THE INSTITUTE FOR PUBLIC RELATIONS
COMMISSION ON PR MEASUREMENT AND EVALUATION
University of Florida * PO Box 118400 * Gainesville, FL 32611-8400
(352) 392-0280 * (352) 846-1122 (fax)
www.instituteforpr.com

GUIDELINES FOR MEASURING THE EFFECTIVENESS OF PR PROGRAMS AND ACTIVITIES

This booklet was first published in 1997 under the title, "Guidelines and Standards for Measuring and Evaluating PR Effectiveness." It was originally written by Dr. Walter K. Lindenmann, based on guidance, input and suggestions from a task force of PR practitioners, counselors, academicians and research suppliers that included the following individuals: Forrest W. Anderson … Albert J. Barr … Dr. Mary Ann Ferguson … Dr. James E. Grunig … Thomas Martin … Geri Mazur … Willard Nielsen … Charlotte Otto … Katharine D. Paine … David Silver … Kathleen Ward … Mark Weiner … and Dr. Donald K. Wright.

The booklet was updated and revised in 2002 and given a new title, "Guidelines for Measuring the Effectiveness of PR Programs and Activities," to more accurately reflect its contents. The revised version is primarily the work of Dr. Walter K. Lindenmann, with input and suggestions from Fraser Likely.

CONTENTS

FOREWORD

For years we have been told that we can never expect to get proper credit for what we do in public relations until we can find an effective way to measure our effectiveness.

Most other professions have recognized measuring tools -- engineering devices, chemical reactions, case law, charts and figures. But public relations efforts have always been gauged in a variety of ways -- each using a different kind of measuring stick.

In an attempt to begin to find a uniform "ruler" we can all use for measurement, a Public Relations Evaluation Summit was called in October, 1996 in New York City. This gathering of top leaders interested in public relations research was sponsored by the Institute for Public Relations, *INSIDE PR*, and the Ketchum Public Relations Research and Measurement Department.

As a result of that Summit, a booklet was published and distributed in 1997 under the title, "*Guidelines and Standards for Measuring and Evaluating PR Effectiveness"* as a first attempt to establish guidelines for how we might begin to agree on uniform ways to measure public relations by using the same measuring sticks. In the light of new developments relating to PR measurement overall, this booklet was revised in 2002 and given a new title, *"Guidelines for Measuring the Effectiveness of PR Programs and Activities,"* to more accurately reflect its contents. We view the revised version of this book as a companion to *"Guidelines for Measuring Relationships in Public Relations,"* and *"Guidelines for Setting Measurable Objectives,"* both of which were published in 1999.

We believe you can use the ideas and suggestions in this booklet as a working document of ways we can continue a dialogue on measuring the effectiveness of public relations.

Jack Felton
President & CEO

OVERVIEW

What is public relations measurement and evaluation?

Basically, it is any and all research designed to determine the relative effectiveness or value of what is done in public relations. In the short-term, PR measurement and evaluation involves assessing the success or failure of specific PR programs, strategies, activities or tactics by measuring the *outputs, outtakes and/or outcomes* of those programs against a predetermined set of objectives. In the long-term, PR measurement and evaluation involves assessing the success or failure of much broader PR efforts that have as their aim seeking to improve and enhance the *relationships* that organizations maintain with key constituents

More specifically, *PR measurement* is a way of giving a result a precise dimension, generally by comparison to some standard or baseline and usually is done in a quantifiable or numerical manner. That is, when we measure *outputs, outtakes* and *outcomes*, we usually come up with a precise measure -- a number; for example, 1,000 brochures distributed … 60,000 hits on a website … 50% message recall … an 80% increase in awareness levels, etc.

PR evaluation determines the *value* or importance of a PR program or effort, usually through appraisal or comparison with a predetermined set of organization goals and objectives. *PR evaluation* is somewhat more subjective in nature, or softer, than *PR measurement*, involving a greater amount of interpretation and judgment calls.

Interest in public relations measurement and evaluation has surged in recent years, as the public relations field has grown in size and sophistication, and as those who practice in the field have found themselves more often than ever being asked to be accountable for what they do.

Those who supervise or manage an organization's total communications activities are increasingly asking themselves, their staff members, their agencies and consulting firms, and their research suppliers questions such as these:

-- Will those public relations and/or advertising efforts that we initiate actually have an effect -- that is, "move the needle" in the right direction -- and, if so, how can we support and document that from a research perspective?

-- Will the communications activities we implement actually change what people know, what they think and feel, and how they actually act?

-- What impact -- if any -- will various public relations, marketing communications, and advertising activities have in changing consumer and opinion-leader awareness, understanding, retention, attitude and behavior levels?

As questions such as these have increased in number in recent years, many public relations practitioners -- as they seek to justify what they, themselves, do -- have sensed a need to establish guidelines or criteria that the industry can follow, when it comes specifically to public relations measurement and evaluation.

This guidebook, which has been revised and edited under the auspices of the Institute for Public Relations Commission on PR Measurement and Evaluation, seeks to set minimum standards when it comes to measuring and evaluating the effectiveness of specific short-term PR programs, strategies, activities and tactics against pre-determined *outputs*, *outtakes* and *outcomes.* Those interested in measuring and evaluating the effectiveness of PR efforts aimed at enhancing the long-term *relationships* that exist between an organization and its key constituents should consult the companion guidebook, *"Guidelines for Measuring Relationships in Public Relations."* (www.instituteforpr.com)

SOME GUIDING PRINCIPLES

In focusing on PR measurement and evaluation, here are some guiding principles or key factors to consider at the outset. These guiding principles are discussed in more detail in the main sections of this booklet.

- Establish clear program, activity, strategic and tactical objectives and desired outputs, outtakes and outcomes before you begin, to provide a basis for measurement of results. PR goals should tie directly to the overall goals of the organization.

- Differentiate between measuring PR *outputs,* which are usually short-term and surface (e.g. the amount of press coverage received or exposure of a particular message), PR *outtakes,* which are usually more far-reaching and can have more impact (e.g. determining if those to whom the activity was directed *received, paid attention to, comprehended and retained* particular messages) and PR *outcomes,* (e.g. did the program or activity change opinion and attitude levels, and possibly behavior patterns?).

- Measuring media content, while of great value, needs to be viewed as only a first step in the PR measurement and evaluation process. It can measure possible exposure to PR messages and actual press coverage; however, it cannot, by itself, measure whether target audiences actually saw the messages and responded to them in any way.

- There is no one, simple, all-encompassing research tool, technique or methodology that can be relied on to measure and evaluate PR effectiveness. Usually, a combination of different measurement techniques are needed. Consideration should be given to any one or several of the following: media content analysis ... cyberspace analysis ... trade show and event measurement ... polls and surveys ... focus groups ... experimental and quasi-experimental designs ... and/or ethnographic studies that rely on observation, participation and/or role playing techniques.

- Be wary of attempts to precisely compare PR effectiveness to advertising effectiveness. The two forms of communication are quite different from each other and the fact that placement of advertising messages can be controlled, whereas placement of PR messages usually cannot be controlled, needs to be taken into consideration.

- PR effectiveness can best be measured if an organization's principal messages, key target audience groups, and desired channels of communication are clearly identified and understood in advance.

- The PR measurement and evaluation process should never be carried out in isolation, by focusing only on the PR components. Wherever and whenever possible, it is always important to link what is planned, and accomplished, through

PR, to the overall goals, objectives, strategies and tactics of the organization as a whole.

MAJOR PR MEASUREMENT AND EVALUATION COMPONENTS

For any PR evaluation research to be credible, five major components of the process need to be taken into consideration. They are:

1. ***Setting Specific Measurable PR Goals and Objectives***

This has to come first. No one can really measure the effectiveness of anything, unless they first figure out exactly what it is they are measuring that something against. So, to begin, the public relations practitioner, counselor and/or research supplier ought to ask: What are or were the goals or objectives of the specific public relations program, activity, strategy or tactic? What exactly did the program or the activities hope to accomplish -- through its public relations component?

This is not always easy to do, since it is often difficult to separate public relations programs and activities (such as publicity efforts, distribution of informational materials, the holding of special events or shows, etc.) from marketing communications (point-of-purchase promotional activities, coupon redemption programs, special contests and give-away activities, etc.) and from advertising (paid print and broadcast messages, cyberspace commercials, etc.)

In setting PR goals and objectives, it is usually important to recognize that measuring PR effectiveness per se -- that is, the management of an organization's overall communications activities with its target audience groups or publics -- can be quite difficult to do unless the individual elements or components of the program are clearly defined. We suggest that instead of trying to measure PR as a total entity, steps be taken to measure the effectiveness of individual or particular PR activities, such as measuring the effectiveness of specific publicity efforts, or a particular community relations program, or a special event or trade show activity, or a government affairs or lobbying effort, or a speaker's program, or an investor relations activity, and so on.

Additional ideas and suggestions pertaining to the setting of measurable PR goals and objectives can be obtained in the IPR Commission on PR Measurement and Evaluation guidebook, *"Guidelines for Setting Measurable Public Relations Objectives."* (www.instituteforpr.com)

2. ***Measuring PR Outputs***

Outputs are usually the immediate results of a particular PR program or activity. More often than not, *outputs* represent what is readily apparent to the eye. *Outputs* measure how well an organization presents itself to others, the amount of exposure that the organization receives.

In media or press relations efforts, *outputs* can be the total number of stories, articles, or "placements" that appear in the media ... the total number of "impressions" -- that is, the number of those who might have had the opportunity to be exposed to the story ... as well as an assessment of the overall content of what has appeared. Media Content Analysis (see Page 9) is one of the principal methodologies used to measure media *outputs*.

For other facets of public relations, *outputs* can be white papers, speaking engagements, the number of times a spokesperson is quoted, specific messages communicated, or specific positioning on an important issue or any number of quantifiable items that are generated as a result of the effort. *Outputs* also might be assessment of a specific event, a direct mail campaign, the number of people who participated in a given activity, how a CEO handles himself or herself at a press conference, or the appearance and contents of a given brochure or booklet.

In any event, both the quantity and quality of *outputs* can be measured and evaluated. Media can be evaluated for their content; an event, as to whether the right people were there; a booklet or brochure for its visual appeal and substance; and so on.

3. ***<u>Measuring PR Outtakes</u>***

Although it is obviously important to measure how well an organization presents itself to others and the amount of exposure obtained, it is even more important to measure PR *outtakes* -- that is, determining if key target audience groups actually *received* the messages directed at them, *paid attention to them*, *understood* and/or *comprehended* the messages, and whether they *retained* the messages and can *recall* them in any shape or form.

When a PR program is launched or when given PR activities or events are initiated -- such as the distribution of a brochure or a booklet, the placement of announcements on websites, or the delivering of a speech -- it is important to assess what, if anything, did the intended recipients "take-away" from this effort.

The first unit of *outtake* measurement could very well be that of favorability. Was the PR program or effort favorably received? Were the creative design elements or "packaging" received favorably? Was the "language" received favorably? Was the "ease of use" of the PR effort favorably received?

The second unit of *outtake* measurement relates to understanding and comprehension. Did the messages that were being disseminated make sense to the intended recipients? Were those to whom the messages were targeted able to decipher them and put them into appropriate context?

The third unit of measurement at the *outtake* level is message recall and retention. It measures whether the messages we craft for inclusion in our brochures, booklets and related PR programs and activities make enough of an impression on the intended recipients, that they become memorable. Can the intended recipients recall the embedded messages and can they retain them for any length of time?

The final unit of measurement at the *outtake* level is that of attention and possible immediate response. Did the receiver respond positively to the receipt of the messages? Did he or she do something with the information now in hand, for example, by passing on materials or messages to friends or colleagues? Did the recipient request more information, for example, by going to a website?

It is possible to compare the *outtake* measures of one particular PR program or activity to one or more others.

4. *Measuring PR Outcomes*

As important as it might be to measure PR *outputs* and *outtakes,* it is far more important to measure PR *outcomes*.

These measure whether the communications materials and messages which were disseminated have resulted in any *opinion, attitude* and/or *behavior* changes on the part of those targeted audiences to whom the messages were directed.

It is usually much more difficult and, generally, more expensive, to measure *PR outcomes,* and to some extent *PR outtakes,* than it is to measure *PR outputs*. This is because more sophisticated data-gathering research tools and techniques are required. Measuring *PR outputs* is usually a question of counting, tracking and observing, while for *PR outtakes* and *PR outcomes,* it is a matter of asking and carrying out extensive review and analysis of what was said and what was done.

Research techniques often used to measure *PR outtakes* and *PR outcomes* include quantitative surveys (in-person, by telephone, by mail, via fax, via e-mail, via the Internet, in malls, etc.) ... focus groups ... qualitative depth attitude surveys of elite audience groups ... pre-test/post-test studies (e.g. before-and-after polls) ... ethnographic studies (relying on observation, participation, and/or role-playing techniques) ... experimental and quasi-experimental research projects ... and multi-variate studies that rely on advanced statistical applications such as correlation and regression analyses, Q-sorts, and factor and cluster analysis studies.

5. *Measuring Business and/or Organizational Outcomes*

Whatever steps PR practitioners take to measure the effectiveness of what they, themselves, do in PR, it is imperative that they also take steps to seek to link their public relations accomplishments to the ultimate goals, objectives, and accomplishments of the organization as a whole.

What we are talking about here is seeking to relate PR *outcomes* to such desired business and/or organizational *outcomes* as increasing market penetration, market share, sales, and, ultimately, increasing an organization's profitability. It needs to be recognized that this is not easy to do. It requires a careful delineation of what the PR program seeks to accomplish in concert with what the organization as a whole seeks to accomplish. It also requires a good understanding about how and why the two processes are supposed to work together. When one has a good understanding of the impacts that are desired, as well as a good understanding of how the process is supposed to work, there are then many research design tools that can be employed to reliably and validly measure that impact.

For example, the subject of tieing PR to sales is frequently discussed. Some trade publications offer response cards after specific articles have appeared in print. These offer very valuable "lead-generation" tools. With an effective "lead generation" system, those leads can frequently be tracked through to sales. However, it must be remembered that while PR may have generated the lead, the closure was, of course, heavily influenced by such items as the individual's need for or interest in that product in the first place, the quality of the products and services that are offered, the distribution channel, the availability of the product or service, the price, etc. All of these items, or variables, need to be taken into consideration when seeking to measure the effectiveness of what occurred.

Most organizations, be they business for profit, public sector governmental or non-profit groups and associations, nowadays take the position that PR objectives really do not have value, unless they further the goals of the total organization, or of its business units or sectors. It is most important, therefore, to integrate an organization's PR programs and goals with the strategies and objectives of the organization as a whole. Further, this requires that the practitioner understand what is critical to the organization overall and to its specific business strategies and plans.

Our communication objectives must be tied to business unit or central function operational objectives. These operational objectives are, or should be, behavioral. They should state who will change (customers, employees, suppliers, stakeholders, investors, management, etc.) in what way, by how much and when. In a results-based organization, the only result that matters is a change in behavior (market segment x bought more widgets; employee segment y became more productive; stakeholder segment z supported our environmental policy, etc.)

In a results-based organization, the business unit objective of behavioral change is stated as a Key Result. An achieved communication effectiveness outcome is one indicator of performance towards that result. Our communication program planning objective becomes a Performance Indicator statement in the business line document. We restate the same outcome as a measurable objective in our communication plan. Our objectives are then tied directly to business or organizational objectives.

GETTING SPECIFIC: STANDARDS FOR MEASURING PR *OUTPUTS*

There are many possible tools and techniques that PR practitioners can utilize to begin to measure PR *outputs*, but these are the four that are most frequently relied on to measure PR impact at the *output* level: Media Content Analysis ... Cyberspace Analysis ... Trade Show and Event Measurement ... and Public Opinion Polls.

1. Media Content Analysis

This is the process of studying and tracking what has been written and broadcast, translating this qualitative material into quantitative form through some type of counting approach that involves coding and classifying of specific messages.

Some researchers and PR practitioners in the U.S. refer to this as "Media Measurement" and/or "Publicity Tracking" research. In the United Kingdom, the technique is often referred to as "Media Evaluation;" and in Germany as "Media Resonance." Whatever the terminology used to describe this particular technique, more often than not its prime function is to determine whether the key messages, concepts and themes that an organization might be interested in disseminating to others via the media do, indeed, receive some measure of exposure as a result of a particular public relations effort or activity.

The coding, classifying and analysis that is done can be relatively limited or far-reaching, depending on the needs and interests of the organization commissioning the research. More often than not, Media Content Analysis studies take into consideration variables such as these:

Media Vehicle Variables, such as date of publication or broadcast ... frequency of publication or broadcast of the media vehicle ... media vehicle or type (that is, whether the item appeared in a newspaper, magazine, a newsletter, on radio, or on television) ... and geographic reach (that is, region, state, city, or ADI markets in which the item appeared).

Placement or News Item Variables, such as source of the story (that is, a press release, a press conference, a special event, or whether the media initiated the item on their own) ... story form or type (a news story, feature article, editorial, column, or letter to the editor) ... degree of exposure (that is, column inches or number of paragraphs if the item appeared in print, number of seconds or minutes of air time if the item was broadcast) ... and the story's author (that is, the byline or name of the broadcaster.)

Audience or `Reach' Variables. The focus here usually is on total number of placements, media impressions and/or circulation or potential overall audience reached -- that is, total readers of a newspaper or magazine, total viewers and listeners to a radio or television broadcast. The term "impression" or "opportunity to see" usually refers to the total audited circulation of a publication. For example, if *The Wall Street Journal* has an audited circulation of 1.5 million, one article in that newspaper might be said to generate 1.5 million impressions or opportunities to see the story. Two articles would generate 3 million impressions, and so on. Often more important than impressions is the issue of whether a story reached an organization's target audience group, by specific demographic segments. These data often can be obtained from the U.S. Census Bureau or from various commercial organizations, such as Standard Rate and Data Services. In addition to considering a publication's actual circulation figures, researchers often also take into consideration how many other individuals might possibly be exposed to a given media vehicle, because that publication has been routed or passed on to others.

Subject or Topic Variables, such as who was mentioned and in what context ... how prominently were key organizations and/or their competitors referred to or featured in the press coverage (that is, were companies cited in the headline, in the body copy only, in both, etc.) ... who was quoted and how frequently ... how much coverage, or "share of voice" did an organization receive in comparison to its competitors ... what issues and messages were covered and to what extent ... how were different individuals and groups positioned -- as leaders, as followers, or another way?

Judgment or Subjective Variables. The focus here usually is on the stance or tone of the item, as that item pertains to a given organization and/or its competitors. Usually tone implies some assessment as to whether or not the item is positive, negative or neutral; favorable, unfavorable or balanced. It is extremely important to recognize that measuring stance or tone is usually a highly subjective measure, open to a possibly different interpretation by others. Clearly-defined criteria or groundrules for assessing positives and negatives --- and from whose perspective -- need to be established beforehand, in order for stance or tone measures to have any credibility as part of Media Content Analysis.

"Advertising Equivalency" is often an issue that is raised in connection with Media Content Analysis studies. Basically, advertising equivalency is a means of converting editorial space into advertising costs, by measuring the amount of editorial coverage and then calculating what it would have cost to buy that space, if it had been advertising.

Most reputable researchers contend that "advertising equivalency" computations are of questionable validity. In many cases, it may not even be possible to assign an advertising equivalency score to a given amount of editorial coverage (for example, many newspapers and/or magazines do not sell advertising space on their front pages or their front covers; thus, if an article were to appear in that space, it would be impossible to calculate an appropriate advertising equivalency cost, since advertising could never ever appear there).

Some organizations artificially multiply the estimated value of a "possible" editorial placement in comparison to advertising by a factor of 2, 3, 5, 8 or whatever other inflated number they might wish to come up with, to take into account their own perception that editorial space is <u>always</u> of more value than is advertising space. Most reputable researchers view such arbitrary "weighting" schemes aimed at enhancing the alleged value of editorial coverage as unethical, dishonest, and not at all supported by the research literature. Although some studies have, at times, shown that editorial coverage is sometimes more credible or believable than is advertising coverage, other studies have shown the direct opposite, and there is, as yet, no clearly established consensus in the communications field regarding which is truly more effective: publicity or advertising. In reality, it depends on an endless number of factors.

Sometimes, when doing Media Content Analysis, organizations may apply weights to given messages that are being disseminated, simply because they regard some of their messages as more important than others, or give greater credence (or weight) to an article that not only appears in the form of text, but also is accompanied by a photo or a graphic treatment. Given that the future is visuals, organizations are more and more beginning to measure not only words, but also pictures.

It should be noted that whatever groundrules, criteria and variables are built into a Media Content Analysis, whatever "counting" approaches are utilized to turn qualitative information into quantitative form, it is important that all of the elements and components involved be clearly defined and explained upfront by whoever is doing the study. The particular system of media analysis that is applied and utilized by one researcher should -- if a second researcher were called in and given the same brief and the same basic criteria pertaining to the aims of the study -- result in broadly similar research findings and conclusions.

2. Cyberspace Analysis

Increasingly, a key measure of an organization's image or reputation and of how that organization might be positioned is the chatter and discussion about that organization in cyberspace -- specifically in chat rooms, forums and new groups on the World Wide Web. The same criteria used in analyzing print and broadcast articles can be applied when analyzing postings on the Internet.

What appears in print is frequently commented about and editorialized about on the Web. Therefore, one component of *PR output* measurement ought to be a review and analysis of Web postings.

In addition, a second *output* measure of cyberspace might be a review and analysis of Website traffic patterns. For example, some of the variables that ought to be considered when designing and carrying out Cyberspace Analysis might include deconstructing "hits" (that is, examining the requests for a file of visitors to the Internet) ... a review of click-throughs and/or flash-click streams ... an assessment of home page visits ... domain tracking and analysis ... an assessment of bytes transferred ... a review of time spent per page ... traffic times ... browsers used ... and the number of people filling out and returning feed-back forms.

Best practices for this type of research are covered in *"Measures of Success in Cyberspace,"* a paper authored by Katharine Delahaye Paine that is available from the IPR Commission on PR Measurement and Evaluation, www.instituteforpr.com; *"Getting Started On Interactive Media Measurement,"* available from the Advertising Research Foundation, 641 Lexington Avenue, New York, NY 10022, and *"Hits Are Not Enough: How to Really Measure Web Site Success,"* prepared by *Interactive Marketing News* and available from Phillips Business Information, Inc., 1201 Seven Locks Road, Potomac, MD 20854.

3. Trade Shows and Event Measurement

Frequently, the intent of a public relations program or activity is simply to achieve exposure for an organization, its products or services, through staging trade shows, holding special events and meetings, involvement in speakers' programs and the like.

For shows and events, obviously one possible *output* measure is an assessment of total attendance, not just an actual count of those who showed up, but also an assessment of the types of individuals present, the number of interviews that were generated and conducted in connection with the event, and the number of promotional materials that were distributed. In addition, if the show is used as an opportunity for editorial visits, one can measure the effectiveness of those visits by conducting a content analysis of the resulting articles.

4. Public Opinion Polls

Although most surveys that are designed and carried out are commissioned to measure *PR outtakes* and *PR outcomes* rather than *PR outputs*, public opinion polls are often carried out in an effort to determine whether or not key target audience groups have, indeed, been exposed to particular messages, themes or concepts and to assess the overall effectiveness of a given presentation or promotional effort. For example, conducting a brief survey immediately following a speech or the holding of a special event to assess the short-term impact of that particular activity would constitute a form of *PR output* measurement.

GETTING SPECIFIC: STANDARDS FOR MEASURING PR *OUTTAKES*

Just as there are many tools and techniques that PR practitioners can utilize to begin to measure PR *outputs*, there also are many that can be used to measure PR *outtakes*. Some of those most frequently relied on include surveys (of all types) ... focus groups ... before-and-after polls ... and ethnographic studies (relying on observation, participation, and/or role playing techniques).

There are many books available that discuss and describe both qualitative and quantitative research techniques. Here are three that specifically discuss such techniques from a public relations perspective: *"Using Research In Public Relations,"* by Glen M. Broom and David M. Dozier (Englewood Cliffs, NJ: Prentice Hall, 1990) … *Primer of Public Relations Research,"* by Don W. Stacks (New York: The Guilford Press, 2002) … and *"Public Relations Research For Planning and Evaluation,"* by Walter K. Lindenmann (available from the IPR Commission on PR Measurement and Evaluation, www.instituteforpr.com.)

Ultimately, one intent of public relations is to inform and persuade key target audience groups regarding topics and issues that are of importance to a given organization, with the hope that this will lead those publics to act in a certain way. Usually, this involves two different types of *outtake* measures: Awareness and Comprehension Measurements and Recall and Retention Measurements.

1. Awareness and Comprehension Measurements

The usual starting point for any PR *outtake* measurement is to determine whether target audience groups actually *received* the messages directed at them ... paid *attention* to them ... and *understood* the messages.

Obviously, if one is introducing a new product or concept to the marketplace for the first time -- one that has never been seen or discussed before -- it is reasonable to assume that prior to public relations and/or related communications activities being launched, that familiarity and awareness levels would be at zero. However, many organizations have established some type of "presence" in the marketplace, and thus it is important to obtain benchmark data against which to measure any possible changes in awareness and/or comprehension levels.

Measuring awareness and comprehension levels requires some type of primary research with representatives of key target audience groups.

It is important to keep in mind that **Qualitative Research** (e.g. focus groups, one-on-one depth interviews, convenience polling) is usually open-ended, free response and unstructured in format ... generally relies on non-random samples ... and is rarely "projectable" to larger audiences. **Quantitative Research** (e.g. telephone, mail, mall, internet, fax, and e-mail polls), on the other hand, although it may contain some open-ended questions, is far more apt to involve the use of closed-ended, forced choice questions that are highly structured in format ... generally relies on random samples ... and usually is "projectable" to larger audiences.

To determine whether there have been any changes at all in audience awareness and comprehension levels, usually requires some type of comparative studies -- that is, either a *before* and *after* survey to measure possible change from one period of time to another, or some type of "test" and "control" group study, in which one segment of a target audience group is deliberately exposed to a given message or concept and a second segment is not, with research conducted with both groups to determine if one segment is now better informed regarding the issues than the other.

2. Recall and Retention Measurements

Traditionally, advertising practitioners have paid much more attention to recall and retention measurement, than have those in the public relations field.

It is quite common in advertising, after a series of ads have appeared either in the print or the broadcast media, for research to be fielded to determine whether or not those individuals to whom the ad messages have been targeted actually recall those messages on both an unaided and aided basis. Similarly, several weeks after the ads have run, follow-up studies are often fielded to determine if those in the target audience group have retained any of the key themes, concepts, and messages that were contained in the original advertising copy.

Although recall and retention studies have not been done that frequently by public relations practitioners, they clearly are an important form of *outcome* measurement, that ought to be seriously considered by PR professionals. Various data collection techniques can be used when conducting such studies, including telephone, face-to-face, mail, mall, e-mail, and fax polling.

When conducting such studies, it is extremely important that those individuals fielding the project clearly differentiate between messages that are disseminated via PR techniques (e.g. through stories in the media, by word of mouth, at a special event, through a speech, etc.) from those that are disseminated via paid advertising or through marketing promotional efforts. For example, it is never enough to simply report that someone claims they read, heard or saw a particular item; it is more important to determine whether that individual can determine if the item in question happened to be a news story that appeared in editorial form, or was a paid message that someone placed through advertising. Very often, it is difficult for the "average" consumer to differentiate between the two.

GETTING SPECIFIC: STANDARDS FOR MEASURING PR *OUTCOMES*

Some of the same tools and techniques that PR practitioners can utilize to begin to measure *PR Outtakes* -- surveys, focus groups, before-and-after polls and ethnographic studies -- also can be used to measure *PR Outcomes.* In addition, researchers designing and carrying out projects aimed at measuring changes in people's opinions, attitudes and behavior patterns also often rely on experimental and quasi-experimental designs, on multi-variate analysis projects, and on model building.

In addition to those works previously cited, two useful resources for qualitative and quantitative research techniques that can be used at the *PR Outcome* level are the Advertising Research Foundation's two documents: "Guidelines for the Public Use of Market and Opinion Research" and the *ARF Guidelines Handbook: A Compendium of Guidelines to Good Advertising, Marketing and Media Research Practice*. Both are available from the Advertising Research Foundation, 641 Lexington Avenue, New York, NY 10022.

Two different types of research are usually called for, when conducting public relations measurement and evaluation research at the *outcome* level: Attitude and Preference Measurements and Behavior Measurements.

1. Attitude and Preference Measurements

When it comes to seeking to measure the overall impact or effectiveness of a particular public relations program or activity, assessing individuals' opinions, attitudes, and preferences become extremely important measures of possible *outcomes*.

It needs to be kept in mind that "opinion research" generally measures what people say about something; that is, their verbal expressions or spoken or written points of view. "Attitude research," on the other hand, is far deeper and more complex. Usually, "attitude research" measures not only what people say about something, but also what they know and think (their mental or cognitive predispositions), what they feel (their emotions), and how they're inclined to act (their motivational or drive tendencies).

"Opinion research" is easier to do because one can usually obtain the information desired in a very direct fashion just by asking a few questions. "Attitude research," however, is far harder and, often more expensive to carry out, because the information desired often has to be collected in an indirect fashion. For example, one can easily measure people's stated positions on racial and/or ethnic prejudice, by simply asking one or several direct questions. However, actually determining whether someone is in actual fact racially and/or ethnically prejudiced, usually would necessitate asking a series of indirect questions aimed at obtaining a better understanding of people's cognitions, feelings, and motivational or drive tendencies regarding that topic or issue.

Preference implies that an individual is or will be making a choice, which means that preference measurement more often than not ought to include some alternatives, either competitive or perceived competitive products or organizations. To determine the impact of public relations preference *outcomes* usually necessitates some type of audience exposure to specific public relations *outputs* (such as an article, a white paper, a speech, or participation in an activity or event), with research then carried out to determine the overall likelihood of people preferring one product, service, or organization to another.

Usually, opinion, attitude and preference measurement projects involve interviews not only with those in the public at large, but also with special target audience groups, such as those in the media, business leaders, academicians, security analysts and portfolio managers, those in the health, medical and scientific community, government officials, and representatives of civic, cultural and service organizations. Opinion, attitude and preference measurement research can be carried out many different ways, through focus groups, through qualitative and quantitative surveys, and even through panels.

2. Behavior Measurements

The ultimate test of effectiveness -- the highest *outcome* measure possible -- is whether the behavior of the target audience has changed, at least to some degree, as a result of the public relations program or activity.

For most media relations programs, if you have changed the behavior of the editor and/or reporter so that what he or she writes primarily reflects an organization's key messages, then that organization has achieved a measure of behavior change.

However, measuring behavior is hard because it is often difficult to prove cause-and-effect relationships. The more specific the desired *outcome* and the more focused the PR program or activity that relates to that hoped-for end result, the easier it is to measure PR behavior change. For example, if the intent of a public relations program or activity is to raise more funds for a non-profit institution and if one can show after the campaign has been concluded that there has, indeed, been increased funding, then one can begin to surmise that the PR activity had a role to play in the behavior change. Or, to give another example: For measuring the effectiveness of a public affairs or government relations program targeted at legislators or regulators, the desired *outcome* -- more often than not -- would not only be to get legislators or regulators to change their views, but more importantly to have those legislators and regulators either pass or implement a new set of laws or regulations that reflect the aims of the campaign. Behavior change requires some one to act differently than they have in the past.

More often that not, measuring behavior change requires a broad array of data collection tools and techniques, among them before-and-after surveys ... research utilizing ethnographic techniques (e.g. observation, participation, and role playing) ... the utilization of experimental and quasi-experimental research designs ... and studies that rely on multi-variate analyses and sophisticated statistical applications and processes.

What is crucial to bear in mind in connection with PR *outcome* behavior measurement studies is that measuring correlations -- that is, the associations or relationships that might exist between two variables -- is relatively easy. Measuring causation -- that is, seeking to prove that X was the reason that Y happened -- is extremely difficult. Often, there are too many intervening variables that need to be taken into consideration.

Those doing PR *outcome* behavior measurement studies need to keep in mind these three requirements that need to exist in order to support or document that some activity or event caused something to happen: 1) Cause must always precede the effect in time; 2) there needs to be a relationship between the two variables under study; and 3) the observed relationship between the two variables cannot be explained away as being due to the influence of some third variable that possibly caused both of them.

The key to effective behavior measurement is a sound, well thought-out, reliable and valid research concept and design. Researchers doing such studies need to make sure that study or test conditions or responses are relevant to the situation to which the findings are supposed to relate, and also clearly demonstrate that the analysis and conclusions that are reached are indeed supported and documented by the field work and data collection that was carried out.

QUESTIONS THAT NEED TO BE PUT TO THOSE ORGANIZATIONS THAT COMMISSION PR MEASUREMENT AND EVALUATION STUDIES

Here are some of the key questions that those who commission PR measurement evaluation studies ought to ask themselves before they begin, and also the types of questions that those who actually carry out the assignment ought to ask their clients to answer before the project is launched:

-- What are, or were, the specific goals and/or objectives of the public relations, public affairs, and/or marketing communications program, and can these be at all stated in a quantitative or measurable fashion? (e.g. To double the number of inquiries received from one year to the next? ... To increase media coverage by achieving greater "share of voice" in one year than in a previous year? ... To have certain legislation passed? ... To enhance or improve brand, product, or corporate image or reputation?)

-- Who are, or were, the principal individuals serving as spokespersons for the organization during the communications effort?

-- What are, or were, the principal themes, concepts, and messages that the organization was interested in disseminating?

-- Who were the principal target audience groups to whom these messages were directed?

-- Which channels of communication were used and/or deemed most important to use in disseminating the messages? (e.g. the media ... word-of-mouth ... direct mail ... special events?)

-- What specific public relations strategies and tactics were used to carry out the program? What were the specific components or elements of the campaign?

-- What is, or was, the timeline for the overall public relations program or project?

-- What is, or were, the desired or hoped-for outputs, outtakes, and/or outcomes of the public relations effort? If those particular hoped-for outputs, outtakes and/or outcomes could, for some reason, not be met, what alternative outputs, outtakes, and/or outcomes would the organization be willing to accept?

-- How does what is or has happened in connection with the organization's public relations effort relate to what is or has happened in connection with related activities or programs in other areas of the company, such as advertising, marketing, and internal communications?

-- Who are the organization's principal competitors? Who are their spokespersons? What are their key themes, concepts, and messages that they are seeking to disseminate? Who are their key target audience groups? What channels of communications are they most frequently utilizing?

-- Which media vehicles are, or were, most important to reach for the particular public relations and/or marketing communications activities that were undertaken?

-- What were the specific public relations materials and resources utilized as part of the effort? Would it be possible to obtain and review copies of any relevant press releases, brochures, speeches, promotional materials that were produced and distributed as part of the program?

-- What information is already available to the organization that can be utilized by those carrying out the evaluative research assignment to avoid reinventing the wheel and to build on what is already known?

-- If part of the project involves an assessment of media coverage, who will be responsible for collecting the clips or copies of broadcast materials that will have been generated? What are the groundrules and/or parameters for clip and/or broadcast material assessment?

-- What major issues or topics pertaining to the public relations undertaking are, or have been, of greatest importance to the organization commissioning the evaluation research project?

-- What is the timeline for the PR Measurement and Evaluation Research effort? What are the budgetary parameters and/or limitations for the assignment? Do priorities have to be set?

-- Who will be the ultimate recipients of the research findings?

-- How will whatever information that is collected be used by the organization that is commissioning the research?

QUESTIONS THAT NEED TO BE PUT TO THOSE RESEARCH SUPPLIERS, AGENCIES AND CONSULTING FIRMS THAT ACTUALLY CONDUCT PR MEASUREMENT AND EVALUATION STUDIES

Here are some of the key questions that ought to be put to those who actually are asked to carry out a PR measurement and evaluation research project, before the assignment is launched:

-- What is, or will be, the actual research design or plan for the PR measurement and evaluation project? Is there, or will there be, a full description in non-technical language of what is to be measured, how the data are to be collected, tabulated, analyzed and reported?

-- Will the research design be consistent with the stated purpose of the PR measurement and evaluation study that is to be conducted? Is there, or will there be, a precise statement of the universe or population to be studied? Does, or will, the sampling source or frame fairly represent the total universe or population under study?

-- Who will actually be supervising and/or carrying out the PR measurement and evaluation project? What is, or are, their backgrounds and experience levels? Have they ever done research like this before? Can they give references?

-- Who will actually be doing the field work? If the assignment includes media content analysis, who actually will be reading the clips or viewing and/or listening to the broadcast video/audio tapes? If the assignments involve focus groups, who will be moderating the sessions? If the study involves conducting interviews, who will be doing those and how will they be trained, briefed, and monitored?

-- What quality control mechanisms have been built into the study to assure that all "readers," "moderators," and "interviewers" adhere to the research design and study parameters?

-- Who will be preparing any of the data collection instruments, including tally sheets or forms for media content analysis studies, topic guides for focus group projects, and/or questionnaires for telephone, face-to-face, or mail survey research projects? What role will the organization commissioning the PR measurement and evaluation assignment be asked, or be permitted, to play in the final review and approval of these data collection instruments.

-- Will there be a written set of instructions and guidelines for the "readers," the "moderators" and the "interviewers"?

-- Will the coding rules and procedures be available for review?

-- If the data are weighted, will the range of the weights be reported? Will the basis for the weights be described and evaluated? Will the effect of the weights on the reliability of the final estimates be reported?

-- Will the sample that is eventually drawn be large enough to provide stable findings? Will sampling error limits be shown, if they can be computed? Will the sample's reliability be discussed in language that can clearly be understood without a technical knowledge of statistics.

-- How projectable will the research findings be to the total universe or population under study? Will it be clear which respondents or which media vehicles are underrepresented, or not represented at all, as part of the research undertaking?

-- How will the data processing be handled? Who will be responsible for preparing a tab plan for the project? Which analytical and demographic variables will be included as part of the analysis and interpretation?

-- How will the research findings and implications be reported? If there are findings based on the data that were collected, but the implications and/or recommendations stemming from the study go far beyond the actual data that were collected, will there be some effort made to separate the conclusions and observations that are specifically based on the data and those that are not?

-- Will there be a statement on the limitations of the research and possible misinterpretations of the findings?

-- How will the project be budgeted? Can budget parameters be laid out prior to the actual launch of the assignment? What contingencies can be built into the budget to prevent any unexpected surprises or changes once the project is in the field or is approaching the completion stage?

DEFINITIONS OF SELECTED TERMS USED IN PR MEASUREMENT AND EVALUATION

Advertising Equivalency: A means of converting editorial space in the media into advertising costs, by measuring the amount of editorial coverage and then calculating what it would have cost to buy that space, if it had been advertising. Most reputable researchers contend that advertising equivalency computations are of questionable validity, since in many cases the opportunity to "buy" advertising in space that has been specifically allocated to editorial coverage simply does not exist.

Attitude Research: Consists of measuring and interpreting the full range of views, sentiments, feelings, opinions and beliefs which segments of the public may hold toward given people, products, organizations and/or issues. More specifically, attitude research measures what people say (their verbal expressions), what they know and think (their mental or cognitive predispositions), what they feel (their emotions), and how they're inclined to act (their motivational or drive tendencies).

Bivariate Analysis: Examination of the relationship between two variables.

Causal Relationship: A theoretical notion that change in one variable forces, produces, or brings about a change in another.

Circulation: Refers to the number of copies sold of a given edition of a publication, at a given time or as averaged over a period of time.

Communications Audit: A systematic review and analysis -- using accepted research techniques and methodologies -- of how well an organization communicates with all of its major internal and external target audience groups.

Confidence Interval: In a survey based on a random sample, the range of values within which a population parameter is estimated to fall. For example, in a survey in which a representative sample of 1,000 individuals is interviewed, if 55% express a preference for a given item, we might say that in the population as a whole, in 95 out of 100 cases, the true proportion expressing such a preference probably would fall between 52% and 58%. The plus or minus 3% range is called the confidence interval. The fact that we are using 95 out of 100 cases as our guide (or 95%) is our confidence level.

Content Analysis: The process of studying and tracking what has been written and broadcast and translating this qualitative material into quantitative form through some type of counting approach that involves coding and classifying of specific messages.

Correlation: Any association or relationship between two variables.

Correlation Coefficient: A measure of association (symbolized as r) that describes the direction and strength of a linear relationship between two variables, measured at the interval or ratio level (e.g. Pearson's Correlation Coefficient).

Cost Per Thousand (CPM): The cost of advertising for each 1,000 homes reached by radio or television, for each 1,000 copies of a publication, or for each 1,000 potential viewers of an outdoor advertisement.

Cross-Sectional Study: A study based on observations representing a single point in time.

Demographic Analysis: Consists of looking at the population in terms of special social, political, economic, and geographic subgroups, such as a person's age, sex, income-level, race, education-level, place of residence, or occupation.

Ethnographic Research: Relies on the tools and techniques of cultural anthropologists and sociologists to obtain a better understanding of how individuals and groups function in their natural settings. Usually, this type of research is carried out by a team of impartial, trained researchers who "immerse" themselves into the daily routine of a neighborhood or community, using a mix of observation, participation, and role-playing techniques, in an effort to try to assess what is really happening from a "cultural" perspective.

Evaluation: Determines the *value* or importance of a public relations program or effort, usually through appraisal or comparison with a predetermined set of organization goals and objectives. PR Evaluation is somewhat more subjective in nature, or softer, than PR Measurement, involving a greater amount of interpretation and judgment calls.

Experiment: Any controlled arrangement and manipulation of conditions to systematically observe specific occurrences, with the intention of defining those criteria that might possibly be affecting those occurrences. An experimental, or quasi-experimental, research design usually involves two groups -- a "test" group which is exposed to given criteria, and a "control" group, which is not exposed. Comparisons are then made to determine what effect, if any, exposures to the criteria have had on those in the "test" group.

Factor Analysis: A complex algebraic procedure that seeks to group or combine items or variables in a questionnaire based on how they naturally relate to each other, or "hang together," as general descriptors (or "factors").

Focus Group: An exploratory technique in which a group of somewhere between 8 and 12 individuals -- under the guidance of a trained moderator -- are encouraged, as a group, to discuss freely any and all of their feelings, concerns, problems and frustrations relating to specific topics under discussion. Focus groups are ideal for brainstorming, idea-gathering, and concept testing.

Frequency: The number of advertisements, broadcasts, or exposures of given programming or messaging during a particular period of time.

Gross Rating Point: A unit of measurement of broadcast or outdoor advertising audience size, equal to 1 percent of the total potential audience universe; used to measure the exposure of one or more programs or commercials, without regard to multiple exposure of the same advertising to individuals. A GRP is the product of media reach times exposure frequency.

A *gross-rating-point buy* is the number of advertisements necessary to obtain the desired percentage of exposure of the message. In outdoor advertising, GRPs, often used as a synonym for showing, generally refer to the daily effective circulation generated by poster panels, divided by market population. The *cost per gross rating point* (CPGRP) is a measure of broadcast media exposure comparable to the *cost per thousand* (CPM) measure of print media.

Hypothesis: An expectation about the nature of things derived from theory.

Hypothesis-Testing: Determining whether the expectations that a hypothesis represents are, indeed, found in the real world.

Impressions: The number of those who might have had the opportunity to be exposed to a story that has appeared in the media. Sometimes referred to as "opportunity to see." An "impression" usually refers to the total audited circulation of a publication or the audience reach of a broadcast vehicle.

Incidence: The frequency with which a condition or event occurs within a given time and population.

Inquiry Study: A systematic review and analysis, using content analysis or sometimes telephone and mail interviewing techniques, to study the range and types of unsolicited inquires that an organization may receive from customers, prospective customers or other target audience groups.

Inputs: (1) Everything that is involved upfront within the organization in the design, conception, approval, production and distribution of communications materials aimed at targeted audience groups. (2) Also, the research information and data from both internal and external sources that are applied to the initial stage of the communications planning and production process.

Judgmental Sample: A type of non-probability sample in which individuals are deliberately selected for inclusion in the sample by the researcher because they have special knowledge, position, characteristics or represent other relevant dimensions of the population that are deemed important to study. Also known as a "purposive" sample.

Likert Scale: Developed by Rensis Likert, this is a composite measure in which respondents are asked to choose from an ordered series of five responses to indicate their reactions to a sequence of statements (e.g., strongly agree ... somewhat agree ... neither agree nor disagree ... somewhat disagree ... strongly disagree).

Longitudinal Study: A research design involving the collection of data at different points in time.

Mall Intercept: A special type of in-person interview, in which potential respondents are approached as they stroll through shopping centers or malls. Most mall intercept interviews are based on non-probability sampling.

Market Research: Any systematic study of buying and selling behavior.

Mean: A measure of central tendency which is the arithmetic average of the scores.

Measurement: A way of giving a result a precise dimension, generally by comparison to some standard or baseline, and usually is done in a quantifiable or numerical manner.

Median: A measure of central tendency indicating the midpoint in a series of scores, the point above and below which 50 percent of the values fall.

Mode: A measure of central tendency which is the most frequently occurring, the most typical, value in a series.

Multivariate Analysis: Examination of the relationship among three or more variables.

Omnibus Survey: An "all-purpose" national consumer poll usually conducted on a regular schedule -- once a week or every other week -- by major market research firms. Organizations are encouraged to "buy" one or several proprietary questions and have them "added" to the basic questionnaire. Those adding questions are usually charged on a per-question basis. Also, sometimes referred to as "piggyback," or "shared-cost" surveys.

Outcomes: A long-term measure of the effectiveness of a particular communications program or activity, by focusing on whether targeted audience groups changed their *opinions*, *attitudes* and/or *behavior patterns* as a result of having been exposed to and become aware of messages directed at them.

Outgrowths: (1) The culminate effect of all communication programs and products on the positioning of an organization in the minds of its stakeholders or publics. (2) For some, the term used to describe the *outtakes* of a communications program activity (see that definition).

Outputs: (1) The short-term or immediate results of a particular communications program or activity, with a prime focus on how well an organization presents itself to others and the amount of exposure it receives. (2) For some, the final stage in the communications production process, resulting in the production and distribution of such items as brochures, media releases, websites, speeches, etc.

Outtakes: (1) A measure of the effectiveness of a particular communications program or activity, by focusing on whether targeted audience groups *received* the messages directed to them … paid *attention* to the messages … *understood* or comprehended the messages … and *retained* and can *recall* the messages in any shape or form. (2) Initial audience reaction to the receipt of communications materials, including whether the audience heeded or responded to a call for information or action within the messages.

Panel Study: 1) A type of longitudinal study in which the same individuals are interviewed more than once over a period of time to investigate the processes of response change, usually in reference to the same topic or issue. 2) Also, a type of study in which a group of individuals are deliberately recruited by a research firm, because of their special demographic characteristics, for the express purpose of being interviewed more than once over a period of time for various clients on a broad array of different topics or subjects.

Probability Sample: A process of random selection, in which each unit in a population has an equal chance of being included in the sample.

Psychographic Analysis: Consists of looking at the population in terms of people's non-demographic traits and characteristics, such as a person's personality type, life-style, social roles, values and beliefs.

Q-Sort: A personality inventory introduced in the 1950's in which respondents are asked to sort opinion statements along a "most-like-me" to "most-unlike-me" continuum. Q-Sorting allows researchers to construct models of individual respondents' belief systems.

Qualitative Research: Usually refers to studies that are somewhat subjective, but nevertheless in-depth, using a probing, open-end, free-response format.

Quantitative Research: Usually refers to studies that are highly objective and projectable, using closed-end, forced-choice questionnaires. These studies tend to rely heavily on statistics and numerical measures.

Quota Sample: A type of non-probability sample in which individuals are selected on the basis of pre-specified characteristics, so that the total sample will have the same general distribution of characteristics as are assumed to exist in the population being studied.

Range: A measure of variability that is computed by subtracting the lowest score in a distribution from the highest score.

Reach: Refers to the range or scope of influence or effect that a given communications vehicle has on targeted audience groups. In broadcasting, it is the net unduplicated radio or TV audience -- the number of different individuals or households -- for programs or commercials as measured for a specific time period in quarter-hour units over a period of one to four weeks.

Regression Analysis: A statistical technique for studying relationships among variables, measured at the interval or ratio level.

Reliability: The extent to which the results would be consistent, or replicable, if the research were conducted a number of times.

Screener Question: One or several questions usually asked in the beginning of an interview to determine if the potential respondent is eligible to participate in the study.

Secondary Analysis: A technique for extracting from previously conducted studies new knowledge on topics other than those which were the focus of the original studies. It does this through a systematic re-analysis of a vast array of already existing research data.

Situation Analysis: An impartial, often third-party assessment of the public relations and/or public affairs problems, or opportunities, that an organization may be facing at a given point in time.

Standard Deviation: An index of variability of a distribution. More precisely, it is the range from the mean within which approximately 34% of the cases fall, provided the values are distributed in a normal curve.

Statistical Significance: Refers to the unlikeliness that relationships observed in a sample could be attributed to sampling error alone.

Survey: Any systematic collection of data that uses a questionnaire and a recognized sampling method. There are three basic types of surveys: those conducted face-to-face (in-person) ... those conducted by telephone ... and those that are self-administered (usually distributed by mail, e-mail, or fax.)

Univariate Analysis: The examination of only one variable at a time.

Validity: The extent to which a research project measures what it is intended, or purports, to measure.

Variance: A measure of the extent to which individual scores in a set differ from each other. More precisely, it is the sum of the squared deviations from the mean divided by the frequencies.