Designing with Sound to Enhance Learning: Four Recommendations from the Film Industry

MJ Bishop, Lehigh University
David Sonnenschein, Sound Design for Pros

ABSTRACT: While we rely heavily upon sound to understand our environments, instructional designers often overlook using sound to facilitate learners' selection, analysis, and synthesis of material under study. This neglect appears to be a matter of just not knowing how to sonify instructional designs to enhance learning. In contrast, increasingly more advanced and refined degree of film sound use has changed the way audiences experience and understand spectacle and storytelling in contemporary cinema. This paper explores what recommendations the film industry might have for instructional designers about ways sound can be designed to help enhance learning from their products.

Keywords: sound design, instructional design, instructional message design, media production, multimedia instruction

Introduction

Sound technologies have always trailed behind visual technologies. Sound was not recorded until the late 19th century, thousands of years after the first images were recorded (Altman, 1992). Sound was not studied as a physical phenomenon until the 1920s and '30s (Blauert, 1983). Sound's role in film was not even discussed until Hollywood released the first "talkie" in 1927 —a full 25 years after sound cinema was mechanically possible (Williams, 1992). And the technological barriers that had prevented the full integration of sound into all types of computer software were not overcome until the early 1990s.

So, it is hardly surprising that little attention has been paid over the years to sound's role in the user-computer interface (Buxton, Gaver, & Bly, 1987; Mountford & Gaver, 1990) and that even less attention has been paid to sound's potential contributions in computerized instructional environments (Barker, 1986; Mann, 1992, 1995). According to Bishop, Amankwatia, and Cates (2008), recently published instructional software programs are not using sound very extensively to support learning. When sound is incorporated, it appears to be used mostly as an attention-getting device or to narrate what might have otherwise been done just with text.

Further, it appears few guidelines are available for those instructional designers who are interested in finding theoretical and conceptual direction for incorporating sound most effectively. Even newer texts published since sound's use has become more technologically feasible provide little guidance on how to design with sound or its appropriate use. For example, Morrison, Ross, and Kemp's *Designing Effective Instruction* (2001) included a section on designing with graphics, but said nothing about designing with sound. Galitz's *The Essential Guide to User Interface Design: An Introduction to GUI Design Principles and Techniques* (2002) dedicated a mere 4 of 730 pages to sound's instructional use, addressing only sound's potential role in supplying verbal redundancy and facilitating dual coding. And, while Clark and Mayer's *e-Learning and the Science of Instruction* (2003) devoted 29 of 293 pages to sound's use, the only sound type considered in 25 of those pages was speech. The remaining 4 pages focused on avoiding the use of "extraneous" background music and environmental sound effects, without suggesting ways in which non-speech sounds might be used to enhance learning. Generally, the authors of instructional design guidelines seem to recommend that sound's major function --other than supplying occasional bells and whistles to gain attention-- should be either to narrate screen text or to provide stand-alone audio examples (like a musical performance or an historical speech).

In contrast, sound is used extensively in the film industry to enhance the storyline or narrative flow by establishing mood, time, location or period; adding pace, excitement, and impact; completing the illusion of reality or fantasy; creating the impression of spatial depth; and adding continuity between a number of discontinuous shots edited together (Wyatt & Amyes, 2005). Increasingly, audiences have come to expect a more advanced and refined degree of film sound use, which has changed the way they experience and understand spectacle and storytelling in contemporary cinema (Whittington, 2007). As a result, sound editing has become an established career in Hollywood and a number of recognized academic institutions, like the University of Southern California, offer graduate degree programs in the field. Further, numerous books on sound design for cinema, like Sonnenschein's (2001) *Sound Design*, help filmmakers understand the expressive power of music, voice, and sound effects and provide concrete ideas for creatively using sound to enhance the filmgoer's experience.

If successful film sound design requires this much expertise, it appears there may be more to enhancing learning through instructional sound design than simply adding sounds as afterthoughts. Instructional designers need guidance on the psychoacoustical ways humans interact with sound and practical guidelines for how to create/select and integrate sounds in ways that will capitalize on its affordances. Thus, exploring what we can learn from the film industry about how sound can help convey instructional messages more effectively and efficiently has been the motivation behind our recent collaborations. The first author is an instructional technology professor with many years of experience in instructional software design and development and the second author is an experienced musician, filmmaker, and sound designer who has also done extensive research in the areas of psychoacoustics, the human voice, Gestalt psychology, and therapeutic uses of sound. In this paper, we explore the ways in which sounds can support learning, present four recommendations derived from "best practices" of film industry sound design and apply them to the process of designing instructional technologies that make optimal use of sound to enhance learning.

How Sounds Can Support Learning

Our discussion should begin by explaining that when we say *sounds*, we are talking about all kinds of auditory stimuli --music, voice, and environmental sounds. While definitions are difficult and the distinctions between each category can get quite fuzzy, by *music* we are referring to the deliberate organization of sound into longer harmonic, melodic, and rhythmic passages; by *voice* we mean the human articulation of auditory language or any other sounds made with the tongue, lips, and other speech organs; and by *environmental sounds* we mean all the other non-musical and non-voice sounds that things make as part of the actions and events that occur within an environment. Sounds can support learning by facilitating cognitive processing in a variety of ways.

For example, sounds are particularly good at gaining attention because, unlike eyes, ears can never be averted or shut with "earlids" (McDonald, Teder-Sälejärvl, & Millyard, 2000; Wickens, 1984). Sounds are generally more effective than images for gaining and focusing attention (Schmitt, Postma, & de Haan, 2000). Sounds like a far-away siren or the whine of a puppy can be particularly effective in focusing our attention by immediately activating existing images and schemas (Bernstein, Chu, Briggs, & Schurman, 1973; Bernstein, Clark, & Edelstein, 1969a, 1969b, Bernstein & Edelstein, 1971; Bernstein, Rose, & Ashe, 1970a, 1970b). Other sounds —like the wind rustling leaves or inspirational music— can hold our attention by making our environment more tangible or by arousing our emotions (Thomas & Johnston, 1984). Thus, sounds might be used not only to gain attention, but also to help focus attention on appropriate information and engage a learner's interest over time.

Sounds provide a context within which individuals can consolidate, elaborate upon, and organize details about their surroundings, thinking actively about connections between and among new information (McDonald, Teder-Sälejärvl, Millyard, 2000; Schmitt, Postma, & de Haan, 2000; Stein, London, Wilkinson, & Price, 1996). Sounds like the steely clank of a metal door closing or a liquid being close to the top of its container supply us with volumes of complex information that we easily interpret in order to extrapolate important details about the world around us (door is not wooden) and make decisions about how to respond (stop pouring) (McAdams, 1993; Perkins, 1983). Like visuals that form hierarchical clusters organized in space, combinations of sounds also form hierarchical clusters –they are just organized in time (Winn, 1993). According to Bregman (1990), temporal organizational clues within a composite of sounds —like a factory operating, a person speaking, a helicopter flying, a truck idling, and a motorcycle running— allow most people to ascertain almost instantly that five sound sources are present, to determine each source's identity, and to locate the sources spatially. In these ways, sounds might also be used to help learners organize and see interconnections among new pieces of information.

Sounds also help individuals tie into, build upon, and expand existing constructs in order to help relate new information to a larger system of conceptual knowledge. According to Gaver (1986; 1989; 1993a; 1993b; 1993c, 1994), we often compare what we are hearing to our episodic and semantic memories for the sounds objects make in order to draw from and link to existing constructs and schemas to support our understanding of what is happening. The metaphorical language we later use to describe these sounds provides us with the means to discuss the experience with others and to transfer this new knowledge to new situations, which can develop even deeper understandings. Consider, for example, "The baby wailed like a siren;" "the mindless bureaucrat squawked like a parrot;" and "the coward squealed like a pig." Thus, it appears sounds might also be used to provide a familiar setting within which learners can relate incoming information to existing knowledge (Winn, 1993; Yost, 1993).

The Recommendations

Given that sound may hold great promise to support learner's selection, analysis, and synthesis of new information, how might instructional designers begin to think more systematically about sound's appropriate use in instructional products in order to capitalize on its affordances? What follows, below, are four recommendations for designing the sound track of a learning environment, which may include music, voice, and/or environmental sounds in any combination.

#1: Consider Sound's Use from the Start of the Design Process

In the film industry, a sound designer's job is to support the storytelling by invoking myth, suspending reality, and creating emotion –to auditorially enhance the audience's ability to perceive what is happening and to assure that it is, indeed, registering in the brain. With the sound designer's contribution, the audience will be led down the path in an integrated, yet most often subconscious manner toward an experience that is authentic and human, a metaphor for the life experience itself. Achieving this goal requires that the sound designer be totally immersed in the story, characters, emotions, settings, theme, and genre of the film.

In order to create audio/visual environments that engage an audience, therefore, the film sound designer must be involved from the start of production. During shooting and editing, the sound designer can offer extremely beneficial advice. For example, he or she might suggest

removing shots that might otherwise be substituted with dialog; or incorporating a counterpoint ambient sound to help lift a voice into more definition (such as a puppy barking against the miserly growl of a gruff old man); or adding non-verbal vocal sounds to amplify a plot point (like adding a wheeze when the character's back is turned to emphasize his worsening disease); or specifying music, where appropriate, and how it might be integrated with the other audio elements.

Like film sound design, the sounds used in instructional materials could be used to support the storytelling by helping learners acquire, organize, and synthesize the material under study (Bishop, 2000; Bishop & Cates, 2001). But selecting or creating the right sounds to achieve these outcomes requires careful harmonization with the subject matter, learner characteristics, pedagogical strategy, and learning objectives for the instruction being developed. Taking our cue from the film industry then, in order to design auditory elements aimed at enhancing the learning environment, it appears best to plan for them from the start of the design process. As with film, the ideas that emerge from early thinking about sound's potential contribution to the instruction can influence other aspects of the overall production as well, such as identifying ways sounds might be combined with still images to convey a larger, moving concept without the need for animation or video (photo of a rainforest combined with bulldozer sounds to represent disregard for the environment); suggesting the overall theme for instruction (lively '60s dance music inspiring the "mod" theme for a unit on modular mathematics); or specifying environmental sounds coupled with visuals that can then be repurposed to reinforce concepts covered (like the animation and sound of an arrow hitting –and then missing– a target to reinforce the concept of accuracy throughout an information literacy lesson).

#2: Identify Key Storytelling Elements to be Amplified By Sound

Ideally, film sound designers begin their work with an initial reading of the script –well before the film has been shot. Throughout this initial reading of the script, the sound designer is "listening" for objects, actions, environments, emotions, and physical or dramatic transitions that can be fleshed out auditorially using the various sound types. Even if the film has been shot, sound designers still will often avoid viewing the dailies until reading the script in order to prevent being influenced by the impression that the visuals will make, potentially short-circuiting their creative process.

Next, the film sound designer meets with the director in order to confirm his/her impressions after the initial script reading and to learn more about the director's artistic intent for the film. The purpose of this meeting is for the sound designer to come away with a firm understanding of the film's key storytelling elements that might be amplified by sound, such as the subject (love story or war story?), genre (comedy or horror?), theme or message (such as "Crime doesn't pay" or "If you try, you will succeed"), and underlying conflict that will drive the story (honesty vs. dishonesty, good vs. evil, and the like). It is important this meeting be a brainstorming opportunity during which the sound designer and director take turns proposing alternate scenarios that provoke additional creative ideas in the other. Through the collaborative paradigm of thesis-antithesis-synthesis within these dialogs, some of the richest ideas can be generated for using sound to enhance the audience's experience with the film.

Thus, whether you are the sole instructional designer/developer with the freedom to do as you please or a member of a design team taking direction from a lead designer, it appears that thinking about instructional sound design should begin immediately after the initial front-end analysis or scripting of the learning context. It is at this point when most instructional design process models move into a production or design phase that systematically specifies the pedagogical, media, and technical strategies to be employed – essentially, what the learner's experience with the instruction will be (Clark & Lyons, 2004; Smith & Ragan, 1999). The individual(s) responsible for media production should carefully review the entire design document (including the findings from the front-end analysis) and, while reading, consider ways that the various sound types might be employed within the specific learning context to facilitate learning the material under study.

Further delineating media specifications for the instruction, particularly the role sounds might play in the overall production, might then be best accomplished in a brainstorming session where at least two members of the design team, likely the lead instructional designer and instructional media developer, carefully listen to and learn from the other's ideas. If the lead instructional designer and instructional media developer are one in the same person, we suggest asking some other colleague to read through the design document and work through some ideas with you. Similar to film production, the nature of the collaboration between these individuals at this stage of instructional development will vary depending on their expertise, prior experiences, and the specified division of responsibilities within the team (Dick, Carey, & Carey, 2001). However this relationship unfolds, it is important that the person responsible for developing the instructional media for the production comes away with a firm understanding of the overall vision for the instruction, including the nature of the subject matter to be learned (such as declarative knowledge, intellectual skills, cognitive strategies, attitudes, or psychomotor skills) (Gagné, 1985); the primary and any secondary learning objectives for the instruction; the instructional strategies likely to be employed (direct instruction versus exploratory learning); and any areas of potential difficulty that the designer anticipates learners might have with the instruction. With these key storytelling elements for the instruction identified, the media developer can begin matching them up with sounds that might help to amplify key concepts, scaffold organizational structures, and/or relate the material under study to learners' existing constructs.

#3: Capitalize On the Way People Listen to Sounds

But it may be that the type of sound used in an instructional product is less important than the kind of listening it encourages (Gaver, 1993, 1989). Borrowing from Michel Chion, the French film theorist, Sonnenschein (2001) distinguished between four types of listening modes: reduced, causal, semantic, and referential. Each is described in more detail below.

Reduced: Reduced listening involves listening only to the qualities of the sound itself and not the source or its meaning. This is how a sound engineer listens, describes, and manipulates sounds through filtering, processing, and mixing. For example, in reduced listening the sound of an ambulance siren would be described as being loud, varying pitch, simple timbre, long duration, and the like.

Reduced listening is contingent on one's ability to discern even small variations in sound quality. Despite the fact that we are primarily visually oriented and, consequently, the visual sciences have usually dominated the corresponding auditory sciences, we really do also have a very high level of auditory acuity. We can easily differentiate sounds as they vary across sound qualities --intensity, pitch, timbre, speed, rhythm, shape, reverb/echo, directionality, and harmony. Our ability to perceive these parameters and the associations we have with the "bi-polarities" at their extremes is governed by the capabilities and limitations of our hearing (see Table 1). Consequently, novel changes in sounds that encourage reduced listening in an instructional environment can be particularly good at gaining learners' attention, focusing it on particularly important content points, and holding it over time (Bernstein, Chu, Briggs, & Schurman, 1973; Kohfeld, 1971; Posner, Nissen, & Klein, 1976; Thomas & Johnston, 1984).

Causal: Causal listeners are gathering information about a sound's cause, both in terms of its source and environment. In causal listening one identifies the sound's source and places it into a descriptive category that is either personally or culturally significant. We label the sound with a recognizable name or word that is useful for communicating verbally or in text form (my mother, a dog, a Harley Davidson motorcycle). This label for the sound can also include a description of the environment that may be influencing the nature of the sound as in, for example, a shower. Water pouring into a bath with hard walls that reflect the sound will help the listener identify this as a shower, hearing both the water itself and the place where it is falling.

Unlike film, which demands a certain amount of realism (when a bus crosses through the screen from left to right the audience expects to hear a bus sound moving from left to right), there are many elements within a technology interface that have no natural sound, leaving the media developer free to create his/her own sound for that event –just what does a button click on a computer screen sound like, for example? In fact, within the context of your interface a real world sound may be inadequate. It may be necessary to embellish upon the sound chosen to convey the exact idea. Sounds used within a technology interface do not need to be realistic but should facilitate causal listening by being strongly associated (at least initially) with an on-screen event, then consistent and expected (Laurel, 1993). Further, sounds that encourage causal listening can help orient learners in complex learning environments. A sound made in the real world contains cues that help us to localize or judge very precisely its source's distance (using sound's overall volume, for example) and position (relying primarily upon the ratio of the sound's volume between our left and right ears) (McAdams & Bigand, 1993). When possible, try using the spatial cues in sounds to help learners with their visual searches for the sound's source in the interface.

Semantic: Semantic listening involves processing the auditory code systems (like language) that symbolize things, actions, ideas, and emotions in order to determine the meaning of a sound. Semantic listening includes both informational and emotional communication. For example, a voice will transmit information through the symbols of words as well emotion through the melody (or prosody) of the phrase. A child will know when his parent is angry or pleased by both the intonation and verbal ideas presented (and may sense confusion if these are contradictory). An ambulance siren will have different semantics depending on who and where the listener is: a) coming from behind a driver, says "Pull over to the side" b) passing on a cross street far ahead of a driver, says "Slow down" c) driving past a pedestrian, says "You're okay, we're helping someone else in need."

Table 1: Listener's auditory acuity and associations with sound quality extremes (American National Standards Institute, 1973; McAdams & Bigand, 1993; Levitin, 2006).

Sound Quality	Human Perception	Associations
Intensity (soft/loud): Perceived volume of sound.	Intensity is measured in energy increments called decibels (dB), a logarithmic scale of sound energy with each ten points representing ten times the loudness. Humans can hear over a 120dB dynamic range and discern sound volume changes of less than one decibel.	Soft sounds can be soothing (a babbling brook) or imply weakness (the whisper of a terminally ill patient). Loud sounds can be irritating (a wailing siren) or convey great strength (the opening measures of Beethoven's 9^{th} Symphony).
Pitch (high/low): Degree of highness or lowness of a tone governed by the rate of vibrations producing it.	Humans' ability to perceive pitch in normal hearing ranges from about 20Hz to 20,000Hz (Hertz, or cycles per second).	High pitches can suggest small size (the squeak of a mouse) or youth (the happy squeal of a baby); low pitches suggest large size (the rumbling of a ship's engine) and have been known to cause feelings of awe or fear.
Timbre (simple/complex): Quality of sound distinct from its intensity and pitch.	Sound waves pulsing at regular intervals create a pure tonal or simple sound (a flute), as opposed to a noisy or complex sound (an explosion) made of overlapping and intermingling frequencies that produce highly complicated waveforms. With experience, humans identify sound sources at a surprising level of acuity.	Through traditional folk associations, metaphorical interpretations of sound quality, and repeated use, certain timbres have become associated with particular moods, emotions or situations (harp = angel and oboe = pastoral, for example).
Speed (fast/slow): Speed with which acoustic impulses are repeated.	At the upper extreme of 20 beats per second, individual sounds begin to blur into a steady pitch (or low frequency). At the lower extreme, resting cardiac pulse and the lethargic march of a funeral procession are examples of slow forms of sound making.	Fast sounds convey a sense of urgency (the sound of running footsteps) or excitement (the rapid speech of a surprised child); slow sounds can convey a lack of urgency (the sound of leisurely footsteps) or disinterest (the sluggish speech of a bored child).
Rhythm (ordered/chaotic): A strong, regular, repeated pattern of sound over time.	Ranging from an absolutely regular clock tick or resting heartbeat (ordered) to the spastic squeals of feeding pigs or the cacophony of a bicycle crash (chaotic). Humans can detect rhythm changes in the low millisecond range.	Ordered rhythms can lend a certain tranquility and assuredness, or nagging oppression. Chaotic rhythms can keep one alert, frightened, confused, or in fits of laughter.
Shape (impulsive/reverberant): Defined by its attack (onset, growth), body (steady-state, duration), and decay (fall-off, termination).	Ranging from more impulsive beginning rapidly, peaking, and decaying rapidly to more reverberant – gradually rising and falling. Listener's perception of sound shape depends not only on the waveform created by the source, but also on the distance and reverberation properties of the surrounding space.	Listeners tend to think of sounds with more impulsive shapes to be more "spontaneous" and "short-lived" (gunshots, slaps, door slams) whereas sounds with more reverberant shapes are more "deliberate" and "persistent" (a dog growl, slowly tearing a sheet of paper, far-away thunder).
Reverb/echo (dry/wet): Governed by types of surfaces in the physical environment that reflect and absorb the sound waves, and the distance of sound to these surfaces.	Reverberations are diffuse, reflected from complex surfaces and have no distinguishable repetitions (e.g. concert hall, cathedral, hard-walled living room), whereas echoes are discrete repetition of a sound based on simple surface geometry (e.g. stone canyon, sewage drain, exterior building wall). Carpets, curtains, foliage or dirt are highly absorbing and create an acoustically dry or "dead" space (e.g. a bedroom). At the other extreme, a hard surface like stone, glass, concrete or polished wood will have a high degree of reflectivity that will create a very wet or "live" space (e.g. a tiled bathroom).	Because the proximity of the surfaces will determine the rate of decay of the reverb or echo, listeners tend to relate shorter decays to smaller spaces and longer decays to larger spaces. In addition to the physical clues revealed by reverb and echo, these can also indicate a change in subjective space --for example, internal thoughts or dream sequences in a film.
Directionality (narrow/wide): The source of the sound may emanate from a narrow, specific region in the acoustic space (monoaural) or from a widespread area (stereophonic or surround sound).	Humans rely on specific cues within the sound that help to "localize" or judge very precisely a source's distance (using sound's overall volume, for example) and position (relying primarily upon the ratio of the sound's volume between our left and right ears).	A single source (e.g. human voice, bird, car) is normally identified as coming from one narrow point in space, whereas multiple sources (e.g. crowd murmur, forest ambience, traffic) are originating from a wide, non-specific place. The movement of a single source from left to right, for example, can widen the directionality of the sound over time, which makes the listener believe the acoustic event is happening over a larger space.
Harmony (consonant/dissonant): *The relationship between two or more different pitches.*	A pitch of 440 Hz together with its octave 880Hz is perceived a extremely consonant and pleasing to the ear, whereas a 440 Hz with a 450 Hz pitch will be very dissonant. The principle of harmony can also be applied to any cluster of non-musical sounds that have distinguishable fundamental frequencies (e.g. individual human voice, telephone ring, bird songs), but not to those with extremely complex timbres (e.g. crowd murmur, waterfall, rustling leaves).	Consonant sounds tend to be perceived as soothing, ordered, and aesthetically pleasing. Dissonance between sounds can lead to feelings of tension, confusion, and displeasure.

Semantic listening makes it possible to create an *auditory syntax* for particular concepts and variations on that concept within a learning environment. By auditory syntax we mean establishing and repeating a set of consistently used sounds and rules for their use that helps learners more easily understand their connections and their relations. The power of this sound design approach to facilitate learning is even greater if the syntax is consonant with the theme or metaphor for the instruction and the content. For example, if one were to use eating sounds throughout a course on relevance ("ptewie" for irrelevant and "chomp" for relevant), one could vary those sounds according to the sort of source being "eaten." Sources that are clearly relevant might make a satisfying "munch" sound whereas sources that are relevant but need to be "softened up a bit" to be appropriate for the assignment's audience might make a teeth-shattering "crunch" sound. Once the syntax has been established, you can then bring back echoes of it later with new feel and meaning to establish new paradigms (transfer) (Emmert & Donaghy, 1981; Fiske, 1990).

Referential: Referential listening involves listening to the context of a sound, linking not only to the source but principally to the emotional and dramatic meaning as well (internal and external). Referential sounds can be *universal* or applicable anywhere, anytime (such as breathing, heartbeat, or wind) or *specific* to a particular setting, culture, or time period (such as pine forest sounds, Moroccan marketplace sounds, or pre-Industrial sounds). These sounds tend to mentally refer us to the person or objects making them. This referential listening is, perhaps, most clearly evidenced by the language we use to describe these sorts of sounds: "a monotonous speaker," "a screeching violin," "a squeaking door," and the like. One must take care to avoid, however, using sounds that evoke "unintended" references based on personal interpretation, multicultural differences, and other prior experiences. For example, it will likely be some time before sound designers can incorporate rhythmic breathing through a scuba tank into their productions without listeners equating the sound to *Star Wars'* Darth Vader.

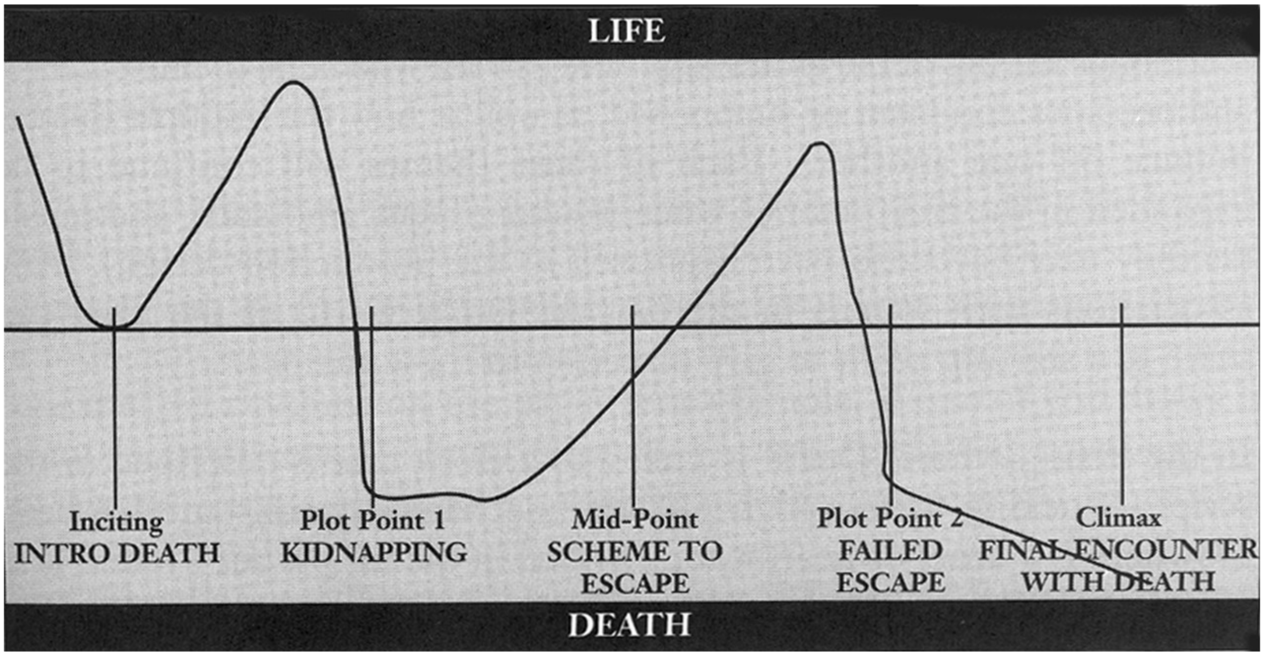
Because referential sounds can be highly image evoking, educators should consider exploiting this dual-coding to help learners process the material under study more deeply (Paivio, 1986). Also, since referential sound effects easily evoke images of familiar things, they might be used to augment or establish a mental model or metaphor for the content under study. Metaphors help us to understand new information by putting it in terms of other, more familiar information, without ever directly stating the comparison (Lakoff & Johnson, 1980). For example, one might consider using the sound of a racing automobile to accompany a cartoon character's hasty retreat in order to build upon learners' understanding of the action. Although mental models more directly state the comparison, they are very much like metaphors in that they help us to understand new information by putting it in terms of other, more familiar information. As with metaphors, referential sounds might be used to help build mental models within the learning environment because they easily evoke images of familiar things.

#4: Be Systematic About How Sounds Are Incorporated

Sound design for film is fairly linear and flows from scene to scene. The story is always dictated by a beginning, a middle, and an end (or climax). Sound designers work systematically within this framework and literally map out across a timeline when and where chosen sound groupings (or voices) will be incorporated as the story unfolds (see Figure 1). While instruction tends not to be quite as linear (not even direct instruction is entirely linear), the nature of human learning and the fundamental order of how we present content within a learning environment does typically flow from "select, analyze, to synthesize" in a way that could be considered to be analogous to film's "beginning, middle, and end."

Recognizing the need to provide a more complete picture of sound's instructional potential, Bishop (2000) has suggested a framework for thinking systematically about designing instruction with sound that is based on this select, analyze, and synthesize flow of instruction (see Table 2, Bishop, 2000; Bishop & Cates, 2001). The framework's nine cells combine information-processing and communication theories to derive strategies for how music, voice, and environmental sounds might be used more effectively at each level of learning (see selection, analysis, synthesis rows) by facilitating information processing (see acquisition, processing, and retrieval columns). Following the cells vertically down the information-processing columns, the framework anticipates deepening attentional, organizational, and relational difficulties at each subsequent phase of learning (top to bottom). When tracing the cells horizontally across the

Figure 1. Visual map for film sound design (Sonnenschein, 2001).



learning phases, the framework similarly anticipates waning interest, curiosity, and engagement at each deeper level of processing (left to right).

Thus, when one traces the first, selection-level row of cells horizontally across the information processing stages, the framework suggests that learner interest may be captured by instruction that employs sound to gain attention with novelty (cell 1), to isolate information through increased salience (cell 2), and to tie into previous knowledge by evoking existing schemas (cell 3). Similarly, learner curiosity might be aroused using sound to focus attention by pointing out where to exert information-processing effort (cell 4), to organize information by differentiating between content points and main ideas (cell 5), and to build upon existing knowledge by situating the material under study within real-life or metaphorical scenarios (cell 6). Likewise, a learner's level of engagement might be increased using sounds to hold attention over time by making the lesson more relevant (cell 7), to elaborate upon information by supplying auditory images and mental models (cell 8), and to prepare knowledge for later use by providing additional knowledge structures that might be useful in subsequent learning (cell 9). When designed systematically into the instruction in this way, sound might supplement instruction by providing the additional content, context, and construct support necessary to overcome many of the acquisition, processing, and retrieval problems one might encounter while learning.

Conclusion

While it appears that human beings rely heavily upon sound to learn about their environments, instructional designers often make little use of auditory information in their computerized lessons. The prevailing attitude seems to be that, after all of an instructional software product's visual requirements are satisfied, the designer might then consider adding a few "bells and whistles" in order to gain the learner's attention from time to time (see, for example, Adams & Hamm, 1994; Brown, 1988; Reiser & Gagné, 1983; Shneiderman, 1998). While instructional designers need to stop ignoring this important channel for communicating instructional messages, this neglect of the auditory sense appears to be less a matter of choice and more a matter of just not knowing how to sonify instructional designs to enhance learning.

That said, it is important that sounds be used only insofar as they reduce the interface's cognitive load and contribute to the instruction —everything else is just noise. If sound has a larger role to play in instructional materials –as it does in film– its use should be planned from the start, well-grounded in key aspects of the material under study, predicated on the way learners' listen to sound, and incorporated systematically in a way that will facilitate learners' selection, analysis, and synthesis of the material under study.

Table 2. Sound-use instructional design strategies framework (Bishop, 2000; Bishop & Cates, 2001).

	ACQUISITION	PROCESSING	RETRIEVAL	
SELECTION	**1. Use sound to gain attention.** Employ novel, bizarre, and humorous auditory stimuli.	**2. Use sound to isolate information.** Group or simplify content information conveyed to help learners isolate and disambiguate message stimuli.	**3. Use sound to tie into previous knowledge.** Recall learner's memories and evoke existing schemas.	**INTERESTED**
ANALYSIS	**4. Use sound to focus attention.** Alert learners to content points by showing them where to exert information-processing effort.	**5. Use sound to organize information.** Help learners differentiate among content points and create a systematic auditory syntax for categorizing main ideas.	**6. Use sound to build upon existing knowledge.** Situate the learning within real-life or metaphorical scenarios.	**CURIOUS**
SYNTHESIS	**7. Use sound to hold attention over time.** Immerse learners by making them feel the content is relevant, by helping to make it more tangible, and by bolstering learner confidence.	**8. Use sound to elaborate upon information.** Supplement the content by supplying auditory images and mental models.	**9. Use sound to integrate with existing knowledge.** Help learners integrate new material into overall knowledge structures and prepare for transfer to new learning contexts.	**ENGAGED**
	ATTEND (CONTENT SUPPORT)	**ORGANIZE (CONTEXT SUPPORT)**	**RELATE (CONSTRUCT SUPPORT)**	

References

Adams, D., & Hamm, M. (1994). *New designs for teaching and learning: Promoting active learning in tomorrow's schools*. San Francisco: Jossey-Bass.

Altman, R. (1992). Material heterogeneity of recorded sound. In R. Altman (Ed.), *Sound theory, sound practice* (pp. 15-34). New York: Routledge.

American National Standards Institute (1973). *American national psychoacoustical terminology.* S3.20. New York: American Standards Association.

Barker, P. (1986). A practical introduction to authoring for computer-assisted instruction. Part 6: Interactive audio. *British Journal of Educational Technology, 17*, 110-128.

Bernstein, I. H., Chu, P. K., Briggs, P., & Schurman, D. L. (1973). Stimulus intensity and foreperiod effects in intersensory facilitation. *Journal of Experimental Psychology, 25*, 171-181.

Bernstein, I. H., Clark, M. H., & Edelstein, B. A. (1969a). Effects of an auditory signal on visual reaction time. Journal of Experimental Psychology, 80, 567-569.

Bernstein, I. H., Clark, M. H., & Edelstein, B. A. (1969b). Intermodal effects in choice reaction time Journal of Experimental Psychology, 81, 405-407.

Bernstein, I. H., & Edelstein, B. A. (1971). Effects of some variations in auditory input upon visual choice reaction time. Journal of Experimental Psychology, 87, 241-247.

Bernstein, I. H., Rose, R., & Asche, V. (1970a). Energy integration in intersensory facilitation. Journal of Experimental Psychology, 86, 196-203.

Bernstein, I. H., Rose, R., & Asche, V. (1970b). Preparatory state effects in intersensory facilitation. Psychonomic Science, 19, 113-114.

Bishop, M. J. (2000). *The systematic use of sound in multimedia instruction to enhance learning.*

Bishop, M.J., Amankwatia, T.B., & Cates, W.M. (2008). Sound's use in instructional software: A theory to practice content analysis. *Educational Technology Research & Development, 56*, 467-486.

Bishop, M.J., & Cates, W.M. (2001). Theoretical foundations for sound's use in multimedia instruction to enhance learning. *Educational Technology Research & Development, 49*(3), 5-22.

Blauert, J. (1983). *Spatial hearing: The psychophysics of human sound localization* (J. S. Allen, Trans.). Cambridge, MA: MIT Press (Original work published 1974).

Bregman, A.S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

Brown, C.M. (1988). *Human-computer interface design guidelines*. Norwood, NJ: Ablex.

Buxton, W., Gaver, W., & Bly, S. (1987). The audio channel. In R. Baecker & W. Buxton (Eds.), *Readings in human-computer interaction: A multidisciplinary approach* (pp. 393-399). Los Altos, CA: Morgan Kaufmann.

Clark, R.C., & Mayer, R.E. (2003). *e-learning and the science of instruction*. San Francisco: Pfeiffer.

Clark, R.C., & Lyons, C. (2004). *Graphics for learning: Proven guidelines for planning, designing, and evaluating visuals in training materials*. San Francisco: Pfeiffer.

Dick, W., Carey, L., & Carey, J.O. (2001). *The systematic design of instruction* (5th ed.). New York: Longman.

Emmert, P., & Donaghy, W.C. (1981). *Human communication: Elements and contexts*. Reading, MA: Addison-Wesley.

Fiske, J. (1990). *Introduction to communication studies* (2nd ed.). London: Routledge.

Galitz, W.O. (2002). *The essential guide to user interface design: An introduction to GUI design principles and techniques*. New York: Wiley.

Gagné, R.M. (1985). *The conditions of learning and theory of instruction* (4th ed.). New York: Holt, Rinehart, and Winston.

Gaver, W. (1986). Auditory icons: Using sound in computer interfaces. Human-computer Interaction, 2, 167-177.

Gaver, W. (1989). The SonicFinder: An interface that uses auditory icons. Human-computer Interaction, 4, 67-94.

Gaver, W. (1993a). What in the world do we hear? An ecological approach to auditory source perception. Ecological Psychology, 5, 1-29.

Gaver, W. (1993b). Synthesizing auditory icons. INTERCHI '93 conference proceedings: Conference on human factors and computing systems (pp. 228-235). Reading, MA: Addison-Wesley.

Gaver, W. W. (1993c). "Class, you're not making enough noise!" The case for sound-effects in educational software. In M. D. Brouwer-Janse and T. L. Harrington (Eds.), Human-machine communication for educational systems design (pp. 139-150). New York: Springer-Verlag.

Gaver, W. W. (1994). Using and creating auditory icons. In G. Kramer (Ed.), Auditory display sonification audification and auditory interfaces (pp. 417-446). Reading, MA: Addison-Wesley.

Kohfeld, D.L. (1971). Simple reaction time as a function of stimulus intensity in decibels of light and sound. *Journal of Experimental Psychology, 88*, 251-257.

Levitin, D.J. (2006). *This is your brain on music: The science of a human obsession*. New York: Penguin.

Mann, B. (1992). The SSF model: Structuring the functions of the sound attribute. *Canadian Journal of Educational Communication, 21*(1), 45-65.

Mann, B. (1995). Enhancing educational software with audio: Assigning structural and functional attributes from the SSF Model. *British Journal of Educational Technology, 26*(1), 16-29.

McAdams, S. (1993). Recognition of sound sources and events. In S. McAdams & E. Bigand (Eds.), *Thinking in sound* (pp. 146-198). New York: Oxford University.

McAdams, S., & Bigand, E. (1993). Introduction to auditory cognition. In S. McAdams & E. Bigand (Eds.),*Thinking in sound: The cognitive psychology of human audition*, (pp. 1-9). Oxford: Clarendon.

McDonald, J., Teder-Sälejärvl, W., & Millyard, S. (2000). Involuntary orientating to sound improves visual perception. *Nature, 407*, 906-908.

Mountford, S., & Gaver, W. (1990). Talking and listening to computers. In B. Laurel (Ed.), *The art of human-computer interface design* (pp. 319-334). Reading, MA: Addison-Wesley.

Morrison, G.R., Ross, A.M., & Kemp, J.E. (2001). *Designing effective instruction* (3rd ed.). New York: Wiley.

Paivio, A. (1986). *Mental representations: A dual encoding approach*. New York: Oxford University.

Perkins, M. (1983). *Sensing the world*. Indianapolis: Hackett.

Posner, M.I., Nissen, M.J., & Klein, R.M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review, 83*, 157-171.

Reiser, R.A., & Gagné, R.M. (1983). *Selecting media for instruction*. Englewood Cliffs, NJ: Educational Technology.

Schmitt, M., Postma, A. & de Haan, E. (2000). Interactions between exogenous auditory and visual spatial attention. *Quarterly Journal of Experimental Psychology, 53*, 105-130.

Shneiderman, B. (1998). *Designing the user interface: Strategies for effective human-computer interaction* (3rd ed.). Reading, MA: Addison-Wesley.

Smith, P.L., & Ragan, T.J. (1999). *Instructional design* (2nd ed.). New York: Wiley.

Sonnenschein, D. (2001). *Sound design: The expressive power of music, voice, and sound effects in cinema*. Studio City, CA: Wiese.

Stein, B.E., London, N., Wilkinson, L.K., & Price, D.D. (1996). Enhancement of perceived visual intensity by auditory stimuli: A psychophysical analysis. *Journal of Cognitive Neuroscience, 8*, 497-506.

Stein, B.E., & Meredith, M.A. (1993). *The merging of the senses.* Cambridge, MA: MIT Press.

Thomas, F., & Johnston, R. (1984). The Disney sounds. In W. Rawls (Ed.), *Disney animation: The illusion of life* (pp. 145-161). New York: Abbeville.

Wickens, C. (1984). *Engineering psychology and human performance*. Columbus, OH: Merrill.

Winn, W.D. (1993). Perception principles. In M. Fleming & W. H. Levie (Eds.), *Instructional message design: Principles from the behavioral and cognitive sciences* (2nd ed., pp. 55-126). Englewood Cliffs, NJ: Educational Technology.

Williams, A. (1992). Historical and theoretical issues in the coming of recorded sound to the cinema. In R. Altman (Ed.), *Sound theory, sound practice* (pp. 126-137). New York: Routledge.

Whittington, W. (2007). *Sound Design and Science Fiction*. Houston: University of Texas.

Wyatt, H., & Amyes, T. (2005). *Audio post production for television and film: An introduction to technology and techniques* (3rd ed.). Boston: Elsevier.

Yost, W.A. (1993). Overview: Psychoacoustics. In W. A. Yost, A. N. Popper, & R. R. Fay (Eds.), *Human psychophysics* (pp. 1-12). New York: Springer-Verlag.