characteristics (such as age, education, socio-economic status), the time of year, the current world/area conditions, etc.?
2. **Introduction**

“The speaker will formulate an introduction that orients the audience to topic and speaker”

(The attention getter should be first and the preview last, but other elements can be in a different order)

a. **Attention getter:** The speaker should start off with something that grabs the audience’s attention such as a quote, a story, a song, a video clip, a startling statistic – something that pulls the listener in. Stating the speaker’s name and/or lab affiliation makes for a dull attention getter. If the speaker was not introduced by someone else, then stating their name is fine, but it should be followed immediately by an attention-getting device.

b. **Establish Credibility:** The speaker lets the audience members know why (s)he is worth listening to on this topic. The simple fact that they are a scientist working in a lab is not adequate for establishing credibility. Speakers should express why they’re an expert on this particular topic. Ideally, they’ll say something about how long they’ve studied the issue or how it personally affects them. The best expressions of credibility include a personal connection to the topic as well as authoritative credibility such as a degree in the field or extensive study.

c. **Orientation to the topic:** The speaker makes the topic clear to the audience and gives them reasons to listen to this topic.

   i. **Example 1:** “You may not realize it, but there are polymers all around you every day of your life. Your water bottle is a polymer, the pan you use to cook your eggs probably has a polymer on it, the toothpaste you use contains polymer. Polymers are all around you.”

d. **Clear thesis:** The speaker clearly states the main idea of the presentation at some point in the introduction. In an ideal thesis statement, the speaker tells the audience generally what the presentation is going to be about and clearly defines any complex terms so that the audience clearly understands what they’re discussing.

   i. **Example 1** “So today I’ll discuss how computers help us develop better polymers …

e. **Preview of main points:** The speaker should let the audience members know the basic direction the speech is going to take by ending the introduction with a clear and direct statement of the main points that are coming up in the speech. The preview will forecast the organizational pattern of the speech as well. The speaker should not give out details in the preview, but should clearly summarize each of the upcoming main points.
i. **Example 1:** To explain how computers help with developing polymers, I’ll be explaining the computer programs that help create polymers, the kinds of products we can make, and why those products are helpful.”

ii. **Example 2:** “So today I’ll step you through the process we take in getting a fuel from leaves – which gives us a product that is called biofuel. We’ll talk about the plants we work with and how we choose them, then we’ll find out how we learn to squeeze oil from the plants, then how we make large amounts of oil, and then how we can turn that into fuel for your cars.”
3. Organization

“The speaker will use an effective orientation pattern”

Often presentations are organized by time (this comes first, this comes second…) or by topic (here’s one important thing you should know, here’s another important thing you should know…) There are other ways to organize a presentation, such as problem and solution (here’s something that was really causing difficulty for a group of people; here’s what we did or what we are doing to fix it).

Speakers can use any pattern of organization, but it should be clear, make sense to the audience, and be clearly stated. If a speaker seems to be rambling around different ideas related to a topic, then they are not demonstrating good organization.

a. **Main points are clear:** Ideally a speaker will state each main point as they get to it in a way that makes it clear to the audience that the speaker is moving to the next point. To make listening to the speaker as easy as possible, the speaker could structure each main point similarly.

   i. **Example 1:** “The first step in making biofuel is choosing the plants or plant products we’re going to use…. (supporting information comes here). The next step in making biofuel is squeezing the oil out of the plants we’ve chosen… (supporting info goes here). The next step in making biofuel is turning that oil into large amounts of oil… (supporting info goes here). The last step in making biofuel is processing that oil so it can be used in cars… (supporting material).

b. **Mutually exclusive:** Each section of a presentation should address a single issue or process. The speaker can certainly refer to something they’ve already discussed, but should avoid the wandering from point to point or mixing up supporting information.

c. **Related to the thesis:** Main points should be on point as previewed in the introduction. There are always many interesting things to talk about, but a speaker needs to choose the handful of points they are going to make (usually from two to five main points in a 10 to 30-minute presentation) and stick with those ideas. Too much information can confuse the listeners.

d. **Effective transitions and signposts:** Transitions and signposts lead the listener from one point to the next in a speech. Transitions may show a causal relationship between main points, a complementary relationship, or a contrasting relationship. They may indicate a chronological relationship as well. A speaker should orally speak transitions and signposts, not rely on a PowerPoint slide to make the move, nor should they simply jump to the next point. Signposts can be as simple as one word.
i. **Examples of signposts**: next, then, first, now, move on to, turn to, elaborate on, most importantly, to begin

ii. **Examples of transitions**:
   1. **Internal**: “Now that we’ve talked about choosing the plants, let’s look at what we do with the plants once we get them”
   2. **Internal preview**: “Now that we’ve discussed choosing the plant, we’ll discuss getting oil through breaking down the proteins, extracting the oil, and collecting the oil.”
4. Conclusion

“The speaker will develop a conclusion that reinforces the thesis and provides psychological closure”

Each presentation needs a conclusion. Speakers should avoid comments such as “that’s it” or “that’s all I got” or anything similar. A good conclusion restates the main points or summarizes them generally, connects to the thesis, and ends with a strong take-away/clincher/call to action. There should be a transition of some kind between the last main point and the conclusion, even if it is as simple as “So, in conclusion…” or “To wrap up…” or “So let’s review what we’ve learned…”

a. Example 1: “So to summarize, there are many complicated steps in creating biofuel, including find the best plants for the job, getting the oil out of those plants and making big supplies of it, and then turning that oil into something our cars can use. The process takes time, but is a smart way to use the plant waste that we are always going to have. Your car and the farmer’s tractor really can run on Plant Power!”

5. Use of Voice (Vocalics)

“The speaker will effectively use vocal expression to engage the audience”

Effective speakers are expressive with their voice. They avoid speaking in monotone or having a lift or drop in their voice at the end of every sentence. They exhibit an ability which, in public speaking circles, is called “vocal variety.”

Effective speakers vary their tone, pitch, rate of speaking, and intensity, and they avoid vocal fillers (words such as “uh,” “um,” “like,” “you know,” and “well” that are often said unconsciously, indicate nervousness or lack of forethought, and serve as a distraction to the listener).

Ideally, the speaker sounds like she/he is having an animated conversation with a good friend, telling the friend a story. They should sound enthusiastic about the topic and natural in their speaking.
6. **Non-Verbal Elements**

“The speaker will use eye contact, facial expressions, and body movement to support the verbal message”

A speaker’s facial and body movements should enhance their speaking. Ideally, the speaker will show the following types of body movement:

a. **Posture**: Erect, natural and balanced (equal weight on each foot). Avoids slouching over a podium or leaning to one side or the other.

b. **Eye Contact**: Frequent, if not constant, eye contact with all audience members. Looks around at the whole audience, not just one or two people. Glances at notes or PowerPoint only occasionally, not constantly. Body is facing forward, toward the audience, not angled toward a screen or visual aid.

c. **Facial Expressions**: Varies facial expressions as appropriate to what is being said. Avoids keeping the same expression throughout the talk. Smiles, raises eyebrows, shows surprise, or uses whatever expressions are appropriate for the talk.

d. **Movement and Gestures**: Avoids standing in one spot the entire presentation. Instead, moves naturally around the speaking area. Moves hands and arms in natural gestures, indicates important items on visual aids, uses gestures to illustrate points.

e. **Confidence and poise**: Is competently incorporating most of the above behaviors. Seems at ease and happy to be speaking. Looks at ease and confident.
Science Communication Elements-

7. Relevance and Importance

“Shows importance and relevance of science to audience members’ lives; connects science to audience.”

Speakers will ideally make the relevance of the science they are discussing explicitly clear to the audience members. As an assessor, it might help you to ask yourself, “After hearing this scientist speak, would audience members feel that this scientist is trying to positively impact their lives?”

The speaker should give at least one specific reason that the topic being discussed should be important to the audience member. To receive the highest points in this category, this importance/relevance will be implied or stated repeatedly during the presentation. Ideally, the speaker gives examples of values and interests they likely have in common with the audience, such as a desire to help those with disabilities, a common concern about aging parents, or a concern about the high cost of living.

While neither assessor nor speaker will know audience members’ interests with absolute certainty, assessors will rate the speaker on how well they seem to have connected their science information to ideas, interests, and values the audience members might reasonably have. If they are speaking largely to college students, speakers might connect their topic to making money, social lives, jobs or families students may have in the future, and so on. If they are talking largely to families with children still at home, they might mention safety, product costs, making life easier, budgeting, education, or transportation. If they are talking to older adults, they might emphasize security, or talk about how the science might help grandchildren, how an innovation could save them money, or how the topic relates to a hobby the audience members may enjoy.

An excellent speaker earning the highest points will be explicit in explaining why the topic is relevant, will make connections to values and interests that the audience might reasonably have, and will make these connections more than once through the presentation.

Example 1 – Relevance/Importance Based on the Common Value of Preserving Species of Wildlife:

Good: “The population of these birds is decreasing, but thanks to our research, we know more accurately how many birds in this species there actually are.” (common value – preserving wildlife)

Better: “Because of our research, we know there are more birds of this species left that we originally thought, and we are better able to preserve the species because we know what habitats they prefer and what their migration routes are.” (same value – better explanation of why the research is relevant)
Example 2 – Relevance/Importance Based on the Common Value of Having a Consistent Source of Electricity:

**Good:** “We want to provide a different source of energy than fossil fuels, so this solar technology is an exciting development.” *(common value: alternative energy sources in case fossil fuels run out)*

**Better:** “Research shows that 89 percent of the power used in the world now comes from fossil fuels, a source that is being depleted as we speak. This solar technology is developed using biofuels, which are constantly renewable, and generates electricity through the power of the sun. The better our technology gets in this area, the more inexpensive and clean solar energy we will have. Solar also works better in rural areas or third world countries that don’t have power plants.” *(better explanation+ helping the poor)*
8. Language

“The speaker will use language to make complex ideas clear and understandable to the public audience.”

Scientists work with complex ideas and processes that are difficult for non-experts to understand. Ideally, scientists giving public presentations should abandon complex or specialized terms, and should instead use language that favors directness and clarity. There are several communication techniques that have been shown to help scientists and others communicate clearly.

Use simple language: Speakers should avoid jargon and complex sentences. They should show an awareness of their innate expertise and should adjust their language to fit their audience.

Define terms: Speakers should define scientific terms that they do use, and should do so using language that is appropriate for their audiences. The use of examples can be helpful in definitions.

Make thoughtful use of metaphors/analogies: One technique that increases clarity is the use of metaphors and other comparisons. Research identifies several different types of metaphors and comparisons; however, the important thing for this assessment is that, when comparisons, metaphors, or analogies would help explain a complex point, the speaker uses them, and does so deliberately, as a way to help the audience understand.

Essentially, when scientists use comparisons, they should use deliberate comparisons that are developed to increase audience understanding and help with the clarity of their speaking. They should avoid metaphors and comparisons that are based on knowledge inherent to being a scientist and are not helpful to people who do not have scientific expertise (called cognitive metaphors).

a. Example 1 – Better (pedagogical, or teaching, metaphors) – messenger RNA; proteins are building blocks for the cell
b. Example 2 – Better (theory-building metaphor that explains a process) – The cell contains what we call a genetic code, a blueprint for building a new cell…

c. Example 3 – Worse (cognitive metaphor) – so the polymer chains make substances that are plastic-like or these organisms all come from the same family, so… (metaphors are used instinctively, not deliberately, and are not used to clarify a concept or practice)

An expert speaker will use a few metaphors/comparisons throughout the presentation rather than offering up multiple and/or contrasting metaphors in the same presentation. The best metaphors/comparisons compare the new object or concept to something that is familiar to audience members, and they give an accurate picture of the scientific concept.
Use concrete, clear, vivid language: Makes use of descriptive words, precise (not generic) nouns and verbs, and simple, plain language. Speakers avoid scientific jargon or “inside” terminology in favor of crisp, to-the-point declarative sentences.

d. **Example 4** – (jargon) “The biota exhibited a one hundred percent mortality response.”

(plain, crisp English) “All of the fish died during the season.”

e. **Example 5** – (precise language) Generic noun: house; More precise nouns: cabin, shack, cottage, mansion, basic ranch. Generic verb: leave behind; Stronger verb: abandon
9. **Visuals**

“The speaker will use clear, simple visual aids employing good science communication design techniques”

The best research on visual aids applies to science communication, but there are additional techniques that have been shown to help with scientific and technical presentations specifically. Since PowerPoint is the visual aid of choice for nearly all science presentations, most of these comments apply to PowerPoint. Here are some of the basic findings:

- Slides should be clear, crisp, and uncluttered.
- One of the strongest techniques science speakers can use is to have a descriptive sentence on the presentation slide with one or possibly two images that create a visual argument supporting their statement. There should be no (or very little) additional text in that type of slide. Slides are most memorable when the visual supports the text rather than simply being a decorative element.
- Text should be written in an active voice with a positive, rather than negative, tone.
  - **Good Example:** Rotors create sandstorms, limiting the sight of the pilots (Active, positive)
  - **Bad Examples:** The storms are not helpful to pilots. (negative) Or The sand is blown by rotors; this is not helpful to pilots (passive, negative)
- Any charts or graphs should be simple and should avoid small labels. Charts should be easily legible and understandable.
- Speakers should limit the complexity of diagrams, charts, and other visual aids. This often means simplifying these charts and diagrams from what would be used in a presentation to expert peers. Avoid scatter plots, three-dimensional graphs or charts, and stacked bar charts.
- If the speaker is using an object rather than a PowerPoint slide or to supplement a PowerPoint, the object should be large enough to be seen easily by the audience or should be projected electronically, if possible.
- General rules about visuals apply to science presentations:
  - The amount of text on slides should be limited
  - Fonts can be san serif or serif, but should be chosen for their clarity and professionalism. Tahoma, Arial, and Century are respected fonts for visual design work.
  - Text should be larger than 18 points so it is legible to all in the audience.
  - Colors and animations should be limited.
10. Increase Understanding of Science Processes

“The speaker will explain procedures, methods, and/or uncertainties”

Ideally, science speakers increase the audience’s knowledge not only of a particular topic, but also of the general processes involved in science. **Scientists do not need to address all processes, values, or methods, but should address at least some.**

Explanations should be given at the level of detail appropriate for a the specific general, public audience to whom the scientist is presenting. A speaker earning the highest points in this category will **explain several processes**, or will give a more in-depth and useful **explanation of one of the methods, processes, or values** of science. Speakers earning lower points will explain such ideas only briefly or with too much detail and complexity for a public audience.

Keep in mind that the **level of detail should be appropriate for a general audience**. There does not need to be enough detail to replicate the experiment, nor the speakers present complicated scientific support for their choices. Instead, the explanations should generally explain processes that give public audiences a better understanding of how the process of science works. For example, scientists build on past research, they ask questions, they often work in large teams, research can take many years, and data is meticulously gathered and recorded.

Specific aspects of science and science processes that are often emphasized in the literature as being valuable for publics to understand and useful to emphasize in public presentations by scientists are:

- **Data Gathering**: the speaker generally explains how the data is gathered, tools that were used to gather data, decisions that were made about data gathering, analyses that were done on the data, how tools were built in order to gather desired data, why certain data was gathered, or what the data indicates.
  - **Example**: “We heat the material to about 550 degrees Fahrenheit, since that is the temperature when it starts to break down into its smaller parts, like gases and solids.”
  - **Example**: “We use the tool shown in the picture to measure the eggs in mid-July, since by that time the mothers have left the nests, so we aren’t disturbing any of the nesting patterns.”

- **Choices and Reasons**: the speaker may explain choices that were made as they developed experiments or carried out research, and they may explain those choices. For example, the speaker may why an experiment was conducted in a certain location, why a specific enzyme or element was chosen for testing, why a particular plant was used for testing. Explaining choices is good; providing
reasons for those choices is even better. These reasons can be brief, and should be appropriate for the audience.

- **Example:** “E. coli is a common bacterium that we know a lot about, so we chose to use it as a host to grow our new enzymes. You may think that sounds dangerous, but there are many types of coli, and most of them don’t make people sick.”

- **The Cooperative Nature of Science:** a speaker may describe a how a group works on a project together, how (s)he is working on one aspect of a project while another researcher or research group is working on another aspect, how (s)he has taken up a line of research that someone else started, that (s)he has taken a certain project to a stopping point and another researcher is taking up the next questions, or that (s)he is assisting a group member on a project.
  - **Example (building on research):** “A number of people have been working on ways to limit weeds through chemical means, and others have looked at ways to limit weeds through cover crops. I wanted to see if there was another non-chemical way we could reduce the number of weeds in fields.”
  - **Example (working together):** “My major professor has been working with this group of prairie dogs for 20 years, and each summer our team goes out to collect data about numbers of animals, burrow development, and food sources so that we can keep learning more about these interesting animals.”

- **Uncertainty and How Uncertainty is Managed:** the speaker is confident and knowledgeable about the uncertainties of science. The speaker may acknowledge what is not known or what still needs to be worked out. Alternatively, (s)he may express how certain the community of scientists in the field are about a particular finding, or acknowledge that ideas and understandings change as more information is gathered. They may express how they think about uncertainty, how they adjust for it, how it affects results of scientific tests.

  For example, scientists can explain that “theory” does not mean “guess,” but rather is a logical, tested attempt to explain a phenomenon. Confirmed theories become important scientific tools, although they may be clarified or superseded by more accurate theories. **A scientist could also explain why they hesitate to use the word “fact,” even when an issue is largely settled.**

  - **Example:** “Doctors have been using this type of cement in hip replacement surgeries for decades, but the cement fails 20 percent of the time. We haven’t been able to figure out a good replacement for that...”
cement, but my research will hopefully provide a possible path to better bonds in hip replacement surgeries.”

- **Example:** “Now we’re trying to figure out how to straighten those pathways so that the electrons can move more quickly. We’ve figured out some ways to straighten them, but we still have more we can’t unravel yet.”

- **Example:** While we still call it “cell theory,” cellular biologists generally accept that all living things, including the plants we’re studying, are made up of cells. There is still some discussion about cells when it comes to things like viruses, but when it comes to plants (and people!), the issue is settled. We’re all cells. Knowing this helps us predict how these plants and animals are going to act.
11. Scientists as Trustworthy, Personable, Inspirational about science

"Speakers will portray scientists as excited and passionate about science, approachable, and knowledgeable."

According to research, scientists engaging with the public should express an excitement for the work they do and inspire their audience with what is being done in their field of science. Additionally, scientists should portray themselves and other scientists as personable, similar to the audience (non-scientists), and trustworthy and should also build a connection with the audience. This can be done using techniques that build "immediacy," or a sense of connection and closeness to an audience. An assessor may think to themselves, "Would an audience member want to sit down and have a cup of coffee with this scientist?" or "If an audience member had the money to fund a scientist in this area, would they choose this particular scientist?"

Research done by organizations such as the Pew Charitable Trust indicate that U.S. citizens largely respect scientists, but often hold different opinions than scientists on matters based in science knowledge (Funk & Rainey, 2015). Therefore, one of the competencies science stakeholders want public science communicators to have is the ability to connect with their audiences, to seem trustworthy and personable to their audiences rather than authoritative or concerned with issues beyond the interest of the "regular" folk.

To show excitement for, their interest in, and a passion for science, the speakers may make statements of awe, enjoyment, or excitement.

**Example:** “I get to swim with turtles on my spring break, and I can’t think of anything I’d rather be doing.”

**Self-disclosure** is identified in communication research as being one of the best ways to increase trust. Self-disclosure is simply the revealing of personal information to others. This can be information about things that happen in the lab or in other areas, but the comments should be personal in nature and touch on the science work. Ideally there are instances of self-disclosure that include vulnerability – indications that the scientist has challenges or weaknesses in their science work. This tends to create a sense within the audience that scientists are like them, which helps them see scientists as trustworthy.

Here are some statements of self-disclosure:

a. **Example 6** – “We couldn’t figure out why this stuff was so sticky; I showed my major professor what was happening and he couldn’t figure it out either.”

b. **Example 7** – “My daughter is afraid of birds, so I don’t bring her to the lab very often...”

c. **Example 8** – “I think one of the perks of my job is crawling around in the river with turtles”
Additionally, scientists can tell stories about themselves and their work, and portray themselves and other scientists in a positive light by talking about themselves, their colleagues and their work in a way that meshes with the values of their audience.

**Concrete language:** Scientists can also increase the trust the audience has in them by using concrete language. Studies show that, in addition to increasing clarity for listeners, a speaker using concrete rather than abstract terms gives listeners a sense of confidence in the speaker and trust in what they say. Often there is a continuum between abstract and concrete, and a scientist’s language will fall somewhere on the continuum. In general, it is better to be closer to the concrete than to the abstract.

**Inclusive Language:** Speakers build relationships of trust with their listeners when they use inclusive language. Inclusive language frames the speaker as part of the group, and can even include the audience in that group. For example, scientists who use “we,” “our,” and “us” rather than “I” or “you” will do a better job at connecting with and building a relationship of trust with the audience.

d. **Example 9:** *Abstract language:* “To complete this research, we really worked hard.”  
*Concrete language:* “To complete this research, we had to read lots of studies, and then design our own project that would take a step forward. We kept reviewing the prior studies as we worked.”

e. **Example 10:** *Non-inclusive language:* “I identified the bacteria by isolating it in the lab, and then I tested each of the samples.”  
*Inclusive language:* “We worked on the project from May through November. First, we isolated the bacteria and then we tested all of the samples.”

**Immediacy:** Another measure of connection between audience and speaker is immediacy, or a feeling of closeness and connection rather than distance and detachment. Some behaviors that can increase this feeling of connection include the speaker:

- Using appropriate humor
- Moving closer to the audience
- Smiling at the audience
- Looking directly at the audience

The use of these techniques by science speakers may also improve their success in this category.
12. Scientists Engage in Interactions, Dialogue, and Learning

Speakers will provide opportunity for interaction with the audience throughout the presentation, and will engage in a question and answer dialogue at the end of the session. To earn highest points for this category, speakers will give audience members opportunity to engage at various times throughout the presentation as well as engage thoughtfully at the end of a presentation.

Ideally, the speaker is open to learning from the audience, such as learning about the audience’s firsthand experiences with an issue or hearing audience ideas for what might be studied.

During the presentation, the speaker may do something as simple as ask a rhetorical question to which (s)he doesn’t expect a response, or (s)he may ask a question and solicit responses, or (s)he may do something more engaging and complex, such as ask the audience to engage in an action or write/draw something that will contribute to the presentation. The more expert, frequent, and engaging the interaction is, and the more open to gathering knowledge from the audience that the speaker is, the higher the points the speaker will earn.

Example (during the presentation): “How many of you have ever used building bricks like Legos®? Let’s see your hands.”

Example (after the presentation): “So what questions can I answer for you about either the prairie dogs or about my research with them?”
# Rubric for Assessing Public Communication by Scientists

<table>
<thead>
<tr>
<th>Performance Standard:</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Topic:</strong> Appropriate for speaker and audience</td>
<td>Topic engages audience, topic is worthwhile, and of interest and of an appropriate complexity for the audience and situation.</td>
<td>Topic is appropriate to the audience and situation and is of some interest to many members of the audience.</td>
<td>Topic is somewhat complex or not entirely relevant to many of the audience members. It is adequate for the situation.</td>
<td>Topic is too trivial, too complex, or inappropriate for audience; topic not suitable for the situation.</td>
<td>The topic is entirely too complex, is not stated clearly, or is inappropriate for the audience and situation.</td>
</tr>
<tr>
<td><strong>2. Introduction:</strong> Orients audience to topic and speaker</td>
<td>Excellent attention getter, firmly established credibility, sound orientation to topic, clear thesis, preview of main points cogent and memorable</td>
<td>Good attention getter, generally establishes credibility, provides some orientation to topic, discernible thesis, previews main points</td>
<td>Attention getter is mundane; somewhat develops credibility; Awkwardly composed thesis; provides little direction for audience</td>
<td>Irrelevant opening; little attempt to build credibility; abrupt jump into body of speech; thesis and main points can be deduced but are not explicitly stated.</td>
<td>No evidence of opening technique; no credibility statement; little to no background on topic; thesis/statement of topic is unclear; no or unclear preview of points</td>
</tr>
<tr>
<td><strong>3. Organization:</strong> Effective for topic and audience</td>
<td>Very well organized; main points clear, mutually exclusive and directly related to thesis; effective transitions and signposts</td>
<td>Organizational pattern is evident, main points are apparent; transitions present between main points; some use of signposts</td>
<td>Organizational pattern somewhat evident; main points are present but not mutually exclusive; transitions are present but are minimally effective</td>
<td>Speech does not flow well; speech was not logically organized; main points were not clear, transitions present but not well formed</td>
<td>Organizational pattern not clear or nonexistent, few to no transitions; information sounds as if it is being randomly presented</td>
</tr>
<tr>
<td><strong>4. Conclusion:</strong> Restates main points, provides closure</td>
<td>Provides a clear and memorable summary of points, some reference back to thesis/big picture, ends with strong clincher or call to action</td>
<td>Appropriate summary of points, some reference back to thesis, clear clincher or call to action</td>
<td>Provides some summary of points, no clear reference back to thesis, closing technique can be strengthened</td>
<td>Conclusion lacks clarity, trails off; ends in a tone at odds with the rest of the speech or brings in new information.</td>
<td>Little in the way of a conclusion; speech ends abruptly and without closure</td>
</tr>
<tr>
<td><strong>5. Voice:</strong> Vocal expression engages the audience</td>
<td>Excellent use of vocal variation, intensity and pacing; vocal expression natural and enthusiastic; avoids vocal fillers</td>
<td>Good vocal variation and pace; vocal expression suited to the situation, few if any vocal fillers</td>
<td>Demonstrates some vocal variation; enunciates clearly and speaks audibly; generally avoids fillers</td>
<td>Sometimes uses a voice too soft or articulation too indistinct for listeners to comfortably hear; little vocal variety, often uses fillers</td>
<td>Speaks much too loudly or softly, enunciation is lacking, speaks in monotone, poor pacing, distracts listeners with vocal fillers</td>
</tr>
<tr>
<td><strong>6. Non-Verbal:</strong> Eye contact, facial expressions, and body movement to support the verbal message</td>
<td>Posture, gestures, facial expression and eye contact well-developed, natural, and display high levels of poise and confidence</td>
<td>Postures, gestures and facial expressions are suitable for speech, speaker appears confident</td>
<td>Some reliance on notes, but has adequate eye contact, generally avoids distracting mannerisms</td>
<td>Speaker relies heavily on notes; nonverbal expression is stiff and unnatural</td>
<td>Usually looks down and avoids eye contact; nervous gestures and nonverbal behaviors distract from or contradict the message</td>
</tr>
</tbody>
</table>

*These elements are adopted with permission from the PSCR (Schreiber, et al, 2012).*
<table>
<thead>
<tr>
<th></th>
<th>Examples and Notes</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Topic: <strong>Worthwhile</strong> Engaging Appropriately complex</td>
<td><strong>Example:</strong> You may not think about dirt very often except to clean it off of you, but today that’s what I’m going to talk with you about – dirt. More specifically, topsoil.</td>
<td>Engages Interests</td>
<td>Of some interest</td>
<td>Not relevant Adequate for situation</td>
<td>Too trivial or too complex</td>
<td>Unclear No mention Inappropriate</td>
</tr>
<tr>
<td>2. Introduction: <strong>Attention</strong> Topic Credibility Preview</td>
<td><strong>Example:</strong> We’ll look at what our topsoil was like in the past, what our topsoil situation is now, and what it could be like in the future.</td>
<td>Excellent Firmly Established Sound Cogent Memorable Preview</td>
<td>Good Generally Some Discriminable Preview</td>
<td>Mundane Somewhat Awkwardly Thesis with little direction</td>
<td>Irrelevant Little attempt Abrupt Deduce thesis, not explained</td>
<td>No evidence Little to no Unclear Topic and/or Preview</td>
</tr>
<tr>
<td>3. Organization Appropriate Clear Main Points Transitions</td>
<td><strong>Example:</strong> (transition, main point) So when the pioneers walked these plains, they walked on five feet of rich topsoil. Let’s discover what our topsoil looks like today.</td>
<td>Very well Clear Mutually exclusive Effective</td>
<td>Evident Apparent Present Some use</td>
<td>Somewhat evident Present Minimally effective</td>
<td>Does not flow well Not logical Not clear Not well formed</td>
<td>Not clear Doesn’t exist Few to no Random</td>
</tr>
<tr>
<td>4. Conclusion: Summary Thesis Take away or call to action</td>
<td><strong>Example:</strong> So, topsoil is vital to healthy agriculture, but we are losing more each year. Where we used to have several feet of topsoil, we now have only a few inches. Cover crops and no-till practices can help.”</td>
<td>Clear Memorable Refers back Strong</td>
<td>Appropriate Some reference Clear</td>
<td>Summary No clear reference Weak</td>
<td>Lacks clarity Trails off Tone at odds</td>
<td>Little Abrupt end No closure</td>
</tr>
<tr>
<td>5. Voice: Vocal Variation Good Pacing No fillers</td>
<td><strong>Example:</strong> (Speaking sounds natural, conversational, tone and pace varied)</td>
<td>Excellent Enthusiastic Avoids Fillers</td>
<td>Good Well Suited Few Fillers</td>
<td>Some Enunciates Generally avoids Fillers</td>
<td>Too loud or soft, Indistinct Little Fillers</td>
<td>Monotone Poor pacing Distracts</td>
</tr>
<tr>
<td>6. Non-Verbal: Erect posture Natural gestures Strong eye contact Balanced, natural stance</td>
<td><strong>Example:</strong> (Avoids reading, looks audience in the eye, makes natural gestures, moves around naturally)</td>
<td>Well developed Confident Natural High Levels</td>
<td>Suitable Appropriate Confident</td>
<td>Adequate Avoids distractions</td>
<td>Relies on notes Stiff Unnatural</td>
<td>Nervous Looks down Reads Distracting</td>
</tr>
<tr>
<td>7. Relevance &amp; Importance: Connects science to lives, shows importance and relevance of science</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Makes a compelling case for the relevance and importance of the science topic to the audience's lives.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Creates common ground by strongly emphasizing common goals, values, and/or experiences, clearly and consistently connects concepts to impacting audience members' lives.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Clearly discusses the importance and relevance of this science topic to the audience, tries to create common ground with the audience through common values and goals, usually connects science concepts to ideas that are impactful to lives of audience.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>May mention importance of topic but not make it explicitly relevant to the audience, inconsistent connection to common goals/values, may make some connection between science concepts and audience lives.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>Only brief mention of topic's importance and/or relevance to the audience, few, if any, mentions of common goals or values, makes little connection between the topic and impacting the audience.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 8. Language: Uses language to make the complex clear, definitions, metaphors, and comparisons |
|---|---|---|---|---|
| 4 | Uses clear, concrete, understandable language that considers audience's current knowledge, particularly about science; defines any complex science terms used; when appropriate, uses comparisons and analogies to make complicated ideas clearer; any analogies that are used are interesting, deliberate, audience-focused, teaching analogies that clarify ideas. |
| 3 | Uses generally clear, concrete and understandable language that is mostly appropriate to audience’s, science terms usually defined, any comparisons or analogies seem deliberate and are mostly helpful to the audience at their understanding, analogies are appropriate for audience and thoughtfully expressed. |
| 2 | May sometimes use complex or unclear language, may fail to define all science terms. Some analogies may not always be appropriate for or make sense to the audience, may not use comparisons or examples when they would help explain and clarify the topic. |
| 1 | Uses several complex terms without defining them; language is often inappropriate for audience's level of understanding. Any comparisons may not be less relevant to audience or not fully developed, may show lack of awareness of or lack of adaptation to audience. |
| 0 | Language is often too complex or unclear, frequently uses scientific terms that are outside the audience understanding. Uses few or no metaphors or other comparisons, or comparisons are not relevant to audience or not fully developed. |

| 9. Visuals: Clear, simple visual aids using good science communication design techniques |
|---|---|---|---|---|
| 4 | Nearly all visual aids use simplicity, clarity, and good design principles. Visuals focus on images or other graphic depictions, are easy to understand, use minimal text, only use clear graphs and charts, non-electronic objects are clearly visible. |
| 3 | Visual aids are usually easy to see, simple, and helpful, but may use several images or graphics per slide or smaller than ideal font size. Slides may use more than minimal text, most charts/graphs are legible, any objects visible. |
| 2 | Visual aids somewhat mostly understandable, but a few may use too many visuals, be cluttered with graphics or too much text, may include some complex charts/graphs, physical objects hard to see. |
| 1 | Visual aids may be cluttered, images or graphics are illegible, use too much text; may have too many charts/graphs, may be too complex, non-electronic objects may not be visible. |
| 0 | Visual aids are confusing, cluttered, and unclear. Most have too many elements, text may be small. Slides may be text-heavy, poor design of charts, graphs, images. |

| 10. Explain Science Processes: Explains procedures, data, choices, and/or uncertainties. |
|---|---|---|---|---|
| 4 | Explains multiple scientific processes at a level appropriate to audience. May give general explanations of steps in an experiment or in gathering data, may explain particular choices made, the team nature of science, and/or navigating uncertainties and/or risk in science. |
| 3 | Describes several scientific processes, will often describe for the audience steps taken and generally why they were needed, and/or make mention of teams, uncertainties, choices, and/or other processes like data gathering, but do not so clearly. |
| 2 | May attempt to describe processes without complete success, given the audience, may discuss data gathering, uncertainties, teams, and/or choices, but explanations are incomplete, brief, or too complex for audience. |
| 1 | Describes scientific processes only occasionally, makes only brief mention with little to no description of teams, uncertainties, choices, or any other processes, or uses too much complicated detail. |
| 0 | Fails to clearly describe any scientific processes. Little to no description of, data gathering, uncertainties, choices, teams, or other scientific processes. Ignores audience understanding. |

| 11. Trustworthy & Personal: Scientists seem approachable, excited, inclusive, confident |
|---|---|---|---|---|
| 4 | Expresses an excitement for and/or is inspiring about science, especially the topic of the presentation, engages in frequent, appropriate self-disclosure, may use personal stories and/or humor, frequently uses inclusive terms such as we, us, and our, seems positive, confident. |
| 3 | Demonstrates an excitement for and/or is inspiring about science. Engages in regular self-disclosure and/or personal stories, may use some humor, mostly uses inclusive pronouns rather than first person (I, me), mostly positive, confident. |
| 2 | Excitement for and/or inspiring about science at some level. Self-disclosure, personal stories, and/or humor used at times, may sometimes use inclusive pronouns, may or may not seem positive/confident. |
| 1 | Expresses little excitement or is not inspirational about science. Infrequent self-disclosure, personal stories, or humor used, “I” and “me” than inclusive pronouns seems distant or aloof, not positive or confident. |
| 0 | Demonstrates no excitement for science, does not inspire interest in science. Little to no self-disclosure, personal stories, or humor, may use mostly “I” language, lacks confidence. |

| 12. Engagement: Have dialogue and interactions about science with public audiences. |
|---|---|---|---|---|
| 4 | Audience members have multiple opportunities to ask and respond to questions, audience makes comments, speaker is also a learner, there is active dialogue between audience and speaker during and after the presentation. |
| 3 | Audience members have opportunities to ask questions and make comments, speaker listens to audience but devotes less time to questions and interactions, has a Q&A session, less dialogue with audience. |
| 2 | Audience members have opportunity to ask but not respond to questions, audience makes few if any comments, speaker gives no indication of learning from audience, little dialogue with audience members. |
| 1 | No questions asked or interactions had with audience during presentation, audience and speaker interaction limited to a Q&A period at the close of the presentation, no dialogue w/ audience. |
| 0 | Speaker consistently assumes an authoritative air, does not ask questions, gives no indication of learning, no dialogue or interaction with audience, may not invite any questions. |
| Examples and Notes | | | | | | |
|-------------------|---|---|---|---|---|
| **7. Relevance & Importance** | Example: “Because of our research, we know there are more birds of this species left that we originally thought, and we are better able to preserve the species because we know what habitats they prefer and what their migration routes are.” | Compelling | Strongly | Clear | Consistent | Relevant | Clearly | Usually | Mention | Not explicit | Some | Inconsistent | Brief | Mention | Little connection | Unclear | No mention | Minimal | Nonexistent |
| **8. Language** | Example: “The fish all died in the experiment” “Proteins are like building blocks for the cell” | Clear | Defines | Comparison when appropriate | Deliberate | Teaching | Audience | Generally | Understandable | Mostly | Seem deliberate | Appropriate | |
| **9. Visuals** | Example: Large image, little text. Clear, legible text. Simple charts and graphs, easy to follow layout | Easy to see | Large image | Limited text | Simple | Visible | Uncluttered | Usually | Mostly clear | Smaller than ideal | Mostly | Minimal | text |
| **10. Science Processes** | Example: “We measure the eggs in mid-July, since by that time the mothers have left the nests, and we aren’t disturbing any of the nesting patterns.” | Explains | A number | Right for audience | Emphasize | Discuss | Various processes | Describes | Several | Good efforts | Mention | Not as much depth | |
| **11. Trustworthy & Personable** | Example: “I get to swim with turtles on my spring break, and I can’t think of anything I’d rather be doing.” | Expresses | Engages | Frequently | Positively | Inspiring | Confident | Excitement | Demonstrate | Regular | Mostly | |
| **12. Engagement** | Example (during the presentation): “How many of you have ever used building bricks like Legos®? Let’s see your hands.” | Multiple | Active | Learning | Answers | Questions | Listen | Multiple | Active | Learning | Answers | Questions | Not respond | Few | Little | No Limited | No | Authoritatively | No | Not |