

Standardization in Public Relations Measurement and Evaluation

David Michaelson and Don W. Stacks

As the public relations profession continues to focus more and more on outcomes associated with campaigns or public relations initiatives the question of standards has shifted to the forefront of discussions among and between professionals, academics, and research providers. Making this shift even more important to establishing impact on business goals and objectives is the fact that standardized measures for public relations activities have never been recognized. Unlike other marketing communications disciplinesⁱ, public relations practitioners have consistently failed to achieve consensus on what the basic evaluative measures are or how to conduct the underlying research for evaluating and measuring public relations performance.

In 2007, Michaelson wrote:

...significant variations continue to exist with the varying range of approaches to public relations measurement and evaluation. These variations result in a lack of standard measures that can be used to gauge the success of a public relations program as well as in an uneven overall quality of the research being conducted.ⁱⁱ

During the intervening four years, little has changed. The concept of standard measures is increasingly debated within the public relations universe, but attempts to develop these measures remain primitive and possibly misunderstood by significant proportions of public relations professionals and academics, as well as by the measurement and evaluation community itself. At best, the public relations measurement and evaluation community has developed “soft” guidelinesⁱⁱⁱ for consideration rather than a definitive model that can be adopted by and adhered to by the profession.

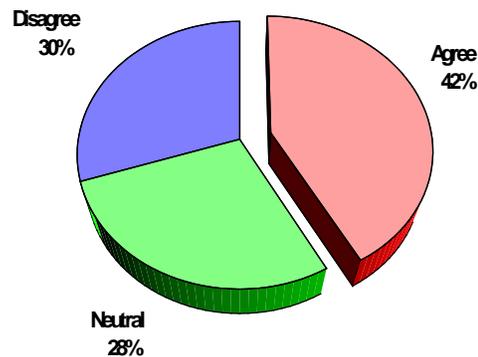
There have been a number of explanations and excuses as to why a consensus on standard measures for public relations performance has not been achieved. These explanations range from a general belief that public relations is somehow “different” and thus no standard can be applied to a desire by many communication consultants to own proprietary or “black box” measures that cannot be independently replicated and tested for reliability and validity, thus cannot be effectively expanded to or compared with other public relations activities.

The Challenge within the Practice of Public Relations

Compounding this debate is an additional challenge. Public relations practitioners cannot even agree on a basic vocabulary to describe communication activities.

The results of a May 2011 survey across a broad cross section of public relations professionals^{iv} adds to the conundrum (see Figure 1). The survey found that 42 percent of the respondents agreed that common terms and definitions for the measurement of public relations “do not exist at all” with an additional 28 percent reporting they are neutral on the issue (most likely a reflection of an overall lack of information).

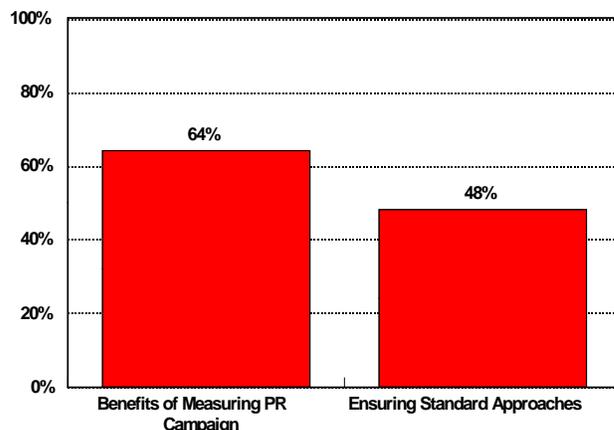
Figure 1
Common Terms and Definitions for the Measurement of Public Relations Do Not Exist At All
Agreement with Statement



Typical of this situation is the multiplicity of uses and misuse of concepts as basic as *outputs*^v and *outtakes*^{vi}. These terms, when used to describe public relations activities and their effects, are used inconsistently, if at all. This ambivalence toward standardized language is a reflection of the overall attitudes of the public relations profession on the broader question of standardized measures that allow for a comparative analysis of the performance of public relations programs regardless of client or agent.

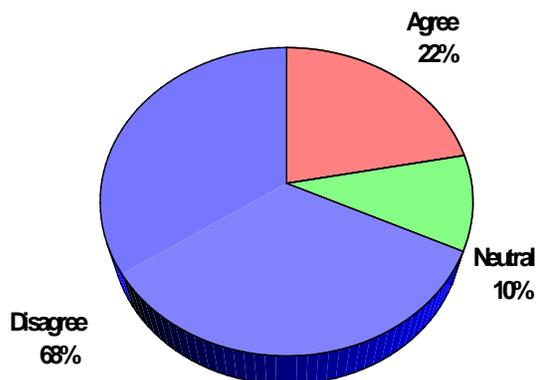
In the same survey of public relations professionals, only 64 percent give the highest levels of importance in the next five years (rated 9 or 10 on a ten point scale of importance) to the benefits of measuring a public relations campaign or program being well understood by the profession (see Figure 2), and, only 48 percent of respondents gave an equal level of importance in the same time frame to ensuring standard approaches for public relations measurement and evaluation were currently understood throughout organizations.

Figure 2
Most Important Issues in the Next 5 Years for Public Relations Practitioners to Understand Those Rating Issue as Highest Level of Importance Rated 9 or 10 on 10 Point Scale



Of greater concern is that one in three public relations professionals appear to believe that a common set of standards for public relations measurement is not necessary. Thirty-two percent either agreed (22%) or were neutral (10%) on the statement “a common set of standards for public relations measurement is not necessary” (see Figure 3). While two in three practitioners disagreed with this statement, it is clear that there is a significant lack of consensus with the profession on the importance and value of standards.

Figure 3
A Common Set of Standards for Public Relations Measurement Is Not Necessary
Agreement with Statement



A Working Definition of Standards

We constantly hear among public relations professionals that the primary reason for measuring public relations is to demonstrate the value of public relations programs. They often say they are looking to measure the “ROI” or return-on-investment of programs^{vii}. This has led to debates about *financial* versus *non-financial returns* and has fostered an industry devoted to the sale of advertising value equivalence (AVE) as a proxy for public relations ROI^{viii}. This focus on ROI, while important when taking into consideration how public relations outcomes impact on business goals and objectives^{ix}, misses the fundamental point of *why* we need to measure public relations activities. This has likely contributed to the lack of adoption of standard measures.

The question that is most basic and needs to be answered most acutely is “Why are standards in public relations measurement and evaluation needed?”

Answering this question starts with defining the concept of standards itself. A standard as defined by the *Oxford English Dictionary* is “an idea or thing used as a measure, norm, or model in comparative evaluations.”^x The operative words in this definition are “comparative evaluations.” This gets at the heart of why standard measures are an essential element in public relations.

Comparative evaluation is the basis of all standardized measurement systems. With a standardized system of comparative evaluation, public relations professionals would be able to gauge the *absolute*^{xi} performance of specific programs and specific program elements. In addition it will allow the comparative performance of prior and competitive programs. And, finally, allow us to compare that performance within industry and category, as well as the performance of the program *relative*^{xii} to other industries or categories. The value of comparative evaluation is found when the ability to determine if specific communication goals are being met (*absolute* measures) and if these changes in specific measures are significant based on the known performance of similar programs or campaigns that have been deemed successful (*relative* measures). These comparative evaluations allow public relations professionals the ability to measure progress and take corrective actions if needed to assure that communications goals are being achieved during the campaign (see Stacks, 2010; Stacks & Michaelson, 2010 for more on campaign phases and measurement).^{xiii}

Starting with this definition, this article is an extension of earlier work that presented “a set of best practices” for the measurement and evaluation of public relations activities. This work on best practices has evolved, as noted in Michaelson’s (2007) article, “into a standard set of measures against which programmatic diagnoses can be made and the success or failure of a program can be judged” (p. 1)^{xiv}. Best practices^{xv} however should not be confused with or used as a substitute for standards. Best practices are defined as: “A method or technique that has consistently shown results superior to those achieved with other means, and that is used as a benchmark.”^{xvi} Standards define and

determine *what* needs to be measured. Best practices illustrate *how* to best meet the objectives of the standard.

Setting the Context for Standardized Measures of Public Relations Activities

Applying a definition of standards requires understanding how standard measures have historically been applied in other communication disciplines. Starting with the need for comparability as the foundation for standardized measures, communicators and marketers have consistently looked to communication and persuasion theory as the foundation for establishing these standards. One of the models commonly used is based in an 1898 theory developed by E. St. Elmo Lewis.

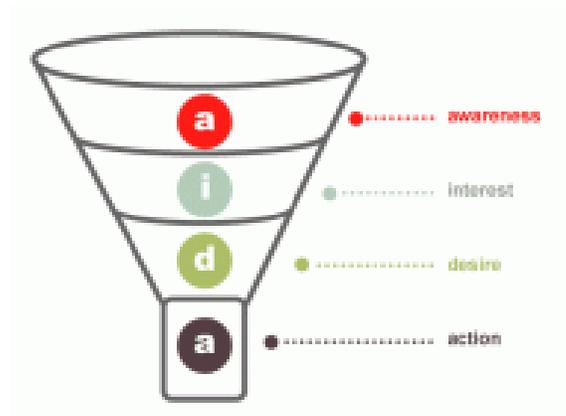
Born in 1872, E. St. Elmo Lewis was an early pioneer in advertising. In 1898, he proposed a theory of communication commonly referred to by its acronym A-I-D-A or awareness, interest, desire, action. This theory, based on studies of the life insurance industry, described “four cognitive phases that buyers follow when accepting a new idea or purchasing a new product”^{xvii} These phases are hierarchical and start with the presumptions that in order for a motivation to act or purchase to occur, several antecedent conditions have to be met that are typically dependent on external stimuli such as advertising, public relations, or direct sales.

These necessary and sequential stages are:

- *Awareness* that the product or service or service exists
- *Interest* in the product and the benefits the product or service offers the buyer
- *Desire* for the product or service

It is only by moving through these first three stages, that the fourth stage – *Action* or the actual purchase of the product or service – takes place. This four stage process has traditionally been represented as a “sales funnel” as illustrated in Figure 4^{xviii}.

Figure 4
The Sales Funnel



This model has evolved over the past century. Specifically, variations in the model have examined how new product introduction, product trial, repeat purchases, and brand loyalty based on long-term product satisfaction can impact action. Nonetheless, and in spite of these modifications, the basic principals which this model is based are still sound. Even in a world of interactivity, online communications and social networking, consumers still need to be *aware* of a product, express *interest* in it based on some degree of knowledge about the products attributes and benefits, *desire* the product because it is relevant to their needs, wants or interests before they will make a purchase decision or take another *action*.

Adapting the Lewis Model to Public Relations Activities

Adapting this model to public relations activities, however, is not as direct and obvious as it might seem. In a paper published in 2005 by the Institute for Public Relations' Commission on Public Relations Measurement and Evaluation, Michaelson and Griffin noted that one of the key challenges and most common flaws in public relations measurement "is linking communications objectives with the actual message analysis."^{xix} They go on further to state:

Typically communications [sic] objectives are directly related to the information needs dictated by a communications [sic] lifecycle. In public relations, there are two key recipients of these messages. The initial recipient is the media, who, in turn, serves as the conduit for transmitting messages to [the] intended recipient (i.e. their readers).^{xx}

The communication lifecycle in public relations described by Michaelson and Griffin is a structure that is consistent with the model developed by Lewis. Effective communication is a process that requires the recipient to go through four stages before a desired action takes place. These four stages of communication

effects include establishing awareness of the brand, the category or the issue, building sufficient knowledge and understanding about the brand, category or issue in order to make an informed decision, creating a level of interest in and preference for the brand, category or issue or at least a recognition of its relevance to the message recipient and finally a change in behavior or intent or commitment to take a specific action based on the received messages.^{xxi}

The Michaelson/Griffin model, however, differs in some key respects from the established AIDA model developed by Lewis. The Lewis model, with its advertising-based focus, emphasizes sales as the ideal outcome of the model if it is properly executed. Even later modifications of the model that added satisfaction as one of the measures (see Figure 4) focus on sales, either initial or continuing, as the ultimate outcome to be measured. The end result of public relations activities, by contrast, does not always look to sales or a similar type of conversion as a desired result. As Michaelson and Griffin further point out in their analysis:

Simply communicating the desire to have a message recipient take an action is unlikely to have the impact a communicator is hoping to achieve. In most cases the analysis fails to account for the stage of the communication [sic] lifecycle that needs to be addressed. For example, at the initial stage of the communication [sic] lifecycle, communicators should be measuring the proportion of messages that are strictly designed to develop awareness. At later stages in the lifecycle, the analysis needs to shift to determine the proportion of messages that communicate knowledge, interest or intent to act.”^{xxii}

For example, public relations programs, particularly those that work in conjunction with other forms of marketing communication (e.g., advertising), may focus on education of the target market or knowledge building. This sometimes narrower emphasis of public relations within the communication mix is consistent with findings from research conducted by Stacks and Michaelson on the relative contributions of advertising and public relations.^{xxiii}

Like the AIDA model, this application for public relations can extend beyond “intent to purchase”, “initial sale” or “product trial.” Changes in behavior, stated intent or commitments to future actions are also potential outcomes of this model. However, the role of public relations often extends beyond these behaviors or intents.

In a world of interactive communication, the target audience for a communicator can fulfill multiple roles. Traditionally, the communicator wants to change or modify behavior or intent to behave. But in a world of interactivity, the communicator also has the ability to employ a target audience as an extension of their communication efforts. This is what is referred to as the “third-person effect”

model of public relations.^{xxiv} For example, a buyer who is highly satisfied with a purchase has access to web sites and social media outlets where they have the ability to post product reviews and recommendations as well as negative commentary or criticisms, of the product or service and increase both the credibility of the message and the engagement with the product with others. In much the same way a reporter can write a positive review, the target audience can function as an advocacy group that can extend messages significantly beyond traditional communications that rely on an intermediary to deliver a desired message. This public relations focused model extends beyond merely satisfaction, but includes extending the role of the target audience to become part of the communication infrastructure.

Applying this communication lifecycle to the public relations process includes measuring the stages of communications among the target audience for the specific communication program, as well as among any intermediaries who may be employed to deliver messages. As Michaelson and Griffin noted:

The communications [sic] lifecycle must be understood for both the media as well as from the target audience. In many cases, each of these groups can be at a different level of understanding and may have different communications needs. An example is when the media may be completely knowledgeable about a product or service, but the target audience has only marginal awareness and little or no understanding about it. In these instances, the media may make assumptions about the level of knowledge held by the target audience and not report as completely and thoroughly as they should. This can inadvertently create a gap or omission in an article that needs to be corrected.^{xxv}

Procedures for Measuring Public Relations Activities

Assuming this structure is measuring the key communication variables, the question that needs to be answered is: “What are the specific measures that need to be employed and the processes and procedures that should be used that will assure comparative results that meet the objectives of a standardized measurement protocol?” In essence, we need to assure that the measures used to gauge public relations effectiveness are valid^{xxvi} as well as reliable^{xxvii}. Note: What is standard practice is not the proprietary results, but the psychometric results demonstrating reliability and validity thus allowing for competitive advantage.^{xxviii}

Prior experiments have attempted to determine both the validity and reliability of specific public relations measures. The most noteworthy of these experiments was conducted in 2009 by the Central Office of Information (COI) of the United Kingdom. In that study, COI tested five evaluation agencies^{xxix} using an identical briefing document that outlined specific measures to be included in the analysis.

In July 2009, COI sent an identical brief comprising 138 items of media coverage to five companies for evaluation. We wanted to know how many people consumed the coverage, how much it cost per 1,000 reached and what the favourability and tone of the coverage was. Not only did we get five different sets of results, but the range within each was very large.^{xxx}

Variations in the results were significant. Estimated reach of the coverage varied between 46 and 93 million reader and while the proportion of positive coverage based on a tonality analysis varied between 17 percent and 100 per cent for the 138 articles included in the analysis.^{xxxii}

The basic validity of the measures COI requested appears to be quite clear. Reach (readership), cost per thousand reached (CPM), opportunities to see, and volume of coverage are common measures used in public relations and COI has been quite rigorous in defining these measures^{xxxii}. However, the reliability of these measures, or the ability of these measures to be independently replicated indicates that standardization of public relations measures requires significantly more than a description of the measure to be included in the analysis, but the implementation of specific research procedures and protocols that will be applied uniformly and consistently. Hence, the measures of validity, as established by advanced statistical testing, may actually bring validity of these measures by audience in question.

Recommendations for Specific Measures for Public Relations Activities
Public relations measurement falls into two major categories. The first is measuring the impact of communication activities on the target audience. The second is measuring the delivery of those messages through third parties or intermediaries such as the media.

TARGET AUDIENCE MEASURES

Using the communication lifecycle as the starting point, the following are specific measures for determining the effectiveness of public relations activities at each stage of the lifecycle among a target audience for the reception of public relations messages. These standard measures are recommended because over the course of numerous studies, they have been shown to be valid in that they actually measure what is intended, as well as reliable in that they produce replicable findings^{xxxiii}.

These measures, however, make two presumptions:

- Appropriate forms of each measure be used consistently
- Appropriate data collection methods based on best practices are employed

Awareness

Awareness measures can take several different forms. These measures include unaided awareness and aided awareness. The most fundamental measure for public relations is a variation of an awareness measure known as *recall*. This measure has its foundation in “day after recall”^{xxxiv} testing that measured if the viewer or reader had any “related” or correct recall of the message elements included in the communication. The most basic level of “related recall” is recall or recollection of the name of the product, service or concept included in the communication being tested. Depending on the data collection method used, the following questions are prototypes for measuring unaided recall (see Table 1).

**Table 1
Recall Measures**

Data Collection Method	Prototype Question	Response Categories
Interviewer Administered^{xxxv} (unaided)	Thinking back to what you have just (<i>read/ observed/ reviewed/ saw</i>), tell me the (<i>brands/ products/services/issues/ topics</i>) that you remember (<i>reading/ observing/ reviewing/ seeing</i>).	Open ended responses with prelist of likely responses and an open response field
Self-Administered^{xxxvi} (unaided)	Thinking back to what you have just (<i>read/ observed/ reviewed/ saw</i>), place an X in the boxes for the (<i>brands/ products/services/ issues/ topics</i>) that you remember (<i>reading/ observing/ reviewing/ seeing</i>).	Open response field
Interviewer Administered (aided)	Thinking back to what you have just (<i>read/ observed/ reviewed/ saw</i>), tell me if you remember (<i>reading/ observing/ reviewing/ seeing</i>) about any of the following (<i>brands/ products/services/issues/ topics</i>)	List of brands, products, services, issues or topics that are or could have been included in the communication. These are typically presented in a random order.
Self-Administered (aided)	Thinking back to what you have just (<i>read/ observed/ reviewed/ saw</i>), place an X in the boxes if you remember (<i>reading/ observing/ reviewing/ seeing</i>) about any of the following (<i>brands/ products/services/issues/ topics</i>).	List of brands, products, services, issues or topics that are or could have been included in the communication. These are typically presented in a random order.

Similar question structures are also employed for measuring aided recall. In most studies, these measures are used in combination to determine over levels or depth of recall, with unaided recall being the more sensitive and therefore more important measure.

Knowledge

The most basic and fundamental challenge in assuring the effectiveness of public relations is exposure of key messages about the brand, product, issue, or topic to the target audience (see Table 2). Many of these key messages are basic facts about the brand, product, issue, or topic that serves as the essential level of knowledge that is critical for a target audience to understand. Levels of agreement with statements that present factual knowledge is a highly effective tool that determines if exposure to the messages occurred and if there is initial acceptance of the messages. Knowledge testing can be supplemented with a credibility measure that determines if the overall story about the brand, product, service, topic or issue is believable.

**Table 2
Knowledge Measures**

Data Collection Method	Prototype Question	Response Categories
Interviewer Administered	Next, I am going to read you a series of statements about a (<i>brand/ product/ issue/ service/topic</i>). That (<i>brand/product/ service/ issue/topic</i>) is a (<i>insert category</i>) called (<i>insert name</i>). After I read you each statement, please indicate if you “strongly agree,” “somewhat agree,” “neither agree nor disagree,” “somewhat disagree,” or “strongly disagree,” with each statement about (<i>insert name</i>).	List of attributes that describe the brand, product, services, issues or topics that are or should have been included in the communication. These attributes are typically read to respondents in a random sequence.
Self-Administered	Next, you are going to read a series of statements about a (<i>brand/ product/service/ issue/ topic</i>). That (<i>brand/product/service/ issue/topic</i>) is a (<i>insert category</i>) called (<i>insert name</i>). After you read each statement, please indicate if you “strongly agree,” “somewhat agree,”	List of attributes that describe the brand, product, service, issues or topic that are or should have been included in the communication. These attributes are typically presented to respondents in a random sequence if an online

	“neither agree nor disagree,” “somewhat disagree,” or “strongly disagree,” with each statement about <i>(insert name)</i> .	survey method is used. Answer categories are shown with each statement.
Interviewer or Self-Administered	Based on everything you have read, how believable is the information you just saw about the <i>(brand/product/service/ issue/ topic)</i> ? By believable we mean that you are confident that what you are <i>(seeing/reading/ hearing/observing)</i> is truthful and credible.	The response categories for this question are typically a scale that measures an overall level of credibility or believability. One of the most common and reliable scales consists of five points ranging from “very believable” to “very unbelievable” with a neutral midpoint ^{xxxvii}

INTEREST AND RELEVANCE

These measures constitute direct questions about interest in the brand, product, service, issue, or topic as well as broader measures that examine how they are perceived by the target audience. When the target audience is closely aligned with the brand, product, service, issue, or topic that is the subject of communication, there is an increased likelihood that they will take an intended action to purchase, support, or recommend. Without interest and relevance there is little or more *motivation*^{xxxviii} by the target audience to take any form or action that is aligned with business or program objectives.

The basic question on interest is an overall or global question on interest in the brand, product, service, issue, or topic. This question is asked on a measurement *scale* to determine an overall intensity of interest (see Table 3). This question can also serve as the “dependent variable”^{xxxix} in an analysis that predicts outcomes. This is commonly called a regression or leverage analysis.

Table 3
Measures of Interest

Data Collection Method	Prototype Question	Response Categories
Interviewer or Self-Administered	After (<i>seeing/reading/hearing/observing</i>) this material would you say you are “very interested”, “somewhat interested”, “neither interested nor uninterested,” “somewhat uninterested” or “very uninterested” in this (<i>brand/product/service/ issue/ topic</i>)?	The response categories for this question are typically a scale that measures an overall level of interest. One of the most common and reliable scales consists of five points ranging from “very interested” to “very uninterested” with a neutral midpoint. The scale is similar to that used in the credibility or believability measure described in Table 1

Supplementing this overall or global question is a series of statements that measure the *relationship* that the target audience has with the brand, product, service, or issue (see Table 4). These statements gauge the degree to which the brand, product, service, or issue is seen to be relevant to or *homophilous* with the needs and interests of the target. Homophily, defined as the state in which a person shares the same values, ideas, beliefs, and so forth as the person with whom they are interacting, is often a key measure that is overlooked in communication research. However it is often a central factor in determining the social acceptability of specific actions or purchases.^{x1}

Typical statements that are included in this measure include:

- This product is a value for its price
- The product has been presented honestly
- Based on what I know of it, this product is very good
- This product is something that is like me
- Based on what I know of it, this product is an excellent choice for me
- Based on what I know of it, I find this product quite pleasant to use
- This product is used by people in my economic class
- I think the product is very consumer unfriendly
- People who buy this product are very much like me
- I think this product is very reliable
- This product reflects my social background
- I would purchase this product because it reflects my lifestyle
- This product is awful
- People who use this product are culturally similar to me

Items are adapted as needed for the specific category of product, service, issue or topic in question and are administered using the following question structure.

Table 4
Measures of Relationship

Data Collection Method	Prototype Question	Response Categories
Interviewer Administered	<p>I am going to read you a series of statements about the <i>(brand/product/service/ issue/ topic)</i>. There are no right or wrong answers, we are interested in how much you agree or disagree with the statements. Do you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree or strongly disagree?</p> <p>Place an X in the box that best represents your answer for each statement.</p>	<p>The response categories for this question are typically a scale that measures an overall level of agreement. One of the most common and reliable scales consists of five points ranging from “strongly agree” to “strongly disagree” with a neutral midpoint. The scale is similar to that used in the interest measure described above.</p>
Self-Administered	<p>Please respond to the following statements about the <i>(brand/product/service/ issue/ topic)</i>. There are no right or wrong answers, we are interested in how much you agree or disagree with the statements.</p> <p>Place an X in the box that best represents your answer for each statement.</p>	<p>The response categories for this question are typically a scale that measures an overall level of agreement. One of the most common and reliable scales consists of five points ranging from “strongly agree” to “strongly disagree” with a neutral midpoint. The scale is similar to that used in the interest measure described above.</p>

INTENT

Intent covers a broad range of measures. It is an attitudinal measure not behavioral and typically includes preference for a brand, product, service, issue, or topic, as well as intent to take a specific action (see Table 5). These actions can include purchase of a product service or brand, support for an idea or concept, willingness to try a product or service or to make an inquiry.

The questions used to measure intent start with preference. In most instances, a preference measure determines the choice of a single brand, product or service to the exclusion of others. The following is the recommended structure for that question.

**Table 5
Measures of Preference**

Data Collection Method	Prototype Question	Response Categories
Interviewer Administered	I am going to read you a list of different (brands, products, services) that you can buy at your local store follows. Which <u>one</u> of these (brands, products, services) do you prefer most?	List of brands, products, services, issues or topics that are or could have been included in the communication. These are typically presented in a random order.
Self-Administered	A list of different (brands, products, services) that you can buy at your local store follows. Which <u>one</u> of these (brands, products, services) do you prefer most? Place an X in the box that best represents your answer	List of brands, products, services, issues or topics that are or could have been included in the communication. These are typically presented in a random order.

Intent to take a specified action, however, differs considerably from overall preference (see Table 6). Members of a target audience may prefer one brand, product or service over others. But, in many instances, this preference does not convert into a likely action. For example, a consumer may prefer one brand of snack chips over another. However, that same consumer may be unlikely to purchase that preferred brand because of price, availability or other product attributes. This question is asked on a scale to measure intensity of the intent to take an action. The question can be asked for multiple brands, products, services, issues or topics in order to gain an understanding of comparative intent.

Table 6
Measures of Specified Action

Data Collection Method	Prototype Question	Response Categories
Interviewer or Self-Administered	Based on everything you have (<i>seen/read/ heard/observed</i>) about this (<i>brand, product, service, issue, topic</i>), how likely are to (<i>purchase/try/support</i>) this (<i>brand, product, service, issue, topic</i>). Would you say you are “very likely”, “somewhat likely”, “neither likely nor unlikely,” “somewhat unlikely” or “very unlikely” to (<i>purchase/try/support</i>) this (<i>brand/product/service/ issue/ topic</i>)?	The response categories for this question are typically a scale that measures an overall level of intent to take a specific action. One of the most common and reliable scales consists of five points ranging from “very likely” to “very unlikely” with a neutral midpoint. The scale is similar to that used in the credibility or believability measure described in Table 1

This question is typically supplemented with an open ended or free response question that asks respondents to explain why they hold this attitude.

ADVOCACY

As noted earlier in the discussion of applying the Lewis model to public relations, public relations often differs from other forms of marketing communication because the end result of a communication program is not necessarily the sale of products or services . One of the key measures for the success of public relations programs is the ability of a program to create advocates among the target audience for a brand, product, service, issue or topic.

A recent study by The Nielsen Company reinforces this axiomatic belief in the value of advocacy.^{xii} In that study, 90 percent of respondents cite recommendations from a person known to them as information they “completely” or “somewhat trust.” According to this study, a recommendation from a known person is the most trusted source of information on products and services. Consumer opinions posted online is the source that is ranked second in trust with 70 percent of respondents saying they “completely” or “somewhat trust” this source.

The questions used for measuring advocacy are similar in structure to those used to measure relevance (see Table 7).

The items used to measure advocacy include:

- I will recommend this (brand, product, service, issue, topic) to my friends and relatives
- People like me can benefit from this (brand, product, service, issue, topic)
- I like to tell people about (brands, products, services, issues, topics) that work well for me
- Word-of-mouth is the best way to learn about (brands, products, services, issues, topics)
- User reviews on websites are valuable sources of information about (brands, products, services, issues, topics)

Like the homophily and relevance measures, these items are adapted as needed for the specific category of product, service, issue or topic in question and are administered using the following question structure.

Table 7
Measures of Advocacy

Data Collection Method	Prototype Question	Response Categories
Interviewer Administered	<p>I am going to read you a series of statements about the <i>(brand/product/service/ issue/ topic)</i>. There are no right or wrong answers, we are interested in how much you agree or disagree with the statements. Do you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree or strongly disagree?</p> <p>Place an X in the box that best represents your answer for each statement.</p>	<p>The response categories for this question are typically a scale that measures an overall level of agreement. One of the most common and reliable scales consists of five points ranging from “strongly agree” to “strongly disagree” with a neutral midpoint. The scale is similar to that used in the interest measure described above.</p>
Self-Administered	<p>Please respond to the following statements about the <i>(brand/product/service/ issue/ topic)</i>. There are no right or wrong answers, we are interested in how much you agree or disagree with the statements.</p>	<p>The response categories for this question are typically a scale that measures an overall level of agreement. One of the most common and reliable scales consists of five points ranging from “strongly agree” to</p>

Place an X in the box that best represents your answer for each statement.

“strongly disagree” with a neutral midpoint. The scale is similar to that used in the interest measure described above.

THIRD PARTY/INTERMEDIARY MEASURES

There are a myriad of commercially available systems for the measurement of public relations placement activities. These systems predominantly employ content analysis using manual as well as automated coding and typically include a series of commonly used measures including volume and sources of coverage, audience reach, and presence of key messages. Tonality is also a common measure in many of these systems. These are, for the most part, important measures of media relations activities. However, as Michaelson and Griffin pointed out in, “A New Model for Media Content Analysis”:

As diverse as these methods of content analysis appear, each contains two commonly held fatal flaws. The first flaw is the absence of a basic analytic structure that determines the accuracy of coverage overall and more specifically determines the accuracy of specific messages included in the content of articles under analysis. The second flaw is the failure to link analysis to communications [sic] objectives and public relations messages.^{xliii}

The most appropriate measure of third party or intermediary outcomes of public relations activities is determining the accuracy of key messages in light of the fact that “The initial recipient is the media, who, in turn, serves as the conduit for transmitting messages to intended recipient (i.e. their readers).”^{xliiii}

Message accuracy in this regard involves three specific measures:

- The *presence of basic facts* in the third party or intermediary story/message
- The *presence of misstatements* or erroneous information
- The *absence or omission of basic facts* that should be included in a complete story

The basic facts at the core of this analysis structure are driven by the communication objectives of the public relations program in conjunction with its lifecycle stage. For example, if based on the understanding of the target audience levels of information, the communications objective of the program is to generate awareness of a brand or product, then the message most essential to communicate through third parties or intermediaries is the name of the brand or product. Content is analyzed to determine if that message is present in the third party story, absent in the story, or appears in an erroneous form.

Unlike research among a target audience to measure if a message has been received, there is no formal question structure to be applied in this analysis. The analysis is based on predefined coding lists that are tied to the communications objectives of the program. Two specific lists are developed. One list contains the messages that should appear in the third party or intermediary article or story. The second list contains messages that are known to be erroneous or incomplete and have appeared or are likely to appear in these articles or stories.

Coding of messages falls into three categories:

- Presence of key messages
- Presence of erroneous messages
- Absence of key messages

The code list for the presence and absence of key messages is identical.

Appropriate Data Collection Methods – Applying Best Practices

Applying these standard measures to public relations activities will create a basis for comparability from program-to-program and from category-to-category. As the use of these standards become widespread, communicators will have access to normative data that can be used to gauge both relative and absolute performance of their communications efforts. However, these standards will only achieve the objective of establishing a basis for comparative evaluation if proper research protocols are employed. These research protocols include using rigorous study design, reliable and replicable data collection methods and appropriate analytic and statistical tools. Without applying these best research practices, the benefits of using standard measures cannot be achieved.

Much has been written on these research protocols and best practices. Consequently, we will not discuss these practices in detail. Readers should review two specific volumes for detailed discussions on these best practices. These volumes are “The Primer of Public Relations Research, 2nd edition”^{xliv} and “A Practitioners Guide to Public Relations Research, Measurement, and Evaluation”^{xlv}.

CONCLUSION

For the past several decades, public relations practitioners have been seeking the “holy grail” of measurement for public relations activities. We have seen the rise of questionable research methods and the applications of “black box” approaches. However, it is quite clear that the approach to developing effective public relations measurement includes several key elements – the acceptance of measures that are valid for determining the impact of public relations and research methods that will produce reliable and replicable results. We must meld *what* we need to know (standards) with the best approaches of *how* to collect this data (best practices). In order to succeed, the industry has to go beyond soft guidelines and accept specific measures that will be universally applied.

A final parting comment is necessary. In most industry discussions of standards, the focus is not on developing them, but rather in the best practices concept. What we find missing in any discussion, however, is a specific outcome that is truly public relations. We believe that this outcome is engagement, the ability of the public relations professional to use the adopted AIDA model and variations of the model (e.g., Michaelson and Griffin) for defining engagement and the best messages to employ when initiating, maintaining, and solidifying that engagement. As noted elsewhere, public relations' function in relation to ROI is to establish brand, product, service, or issue engagement outcomes dealing with credibility, relationship, reputation, and trust target audience perceptions of that brand, product, service, *or issue*.

Endnotes

ⁱ Advertising and direct marketing

ⁱⁱ Michaelson, David "The Application of Best Practices in Public Relations Measurement and Evaluation Systems", *Public Relations Journal*, Volume 1, Number 1; October 2007 page 1

ⁱⁱⁱ Since 1997, there have been several attempts to develop standardized measures for public relations activities. This includes work by Walter K. Lindenmann ("Guidelines and Standards for Measuring the Effectiveness of PR Programs and Activities" -- http://www.instituteforpr.org/wp-content/uploads/2002_MeasuringPrograms.pdf), The Central Office of Information (UK) ("Standardisation Of PR Evaluation Metrics" -- <http://coi.gov.uk/documents/guidance/pr-evaluation-metrics.pdf>) and the Association of Measurement and Evaluation Companies ("Valid Metrics for PR Measurement" -- <http://ameceuropeansummit.org/amecorgdocs/ValidMetricsFramework7June2011PrintVersion.pdf>).

Each of these models presents general guidelines for measurement but do not offer specific processes or procedures for measuring the effectiveness and impact of public relations.

^{iv} Survey of 141 public relations professionals conducted online from May 5 to June 1, 2011. The survey was conducted by the Association of Measurement and Evaluation Companies (AMEC) among attendees (80 respondents) of the 3rd European Summit on Measurement (June 2011) and other public relations practitioners (61 respondents). 48% of respondents were from public relations agencies and 12% were from research companies. 35% of respondents were from Europe, 38% from the UK and 14% from the Americas. Full details of the survey are available from AMEC (www.amecorg.com)

^v "Output — what is generated as a result of a PR program or campaign that impacts on a target audience or public to act or behave in some way ... the final stage of a communication product, production, or process resulting in the production and dissemination of a communication product (brochure, media release, Web site, speech, etc.); s. the number of communication products or services resulting from a communication production process; the number distributed and/or the number reaching a targeted audience." (Don W. Stacks (ed.) "Dictionary of Public Relations Measurement and Research", Institute for Public Relations, 2006, page 14)

^{vi} "Outtake — measurement of what audiences have understood and/or heeded and/or responded to a communication product's call to seek further information from PR messages prior to measuring an outcome; audience reaction to the receipt of a communication product, including favorability of the product, recall and retention of the message embedded in the product, and whether the audience heeded or responded to a call for information or action within the message." ("Dictionary of Public Relations Measurement and Research", page 15)

^{vii} In June 2011, 89 percent of attendees at the AMEC European Summit on Measurement in Barcelona voted 'how to measure the ROI of public relations' is an important goal to be achieved. (SOURCE: <http://www.holmesreport.com/news-info/10490/Education-Social-Media-Measurement-And-ROI-Top-Priorities-At-AMEC-Summit.aspx>)

^{viii} Recent industry initiatives have attempted to minimize the use of advertising value equivalence measures. The most prominent of these initiatives has been the "Barcelona Principles" that were adopted at the 2nd Annual European Summit on Public Relations Measurement in 2010. Those principles stated that "AVEs were not the value of public relations." Further information on these principles can be found on the Institute for Public Relations website (www.instituteforpr.org) or on the Association of Measurement and Evaluation Companies (AMEC) website (www.amecorg.com).

^{ix} Stacks, Don W, and Michaelson David; A Practitioners Guide to Public Relations Research, Measurement, and Evaluation; New York: Business Expert Press, 2010

^x <http://oxforddictionaries.com/definition/standard?region=us>

^{xi} Absolute performance is an independent measure that determines how well a program performs on key measures. These measures include outtake measures such as awareness, knowledge and purchase intent as well as outcome measures that include number of placements and presence of key messages.

^{xii} Relative measures are comparisons with programs or campaigns that are outside the influence of a specific campaign or program. These measures can be comparisons with competitors, other industries or even established performance measures across industries, categories or competitors.

^{xiii} Stacks, Don W.; Primer of Public Relations Research; New York: Guilford Press, 2011

^{xiv} Michaelson, David "The Application of Best Practices in Public Relations Measurement and Evaluation Systems", Public Relations Journal, Volume 1, Number 1; October 2007, page 1

^{xv} For a detailed understanding of best practices in public relations research, refer to Michaelson, David, "The Application of Best Practices."

^{xvi} <http://www.businessdictionary.com/definition/best-practice.html>

^{xvii} SOURCE: "The AIDA Sales Funnel", <http://www.provenmodels.com/547>

^{xviii} SOURCE: "The AIDA Sales Funnel", <http://www.provenmodels.com/547>

^{xix} Michaelson, David and Griffin, Toni; "New Model for Media Content Analysis"; Institute for Public Relations, 2005, Page 7 (<http://www.instituteforpr.org/wp-content/uploads/MediaContentAnalysis.pdf>)

^{xx} Michaelson, David and Griffin, Toni, 2005, page 7

^{xxi} Michaelson, David and Griffin, Toni, 2005, page 7

^{xxii} Michaelson, David and Griffin, Toni, 2005, pages 7 - 8

^{xxiii} Stacks, Don and Michaelson, David.

^{xxiv} Stacks, Don & Michaelson, David, "Exploring the Comparative."

^{xxv} Michaelson, David and Griffin, Toni, 2005, page 8

^{xxvi} "Validity – The extent to which a research project actually measures what it intended, or purports to measure." ("Dictionary of Public Relations Measurement and Research", page 23)

^{xxvii} "Reliability – the extent to which results would be consistent, or replicable, if the research were conducted a number of times." ("Dictionary of Public Relations Measurement and Research", page 18)

^{xxviii} It should be noted that many measures employed in public relations are adapted from scales created by academics and many are available at no cost.

^{xxix} Echo Research, Metrica, Millward Brown Precise, Panarc International and TNS Presswatch

^{xxx} “Standardisation Of PR Evaluation Metrics”, Central Office of Information, 2009, page 1

^{xxxi} Source: PR WEEK, June 2009

<http://www.prweek.com/uk/channel/PublicSector/article/915710/Central%20Office%20of%20Information%20turns%20spotlight%20on%20evaluation/>

^{xxxii} Central Office of Information, “Standardisation Of PR Evaluation Metrics”, 2009 Pages 3-4

^{xxxiii} These measures should be understood only as recommended question structures that should be adapted for the specific study being conducted.

^{xxxiv} An advertising research test to see how much someone can remember of an advertisement the day after it appeared or was broadcast. This test is used rarely today as an exclusive measure of communications performance.

^{xxxv} Interviewer administered studies include telephone surveys, in person surveys and intercept studies where a trained research asks questions of a respondent and records their responses.

^{xxxvi} Self-administered studies include online studies and any other type of study where the respondent records their own answers to questions. The recording medium can include paper questionnaires as well as computer based recording.

^{xxxvii} This type of scale is often referred to as a Likert scale. The scale was developed by Rensis Likert at the University of Michigan Institute for Social Research. The scale is noted for its high degree of reliability in survey research.

^{xxxviii} There are three objectives that all public relations campaigns or initiatives must take into account in order: informational objectives (did the target audience receive, recall, and understand the message), motivational objectives (did the messages motivate the target audience toward action), and behavioral objectives (how does the target audience plan on behaving (adopting, recommending, etc.); see Stacks, Don, Primer, chapter 2, and Stacks, Don, and Michaelson, David, A Practitioner’s Guide, chapter 2.

^{xxxix} The dependent variable is the variable that is being measured in an experiment

^{xl} For a more complete discussion on homophily, see Stacks, Don, and Michaelson, David, A Practitioners Guide; and Stacks, Don, and Michaelson, David, “Examining the Comparative.

^{xli} Nielsen Global Online Consumer Survey of over 25,000 Internet consumers from 50 countries, conducted April 2009. <http://blog.nielsen.com/nielsenwire/consumer/global-advertising-consumers-trust-real-friends-and-virtual-strangers-the-most/>

^{xlii} Michaelson, David and Griffin, Toni, page 6

^{xliii} Michaelson, David and Griffin, Toni, page 7

^{xliv} Stacks, Don W.; Primer of Public Relations Research

^{xlv} Stacks, Don W, and Michaelson David, A Practitioners Guide to Public Relations Research, Measurement, and. Evaluation.