An instrument for assessing scientists' written skills in public communication of science

Ayelet Baram-Tsabari

Department of Education in Technology and Science, Technion – Israel Institute of Technology, Haifa 32000, Israel, ayelet@technion.ac.il

Bruce Lewenstein

Departments of Communication and of Science & Technology Studies, Cornell University, Ithaca, New York 14853, USA

b.lewenstein@cornell.edu

Keywords. media training, science in the media, communication skills, standardised evaluation.

Running head: Assessing scientists' written skills

An instrument for assessing scientists' written skills in public communication of science

Abstract.

This article describes the development of a first of its kind tool for measuring scientists' written skills in public communication of science. It includes the rationale for establishing learning goals in seven areas: clarity and language, content, knowledge organization, style, analogy, narrative, and dialogue, as well as the questions designed to assess these goals. The skills testing is primarily designed for assessing written communication skills, and can be used in many science communication training contexts. It can serve as a baseline survey, formative assessment or in summative pre-test/post-test evaluations. The article provides detailed criteria for analysing the results of the instrument as well as findings from baseline data collected from science graduate and undergraduate students.

Keywords. media training, science in the media, communication skills, standardised evaluation.

Introduction

Several studies have described motivations and challenges for scientists who wish to speak with the media (e.g. Dunwoody & Ryan, 1985; Martin-Sempere, Garzon-Garcia, & Rey-Rocha, 2008; Sturzenegger-Varvayanis et al., 2008; Treise & Weigold, 2002), as well as the abundance of such interactions (e.g. Jensen, 2010; Peters et al., 2008; The Royal Society, 2006). However, few studies have systematically examined scientists’ ability to communicate with the media and the public. Nonetheless, many organizations and institutions have created training opportunities to help scientists become better at public communication. Moreover, science communication scholars are largely in agreement that bench scientists and engineers as well as science and health regulators would benefit from both media training and training in communicating with the public (Besley & Tanner, 2011).

Before one asks 'What do scientists need to learn in order to communicate better?' one needs to attend to scientists’ knowledge, and more importantly, current expectations and norms related to communicating with the public (either directly or through the media). From a sociocultural perspective, which sees knowing and understanding as anchored in cultural practices within communities (Wenger, 1998), it helps to consider what scientists learned while *becoming* scientists. From this perspective, learning science means learning to talk science, with its own semantic patterns and specific ways of making meaning. "It means learning to communicate in the language of science and act as a member of the community of people who do so" (Lemke, 1990, p.1). This view of learning science is in alignment with the theory of situated learning, which views learning as a process of enculturation into a community of experts, with its own belief system, jargon, and norms (Brown, Collins, & Duguid, 1989). This view is also in line with ideas from the social studies of science, which highlight the degree to which knowledge exists only in the form in which it is expressed (Latour & Woolgar, 1979; Shinn & Whitley, 1985; Ziman, 1968, 1978).

In this sense, the process of becoming a scientist inevitably involves learning to talk and write science according to the norms of the scientific community. These would include generalizing and abstracting rather than building on examples, stories and anecdotes, and using accurate descriptions rather than analogical approaches. Although many specific counter-examples can be identified – see, e.g., analyses of Stephen Jay Gould's "Spandrels of San Marco" article in Selzer (1993) – we are here describing the bulk of writing in the sciences, including the advice given in various handbooks (Day & Gastel, 2011; Wilkinson, 1991). In scientific writing there is much use of the passive voice, of abstract nouns in place of verbs, and of verbs of abstract relation in place of verbs of material action (Lemke, 1990). According to a statistical formula for the measurement of readability, scientific magazines achieved the lowest score on the ‘‘reading ease" and "human interest" score, ranking them as "very difficult" and "dull," respectively (Flesch, 1948). Furthermore, using jargon has the effect of making people from other communities feel excluded and alienated (Halliday, 1993; D. P. Hayes, 1992).

Since part of learning science is learning to talk science, communicating science to the lay public demands yet more learning – the ability to use non-technical language and norms when discussing science beyond the scientific community. Furthermore, recent exhortations of scientists not only call for them to be clearer speakers, but also call for engaging in a respectful dialogue with the public, as public values make important contributions to science-related policy issues (Leshner, 2009). Instead of 'training' in science communication (which usually refers to acquiring specific practical skills), the terms 'teaching' and 'learning' better fit the complex process in which scientists develop new knowledge, behaviors, skills and attitudes.

It is thus within the framework of sociocultural learning theory, rather than training, that we set out to develop a measurement for scientists' skills in science communication. We were motivated to undertake this after inquiring of science communication educators worldwide whether or not they had formal evaluations of their programs. Personal Emails were sent in June 2010 to 134 leaders of science communication training programs from Europe, Australia and North America, identified through directories maintained by the European Union and by the PCST Academy (Directorate General Research, 2010; PCST Academy, 2010). Based on 36 responses received, no systemic evaluation of the courses' learning outcomes is being performed. Of the fourteen programs that conduct some kind of assessment, many used generic student evaluation questionnaires looking at issue of instructor skill and organization.

Therefore, the data we received suggested that no standard measure yet exist. Indeed, methods and methodology for analyzing communication skills and programmatic efforts in promoting communication skill development are essential to the mission statement of these programs, yet program directors and program evaluators are left with unspecific directions for defining, let alone promoting and analyzing, science communication skills (Fennewald, 2011).

Constructive evaluation occurs when specified outcome measures are conceptually related to intended learning objectives (Kraiger, Ford, & Salas, 1993). Since no list of learning objectives for science communication teaching yet exists to our knowledge, our task is twofold: first, to compile such a list of potential learning goals, and then to derive a conceptually based scheme for evaluating whether they have been achieved. In order to create this list and scheme, we are introducing into the science communication community traditions borrowed from the field of science education. We were inspired by the conceptualization and measurement of students' and teachers' views of the "Nature of Science" (Aikenhead & Ryan, 1992; Norm G. Lederman, Abd-El-Khalick, Bell, & Schwartz, 2002). The development and assessment of students' and teachers' conceptions of nature of science have been concerns of science educators for over 40 years and arguably constitute a line of research in their own right (Norman G. Lederman, 2007). This type of work might inspire the beginning of a similar line of research within science communication, similar to the quantitative and qualitative work that was invested in conceptualizing and measuring public knowledge of and attitude to science.

This study is part of a bigger project to develop a measurement for scientists' views and knowledge of, and skills in, public communication of science, especially in the context of science communication education programs. This paper focuses on the development, piloting and analyzing of a measurement for scientists' science communication written skills.

Naturally, many different skills could have been evaluated. We could have studied role play of oral presentations to policy makers or televised interviews. Written skills were chosen in order to allow analysis of larger samples, without the need to record and transcribe the responses, as well as to ease comparison across samples. Furthermore, analyzing oral communication would add another level of complexity, relating to tone, body language, charisma, etc. That said, we believe the questions described below can be used as a basis for evaluation of oral communication as well, for example by replacing the number of words asked for with a request for specific timing.

We argue that the instrument can be used with a wide range of science communication education programs, either as a baseline survey, as a pre/post evaluation of the learning outcomes for particular programs, or as a form of formative assessment which promotes teaching and learning alike. Principally, we ask: what should scientists learn in order to write better for the public? And once these learning outcomes are identified, how can we assess them?

Learning goals

Reviewing the literature, we developed a framework of learning goals for public communication of science (table 1), with particular attention to goals for written communication (which we believed would be most amenable to systematic assessment). This framework builds on work that looks at different contexts and types of public communication (Miller, Fahy, & The ESConet Team, 2009); on theoretically-informed analyses of pedagogical presentations by scientists and science students (S Kapon, U Ganiel, & B S Eylon, 2009; S Kapon, U Ganiel, & B S Eylon, 2009; Sevian & Gonsalves, 2008); and on our own reading from a wide range of practical advice books for scientists (e.g. Baron, 2010; Christensen, 2007; Cribb & Hartomo, 2002; Dean, 2009; R. Hayes & Grossman, 2006;

Meredith, 2010). The advice was extracted from the various books, grouped by theme and given a common title.

[table 1 about here]

Learning goals must be tied to the specific context for which the learning is intended. The most explicit set of categories we found came from the EU-funded ESConet training program, which divides its study modules into "basic" and "advanced" (Miller et al., 2009). The advanced modules, in particular, deal extensively with "dialogue-based" science communication in which scientists and publics interact with regard to issues that have high policy relevance or social controversy associated with them. Although full dialogue settings are beyond the scope of the instrument we were trying to develop, we included the advanced category as certain written materials may use such an approach. However, we divided the ESConet "basic" category into "basic" and "intermediate" categories, to better align with the material we drew from the pedagogical literature.

The pedagogical literature looks at issues such as the presentation techniques of leading physicists who are also highly successful popular physics public lecturers (S Kapon et al., 2009; S Kapon et al., 2009) and the practices used by science graduate students to explain their research to nonscientists (Sevian & Gonsalves, 2008). We found the clusters of goals identified by Kapon et al. to be especially useful:

(1) *Content features*: includes elements that reflect a judicious choice of content: What to include, what to omit, and means to achieve this goal (e.g. selection of topic or level).

(2) *Knowledge organization features*: includes elements that manage knowledge (e.g. structure, repetition).

(3) *Analogical approaches:* includes elements that explain the novel in terms of the known (e.g. analogy, metaphor).

(4) *Stories*: includes elements that construct scientific ideas through means that are common in fiction (e.g. narrative).

To these clusters, we added "dialogic," to provide space for goals associated with the dialogue model of science communication. We also considered separately issues of style (taken from the "story" cluster) and combined the issues of clarity and language (handled in "knowledge organization" and "story," respectively, by Kapon et al.). Following Kinneavy (as cited in K. E. Rowan, 2003) we conceptualize analogy, narrative, and dialogue as forms of organization, rather than as a desired outcome, such as persuasion or curiosity.

Finally, we considered the many detailed suggestions provided by practical science advice books. This advice, we found, fell into four clusters: preparation, content, language, and style (Table 2). We found that these categories could easily be aligned with the clusters derived from the other works, as shown in Table 1.

[Table 2 about here]

In the process of identifying potential resources, we examined a wide range of other sources, such as the guidelines provided by the British Council for participants and judges of the international science communication competition "Famelab" (British Council, 2011) and the judging criteria provided by the "Intel International Science and Engineering Fair" (Society for Science and the Public, 2011). These guidelines can easily be incorporated within the clusters, suggesting that the clusters listed in Table 1 are both sufficiently broad and sufficiently differentiated to capture the range of possible learning goals.

Because of the "basic, intermediate, advanced" structure, the learning goals appear to be hierarchical, with each additional goal building on the previous one. While we do not wish to be dogmatic regarding the ordering of the goals, we do call attention to the inherent dependency of higher learning goals on earlier ones.

Clearly, the potential learning outcomes of an intervention depend on the conceptualization of science communication; the agenda of the organizers; and the specific learning outcomes of the specific workshop/course (different trainings by the same trainer can have different learning goals). A recent web survey involving 255 science communication experts indicated that training focus differs greatly. For example, 30% of the responders indicate that public speaking is not at all a focus of the training, while 27% designate it as a primary focus. Training focused on public engagement and news values were relatively common for bench scientists, engineers and science and health regulators (Besley & Tanner, 2011).

The question of what learning outcomes to measure is therefore tied to the question of what the learning goals were to begin with. In this work we suggest ways to assess a wide range of learning goals in written science communication skills. Teachers and researchers can chose those assessment which are associated with their goals.

Instrument development and testing

Questions were developed to address each of the goals listed above (table 3). In a close-ended question, responders were presented with a list of science concepts and were asked to mark those that should be defined when writing to a non-technical audience. Then they were presented with three open-ended practical tasks: describing one's research, responding to a question about science in everyday life, and responding to a question about science's role in society. The three questions ask scientists to describe science at the fundamental level (explaining research), contextual level (relevancy to everyday phenomena) and finally as a social institution (science's role in society). The third level will often depend on national context; therefore some of the possible issues might be more relevant in some countries than in other. A list of issues is provided to cater for concerns of citizens from different developed societies. However, further adaptation to local concerns might be needed (table 3).

The decision to include several options within questions 2 and 3 was motivated by the varied scientific and cultural background of the potential participants. Some users may wish to force only one question in order to enhance standardization.

There are (at least) two options for using the questions in a post-course questionnaire: (a) ask participants to answer again the same questions, or (b) provide responders with their own original answers, and ask them to edit their essays to make them more clear and useful to readers. The second option is the one used here with a control group.

The analysis of the answers will be described in detail in the section 'written skills analysis'.

[table 3 about here]

Face validity. To establish face validity for the questionnaire, 15 interviews were conducted with people from relevant groups in the research population - active faculty and graduate students in the life and natural sciences. Interviewees were recruited using snowball sampling procedure, and were heterogeneous with regard to their experience with and attitudes towards science in the media. The interviews established cognitive validity of self-report items (Karabenick et al., 2007), and were also used as pre-clinical interviews. These interviews informed and enriched the questionnaire and provided samples for the development of the coding scheme.

Instrument testing: sample and procedure. The instrument was tested as a baseline survey with 51 science, technology, engineering, mathematics [STEM] graduate students and 35 undergraduate students from a northeastern United States research university.

The 35 undergraduate students, 15 of them majoring in a STEM field, were all enrolled in the elective course "Science writing for the media". Data was collected during August 2010 using an electronic version of the instrument.

Data from the graduate students was collected in two ways: (1) Twenty-eight responders were solicited by electronic invitations sent to graduate students in five randomly selected science and engineering departments. The invitations were sent in September 2010 via departmental lists of graduate students to approximately 476 addresses. About 50 people opened the link, and 28 completed the survey; (2) Twenty three additional responders completed the survey electronically prior to participating in two science communication workshops in December 2010 and March 2011. In the latter workshop the instrument was used as a form of formative assessment, directing discussion.

The low response rate and the choice to take part in the workshops and course imply that our sample is more interested in the topic of science communication than the general science student cohort. In this sense, the graduate sample is representative of the general audience for science communication courses and workshops.

Test/retest reliability. In order to establish test/retest reliability, 19 STEM graduate students of the sample that was solicited by electronic invitations completed the questionnaire twice, with an interval of approximately two weeks. Participants were allowed a week to complete the online questionnaire. Approximately two weeks after completing it they received a copy of their own answers and were instructed to make any changes that might make their essays clearer and more useful to a non-technical audience. Participants were compensated with $20.

The purpose of the procedure was to find out if the opportunity to revise ones' answers results in better performance without any other intervention. This control is important in order to attribute potential pre-post changes to communication training and not to the measurement tool itself. Two weeks was chosen as the interval time between the questionnaires since it allows time to forget the details of one's answer, while giving enough time for new ideas to sink in, and for people to talk and think about them.

No systematic changes were seen in the test/retest data of participants' answers. Of the 19 participants nine took the opportunity to make changes to their essays. However, these changes did not always make the answers more understandable to a lay person. While some made a change in an attempt to increase clarity (e.g., with changes shown in italics: "I plan to do an association between the trait phenotype *(the observable character i.e colonization of kernels, and aflatoxin amounts)* and the genotype"), others added more scientific information (e.g., with changes shown in italics: "Bayesian networks have been widely applied to the area of artificial intelligence. *It consists of two basic elements: the structure and parameters. Its structure is a direct graph, each node represents an event and a direct edge point from a cause to its effect. The parameter is the conditional probability of cause-effect alone its edge*"). At least one created a typo. Only one person significantly improved the essay's readability[i].

Inter coder reliability. In order to establish inter coder reliability, over 10% of the data were coded independently by two researchers. Disagreements were discussed and resolved by refining the coding scheme. The average free marginal Cohen's Kappa was 0.83 for the two coders.

Written Skills Analysis

Our responders were presented with four practical tasks: identifying jargon, describing one's own research, responding to a question about science in everyday life, and responding to a question about science's role in society (table 3). In this section we will describe various ways in which their answers could be meaningfully analyzed. As described above, the analytical scheme addresses seven clusters of learning goals, divided into basic, intermediate and advanced levels (table 1). Each cluster can be assessed using multiple criteria; below, we present examples of how we approached the analysis of the questions and findings regarding the prevalence of the different categories in the answers we collected. It is important to note that these data is based on a small group of participants. Therefore, it should be viewed as a demonstration and a baseline point of reference for the instruments' use in practice.

Basic Level: This level includes assessment of skills related to clarity, content and knowledge organization.

(1) Clarity

Language. A scientist who wants to communicate directly with a general public about issues of science faces several important hurdles. Perhaps the most basic of these is language (Weigold, 2002). In order to assess the appropriateness of language used, we developed the "Science concept familiarity index" and “Jargon index” to code for the use of jargon (specialized vocabulary).

Academic linguistic databases such as PERC Corpus Online[ii], allow searches within corpora (systematic aggregations of texts) of scientific research papers. This type of corpus might not always be freely available and not updated after being sealed. Therefore, two publicly available ways to add objectivity to the classification of the level of jargon were explored.

First, science concepts were classified using the *Benchmarks for Science Literacy* (American Association for the Advancement of Science, 1993); we identified concepts as mapping to the K-8th, 9th-12th grade curriculum, or being more advanced (Over K-12). The main weaknesses of this method are the limited number of science concepts included in the publication and the lack of contemporary concepts.

A second classification was based on *Google News* (news.google.com), an automated news aggregator available to the public since January 2006. The exact list of news sources is not known outside of *Google*, but *Google* itself reports that it indexes over 4,500 English-language news sites, including blogs. It aggregates several million articles a day and sends about 1 billion clicks each month to news publishers worldwide, which makes it a reasonable proxy for broader media coverage of news (Segev & Baram-Tsabari, in press). Science concepts were classified according to the number of hits on *Google News* in the past three years since the day of measurement[iii]:

- Not jargon, if the word/phrase received over 80,000 hits (e.g. virus, galaxy, atom).
- Familiar science concept, if the word/phrase received between 8,000 and 79,999 hits (e.g. gravity, Nitrogen).
- Recognizable science concept, if the word/phrase received between 800 and 7,999 hits (e.g. phosphorous, magnetic field).
- Unfamiliar science concept, if the word/phrase received between 80 and 799 hits (e.g. bioremediation, uric acid).
- Strictly professional concept, if the word/phrase received less than 80 hits (e.g. meiosis, baryonic).

The main weaknesses of this method are its lack of transparency and instability. *Google News* does not report the actual number of hits, but an estimate. The exact way in which this estimate is calculated in not published. Furthermore, *Google News* occasionally changes its algorithms and data sources without notifying its users, which may result in changes in concepts' scoring and impair reliability.

In order to address this problem one may compare a few search words using the same time frame and corpus or standardize the results by dividing them with a very common search term (such as 'www'). In this study we have used anchors – several terms that were recorded repeatedly in order to pinpoint changes in the measurement[iv]. For certain terms such as plasma and cell, another problem has to do with the validity of the measurement, with some of the hits referring to the concept in a non-scientific context (e.g. cell phone, plasma TV) (Segev & Baram-Tsabari, in press). However, an examination of the data quality of count estimates provided by web search engines that were used for quantitative measurement revealed satisfying quality of data according to objectivity, reliability and validity for simple measurements (Janetzko, 2008).

When asked to mark science concepts that should be defined when writing to a non-technical audience (Table 3), many participants in our sample distinguished between K-8/not jargon terms and less familiar terms (Table 4). The biggest difference between undergraduates and graduate students was with regard to the term DNA, which half of the graduate students felt should be explained, compared with 23% of the undergraduates. The terms 'kinetic energy' and 'polymer' were marked by almost two-thirds of the participants, and the rest enjoyed a wider consensus regarding the need to define them. It is interesting to note, however, that for each of the terms, unfamiliar as it may be (e.g. 'epigenetics') there were some participants who did not view it as jargon that should be defined.

[table 4 about here]

Jargon was also assessed in participants' answers to the open questions. Thirty nine concepts extracted from participants' answers were coded according to both the *Benchmarks* and *Google News,* and the coding showed a significant correlation ($p < 0.01$). Therefore, we developed a "Science concept familiarity index" based on *Google News,* as it provides for greater comprehensiveness, it is continuously updated, and it has greater ecological validity since we are interested in what scientists write for the media, which *Google News* addresses, rather than the formal educational context that the *Benchmarks* address.

After coding each of the concepts, we combined the scores to create a single "jargon index". This index is based on the idea that including undefined jargon in a text for the wide public should be coded based on the level of unfamiliarity of the word and not only by counting the number of concepts. The 'penalty' for increasingly unfamiliar jargon was based on powers of two. Jargon words that were defined in the text were only counted once, regardless of their familiarity level, except for familiar science concepts that were defined – these were not counted at all. The formula we used for calculating the jargon index was let n(familiar,nd) be the number of jargon words at the level of 'Familiar science concept' 'not defined' (nd) in the text, n(recognizable,d) be the number of jargon words at the level of 'Recognizable science concept' which are 'defined' (d) in the text, and so on:

$$\text{Jargon Index} = n(\text{familiar,nd}) + 2n(\text{recognizable,nd}) + n(\text{recognizable,d}) + 4n(\text{unfamiliar,nd}) + n(\text{unfamiliar,d}) + 8n(\text{professional,nd}) + n(\text{professional,d})$$

Let us demonstrate with the following jargon-heavy answer:

- "I study how tissue damage by an intestinal parasite promotes an immune response. Our body is programmed to recognize toxins, cancerous cells, and infections (such as viruses and bacteria) by detecting 'danger signals'. These

> signals consist of molecules that are not found in the human body under normal conditions, such as the endotoxin secreted by bacteria. When an intestinal parasite, such as Trichinella, infects the intestine it destroys some of the intestinal epithelial cells and these cells release their own danger signals, called 'alarmins'. We are interested in a novel alarmin, called IL-33, which is required for the body to develop a potent immune response to the parasite. The same type of immune response, induced by IL-33, is involved in allergic asthma and autoimmune diseases such as ulcerative colitis.“

This answer contains 16 jargon terms: 6 familiar science concepts (molecule, tissue, allergic, secrete, bacteria, parasite), 4 recognizable science concepts (toxin, epithelial, autoimmune, cancerous cells), 4 unfamiliar science concepts (intestinal parasite, endotoxin, Trichinella, ulcerative colitis), none of these were explained except ‘Trichinella’ and ‘ulcerative colitis’. In addition the text contains 2 professional terms (alarmin, IL-33,) which were explained. Therefore its jargon index is equals to:

6*(familiar,nd) + 2*4(recognizable,nd) + 4*2(unfamiliar,nd) + 2(unfamiliar,d) + 2(professional,d) = 6+8+8+2+2 = 26

The responses to the three open questions used in this study yielded distinctively different jargon indexes, with descriptions of research being more ‘jargony’ than explanations about science in everyday life (Table 5). Generally speaking, graduate students used more jargon than undergraduates, but the groups themselves were highly heterogeneous. The range of the jargon index among graduate students describing their research spanned from 1 to 34. This finding also highlights the distance between what people say should be done and what they actually do – between verbal knowledge and the application of that knowledge. While 75%-96% of the graduate students noted that recognizable and unfamiliar science concepts should be defined (Table 4), many of them actually used similar terms without defining them a few minutes later when describing their own work. Similar incoherence or dissociation between declared and actual practice has been described in the science education literature (e.g. Henderson, Yerushalmi, Kuo, Heller, & Heller, 2004). This finding also highlights the importance of evaluating scientists' skills based on actual performance of the relevant skill. *Readability*. Readability was assessed using the Flesch Reading Ease Score (Flesch, 1948). It is calculated based on the average sentence length and the average number of syllables per word. The Flesch Reading Ease Score indicates comprehension difficulty when reading a passage of contemporary academic English on a scale of 0 (very complex) to 100 (extremely simple). A score between 60 and 70 is largely considered acceptable, and understood by 8th and 9th graders. The reason we chose to use this readability formula is its availability in standard *Microsoft Word* software.

In our sample readability varied greatly between individual participants and questions. Generally speaking, readability was lower when people described their own research than when answering the other two questions. However, even within the same group answering the same question, the range was very wide – spanning, for example, from 0 to 67 among STEM graduates' descriptions of their own work, with an average of 35.3 (Table 5).

Type of explanation. The type of explanation was classified based on Rowan (1992) with some adaptations which are described below. Explanations were categorized as:

- Absent
- Definition. Example: "The internet is a virtual network." (A new category added by us).
- Elucidating explanation, which is definition with an example/non-example. Example: "Antibiotics only work on bacteria, which mean that they can only be used for diseases caused by microbes belonging to the bacteria family. Flu, on the other hand, is caused by viruses."

- Quasi scientific, which are explanations that create an image in the mind of the reader, such as using an analogy. Example: "Consider each computer as a node, and the Internet as a web."
- Transformative explanation refers to addressing alternative frameworks which already exist in the learner's mind and aiming for a conceptual change in which one central concept come to be replaced by another. This type of explanation is based on extensive work within the field of science education (e.g. Posner, Strike, Hewson, & Gertzog, 1982). In our analysis any explanation whose starting point was what the audience might think and progressed to point to dissatisfaction with the existing conceptions or explaining why the scientifically accepted theory is more plausible or fruitful was coded as a transformative explanation. Example: "I believe that the Bible must be interpreted in the context in which it was written. When the original text was written, people did not have our understanding of the natural world. They needed an explanation for their existence in terms that they could understand. That took the form of God creating them. Today we have proof that species evolve from one another and there is no reason to think that we are so special that we should not follow the same rules as the rest of nature".

If more than one explanation was found in a single answer the score was assigned to reflect the highest level of explanation. In our sample definitions, elucidating explanations and qausi scientific explanations were more frequent than transformative ones (Table 5). Those were used mainly by graduate students who answered a question about science in society (Table 6). However, when describing their own research, almost 40% did not offer any explanation at all (Table 5).

(2) Content

Presenting correct information. Responses were categorized as either correct or incorrect (e.g. "Antibiotics are remnants of dead flu cells" [false]).

In our sample almost all participants presented correct information (Table 5 and 6).

Science content. The number of "content units" was defined as any stand-alone fact stated about scientific research or its results. For example, the sentence "Two facts motivate my research – first, diverse systems are healthier systems and second, humans are rapidly altering diversity around the globe" would be coded as having 2 content units.

In our sample the number of content units varied widely between participants and between the different questions. Generally speaking, participants used more science content while describing their own research than when answering a question about science in everyday life, and the least science when answering the question about science in society (Table 5 and 6). Graduate students used more science in their answers than did undergraduates. However, even within the same group answering the same question, the range was very wide – spanning, for example, from 0 to 29 among STEM graduate descriptions of their own work.

Nature of Science. When scientists write science they also write *about* science – its dynamics, affordances, limitations and human aspects. These were identified and classified based on an extensive list of 18 ideas compiled by Lederman, Abd-El-Khalick, Bell, and Schwartz (2002). A few examples include:

- Principle: Observations are used to make scientific knowledge. Example from participant writing: "When we count all the material that we observe, we only account for a half of what we know existed after the big bang."
- Principle: Empirical evidence *supports* rather than *proves* scientific claims. Example from participant writing: "So far, no evidence that contradicts this theory [has been] found."

- Principle: Theories are well substantiated. Example from participant writing: "Now, many centuries later, modern astronomy, geology, and cosmology have provided us with several independent lines of evidence which all agree that the universe is at least 10 billion years of age. An age of just 5800 years is completely inconsistent with even the most basic of these observations."
- Principle: Society as an influence on science. Example from participant writing: "Americans have taken the stance that unless we can prove they [GMO's] are unsafe, we will use them. Europeans believe that unless there is solid proof that they are safe, they will *not* use them."

Although the potential maximum number of content units referring to the nature of science in a single answer was 18, many answers in our sample did not make even one such reference. The highest average of content units about the nature of science was found, unsurprisingly in answers to the science in society question (Table 6). Those referred mainly to the principles “Observations are used to make scientific knowledge”, “Science does not rely solely on empirical evidence”, “Empirical evidence supports rather than proves scientific claims“, “Theories change due to new ways of looking at existing evidence”, “Theories are well substantiated”, “There is no one single scientific method”, “The same data can be interpreted in several ways”, “Science is necessarily a mixture of objective and subjective components”, “Peer review limits subjectivity”, and “Society as an influence on science”.

Different ideas were stressed in response to the different questions. For example, when answering the question "Are humans responsible for the Earth getting warmer or not? Why can't scientists agree on that?" the principles “The same data can be interpreted in several ways” and “Science is necessarily a mixture of objective and subjective components” were mentioned frequently. When answering the question "How can you believe that humans developed from monkeys, when the Bible says God made us?" the principle “Theories are well substantiated” was used several times.

Connection to everyday life. Connection to everyday life was defined as an explicit connection made to common knowledge, previous event, or news story that was not already embedded in the question. It was categorized as either absent or present (e.g. "There has been a lot of talk lately in the news about the H1N1 virus. Like they’ve said, it’s a virus and other flus are too").

In our sample most participants did not add an additional connection to everyday life to that suggested in the original question. However, when describing their own research a little more than half of the graduate students used this strategy to explain what they were doing (Table 5). This is interesting since a focus group study of scientists’ discourse about science communication found that establishing relevance was a an almost a universal point brought up by the scientists, many times understood as personally relevant applications (Davies, 2008). This finding points again to the importance of evaluating scientists’ skills according to their actual performance.

(3) Knowledge organization

Framing. The framing of the answer was classified based on Nisbet and Scheufele's (2009) typology of eight frames applicable to science-related policy debates: social progress, economic competitiveness, morality and ethics, science and technology uncertainty, Pandora's box, public accountability, middle way, conflict and strategy. Based on qualitative review of the essay responses, we added an additional frame – science vs. non-science (demarcation, posing science as being at odds with another world view and indicating one is correct and the other nonsensical). Analysis according to framing was only used in the "science in society" question (Table 3). Examples for the use of five of the nine frames include:

- Frame: Social progress. Example: "We live in a nation that was built on the promise of a better and brighter future. That can only happen if we are committed to moving forward. So we invest in science and technology because if we know one thing for sure it is that we don't know everything and while we may not be able to save the world today we can most certainly make it better tomorrow."
- Frame: Economic development and competitiveness. Example: "There are many reasons that Europe chooses not to use genetically modified foods. A primary one that cannot be overlooked is the desire to have a readily accessible trade barrier to prevent American agricultural products from entering Europe. We have more land and we can produce food cheaper than European countries can. They do not want our grains and animal products and genetic modification is an easy trade barrier."
- Frame: Scientific and technological uncertainty. Example: "The vast majority of scientists do agree [with anthropogenic climate change]. The few scientists and many lay persons who deny anthropogenic climate change use the historical climate cycles of the Earth to bolster their position, but they fail to recognize that it's the rate of change that is at issues, not whether or not change has previously and is again occurring."
- Frame: Middle way/alternative path. Example: "The Bible said God made us, but it doesn't say how God made us."
- Frame: Science vs. non-science. Example: "First, the Bible has been written by Man and provides no evidence for the claim that God made us! On the other hand, there is vast amount of scientific evidence about human evolution."

The topic of the question biases the types of frames scientists choose in order to answer the question. In our sample, for example, most participants framed their answers either as 'middle way' or 'science vs. non-science' when answering questions about science and religion. More participants chose the demarcation frame rather than the one trying to find a middle way (Table 6). When answering the question about climate change many chose to frame the discussion as 'scientific and technological uncertainty'. Explaining why funding should be invested in science research when there are hungry people in the world was mainly done using a frame of 'social progress'.

Scaffolding explanation. Scaffolding is present when the paragraph successfully builds understanding incrementally, the size of each step is appropriate, and the audience is able to follow logical development of successively more complicated ideas (Sevian & Gonsalves, 2008). It was categorized as either absent or present. For example, scaffolding is present in the following description: "Two facts motivate my research – first, diverse systems are healthier systems and second, humans are rapidly altering diversity around the globe…My research asks if it matters that species are being gained and lost rapidly from these communities...By manipulating plant number and type (i.e., native plants versus exotic plants) in each community, I can begin to uncover the mechanisms underlying the superior performance of diverse systems."

An example for an answer that does not gradually build understanding is "The dynamic contact angle θ is the angle between a moving liquid/vapour interface and a solid surface, measured within the liquid at the contact line where the three phases (liquid, solid, gas) meet. There is significant evidence that the dynamic contact angle depends on the velocity V of the contact angle. Measurement of its advancing contact angle θA indicates the degree of wettability of the surface".

In our sample over half of the responders gradually build understanding in their answers.

Intermediate Level: This level includes assessment of skills related to style, analogical and narrative approaches.

(4) Style:

Humor. Humor included both explicit jokes and ironic language. We categorized humor as either absent or present (e.g. "Hence, there is no biotic [living form] for the antibiotic to target, other than killing our own cells, which wouldn’t be very wise"). In our sample very few participants took advantage of this efficient strategy to enhance their message.

(5) Analogical approach:

Analogy. Analogies are defined as a systematic mapping between two situations: the *source* (familiar situation) and the *target* (novel situation) (S Kapon et al., 2009). We categorized analogies as either absent or present (e.g. "viruses and bacteria are different organisms, like a cat and dog are different organisms"; "DNA is like the 'How to' manual for your body"). In our sample, over 10% of the answers used analogies to explain the unknown in terms of the already known.

Metaphor. Metaphors structure one concept in terms of another. Unlike analogies, metaphors do not necessarily map directly between source and target; similarities can be associative (S Kapon et al., 2009). We categorized metaphors as absent or present (e.g. "fusion in their cores releases energy and heat which bubbles up to the surface"; "forming the seeds to the galaxies that we see today"). In our sample graduates were more likely to use metaphors, especially for talking about their own science.

(6) Narrative:

Narrative. Narrative was defined and identified based on Norris, Guilbert, Smith, Hakimelahi, and Phillips (2005). It was categorized as either absent or present. For example, narrative is present in the following story: "Grandpa was born with a certain chance of getting cancer and so was Doctor Goody Two-shoes. There are certain factors like conscious decisions that Grandpa made (e.g. smoking and drinking like a fish) and other factors that are out of one's control (e.g. the fact that Doctor Goody Two-shoes lived next to that jerk who made his yard into a superfund site tainting his veggies) that increases one's chances.".

In our sample over 10% of those explaining science in everyday life used some kind of narrative approach in their answer, but less so in the other questions.

[table 5 about here]

Advanced Level: This level includes assessment of skills related to applying a dialogic approach in science communication.

(7) *Dialogic approach*:

References to multiple worldviews. Answers to the science in society question were classified with regard to references they made to multiple worldviews. These could be:

- Absent
- Acknowledging more than one world view. Example: "there are a few scientists that do not believe that humans are at least partially responsible for the Earth getting warmer, but the overwhelming general consensus is…".
- Explaining more than one worldview. Example: "The best way of testing this is to make the modified plant and monitor it for an extended period of time…. Europe and the US have differing ideas about what is an extended period of time and how much testing is sufficient. Although laboratory tests have shown GM foods are safe for human consumption, there is much more testing that could be done, including the effects of GMO fields on nearby crops. Because the EU has stricter regulations, they are waiting for more tests, while the US considers the testing that has already been done sufficient."

In our sample over 15% of responders did not make any reference to an additional view, about third acknowledged such a view. More responders choose to explain more than one world view than the two other options (Table 6).

Respect for other world views. Answers to the science in society question were also classified with regard to the respect they showed to multiple worldviews, except for answers that made no reference to other worldviews. Possibilities included:

- Denying others' basis for beliefs, in either sarcastic or straight forward ways. Example: "The only reason people don't eat genetically modified foods in other countries is because people are scared of them."
- Accepting other's right to believe differently. Example: "Some people are willing to be convinced. Some are not. And that's OK."
- Accepting the possibility that others might be right. Example: "All of these things are debates and it's really crucial for science to not squelch debates. It is important to hear out skeptics, think through, address them."

In our sample over one quarter of the responders chose to deny other's basis of belief, about 15% accepted other's right to believe differently, while the majority of responders accepted the possibility that those others might be right (Table 6).

Argumentation. Argumentation plays a central role in contemporary science education as a means for learning about the social practice of science, and developing knowledge and understanding of the evaluative criteria used to establish scientific theories (Duschl & Osborne, 2002). Placing argumentation in the advanced level of the scale has to do both with its dialogic nature as well as with the level of skill needed to do it well. Answers to the science in society question were also classified with regard to the level of the argument, using a structural scheme based on Schwarz, Neuman, Gil, and Ilya (2003):

- An argument may include an assertion but no reasons to back up the assertion. Example: "The center of the Earth is hot."
- A one sided argument includes an assertion and at least one reason to back up the assertion. Example: "Genetically modified food is safe. All that is occurring when food is genetically modified is that we change the DNA of food very slightly. Because everything we eat has DNA in it, this is not going to hurt us at all. The changes to the DNA typically give plants proteins that help the crops grow better and fight off insects. But these proteins will not hurt people for several reasons: these molecules are specific to the plants and insects and will not interact with human proteins; these proteins have been well studied long before being introduced to crops, because it is important to know how these proteins act; and in the end they are only proteins, we eat lots of proteins and need to eat protein to live."
- A two-sided argument presents two different assertions and at least one reason for each. Example: "Like many other technologies, genetic modification is very complicated. On one hand it does not seem harmful because there are no foreign chemicals being introduced to the crop- we only change the order of building blocks in the DNA. However, because our understanding of how the DNA sequence controls all of a plant's characteristics, changing the sequence can have unintended and unexpected consequences."
- A compound argument presents two assertions with reasons for each connected together by a qualifier (“even if”, “even though”, “it depends”). A Compound argument undertakes the analyses of the pros and cons necessary to solve the issue. Example: "As you can imagine the Earth is a very complicated system, with many different things affecting it all over the world. This complexity can sometimes

> make it difficult to figure out exactly how humans may be impacting the Earth on a broad scale. The complexity also means there may be evidence to support several different theories. However, at this point, the vast majority of climate scientists agree that the Earth is warming and humans are at least contributing to that warming trend. If you look at historical cycles of warming and cooling on the planet, you see that it is common for the planet to go through ice ages and warm periods. What makes this time different is how fast the temperature is changing. In addition, scientists observe that humans are releasing lots of carbon dioxide gases into the atmosphere (from activities like burning gasoline in cars). This carbon dioxide acts like a blanket on the Earth, heating it up. Carbon dioxide data gathered shows that carbon dioxide levels in the atmosphere have increased dramatically in the last hundred years, coinciding perfectly with human explosion in fossil fuel use. All of these pieces put together lead most scientists to be sure that the planet is warming and humans are contributing. If some scientists do not agree it is because the earth is a very complex system and we cannot be 100% sure we understand everything that is happening here."

In our sample most answers gave either a one-sided argument or a sophisticated compound argument, potentially enabling their audience to understand both sides of the argument (Table 6).

[table 6 about here]

The data presented in this section suggest that although our sample was representative of people who are already interested in science communication to some degree, great heterogeneity still exists between the outcomes they produced. Many were unable to apply their declarative knowledge regarding the use of jargon and importance of relevance into practice, and only a few used humor, analogies and narratives in their communication. Science content was emphasized, while discussion of the nature of science was infrequent. Participants were more likely to respectfully accept others’ worldviews when answering them, rather than denying their basis for belief, and 40% were able to offer a complete and two sided argument. The data also suggest that each of the questions does explore a different aspect of science communication written skills.

Conclusion

This article describes the development of a first of its kind tool for assessing scientists' written skills in public communication of science. Our task was to compile a list of potential learning goals for science communication education programs, and then to derive a conceptually based scheme for evaluating them. The outcome of our study is the assessment tool and the analytic framework for understanding the results it produces.

The development of the goals and instrument were guided by existing literature, interviews with active scientists in order to establish face validity, as well as establishment of test/retest reliability. It may be used as a baseline survey, as formative assessment during media training or as a pre-test/post-test summative evaluation of the learning outcomes of a wide range of science communication education programs and courses. For the purpose of this study 16 different characteristics of the text were examined (all summarized in table 6). However, in practice, only those which evaluate the learning goals of the specific intervention ought to be used.

The main contribution of this paper is the detailed description of the written skills analysis, drawn from a variety of content worlds and assessment traditions which includes also novel measurements, such as the jargon index. The analysis pointed to the importance of assessing actual performance, rather than declared knowledge. In particular, we are intrigued by the idea that scientists may need to "unlearn" the communication skills they have acquired as scientists. If learning the discourse of science is essential to becoming a scientist, learning the discourse of public communication of science is essential for scientists engaging with the public. This process may take place in sociocultural environments which value such practices.

Unfortunately, the two discourses are sometimes in tension: one rewards jargon, the other penalizes it; one rewards precision, the other accepts approximation; one rewards quantification, the other rewards story-telling and anecdotes. Using some of the criteria described here, it may be possible to document changes in scientists' ability to develop their messages in ways that will reward them in the public arena.

The small sample sizes reported here prevent us from making strong claims about the science communication skills of science students. However, the results do suggest potential areas for further investigation, such as the frequent mismatch between what scientists say they know about science communication and what they actually do; the limited use of narrative, analogy, and metaphor despite their frequent presence in advice books; and the very high presence of content while simultaneously failing to present the Nature of Science or the connections to everyday life.

Future research in the evaluation of science communication skills could develop in many directions. One may be a theoretical exploration of the possibilities of standardized and quantitative assessment of science communication education initiatives. Another might be an attempt to compare the findings of this type of assessment with a different form of assessment (e.g. evaluations by lay readers). A third direction may quantitatively describe the interactions between different measures, such as content units and jargon while taking into account also the number of words in each answer. This type of study may consider the ‘density of content’ alongside readability and jargon. Similarly, demographic characteristics of the participants (e.g. seniority and gender) may interact with certain measures. Further research may also adapt this type of analysis to analyzing real media messages of scientists on unedited broadcast formats (such as talk show interviews). All of these may help practitioners and scholars alike to identify best practices in science communication education.

Acknowledgements

This research received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ IOF Grant agreement n° PIOF-GA-2009-236283. The authors would like to thank the scientists who helped us develop this measurement and to Karina Knoll, Claire Peters and Michal Nitzany for their skilled and thoughtful assistance with the study.

References

Aikenhead, G. S., & Ryan, A. G. (1992). The development of a new instrument: Views on Science-Technology-Society (VOSTS). *Science Education, 76*(5), 477-491.

American Association for the Advancement of Science. (1993). *Benchmarks for Science Literacy*. New York: Oxford University Press.

Baron, N. (2010). *Escape from the ivory tower, A guide to making your science matter*. Washington: Island Press.

Besley, J. C., & Tanner, A. H. (2011). What science communication scholars think about training scientists to communicate. *Science Communication, 33*(2), 239–263.

British Council. (2011). The Times Cheltenham Science Festival hosts the International Final of the world's greatest science communication competition Retrieved June 20, 2011, from http://www.britishcouncil.org/new/press-office/press-releases/Famelab-proves-science-has-got-talent/

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Christensen, L. L. (2007). *The Hands-On Guide for Science Communicators, A Step-by-Step Approach to Public Outreach*: Springer Science + Business Media, LLC.

Cribb, J., & Hartomo, T. S. (2002). *Sharing Knowledge, A guide to effective science communication*. Collingwood, Australia: Csiro Publishing.

Davies, S. R. (2008). Constructing Communication: Talking to scientists about talking to the public. *Science Communication, 29*(4), 413-434.

Day, R., & Gastel, B. (2011). *How to Write & Publish a Scientific Paper (7th ed.)*. Westport, CT: Greenwood Press.

Dean, C. (2009). *Am I making myself clear? A Scientist's guide to talking with the media*. Cambridge, Massachusetts: Harvard University Press.

Directorate General Research. (2010). *European Guide to Science Journalism Training*. Brussels: EUROPEAN COMMISSION.

Dunwoody, S., & Ryan, M. (1985). Scientific barriers to the popularization of science in the mass media. *Journal of communication, 35*(1), 26-42.

Duschl, R. A., & Osborne, J. (2002). Supporting and Promoting Argumentation Discourse in Science Education. *Studies in Science Education, 38*, 39-72.

Fennewald, T. J. (2011, November 2-5). *NSF Fellows in Classrooms: A Way to Communicate Science?* Paper presented at the Annual Meeting of the Society for Social Studies of Science (4S), Cleveland.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*(3), 221-233.

Halliday, M. A. K. (1993). Some grammatical problems in scientific English. In M. A. K. Halliday & J. R. Martin (Eds.), *Writing science: Literacy and discursive power* (pp. 69-85). London: The Falmer Press.

Hayes, D. P. (1992). The growing inaccessibility of science. *Nature, 356*, 739-740.

Hayes, R., & Grossman, D. (2006). *A scientist's guide to talking with the media, Practical Advice from the Union of Concerned Scientists*. New Brunswick, New Jersey: Rutgers University Press.

Henderson, C., Yerushalmi, E., Kuo, V. H., Heller, P., & Heller, K. (2004). Grading student problem solutions: The challenge of sending a consistent message. *American Journal of Physics, 72*(2), 164-169.

Janetzko, D. (2008). Objectivity, reliability, and validity of search engine count estimates. *International Journal of Internet Science, 3*(1), 7-33.

Jensen, P. (2010). A statistical picture of popularization activities and their evolutions in France. *Public Understanding of Science, doi: 10.1177/0963662510383632*

Kapon, S., Ganiel, U., & Eylon, B. S. (2009). Explaining the unexplainable: translated scientific explanations (TSE) in public physics lectures *International Journal of Science Education, 32*(2), 245 — 264.

Kapon, S., Ganiel, U., & Eylon, B. S. (2009). Goals and design of public physics lectures: perspectives of high-school students, physics teachers and lecturers. *Physics Education, 44*(5), 528-535.

Karabenick, S. A., Woolley, M. E., Friedel, J. M., Ammon, B. V., Blazevski, J., Bonney, C. R., et al. (2007). Cognitive processing of self-report items in educational research: do they think what we mean? *Educational Psychologist, 42*(3), 139-151.

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology, 78*(2), 311-328.

Latour, B., & Woolgar, S. (1979). *Laboratory Life*. Beverly Hills: Sage.

Lederman, N. G. (2007). Nature of Science: Past, Present, and Future. In *Handbook of Research in Science Education* (pp. 831-879).

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching, 39*(6), 497-521.

Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Norwood, NJ: Ablex Publishing Corporation.

Leshner, A. I. (2009). Alan I. Leshner: Commentary on the Pew/AAAS Survey of Public Attitudes Toward U.S. Scientific Achievements. Retrieved November 8, 2010, from http://www.aaas.org/news/releases/2009/0709pew_leshner_response.shtml

Martin-Sempere, M. J., Garzon-Garcia, B., & Rey-Rocha, J. (2008). Scientists' motivation to communicate science and technology to the public: surveying participants at the Madrid Science Fair. *Public Understanding of Science, 17*(3), 349-367.

Meredith, D. (2010). *Explaining Research, How to reach key audiences to advance your work*. New York: Oxford University Press.

Miller, S., Fahy, D., & The ESConet Team. (2009). Can science communication workshops train scientists for reflexive public engagement?: The ESConet experience. *Science Communication, 31*(1), 116-126.

Nisbet, M. C., & Scheufele, D. A. (2009). What's next for science communication? Promising directions and lingering distractions. *American Journal of Botany, 96*(10), 1767-1778.

Norris, S. P., Guilbert, S. M., Smith, M. L., Hakimelahi, S., & Phillips, L. M. (2005). A theoretical framework for narrative explanation in science. *Science Education, 89*, 535– 563.

PCST Academy. (2010). Retrieved January 4, 2010, from www.upf.edu/pcstacademy/Courses

Peters, H. P., Brossard, D., de Cheveigne, S., Dunwoody, S., Kallfass, M., Miller, S., et al. (2008). Science-media interface: it's time to reconsider *Science Communication, 30*(2), 266-276.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*(2), 211-227.

Rowan, K. E. (1992). Strategies for Enhancing Comprehension of Science. In B. V. Lewenstein (Ed.), *When Science Meets the Public* (pp. 131-143). Washington, DC: American Association for the Advancement of Science.

Rowan, K. E. (2003). Informing and Explaining Skills: Theory and Research on Informative Communication. In J. O. Greene & B. R. Burleson (Eds.), *Handbook of Communication and Social Interaction Skills* (pp. 403-438). Mahwah, NJ: Erlbaum.

Schwarz, B. B., Neuman, Y., Gil, J., & Ilya, M. (2003). Construction of collective and individual knowledge in argumentative activity. *Journal of the Learning Sciences, 12*(2), 219 - 256.

Segev, E., & Baram-Tsabari, A. (in press). Seeking science information online: Data mining Google to better understand the roles of the media and the education system. *Public Understanding of Science, Published online before print December 1, 2010, doi: 10.1177/0963662510387560*.

Selzer, J. (Ed.). (1993). *Understanding Scientific Prose*. Madison: University of Wisconsin Press.

Sevian, H., & Gonsalves, L. (2008). Analysing how scientists explain their research: A rubric for measuring the effectiveness of scientific explanations. *International Journal of Science Education, 30*(11), 1441 — 1467.

Shinn, T., & Whitley, R. (Eds.). (1985). *Expository Science: Forms and Functions of Popularisation*. Dordrecht/Boston/Lancaster: D. Reidel.

Society for Science and the Public. (2011). Judging at Your Fair. Retrieved June 20, 2011, from http://www.societyforscience.org/page.aspx?pid=284

Sturzenegger-Varvayanis, S., Eosco, G., Ball, S., Lee, K., Halpern, M., & Lewenstein, B. (2008, June). *How university scientists view science communication to the public.* Paper presented at the 10th International Conference on Public Communication of Science and Technology Malmo, Sweden.

The Royal Society. (2006). *Factors affecting science communication: A survey of scientists and engineers*. London, UK.

Treise, D., & Weigold, M. F. (2002). Advancing science communication: a survey of science communicators. *Science Communication*, *23*(3), 310-322.

Weigold, M. (2002). Communicating Science. In *Management Benchmark Study*: Office of Science, Deptment of Energy.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. UK: Cambridge University Press.

Wilkinson, A. M. (1991). *The Scientist's Handbook for Writing Papers and Dissertations*. Englewood Cliffs, N.J.: Prentice Hall.

Ziman, J. M. (1968). *Public Knowledge: An Essay Concerning the Social Dimension of Science*. Cambridge: Cambridge Univ. Press.

Ziman, J. M. (1978). *Reliable knowledge : an exploration of the grounds for belief in science.* Cambridge [Eng.] ; New York: Cambridge University Press.

---

[i] The same test/retest questionnaire also included questions regarding participant's views of science communication. Systematic changes were recorded with regard to the control group's views.

[ii] http://scn.jkn21.com/~percinfo/

[iii] When plural and singular forms differed, the higher score between the two was used.

[iv] For the purpose of future comparison, the values measured on *Google News* in the period December 29 2008-December 29 2011 were: virus 488,000; cell 306,000; galaxy 128,000; atom 88,300; bacteria 48,100; gravity 38,000; ecosystem 29,500; Synthetic 22,500; Nitrogen 11,800; magnetic field 2,450; phosphorus 2,070; uric acid 209; bioremediation 122; endotoxin 88; eutrophication 81; meiosis 70; baryonic 17.

Table 1. A framework for learning goals in written science communication

Level	**Learning goals**	**Clusters**	**Kapon, Ganiel, & Eylon (2009)**	**Sevian & Gonsalves (2008)**	**Practical advice books**
Basic	Use appropriate language, address readability, use basic explanations as appropriate, avoid jargon, acknowledge prior knowledge (or lack of specific necessary knowledge)	Clarity	[includes clarity within "knowledge organization", includes language within "story"]	Pedagogical content knowledge	Content, Language
	Select appropriate content: engaging, interesting, relevant to particular audience. Include scientific information, as well as nature of science, scientific method, implications	Content	Content	Content knowledge	Preparation, Content
	Organize presentation well, using good pedagogical and communication techniques: main theme, framing, scaffolding, repetition	Knowledge organization	Knowledge organization	Pedagogical knowledge	Content

Intermediate	Use aspects of style creatively: humor, emotions, anecdotes, local references	Style	[includes aspects of style within "story"]	Pedagogical knowledge	Style
	Develop analogic strategies for explaining complex topics	Analogy	Analogical thinking	--	Style
	Use complex narrative tools as appropriate, such as character development, conflict and resolution	Narrative	Story	--	Style
Advanced	Acknowledge and show respect to multiple world views	Dialogue	--	--	--

Table 2: Summary of advice for scientists interacting with the media[1]

Title	A Scientist's Guide to Talking with the Media	Am I Making Myself Clear?	Sharing Knowledge	The Hands-On Guide for Science Communicators	Explaining Research	Escape from the Ivory Tower	Goals and design of public physics lectures
Author (year)	Hayes and Grossman (2006)	Dean (2009)	Cribb and Hartomo (2002)	Christensen (2007)	Meredith (2010)	Baron (2010)	Kapon, Ganiel and Eylon (2009)
Preparation							
Know the basics of journalism	✓	✓	✓	✓	✓	✓	n/a
Get media training	✓	✓	✓	✓	✓	✓	n/a
Know your interviewer	✓	✓	✓	✓	✓	✓	n/a
Build relationship with the journalist	Keep reporters up to date, become on-call expert	✓	✓	Be proactive	✓	Reach out instead of waiting	n/a
Know your audience and shape your message accordingly	✓	✓	✓	✓	✓	✓	✓
Assume minimal prior knowledge	✓	✓	✓	✓	✓	✓	✓
Practice	✓	✓	✓	✓	✓	✓	✓

The message: Content							
Focus on a few main points	3-4	1-3	2-3	1-3	✓	3-4	✓
Repeat your main points (in different ways)	✓	✓	✓	✓	✓	✓	✓
Create clear and concise messages	✓	✓	✓	✓	✓	✓	✓
Include methodology /processes	✓	✓	✓	✓	✓	✓	✓
Avoid too much information	✓	✓	✓	✓	✓	✓	✓
Qualify when necessary	✓	✓	✓	✓	✓	✓	✓
Put message into perspective (the "so what?" question)	connection with day-to-day life	✓	connection to current issues	✓	✓	✓	✓
The message: Style							
Use examples, analogies, metaphors, give meaning to numbers	✓	✓	✓	✓	✓	✓	✓
Use/avoid clichés	use	-	-	avoid	avoid	avoid	-
Visualize: use	✓	✓	Bring the	✓	✓	✓	✓

pictures, graphics, tables, animations and movies			journalist to your lab				
Talk from the heart. Show passion and enthusiasm	✓	✓	✓	✓	✓	✓	✓
Use humor	✓	-	-	-	✓	✓	✓
The message: Language							
Use simple and short words and sentences	✓	✓	✓	✓	✓	✓	✓
Avoid jargon, acronyms, and abbreviations	✓	✓	✓	✓	✓	✓	✓
Use sound bites	✓	✓	Eye-catching headlines, lively quotations	✓	Artful quotes	✓	-

[1] Additional advice concerns the actual delivery of an interview. These include: Stay on message; don’t get angry with the interviewer; never guess answers; offer to check the final draft (but don’t expect the journalist to agree); suggest other scientists who could comment on your work; and offer a written summary so the interviewer will have all the details.

Table 3. Items for assessing written communication skills

Purpose	**Levels**	**Item**
Identifying jargon, recognizing the level of public prior knowledge	Basic	In your opinion, which of the following science concepts should be defined when writing to a non-technical audience? Mitochondria, Angle, Pulsar, quantum, Meiosis, Dark matter, Polymer, Epigenetic, Isotope, Kinetic energy, Density, DNA, Cell, The standard model.
Describing one's research	Basic and Intermediate	Please describe your research, its context and implications for a general audience in 150-200 words (you can pick a specific project in progress or research that has already been completed).
Responding to a question about science in everyday life	Basic and Intermediate	Imagine you are talking to members of your family, who do not have a science background. Knowing that you have general science knowledge, they ask you one of the following questions about science in their lives. Choose one question and answer in 75-150 words. (1) "Why doesn’t the doctor prescribe antibiotics for flu?" (2) "If there is no oxygen in space, how does the sun burn?" (3) "Why can't I use metal in a microwave?" (4) "What is the Internet and how does it work?" (5) "Why does a white shirt becomes transparent when it's wet?" (6) "How do the police identify people based on their DNA?" (7) "How come grandfather, who smoked a pack a day for 72 years, is alive and well at the age of 91, while his vegetarian nonsmoking doctor died of cancer?" [1]
Responding to a question about science's role in society	Basic, Intermediate, and Advanced	Happy with your answer, they now ask one of the following questions, about science’s interaction with society. Choose one question and answer in 100-200 words. (1) "How can you believe that humans developed from monkeys, when the Bible says God made us?" (2) "How can you believe that the universe is 13 billion years old, when the Bible says God created it less than 6000 years ago?" (3) "Are humans responsible for the Earth getting warmer or not? Why can't scientists agree on that?" (4) "Is genetically modified food safe? How come the Europeans don't use it but people in America do?" (5) "Why do we spend all this money on giant particle accelerators and journeys to Mars, when there are hungry people in the world?"

[1] not all questions were equally popular. The most popular questions participants chose to answer were questions number 1,2,6 and 7. Questions 3 and 4 were less popular and no one in our sample answered question number 5.

Table 4. Percentages of graduate and undergraduate students that marked concepts in response to the question “In your opinion, which of the following science concepts should be defined when writing to a non-technical audience?”

Science concept	STEM graduates (n = 48)	Undergraduates (n = 35)	Grade level according to the Benchmarks[1]	Classification based on *Google News* hits[2]
Angle	12	11	K-8	Not jargon
Density	27	23	K-8	Familiar
Cell	29	17	K-8	Not jargon
DNA	53	23	K-8	Not jargon
Kinetic energy	75	63	9-12	Recognizable
Polymer	72	77	9-12	Recognizable
The Standard model	88	97	Over K-12	Recognizable
Quantum	92	89	Over K-12	Familiar
Isotope	84	80	9-12	Recognizable
Epigenetic	92	97	Over K-12	Unfamiliar
Pulsar	94	91	Over K-12	Recognizable
Meiosis	94	89	9-12	Professional
Dark matter	92	89	Over K-12	Recognizable
Mitochondria	96	86	9-12	Unfamiliar

[1] Based on the online version of Project 2061 Benchmarks for Science Education (AAAS, 1993).

[2] Based on the number of hits on Google News (http://news.google.com) during the last three years.

Table 5. Descriptive statistics of the analysis of graduate and undergraduate students' answer to the questions "Please describe your research, its context and implications for a general audience in 150-200 words" [My research] and "Imagine you are talking to members of your family…they ask you one of the following questions about science in their lives". (75-150 words) [Science in everyday life].

			My research	**Science in everyday life**	
Cluster	**Category**	**Measure**	**STEM graduates (n = 46)**[1,2]	**STEM graduates (n = 51)**	**Under-graduates (n = 35)**
Clarity	Language: Science concept familiarity index	Average jargon index and s.d.	8.8 ± 7.7	2.5 ± 2.9	0.8 ± 1.2
	Readability	Fleisch readability ease average and s.d.	35.3 ± 15.2	51.3 ± 14.4	50.3 ± 15.7
	Type of explanation	% absent	39	6.3	11.4
		% definition	22.2	17.8	40
		% elucidating explanation	22.8	40.2	37.1
		% quasi scientific	13.3	31.6	11.4
		% transformative explanation	2.3	3.8	0
Content	Presenting correct information	% correct	100	100	86.6
	Science content	Average no. of units and standard deviation (s.d.)	10.9 ± 6.6	9 ± 4.9	6.5 ± 4.2
	Nature of Science	Average no. of units and s.d.	0.68 ± 0.8	0.02 ± 0.14	0.23 ± 0.6
	Connection to everyday life	% present	53.3	39	20
Knowledge organization	Scaffolding explanation	% present	65.3	88.3	80
Style	Humor	% present	2.3	7.9	2.9

Analogy	Analogy	% present	11.3	15.5	14.3
	Metaphor	% present	15.3	13.6	5.7
Narrative	Narrative	% present	6.5	15.6	11.4
No. of words average and s.d.			141.7 ± 77.3	94.1 ± 54.7	79 ± 38.1

[1] not all graduate students responded to this question

[2] undergraduate students were not asked to describe their research, since many of them were not involved in a research project

Table 6. Descriptive statistics of the analysis of graduate and undergraduate students' answer to the question "Happy with your answer, they now ask one of the following questions, about science's interaction with society. Choose one question and answer in 100-200 words".

Cluster	Category	Measure	STEM graduates (n = 50)	Under-graduates (n = 35)
Clarity	Language: Science concept familiarity index	Average jargon index and s.d.	2.2 ± 2.5	0.7 ± 1.2
	Readability	Fleisch readability ease average and s.d.	49.7 ± 10.9	49.9 ± 12.3
	Type of explanation	% absent	18.6	0
		% definition	17.6	42.8
		% elucidating explanation	26.6	28.6
		% quasi scientific	22.4	28.6
		% transformative explanation	14.1	0
Content	Presenting correct information	% correct	93.3	100
	Science content	Average no. of units average and standard deviation (s.d.)	6.8 ± 5.9	3.8 ± 3
	Nature of Science	Average no. of units average and s.d.	2 ± 2.2	1.3 ± 1.2
	Connection to everyday life	% present	32	14.3
Knowledge organization	Framing	% social progress	19.3	5.9
		% economic competitiveness	7.8	5.9
		% morality and ethics	0	2.9
		% science and technology uncertainty	13.8	20.59

		% Pandora's box	2.2	5.9
		% public accountability	4.2	8.8
		% middle way	9.6	20.59
		% conflict and strategy	2	0
		% science vs. non-science	17.8	32.3
	Scaffolding explanation	% present	50	62.9
Style	Humor	% present	2.2	0
Analogy	Analogy	% present	12.2	5.7
	Metaphor	% present	6.2	2.9
Narrative	Narrative	% present	9.9	2.9
Dialogic approach	References to multiple worldviews	% absent	17.9	14.7
		% Acknowledging more than one world view	39.8	32.3
		% Explaining more than one worldview	41.8	53
	Respect for other world views	% Denying others’ basis for beliefs	28.8	25.8
		% Accepting other’s right to believe differently	15.6	12.9
		% Accepting the possibility that others might be right	55.8	61.3
	Argumentation	% only assertion	7.9	2.9
		% A one sided argument	40	45.7
		% A two-sided argument	6.2	11.4
		% A compound argument	43.9	40
No. of words average and s.d.			133.1 ± 75.6	97 ± 34.7