Ear, eye, and A/V learning

What is the precise relationship of the auditory to the visual mate¬rial in the learning process? Nobody knows. Publications in the meetings field regularly quote as fact a percentage-based ratio; but in all our readings in psychology and education, we have never seen such a ratio expressed in scientific papers.

Multisensual learning is the human norm, of course. But researchers have found that the major source-sense in any learning process varies not only with the topic being learned but also among different learners of the same material.

The much-quoted percentage ratio originated in the 50s in an informal, internal estimate prepared by the Mobil Oil Company for the purpose of evaluating executive presentation methods. Mobil has never considered the studies either scientifically sound or valid. The ratios, heavily weighted toward visual learning, do help to sell slides and visual media.

The field is too complex for that easy answer. Much information is available on various aspects of learning. Studies differentiate between such spheres as knowledge, meaning, comprehension, and understanding; or cognition, information processing, and memory. Because these studies are distributed among the fields of psychology, psycho¬biology or psychophysiology (how the body functions to create the senses and produce the responses), and education, the results have not been correlated and collected into a handy reference book for meeting planners.

Moreover, because scientific evidence depends heavily for its validity on single-variable methodology with replication of results by other researchers, and because there are so many varieties of learn-able topics and billions of persons as potential test subjects, audiovisual learning is not necessarily a first choice of subject for research.

The U.S. Army studies have shown that color and motion are not necessarily aids to comprehension and learning. Other work done on the subject of the interdependence of the two key senses has established that visual cues do indeed enhance the understanding of spoken material when those visual cues link poorly organized or dis-jointed information. That tends to explain the common-sense observation that good A/V presentations tend to outperform verbal-only presentations.

Split-brain studies have demonstrated that the left hemisphere of the brain specializes in sequential processing (as in language) and the right hemisphere in conceptual processing. Although the ears have a bias toward the left side, the eyes feed equally to the left and right hemispheres. Presumably, then, the eye has the capacity to link or translate the two modes of learning--an original conclusion--if the material is prepared so as to facilitate that translation.

In this chapter we will attempt to codify some of the relevant principles so that we can improve the effectiveness of A/V presentations as learning aids. So that those few principles are removed from the area of pure speculation, we will summarize the most significant findings in the several fields of research. You will then understand the reasons for particular methods and choose more securely in borderline cases.

Can we compare the two senses?

Probably not. Visual cues seem to impact on more cells of the brain than does something spoken, which might account for some discrepancy in what is retained. On the other hand, the eyes fatigue or cloy much sooner than the ears do; so for an extended program, the ear presents advantages.

Until the 19th Century there was no "science," and until this century there were few educational theories. Adult education became a phenomenon after World War II because of both the mass of GIs returning to school and the abundant new learning that had taken place in various fields of as psychology a result of government experiments in mass-persuasion techniques for training and propaganda purposes. A landmark book, Experiment on Mass Communication, by Carl Hovland, Arthur Lumsdaine, and Fred Sheffield, was part of the explosion of information.

Next up was role playing (King and Janis; Jansen and Stulurow; Corsini, Shaw, and Blake); the development of training programs (Rose); attitude change (Cohen; Sherif and Shen, PWeick); and media (Rossi and Biddle/Knowles; Brown and Thornton). All the areas were very new, and little was directly useable by the meeting planner. (See the Bibliography for specifics.)

In England, D. E. Broadbent (Perception and Communication, Oxford: Pergamon, 1958) attacked the problem of concentration and dis¬traction in learning. That seminal work provided insights into atten¬tion spans and information overloads.

Recognizing that the brain cannot process large numbers of simultaneously received stimuli equally and instantly, Broadbent introduced a "filter" theory. He hypothesized that to avoid bottlenecks, the brain filters and sequences material and selects the most pressing (our word) for immediate processing. Overload is put into short-term storage until handled and transferred to long-term storage. Broadbent's filter biases accepted as pressing the novel or intense over the familiar; acoustic over visual signals; sounds of high over low frequency; and signals of biological importance. Work by other researchers bore out many of his theses, including loss of information during short-term storage--a key finding.

Subsequently, Melton demonstrated the presence of both short-term and long-term memory but disagreed with Broadbent regarding sequencing delay and loss of information. Melton states that long-term memory is characterized by (a) presence of learning and (b) effects of the established phenomenon of forgetting through interference of new stimuli/learning.

Soon after that, Waugh and Norman contributed the terms primary and secondary memory to clear up confusions of terminology in various writings and to consolidate compatible theories regarding memory storage. They argue that a primary memory receives everything and that thinking about or rehearsing it will transfer some

material to secondary (permanent) memory. The balance will be forgotten, some of it within seconds or minutes.

Working at about the same time, in the mid-1960s, Hernandez-Peon confirmed a filter process but disagreed with Broadbent about the permanence of filter biases. He finds actual neurophysiological blocking of certain classes of inputs when attention is distracted.

We can fairly summarize findings in the field of stimuli filtering to demonstrate that an overload of stimuli results in the loss of much material presented; that distractions cause further blocking; and that lack of opportunity to rethink or rehearse causes still more loss of potential learning. All these counterproductive circumstances are present in the maximedia blitz, which depends on over stimulation for its "excitement."

Even more important to the relationship of the auditory to the visual portions are findings reported in The Psychology of Perception, by William Dember and Joel Warm (New York: Holt Rinehart & Winston, 1979). They believe that vigilance is essential to learning and comprehension; that is, the sustained attention leading to alertness and consequently to peak performance over an extended period of time.

Dember and Warm find that performance efficiency in sustained attention tasks involving auditory signals tends to be superior to that in tasks involving visual eye signals. Related experiments reveal rather low correlations between people's ability to sustain attention in different sensory channels simultaneously.

That suggests that "distractions," in the sense of the findings previously summarized, can include the opposite half of the audiovisual combination. And new evidence indicates that auditory visual correlations can be increased by closely equating the types of discriminations required in the two types of tasks. Clearly, the learning purpose is not served by the willy-nilly pairing of words and pictures.

The legitimacy of all those findings was enhanced in the late 60s when Roger Sperry and Michael Gazzaniga carried the then current research on hemispheric specialization of the brain into new areas. They experimented with patients whose left and right cerebral hemispheres had been surgically separated to eliminate severe epileptic seizures. Separation was achieved by severing the corpus callosum, a bundle of nerve fibers connecting left and right hemispheres in the normal brain.

It was already known that the left hemisphere tends to control motor responses of the right side of the body, and vice versa, and that the eyes feed information to both hemispheres. But the hemispheres cooperate, and with the corpus callosum intact information flows where it is needed.

With the corpus callosum cut, Sperry and Gazzaniga demonstrated, the patient was stymied. The person could recognize in the visual left field/right hemisphere a word in the abstract sense (that is, find its identical letter picture in a list of different words), but could no longer say that word because the language ability in the other hemisphere was estranged. Many related experiments by these and other re¬searchers show that the left side of the brain is specialized to process language, mathematics, and general analytic tasks which process information serially or sequentially. The right hemisphere seems specialized to process abstracts and spatial tasks; artistic creativity is its province, and information is processed in a simultaneous/wholistic/conceptual manner.

Even these findings are not neat and clean. In aural tests, normal persons seem to process music differently depending on whether or not they have previous musical training. Most of us tend to hear melody as a right-hemispheric creative whole--the melodic picture. However, trained musicians process in both hemispheres: they hear the "grammar" of the music in the left side (tonal relationships, time values, and so on) simultaneously with the melody. This would seem to suggest that writers would hear poetry differently from persons with lesser language skills because language and technique would be distinct from the "story."

We could conclude that all creators of learning aids should first determine whether the material to be conveyed is sequential or conceptual in nature and then favor ear or eye, respectively. Workshops must convert sequential information into gestalt understanding.

That sounds simple. However, some learning tasks are in the cognitive domain (knowledge, such as arithmetic) and others in the affective domain (attitudes). Still others involve sensory-motor skills in which practice is at least as important as knowledge or attitudes. Has anyone ever learned to ride a bicycle simply by reading about it or wanting to ride?

Obviously, even a well-prepared visual presentation could fail to produce desired results if the expectations are unrealistic, that is, if the film shows bikes being ridden but the program does not provide breakout sessions to practice riding. And some sales tasks--especially those involving product demonstrations--must provide practice sessions simply to be complete.

We already know that we learn differently with eye and ear, and even that one can interfere with the other. But what is learning? It's a complex process of converting sensory stimuli into memory-preferably long-term memory. There are many related concepts to be mastered.

One of the best texts in this area is Cognition and the Symbolic Processes, edited by Walter Weimer and David Palermo (Hillsdale, N.J.: Lawrence Erlbaum & Associates, 1974). In the chapter "Toward Understanding Understanding," Jeffery J. Franks sets out the following terminology:

Understanding: a cover term for all aspects of cognition, including memory, meaning, language, perception, etc.

Tacit knowledge: a reference to knowledge of which we are not (and possibly cannot be) directly aware.

Knowledge: a reference to static, semi-permanent, long-term memory relationships.

Meaning: a reference to relationships activated or generated as a function of some knowledge relationship within the present environmental context.

Comprehension: a function of the extent to which adequate (coherent, complete) meanings have been generated in a particular con¬text.

Pattern recognition: classifying a particular input into its appropriate perceptual/conceptual class.

According to Franks, "The structure of meanings is determined by (systemic) knowledge just as the structure of sentences is determined by grammar . . . Overt manifestations of meaning (imagery; overt and covert speech; and responses) are analogous to surface structure in sentences. Tacit meanings are like the underlying structure of sentences."

Psychologists hold conflicting theories about whether long-term memory (knowledge) is a generative conceptual system or a storehouse of specific memories of past experiences. Although tradition favors the latter, Franks argues that both are valid at different times and in different contexts.

In short, Franks' research and theories demonstrate that in even the most simple of verbal communications (written or spoken) the language and the abstract conditioning of the individual person regarding the individual words will color the communication. He is in accord with another seminal writer, Arthur L. Blumenthal (Language and Psychology: Historical Aspects of Psycholinguistics; New York; John Wiley & Sons, 1970), as well as with his colleagues John Bransford and Nancy McCarrell whose paper appears with Franks' own under the chapter title "A Sketch of a Cognitive Approach to Comprehension." All three honor the highly regarded findings of Karl Buehler about the need for any two conversing parties to share the same "field" of reference or abstract meaning.

Bransford and McCarrell conclude that people do make cognitive contributions while comprehending; that such contributions are prerequisites for achieving a "click of comprehension"; that knowledge of underlying abstracts helps determine the individuals' contributions; and that meanings derived as the result of such contributions is best viewed as something created rather than retrieved from storage.

Bransford and McCarrell further demonstrate that test subjects make assumptions (and falsely "remember") data that might be implied but was never stated. They conclude that "comprehension occurs only when the comprehender has sufficient alinguistic information to use the linguistic cues specifically in linguistic input to create some semantic content that allows him to understand." Prior knowledge alone is not sufficient to assure comprehension; knowledge must be activated.

That concept was presented by Wundt (quoted by Blumenthal, 1970) as "the mind of the hearer is just as active in transforming and creating as the mind of the speaker."

Briefly, there is scientific data to back up the commonplace observation that different people hear different things in the same words; or that people tend to hear what they prefer to hear in a spoken message. And they prefer to hear whatever will avoid psychic dissonance--that is, they tune out new material at odds with what they already know or believe.

You must support your controversial concepts with concrete if you hope to persuade. Hovland (1953) demonstrates that high credibility of speakers increases the amount of opinion change; low credibility reduces opinion change (and can negate prior acceptance).

One of the great values of any appeal to the eye, therefore, is the identicalness of material presented to every person; even more, the visual can be held before the eye longer than a word can be held in the ear. This serves the rethink/rehearsal sequence believed necessary to convert the information into long-term memory. The fact that a physicist will get different information from a "picture" of an equation than will a layperson is a different concept: a matter of "field" rather than "time."

We must conclude that a prime responsibility of the visual is to narrow the interpretive focus by limiting potential (mis-) interpretations of the spoken word.

In this aspect, the visual portion of the A/V presentation is admirably suited to serve the auditory portion: the eye can present to the right hemisphere (conceptual processing) the "big picture" of the material under discussion. The eye can then fixate repeatedly on different portions of the image to study details. By contrast, the ear hears and processes details that do not finally make sense (or deliver a big picture) until the last word of the sentence has been spoken.

Can the eye truly guide the ear to comprehension?

Definitely, according to findings of Bransford and McCarrell. The team devised an unlikely picture accompanied (not described or explained) by a written paragraph intended to be read aloud to viewers of the picture. The paragraph was so disjointed that no clear idea or progression was apparent--it was the grammatical equivalent of non¬sense syllables. Several variables of presentation were tried.

Recall of the spoken words was predictably poor if the subjects heard only the paragraph, with no visuals. Recall was not appreciably improved even if they saw the picture after having heard the entire paragraph. But their recall was quite good if they saw the picture for 30 seconds before the passage was read to them. Moreover, most subjects felt that the disjointed paragraph was acceptably organized if they had seen the picture first.

In a related, second type of test, a similarly disjointed paragraph read without a title was poorly recalled; but hearing a title in advance gave enough context so that recall was improved.

In a third type of test, a paragraph was ambiguously designed to reflect either a common or an exotic experience. The either/or decision varied according to the title attached, but recall of specific test items was better on an element most "coached" by the title.

We can conclude that not only is the eye able to present concepts to the right hemisphere while the spoken material goes to the left hemisphere by way of the ear, but the eye can present "roadmaps" that tell the ear where the material is headed.

Is that necessary? Probably. The response to auditory signals in language is extremely complex. We have already seen that the listener must make a cognitive contribution to his or her own comprehension using knowledge already in storage.

But the listener makes an additional effort-that of making units of sounds intelligible.

George Miller laid a cornerstone in the foundations for the study of language comprehension in his The Psychology of Communication (New York: Basic Books, 1967). Miller establishes six distinct functions essential to comprehension: (1) The listener must hear the utterance as an auditory stimulus; (2) match it as a phonemic pattern in terms of his own phonologic skills; (3) analyze and accept the utterance as a sentence in grammatical terms; (4) interpret the utterance as meaningful in semantics or field; (5) understand the utterance's meaning; and (6) believe the validity (and the speaker's intent to deliver that meaning with those sounds. We can understand without believing, of course).

With that explanation of process, we can appreciate the miracle of human speech. And we can accept more easily the inability of the listener to understand and comprehend every detail of a complicated thought after a single hearing.

Miller also established the concept of Magic Seven. Apparently, seven items is the outer limit of "chunks" of information the mind can retain without forgetting in short-term processing. To complicate matters, the seven items need not be bits of information--such as a series of single-digit numbers--but can be up to seven internally related concepts, each containing more than one bit: phrases.

No one is recommending that all information be delivered in sets of sevens, but that finding of the "chunking" faculty does argue for think periods, practice, and avoidance of the blitz techniques of maximedia.

We needn't even be overly concerned that we present only the "proper" type of information to either half of the brain because the normal brain is quite capable of shifting information around--over a period of time. It also creates the creative "aha" experience for itself--that is, a sudden insight which correlates previous detailed (pre-sumably left-hemisphere) learning.

But it does matter, on the basis of noninterference and rethinking, rehearse, and the filter biases, that we make an effort not to overload the circuits or present incompatible material to the ear and eye simultaneously or challenge cherished beliefs or prerogatives without al¬lowing time both for the learner to understand the message and comprehend the ramifications. If we meet the physiological strictures of the brain half way, we will get more work out with less input. That's efficiency.

To present the scientific bases for our own argument, we have read the original papers summarized here. Encapsulating a paper or book into a single brief quote is unfair to the authors as well as to their ideas; however, tracing, reading, and interpreting that technical mate¬rial can be time consuming beyond its immediate value.

Therefore, if you are interested in reading more deeply, we recommend your starting with an excellent college text that puts the pieces together: Cognition: Mental Structures and Processes, by David Dodd and Raymond White, Jr. (Newton, Mass.: Allyn & Bacon, 1980.) The extensive bibliography will steer you to other significant work in the field.

We can put the heavy stuff aside and get on with the practical application of the evidence the laboratory has put at our disposal.

Whatever information is presented in a meeting must ultimately be useful to company and meeting participants. That usually requires a conversion process from the data state to meanings state, which could involve cognitive and/or dexterity usage. That conversion process is called either education or training, depending on whether the learning is (a) general preparation for approaching unknown problems in the future or (b) carefully selected to bring the best available knowledge and techniques to bear on the problem immediately at hand.

There have been attempts to define learning as something done voluntarily or in schools and to define training as something required, usually by the company. Those definitions are shamefully narrow, meaningless, and useless.

Various researchers in the education/training disciplines have at¬tempted to argue that liking and learning are closely interlinked; but one author acknowledged that in reading about 800 papers in preparing his own monograph, he could argue his thesis in terms of trends but could provide "no conclusive evidence, pro or con." His readings include the papers of Zimbardo and Ebbesen, who deal in the affective domain (attitude).

In the 1960s Edgar Dale developed what he called the Cone of Learning for Cognitive and Affective Learning. A pyramid rising from the most concrete experiences, his base, to the most abstract, his apex, Dale's all-circumstances/all-media assessment reads: Direct, purposeful experiences; contrived experiences; dramatized experiences; demonstrations; study trips; exhibits; educational TV/motion pictures; still pictures; radio and recordings; visual symbols; verbal symbols.

Researchers building on Dale's thesis have hypothesized that differences in response to concrete/abstract media vary by age of learner. To establish attitudes in adults or change attitudes in young people, use abstracts. To change attitudes in adults or establish attitudes in young people, use the concrete. These postulates, too, are incon-clusive.

The inconclusiveness of the research makes perfect sense when we consider the variables: the topic, the individual learners, the concreteness or abstraction of media, and the cognitive or attitudinal or dexterity nature of the new behavior.

Changed (occasionally, reinforced) behavior is, after all, the end point of all the information passed around in meetings. And behavior springs not only from the animal drives that are so important in the theories of such behaviorists as John B. Watson and B. F. Skinner but from nonissue motives based on values and interests. Values and interests can cause--or motivate--an individual to act against animal nature, and so hopelessly confound the strict behaviorists. Because of the cognitive contribution the listener must make to the material spoken, he or she must maintain an active interest throughout the presentation if he is to comprehend. Motivating the listener is your job too.

But motivation is different in every individual and the priorities of each of us change with circumstances and often over time as well. So in Chapter 7 we will deal with motivation in the context of other human relations theories, and in Chapter 8 we will check the substantial differences between incentive and motivation.

In the balance of this chapter, we will deal with the practical aspects of visualizing for the meeting room.

The message presents management's request/directive to the meeting participant, who on understanding and comprehending might agree to cooperate. If so, he must be provided with the tools with which to do the job requested or directed.

Those tools, which are tangible or otherwise sensate implements, make possible a direct action by the learner upon a material or a concept, such as a problem. Tools, therefore, have an intrinsic value in relation to the message. They might tell the learner how to do some¬thing or describe size or shape or usage, or verbalize or demonstrate a product's features, advantages, and benefits, which are not always obvious, tangible, or sensate.

The tools you might create to support your message could be slot¬ted into Dale's cone of learning. One problem is that he has listed learning sources which cannot necessarily be accommodated in the meeting room--such as field trips. Differentiate between the overall convention and the basic agenda-directed meeting. Another problem is that the hierarchy tends to imply some sort of superiority from one medium to another, as some interpret it.

In March 1970 issue of Advertising & Sales Promotion magazine we presented an assessment of practical meeting room tools (yes, they can be called media) evaluated for their degree of credibility/manipulability/accessibility from the learner's viewpoint. Unlike Dale's cone, the tools are not ranked because their practical value depends on the specific circumstances of usage each time.

Briefly, the categories we favor include these: (1) Products and real objects: highest credibility; but size can be a problem and benefits can be insensate. (2) Charts drawings, chalkboards: lend immediacy and flexibility when generated but are difficult to view by crowds and take valuable time to draw. If prepared in advance, they are static images, in the sense of photographs. (3) Photographs: substitutes for the genuine article when cost or size is too impractical to permit distribution. Static exhibits are 3-D photos, both less credible than reality. (4) Books, booklets, outlines, tape recordings: orderly presentations of selected factual information and opinions delivered complete; but such presentation lacks immediacy and is impossible to control if printed matter is distributed in a large audience. (5) Demonstrations and sociodramas, plus exhibits that contain one of these: highly convincing when performed live as a form of vicarious practice; but need generous lead time and subject to human error; (6) Role playing: unmatched for creating participant insight into other people's behavior, but must be supervised by a trained leader; explosive, unpredictable, powerful. (7) Projected images of any of the above (except role playing) on a screen: excellent for magnifying detail, conveying emotion, and changing time relationships in movie form; otherwise one step removed from the conviction level of the tool for which it is a substitute.

Notice that by eschewing rank order of any type we are free to choose that medium that best serves the dictates of the given message in the context of the particular product and meeting room. There is no other way to choose responsibly.

All of the foregoing lend themselves to audiovisual form of presentation. When used that way, they are subject to the same types of strictures we've established for the traditional film/video A/V presentation, based on the psychological studies previously discussed.

Here are a few of our own premises for the optimum use of audio¬visual materials:

Premise 1: All visuals should present basic concepts when it's possible to do so and roadmaps for the auditory portion when concepts cannot be visualized.

Premise 2: Visuals should interpret spoken material, not merely repeat it. Words should appear on screen if the word is strange or part of a series; but the printed word appeals to the left hemisphere (duplicating the spoken word) while doing absolutely nothing for the conceptualization of the right hemisphere.

Premise 3: Avoid illustrating a script in a mistaken bid to conceptualize. If you are discussing winemaking, the use of grapes, vineyards, and bottles does nothing to aid the understanding of fermentation, the chemical process at issue.

Premise 4: Simplify the visuals. In photos of complex subject matter, reference the exact information needed by masking or high¬lighting specific areas. Overly complex visuals tend to overload the learner 5 circuit and split his concentration.

Premise 5: When conceptual slides are available, open the related commentary with a conceptual statement which agrees with the slide's content. Then go back and fill in details for the ear as the eyes studies selected areas.

Premise 6: Keep the auditory and visual portions wholly compatible so they can function synergistically--each aiding the other for heightened clarity and viewer understanding of intended meanings.

Premise 7: If you can add elements of theme or style without violating any of the other premises, and if you can afford the possible additional cost, add them. But do not overpower your message!

That's it. These simple rules will assure you an audio-visual presentation which has every right to succeed in conveying your message effectively and efficiently.

It all comes down to judgment. Perspective. Understanding of your needs, your people, and your alternatives. While the ads say your program needs gee-whiz media and intimidating hotel decor, the truth is that your people really need honest information about your ideas, your intentions, and your view of how it affects them. The more simply and completely you can state your case, the more likely they are to cooperate.

Now, given the research we've reviewed and the seven rules we've advanced to aid the A/V processing, can you "buy" the pitch that your meeting can't survive without the hype, the glitz, and the blitz of maximedia?

If you can't buy the pitch, why buy the media?