Chapter Two

Review of Related Literature

Participant interactions with the audio components of IMMI and the effects that such interactions may have on music listening are the central concern of this work. Despite widespread evidence that listening is an essential foundation of musical experience (Haack, 1992, Reimer & Wright: 1992), the function of sound and the role of listening in computerised instruction are poorly understood (Aarntzen:1993, Barron and Kysilka:1993, Daniels: 1995). Nonetheless, the evolution of digital music technology and the use of computers in music instruction continue to gather momentum (Higgins: 1992, Berz and Bowman: 1994, Stevens: 1994, Bray: 1997). Concurrent with this growth is the development of new forms of interactive computerised instruction capable of delivering auditory, graphical, and textual information simultaneously. The term *interactive multimedia* is frequently used to describe this new method of computerised information exchange. However, the application of interactive multimedia in computerised music instruction is still at a rudimentary stage of development and the educational implications of delivering computerised music instruction in combined auditory and visual formats are not well understood. Knowledge of how participants interact with sound in conjunction with other IMMI media components is extremely limited, even though the primacy of sound and listening in musical experience suggests that such knowledge may be a critical factor in the development of research-based guidelines for the successful design and implementation of IMMI.

To provide a context for the present investigation, a representative sample of research evidence and theoretical opinion relating to the essential issues of the study will be examined. A discussion of terminology that is critical to the investigation precedes the body of the review of related literature, which is organised in sections that focus on
(1) the importance of listening in musical experience, (2) the influence of audio technology on music listening, (3) audio technology and computerised music instruction, (4) interactive multimedia and music instruction, (5) sound in computerised instruction, (6) human information processing and multimedia stimuli, (7) the effects of instructional flexibility, and (8) previous investigations of IMMI.

Terminology Relating to the Study

In the field of computerised instruction, where rapid technological evolution is the norm, the definition of many concepts is in a state of flux. Thus, a critical early concern of the present investigation was to define the following concepts that are germane to the study.

Interactive Multimedia

The term interactive multimedia is used to describe the form of computerised instruction currently under investigation. Interactive multimedia is a broad term that is closely associated with the related entities of hypertext and hypermedia (Ambron and Hooper: 1990). The terms hypertext and hypermedia are rooted in an original concept attributed to Bush (1945) who described a hypothetical device called a memex (memory extender) that would be capable of storing and indexing information so that associative links could be created between discrete pieces of data. Immediate and automatic transfer between any of the associated pieces of information would be possible when the user pressed a button which would cause linked information to be instantly displayed. Technological limitations prevented Bush from bringing his concept to fruition but, with the advent of computers in the 1960s, Bush's prototype served as a model for the development of computer-based hypertext, a term coined by Nelson (1967, 1987a, 1987b). The development and principal features of hypertext have been comprehensively reviewed by Conklin (1987) who reports that in essence, hypertext is an approach to database management which provides a structural network of linked information that allows users to cross-reference information about related concepts and retrieve it in non-linear
sequences. In early hypertext systems the database of "information" consisted almost exclusively of text. Gradually however, charts, maps, and other graphical items were interpolated into hypertext documents. In 1987, the release of a computer software program known as *HyperCard* heralded the widespread extension of hypertext principles to include moving images and sound as a part of the information database. This augmentation of hypertext principles became known as *hypermedia*.

The terms *hypertext*, *hypermedia*, *multimedia*, and *interactive multimedia* are used inconsistently in the literature and are often defined in contradictory ways (Gygi: 1990, Tolhurst: 1995). Tolhurst points out that while the terms *hypertext* and *hypermedia* are closely related, there are differences between them that relate primarily to the degree of animation and stasis in the media components they convey. According to Tolhurst, *hypertext* includes both text and static images, while the use of animated graphics, video segments, and sounds denotes *hypermedia*. She believes that while *multimedia* may include the same media formats found in *hypertext* and *hypermedia*, it differs from *hypertext* and *hypermedia* in the sense that it is not necessarily computer controlled. However, as *multimedia* is now increasingly computer–controlled and frequently includes non-linear interactive links, Tolhurst suggests that *multimedia* may "overlap the terms hypermedia and hypertext" (p. 25).

While the present study employs the term *interactive multimedia* in preference to the term *hypermedia*, the reader is cautioned that in related literature the term *hypermedia* is frequently encountered. The nomenclature adopted in the present investigation rests on the assertion that when the term *interactive* is used in conjunction with the term *multimedia*, the resulting descriptor (interactive multimedia) is effectively the same entity as *hypermedia*. As Reeves (1992) and Berz and Bowman (1994) observe, the terms *hypermedia* and *interactive multimedia* are often viewed as synonymous and are frequently used interchangeably.
In the current investigation, the term *interactive multimedia* describes a computing environment which results from the integration of diverse media elements such as graphics, moving images, sound, and text. The word *interactive* relates to an approach to computer software design which allows users to choose non-linear pathways through the program material and where there is a capacity for participants to modify elements of the program or the instructional sequence to suit personal needs. Thus, when the terms *interactive* and *multimedia* are combined in an educational context, the resulting descriptor indicates the integration of diverse media elements where users interact with program content and software design to produce computerised instruction that is tailored for the individual. In the present study, computerised music instruction programs that employ *interactive multimedia* are referred to as interactive multimedia music instruction (IMMI) programs.

**Interaction, Interactive and Interactivity**

While there is agreement in related literature about the importance of interaction in learning, there is little consensus about exactly what interaction is in computerised instruction or what it involves. The use of the terms *interaction*, and the related words *interactive*, and *interactivity* is inconsistent in the literature. Arriving at an operational definition of the construct of interaction as it applies to instructional technology is one of the major difficulties surrounding research in the area (Schwier and Misanchuck: 1993, Wagner: 1994).

Two broad categories of interaction are frequently identified in computerised instruction, namely, technological interaction and instructional interaction. Technological interactions are events which, for example, occur between a computer user and a computer. These interactions are sometimes referred to as "delivery system interactions". Wagner (1994) argues that the term *interactivity* should be used to refer to technological interactions. According to this view, *interactivity* relates to exchanges that occur between the user and
computer input devices such as keyboards, pointing devices such as a mouse, and output
devices such as computer monitors and sound cards. Sims (1995) supports this use of
the terminology and further defines *interactivity* as relating to the Human-Computer
Interface (HCI). The second category, instructional interaction, relates to events that
occur between the participant and the instructional content. Wagner (1994) maintains that
the term *interactive* should be used to refer to instructional interactions. Thus, *interactive*
events would encompass the cognitive transactions that occur between learners and the
instructional content of the computer programs that they use.

Many commentators have sought to define the process of interaction at a fundamental
level. For example, Jonassen (1988a) describes interaction simply as activity between two
organisms which results in a dialogue. Wagner (1994) reports that "simply stated,
interactions are reciprocal events that require at least two objects and two actions.
Interactions occur when these objects and events mutually influence one another" (p. 8).
Likewise, Barker (1994) defines an interactive system as

\[
\text{a collection of reactive objects which are able to pass messages to one another. Objects react to incoming messages by means of the various 'scripts' that are embedded within them; they then interact with other objects by means of further message passing activity. (p. 5)}
\]

He notes that this simple model of interactivity can be applied equally to both human-
human interaction and human-computer interaction. In an attempt to clarify the processes
of interaction, Gavora and Hannafin (1995) maintain that "operationally, human-
computer interactions involve four steps: presentation of eliciting stimuli, activation of
mental processes, production of appropriate physical responses, and presentation of
response-contingent feedback" (p. 448).
A functional view, which emphasises the various roles that interaction plays in computerised instruction systems, has been adopted by Hannafin (1989) who describes five functions of interaction: pacing, elaboration, confirmation, navigation, and inquiry. *Pacing* refers to the flow of the instructional session; *elaboration* involves the user in reflecting upon instructional material in a personal manner; *confirmation* interactions are used to check the accuracy of responses to problems posed during the session; *navigation* interactions relate to the choice of the instructional pathway; and *inquiry* interactions allow the user to search a database for information.

Emphasis is placed upon user activity by Ambron and Hooper (1990) who describe interaction as a state in which participants are able to browse, annotate, link, and elaborate within a rich, non-linear database. Likewise Reeves (1997) views interaction from the perspective of the human participant and suggests that

> a learning environment is 'interactive' in the sense that a person can navigate through it, select relevant information, respond to questions using computer input devices such as a keyboard, mouse, touch screen, or voice command system, solve problems, complete challenging tasks, create knowledge representations, collaborate with others near or at a distance, or otherwise engage in meaningful learning activities. (p. 5)

These descriptions of interaction do not, however, adequately address the computer’s capacity to analyse user input and generate appropriate responses. Berz and Bowman (1994) observe that "after actions are made, the user must receive appropriate and understandable feedback in order to maintain meaningful interaction" (p. 42). Jonassen (1985) believes that "interactive lessons are those in which the learner actively or overtly responds to information presented by the technology, which in turn adapts to the learner, a process more commonly referred to as *feedback*" (p. 7). The capacity of instructional software to assess user input and to modify the instructional procedure accordingly is an essential characteristic of successful interactive instructional design. Jonassen maintains...
that "when we speak of interactive instruction designs, we are usually referring to interactive *adaptive* designs" (p. 8).

The computer's role in interaction is further scrutinised by Wagner (1994) who expresses apprehension about the growing acceptance of a causal relationship between system interactivity and instructional interaction. She suggests that there is an unrealistic expectation about the capacity of so-called interactive technologies (such as interactive multimedia) to ensure that instructional interactions occur. This concern is consistent with the misgivings of Livengood (1987) who reports that

> interactivity is not the number of times the student presses the [SPACE BAR]. Interactivity is not X number of keystrokes per minute. Interactivity is not simply asking a question or two after presenting a certain amount of information. Interactivity is not just the communication between two machines (although interaction between system components does occur). (p. 28)

Thus, while enhancement of the level of technological interaction may be important, the critical factor from an instructional standpoint is the level of the learner's engagement or active involvement with the subject material (Dickinson: 1995, Spector: 1995). Jonassen (1985) notes that "the more mentally active learners are as they process information from learning materials, that is, interact with the materials they are trying to comprehend, the more likely they are to comprehend them" (p. 8). Regardless of the complexity or level of interactivity, the extent of the learner’s cognitive processing is a critical, but largely unknown factor that facilitates learning.

Human-computer interaction exists at different levels and while a high degree of technological interactivity may facilitate instructional interaction it does not necessarily guarantee that instructional interaction will occur. Multimedia information presentation is not inherently interactive and instructional designers need to employ complex forms of
interactivity if they are to stimulate instructional interaction and hence learning (Nelson and Palumbo: 1992, Sims: 1995). The level of complexity and the type of interaction is determined by the events that occur between the central components of interaction, namely, the user, the medium, the message, and the environment (Hanssen: 1996).

Commentators have identified various forms, levels, and types of interaction (Rhodes and Azbell: 1985, Thompson and Jorgensen: 1989, Jonassen: 1988c, Lucas: 1992, Schwier and Misanchuck: 1993, Sims: 1995, Tannenbaum, 1998). These observers employ differing nomenclature to describe anything from three to eleven forms of interaction. There is however some consensus in favour of a three-form model of interaction (Rhodes and Azbell: 1985, Lucas: 1992, Schwier and Misanchuck: 1993). The proponents of this model use similar terms to describe essentially the same three forms of interaction. The model developed by Schwier and Misanchuck, which has been adapted and summarised in Figure 2.1., uses the terms *reactive*, *proactive*, and *mutual*.

**Figure 2.1 Forms of Interaction in Computerised Instruction**

![Diagram showing the forms of interaction]

*Note.* Adapted from Rhodes and Azbell (1985) and Schwier and Misanchuck (1993).
According to this model, reactive interactions relate, for example, to situations in which users respond to computer prompts by pressing a key to advance the program. This reactive model of interaction has been widely used in drill-and-practice software to teach aural skills, music notation, and music theory (Peters:1992, Stevens: 1994). A second form of interaction in which users take a more active role in constructing the sequence of learning events is defined as proactive. In this form of interaction, learners go beyond selecting and responding to existing structures and begin to adapt instructional material to suit their individual needs. For example, a proactive level of interaction may allow the user to rearrange the instrumentation or formal structure of a piece of music. A third type of interaction, which Schwier and Misanchuck (1993) label as mutual, relates to the situation where "the learner and the [computer] system are mutually adaptive, that is, capable of changing in relation to encounters with the other" (p. 12). Such forms of interaction occur for example in artificial intelligence and virtual reality systems where the computer may anticipate users’ needs based on an analysis of their previous responses.

As the model shown in Figure 2.1 implies, the three forms of interaction are viewed as hierarchical and thus the quality of interaction at the mutual level is higher than at the proactive level, while a proactive interaction is of a higher level than a reactive interaction. Higher levels of user interaction (e.g., Mutual Interaction - upper right quadrant in Figure 2.1) are assumed to improve the educational potential of computerised instruction where the level of dialogue is seen as important because it encourages users to be continually and productively active while engaged with the instructional content (Schwier and Misanchuck: 1993, Spector: 1995). The three forms of interaction are not necessarily exclusive and interactive multimedia programs often employ combinations of them.

The IMMI programs under investigation in the current study allow users a considerable degree of freedom over the pacing, structure, and content of the instructional transaction.
Thus, according to the above three-form model of interaction, these programs offer a proactive form of interaction. While acknowledging the importance of the differing forms of interaction in computerised instruction, the present study does not attempt to distinguish participant interactions with the audio components of IMMI in relation to the three-form model. Rather, interaction is viewed from the perspective of the aforementioned two category model (technological interactivity - instructional interaction) proposed by Wagner (1994). It is thus interactivity that is the subject of this investigation and the study examines the technological interactions that occur between IMMI participants and computer-delivered music examples. The basis for this position rests in the belief that technological interaction with computer-delivered music examples is the first link in a sequence of events that facilitate instructional interaction. Whether these exchanges take the form of reactive, proactive, or mutual interaction remains the subject of future investigation. Since music listening in IMMI arises out of technological interaction, the current study aims to document and explain these technological interactions and their relationship to music listening.

CD-ROM and the Evolution of Interactive Multimedia Delivery Systems

Introduced in 1985, CD-ROM is a digital recording, storage, and delivery mechanism that is an adaptation of the audio Compact Disc (Galbreath: 1992, Parker:1993, Boumans:1996, Ghislandi: 1996, Tannenbaum:1998). Any information that can be digitised may be stored on CD-ROM discs, which may contain both audio and visual components. Various forms of CD-ROM exist which, when linked to a computer, facilitate the storage and retrieval of the large digital data files necessary to support interactive multimedia. Schwier (1994) reports that CD-ROMs can store up to 660 MB of digital information, and the rapid random access, easy portability, and relatively large storage capacity of CD-ROM have been instrumental in the early development of interactive multimedia.
The continuing evolution of digital technology means that CD-ROM is currently being superseded by enhanced CD formats such as Digital Versatile Disc (DVD), while network communication technologies such as the Internet are also beginning to play an important role in the delivery of interactive multimedia. Hayward and Orrock (1995) point to the increasing influence of the Internet, in particular the World Wide Web (WWW), and view CD-ROM as a transitional medium that will "soon assume the status of an historical curio…" (p. 76). Nevertheless, observers such as Boumans (1996) predict that CD-ROM technology will survive, albeit in enhanced formats that will function in complementary ways with the Internet.

Current limitations in the speed of data transmission restrict the delivery of interactive multimedia over the Internet, but marked improvements in data compression techniques and software capabilities, combined with on-going increases in data transmission speeds, continue to enhance the delivery of interactive multimedia on the WWW (Patterson and Melcher: 1998). Confirming evidence of progress in the delivery of interactive multimedia over the Internet can be seen with the recent arrival of hybrid technologies such as Web CD (Czencz and Hayward: 1997). The delivery of interactive multimedia using Web CD cleverly exploits the best features of both the Internet and CD-ROM. Memory intensive data can be accessed from a local CD-ROM, while time sensitive program information can be delivered and updated dynamically from the Internet. This arrangement helps avoid the long download times that are required for the delivery of high quality audio and video data across the Internet, while providing the potential to keep the product up to date through changes to program content which can be delivered dynamically from a Web site. Advances such as Web CD suggest that emerging network technologies will soon replace CD-ROM as the principal delivery mechanism of interactive multimedia.
Observers such as Clark (1983) and Hayward and Orrock (1995) reflect a commonly held belief that the mechanism by which interactive multimedia is delivered is of limited importance in instructional transactions, and that CD-ROM is therefore best regarded as a storage and retrieval mechanism rather than as an instructional device. The current study embraces the view that whether IMMI is delivered from CD-ROM or Web CD, the characteristics of participant interactions with its audio components will remain essentially unaltered. Hence, while the current investigation examines IMMI that is delivered from CD-ROM, the assessment and analysis procedures developed for the study are likely to produce results with substantial relevance in circumstances where IMMI is presented using emerging technologies such as Web CD.

The Importance of Listening in Musical Experience

Literature relating to the role of listening in musical experience is extensive and a broad diversity of opinion exists about what listening is and precisely how it occurs (e.g., Copland: 1939, Mursell: 1948, Schoenberg: 1951, Hindemith: 1953, Langer: 1953, Meyer: 1956, Bernstein: 1968, Sessions: 1971, Hedden: 1980, Sloboda: 1985, Gordon: 1989, Reimer and Wright: 1992, Elliott: 1995, and Kemp: 1996). The importance of listening in musical experience is reflected by Haack (1992) who maintains that listening is the first communication skill humans develop, and it is the foundation of all other communication skills, including music. Just as we cannot learn to speak well without learning to listen well, we cannot fully enjoy or make music well without learning to listen well. Music is essentially an aural art and we cannot use or appreciate it effectively without well-developed listening skills – skills relating to perception, cognition, memory, understanding, discrimination, uses and functions, judgment, and valuing. (p. 463)

While there is consensus about the central importance of listening in musical communication, the subjective nature of the listening experience presents significant challenges for researchers who seek to understand the process of music listening.
Bamberger (1994) summarises the problem cogently: "Since hearing is by its nature a necessarily private, internal experience, an external description in whatever mode can provide only impoverished clues to a subject's momentary organizing of a melody or rhythm - that is, to his or her hearing" (p. 133). Aiello (1994) maintains that the lack of understanding stems from the fact that listening to music does not yield "palpable results" of what actually happens between the individual and the music. Moreover, in an attempt to enhance the reliability of their studies of music listening, investigators have frequently used "artificially controlled laboratory stimuli" which Aiello maintains represents a weakness in controlled experimental investigation. Such controls limit the complexity and richness of musical stimuli and reduce the choice available to the listener. Aiello contends that due to artificial experimental conditions, the findings of many listening studies have limited external validity. While there is a growing interest in the use of "real" musical stimuli in music listening research, she points out that even when "real" music is used in experimental settings, investigators can not guarantee that listening responses to music will be the same as in non-experimental situations. Therefore, the process of listening and its relationship to musical experience are difficult to quantify (Radocy and Boyle: 1979).

Despite (and perhaps partly as a result of) the research difficulties, the topic of music listening has received widespread attention over an extended period. The longstanding interest in music listening is reflected by the temporal range of the literature cited in the present review. One of the most striking features of this literature is that the discussion of music listening and its importance in musical experience by early commentators such as Copland (1939, 1952) is often closely paralleled by the concerns of contemporary observers such as Kemp (1996) and Jourdain (1997). In a comprehensive overview of the output of authors who place listening at the core of musical experience, Reimer and Wright (1992) confirm the breadth of literature. They offer the perspective that music listening should be both active and creative, and note that at its richest, music listening
involves a transformational process in which the listener draws upon perceptual, technical, and affective resources to create a coherent musical experience. Serafine (1988) who likewise stresses the importance of an active approach, defines music listening as "an active organizing and construing of the temporal events heard in a composition" (p. 71). Elliott (1995) also depicts music listening as an active process which involves scanning, constructing, interpreting, and comparing cohesive musical patterns. While Bamberger (1994) supports an active approach to music listening, she notes a further creative role for the listener by asserting that "a hearing is a performance ; what the hearer seems simply to find in the music is actually a process of instant perceptual problem solving - an active process of sense making..." (p. 133). Jourdain (1997) also believes that listening is "... a performance skill in which the listener inwardly reproduces many features of a piece by anticipating them, and thereby better prepares himself (sic) to perceive them" (p. 264). He argues that skilful listening is based on anticipation, which is derived from previous listening experience. According to Jourdain, hearing music in the background is a passive pursuit, while listening to music is an active process. He believes that active listening is "always effortful" and compared to passive listening is like the "difference between watching a dance from the sidelines and taking part" (p. 266).

In his influential early treatise on musical experience, Sessions (1950) extended the notion of active music listening to include a creative partnership between the composer, performer, and listener. He suggested that the performer and listener are in a sense true collaborators by adding to the work of the composer. Furthermore, he argued that it is possible for the listener to "visualise" a composition (with ears and not with eyes) and go beyond what is conceived by the composer and performer. Ulrich (1970) endorsed this notion of collaborative interaction by pointing out that while performers bring music to life, the composer and listener are the "terminals between which the music flows" (p. 3).
While many commentators emphasise the importance of an active approach to music listening, it is the quality of listening that is often viewed as critical in music learning. Dewey's (1938) observation of the relationship between experience and effective educational intervention supported this contention. He suggested that "it is not enough to insist upon the necessity of experience, nor even of activity in experience. Everything depends on the quality of the experience which is had" (p. 27). The quality of music listening is at the core of a debate in which many observers hold that listening occurs on different levels and that the quality of musical experience is dependent upon the level of listening. For example, in reflecting Dewey's view that it is the quality of the experience that is important, Hindemith (1953) maintained that music "remains meaningless noise unless it touches a receiving mind. But the mere fact that it is heard is not enough: the receiving mind must be active in a certain way if a transmutation from mere acoustical perception into a genuine musical experience is to be accomplished" (p. 14). Hindemith alluded to a central concern: the difference between hearing music and listening to music. While Mursell (1948) was equivocal about the distinction often drawn between hearing and listening, and contended that the difference is "far from clear-cut" (p. 139), he nonetheless observed that "in all cases it may truly be said that we hear with our minds rather than with our ears" (1943: 147). Many other commentators including Langer (1953), Ulrich (1970), Handel (1989), Heinich, Molenda and Russell (1993), Alten (1994), and Elliott (1995) share a belief that hearing is not the same as listening. For example, Langer differentiated between two types of listening; namely physical or perceptual hearing and inner or mental hearing. Physical hearing is the sensory perception of sound, whereas inner hearing relates to the work of the mind in the conceptualisation of sound. Ulrich argued that hearing music does not necessarily imply attention or application, whereas listening to music does. Handel reports that "the physical pressure wave enables perception, but does not force it. Listening is active; it allows age, experience, expectation, and expertise to influence perception" (p. 3). Heinich et al. suggest that in simple terms, hearing refers to a physiological process of
perception, whereas \textit{listening} refers to a psychological process of interpretation. As previously noted, Elliott distinguishes between listening \textit{to} music, which indicates a passive approach, and listening \textit{for} music, which requires the active participation of the listener. Alten characterises active listener involvement by suggesting that "listening is perceiving sound with careful and responsive discrimination. It is thinking about sound - analyzing its quality, style, interpretation, and nuance" (p. 7).

In discussing the quality of music listening, Sessions (1950) reinforced the distinction between hearing and listening and commented on a further process of internalisation.

In the primary sense, the listener's real and ultimate response to music consists not in merely hearing it, but in inwardly reproducing it, and his (sic) understanding of music consists in the ability to do this in his imagination...the really understanding listener takes the music into his consciousness and remakes it actually or in his imagination, for his own uses. He whistles it on the street, or hums it at his work, or simply 'thinks' it to himself. (p. 97)

The "thinking" of music to oneself is a concept that has been widely discussed by observers such as Sessions (1950), Langer (1953), Meyer (1973), Gordon (1989), Elliott (1995), and Jourdain (1997). Gordon, one of the principal proponents of the concept of inner musical thought, coined the term \textit{audiation} to describe this process of inner hearing. He asserts that "audiation takes place when one hears and comprehends music silently, that is, when the sound of the music is not physically present. In contrast, aural perception takes place when one hears music of which the sound is physically present " (p. 7). Gordon argues that the development of audiation ability is essential for skilled listening. The concept of inner hearing was raised earlier by Hindemith (1953) who referred to an "inner ringing and singing" which he argued provides the conceptual and creative foundation for understanding a musical composition. He suggested that listeners
construct a "parallel and simultaneous" image of music as they listen, a process which requires "active transformation" of musical stimuli. By comparing the sound stimulus with an inner musical image, listeners construct musical meaning.

In affirming constructivist models of knowledge representation espoused by theorists such as Kant (1951), Serafine (1988) describes a creative view of music listening in which musical meaning is constructed by the individual. Music listening involves unique activities and does not, she stresses, depend upon a process of merely receiving set musical properties which have been "injected" into the composition by the composer. Likewise, Hindemith (1953) maintained that skilful listening depends on personal qualities of the listener and suggested that "everyone who wants to listen understandingly to musical structures builds up within his (sic) mind his own technique - the musical specialist as well as the unsophisticated recipient … the innermost physical and mental adaptation is the individual's own personal achievement" (p. 17). In seeking to clarify this creative view of the listening process, Reimer and Wright (1992) report that "listeners construct their own experience by not only noticing musical events as they occur, but by listening forward, anticipating what might occur, and by listening backward, checking what is now occurring against what has already occurred" (p. 238). Meyer (1973) believes that meaningful musical experience depends upon a listener's capacity to generate musical expectation. He notes that listeners assess musical stimuli for patterns from which they develop expectations that create musical meaning. Smith (1987) aptly characterises the stimulation of expectation and resolution in music listening as a process in which "the delicate syntactic teasing by the composer evokes emotions of forepleasure, anticipation, and restful satisfaction" (p. 374). Jourdain (1997) confirms the importance of expectation in music listening and suggests that even when listeners encounter a piece of music for the first time, they derive meaning from it by making reference to elements of the music that they know well from previous experience. He concludes that "a musical
object is not so much something that strikes our brains as something that our brains reach out and grab by anticipating it" (p. 246).

Some observers report that the quality of musical experience relates directly to the level at which music listening occurs. These commentators have identified levels of music listening which they describe variously as analytic, creative, expressive, perceptual, sensual, syntactic, technical and so forth. Copland (1939) for example, specified three "hypothetical" listening planes, namely the sensual plane, the expressive plane and the sheerly musical plane. Sensual listening is the simplest form in which listeners hear music for the pleasure of the sound itself. This sensual plane of listening is analogous to the concept of hearing music as described by Ulrich (1970), Handel (1989), Elliott (1995) and Jourdain (1997). Copland emphasised, however, that while the sensual plane is important, it is not the "whole story". A second, expressive plane relates to the meaning of music. He suggested that while music can convey meaning, the precise meaning is elusive and differs between individuals and between hearings of a piece of music. According to Copland, the third category of listening is on the musical plane, which relates to "the notes themselves and of their manipulation" (p. 21). While he stressed the importance of the development of listening on this musical plane, Copland noted that "actually, we never listen on one or the other of these planes. What we do is to correlate them - listening in all three ways at the same time. It takes no mental effort, for we do it instinctively" (p. 22).

Like Copland, Mursell (1948) adopted a tripartite view when considering approaches to music listening. He defined three types of listeners, namely, intellectual listeners who focus on technical factors, motor listeners who respond to the rhythm and volume of the sound, and emotional listeners who ascribe moods and feelings to the music which often have no direct relationship to it. Mursell pointed out that the development of music listening skill requires a disciplined approach incorporating each type of listening;
therefore each of the three listening types has a legitimate place and is desirable. However, he argued that developing a high level of analytical music listening expertise should not be the sole objective in the development of music listening skill. While knowledge about the technical aspects of music is necessary to conceptualise music fully, Mursell believed that such understanding needs to be lowered to a subconscious level during active musical experience. He pointed out that while non-musicians may lack analytical resources, they nonetheless have valid responses to music. Conversely, the average musician is "too consistently analytic" when listening and tends to believe that there is only one "right" way to listen to music. Mursell suggested that listeners need to develop and draw upon "a repertoire of types of responsiveness to music" (p. 139). According to Mursell, as listeners expand their repertoire of responses to music they begin to find "many facets of interest and appeal" and thus become "selective" listeners.

Ulrich (1970) endorsed Mursell’s notion of "selective" listening. In affirming the belief that music listening occurs on different levels he maintained that at each level, listeners accumulate more information which enhances their understanding and enjoyment. For Ulrich, the process of selective listening involves listener concentration on different "layers" of the music, for example, the melodic line, the rhythm or the harmony. The greater the number of details a listener notices in each layer, the more meaning will be found in the music. As listeners gain selective listening experience, they are able to listen to several layers simultaneously and advance through levels of musical understanding to become more competent listeners.

While some commentators maintain that music listening occurs at different levels, others suggest that perceptual and psychological characteristics of the listener also play an important role in music listening. For example, Kemp (1996) believes that music listening is influenced by the perceptual style of listeners which he conceptualises in two broad, but distinct, categories. He points out that researchers frequently adopt differing
Kemp (1996) therefore shares Copland’s (1939) belief about a dichotomy in listening styles, but views the issue from a different perspective. In acknowledging the complexity of the process of music listening, he suggests that both intrinsic and extrinsic components of the music listening experience can coexist. He maintains that

in a sense, the ideal listener is both inside and outside the music at the same moment, judging it and enjoying it, wishing it would go one way and watching it go another - almost like the composer at the moment he (sic) composes it; because in order to write his music, the composer must be inside and outside his music, carried away by it and yet coldly critical of it. A subjective and objective attitude is implied in both creating and listening to music. (p. 23)

Believing that prior musical experience is an important factor in the development of music listening skills, many researchers have investigated differences in the ways in which trained and untrained musicians listen to music. Smith (1987), for example, has
shown that music experts and novices respond to different aspects of a piece of music, rather than the same ones at different levels of efficiency. Likewise, Madsen and Geringer (1990) demonstrated that trained musicians focus upon different elements of the musical experience than untrained musicians. They found that musicians focus primarily on melody while non-musicians focus on dynamics. In a study of melodic memory in subjects of differing ages and prior musical experience, Oura and Hatano (1988) found that subjects with previous musical experience performed better on listening recognition tasks that those without prior music training.

Debellis (1995) maintains that skilled listening depends upon musical training, and argues that it is not merely the degree of conscious awareness that determines the level of musical understanding, but the level of the listener's conceptualisation of music. He suggests that

the way an ordinary listener, untrained in music theory, hears music
is nonconceptual... [Nonconceptual hearing] is to be contrasted with
that of the trained musical analyst, whose hearing is typically conceptual
and theory-laden; musical training thus characteristically advances one's
listening from nonconceptual to a conceptual level. (p. 1)

DeBellis believes that the level of conceptualisation is fundamental to the appreciation of music and a key element in the development of listening skills.

In identifying differing levels of music listening skill, commentators frequently make reference to the concept of the competent, intelligent, or skilled listener. There is a consensus that listening skill can be developed through experience and practice, and as Haack (1992) reports, a wide variety of practical and theoretical advice about the development of listening skills has been published. Meyer (1973) for example, observes that "by calling attention to patterns and relationships which might otherwise have been missed, it [music instruction] refines the aural imagination and increases the sensitivity of
the cognitive ear” (p. 17). Reimer and Wright (1992) hold that the acquisition of "perceptual discernment and affective sensibility" is fundamental to skilled music listening. They maintain that competent listeners transform musical stimuli into "coherent experience" through an active and creative process and point out that

this transformation involves several mental-emotional operations that, while explained variously by different writers, include attention to aspects of sound (sensuous, technical, structural) and investment of the self in experiencing the expressive implications of what is heard. Music listening in its richest sense is a work of mind in which mind includes and depends on both perceptual and affective dimensions of experience. (p. 231)

The prerequisites for competent listening were expressed in characteristically straightforward terms by Copland (1939) who noted that intelligent listeners need to develop an "awareness of the musical material and what happens to it" so that they can "hear the melodies, the rhythms, the harmonies, the tone colours in a more conscious fashion" (p. 22). Meyer (1973) extends Copland’s description by suggesting that "a competent listener understands not only the implications generated by tonal, melodic, and rhythmic relationships within a pattern, but the functional potential of the pattern as a whole" (p. 207). He maintains that skilled listeners understand how different kinds of musical gestures tend to behave and are therefore able to reflect upon their implicative relationships. For Meyer,

an implicative relationship is one in which an event - be it a motive, a phrase, and so on - is patterned in such a way that reasonable inferences can be made both about its connections with preceding events and about how the event itself might be continued and perhaps reach closure and stability. (p. 110)

Meyer argues that since understanding patterns and implicative relationships is an integral part of daily human activity, listeners have inherent interpretive skills which can
be applied to the understanding of musical patterns. Meyer believes that musical patterns are implicative signs which experienced observers know how to interpret. He describes a holistic process by which competent listeners respond to the implicative patterns in music with their total being: "As tonal stimuli, filtered and processed by a selective auditory nervous system, are related to one another by the patterning proclivities and habits of the human mind, every facet of behaviour - physiological and psychological, motor and mental - becomes attuned to and congruent with the process and structure of musical events" (p. 242).

Gordon (1971) also favours a holistic approach in the development of music listening skill and contends that listening expertise needs to develop as an integral part of a comprehensive music training. He observes that "achievement in music constitutes a systematic study of tonal and rhythmic elements, timbre, musical form, and musical style. Singing and rhythmic activities, creativity, the reading and writing of music, and listening to music, one and all, represent facets of musical achievement which contribute to musical understanding" (p. 115). For Gordon, each of these musical components complements the other to produce what he refers to as musical "readiness" which is an informed preparedness to engage in any musical activity, including listening.

Reimer (1970) argues that an analytical approach is critical in the development of skilful music listening. He maintains that the study of music should include "a large element of conscious, careful exploration of the inner workings of music" (p. 121). Ulrich (1970) supports the contention that various types of theoretical knowledge are required for skilful music listening and argues that "in order to progress beyond the level of sensory enjoyment, the listener must know something about the purposes, structures, and techniques of music. Music is not only a collection of pleasant or exciting sounds; it, too, has principles, materials, subject matter, and history" (p. 7).
However, there is considerable debate about the importance of formal musical knowledge in the development of listening skill. Meyer (1973) calls into question the need to understand the inner workings of music in order to be a skilled listener. He suggests that "...experienced listeners can respond sensitively to music without knowing anything about what makes music work: without knowing about the theory or history of music...We can perceive and comprehend actions and relationships - musical as well as nonmusical - without the explicit conceptualization necessary for explanation" (p. 15). Meyer points to the fact that keen listeners sometimes acknowledge that even though they value music highly, they don't understand it. This, he maintains, is nonsense because people rarely enjoy what they don't understand. According to Meyer, what listeners mean when they say they don't understand music is that they do not have theoretical knowledge, but he maintains that musical understanding does not depend upon theoretical musical knowledge.

Likewise, Kemp (1996) reports that the detailed musical analysis often associated with expert listening can hamper progress. He notes that undergraduate students sometimes complain that "some methods of musical analysis actually destroy well-loved music for them. Adopting a Gestalt stance that 'the whole is greater than a sum of the parts', they feel that the essential ingredients of a piece tend to become overlooked in the minutiae of the dissection" (p. 129-130). Earlier critics, such as Schoenberg (1951), have also questioned the value of tangentially related information in the development of musical understanding by suggesting that "historical facts, biographies of authors and performers, anecdotes of their lives, pathetic, humorous, interesting or instructive, may be of some value to people who are otherwise deaf to the effects of music. But all this cannot help anyone to absorb and remember the content" (p. 147).

Elliott (1995) attempts to reconcile the opposing views by noting that all musical activity is context dependent. While he acknowledges the importance of analytical and technical
understanding of music, he points out that "studied out of context, as an end in itself (and in isolation from the sounds and actions of music making), formal musical knowledge is inert and unmusical. To be of musical value (as opposed to purely scholastic value), and to be transformed into procedural knowledge, verbal knowledge ought to be employed parenthetically - as an adjunct to music making and listening" (p. 97).

Another issue that has been identified as an important factor in the development of listening skills and musical understanding is that of repetition. For example, Sessions (1950) observed that for a listener "the main source of (musical) understanding will be through hearing music in general, and specific works in particular, repeatedly, and making them his own through familiarity, through memory, and through inner re-elaboration" (p. 98). Likewise, Hindemith (1953) implied that the development of listening skills and musical understanding are achieved in part through repetition. He observed that "all musical structures that stand entirely without his (the listener's) previous experience will have to exert their impact many times on his physical and mental receptivity if they are to be added to his stock of accumulated knowledge" (p. 20). Schoenberg (1951) maintained that "it is perhaps true that one starts to understand a piece only when one can remember it at least partially. But memory must be nursed and given an opportunity to function" (p. 184). Likewise, Elliott (1995) reports that "most children seem to achieve a novice level of listenership (or music-listening know-how) through repeated contact with and casual exploration of musical sound patterns" (p. 94).

While the aforementioned observers emphasise the value of repetition in the development of music listening skills, other critics express concern about the dominance of repetitive structures in contemporary popular music and the influence that popular music may have on the development of listening skills (Abeles et al.:1984, Etzkorn: 1990, Haack: 1992). For example, Jourdain (1997) emphasises that the quality of music has a significant influence on the development of listening skill and musical understanding but warns that
the simplistic and repetitive character of much popular music may be undermining listening habits. Lamenting the overall standard of current popular music he writes:

> What is rare is music that tells a story, that brings a multitude of themes and devices into elaborate and unpredictable interplay, like characters in a good novel - in a word, music that is literature and not mere genre writing. Lacking long exposure to such music, many people remain unaware of the limitations of the music they listen to, and haven't a clue about what music can be. Their unskilled ears make so little sense of complex music that they can only determine that their own music must be superior. (p. 267)

Considerable disquiet exists not only about the variable quality of popular music, but also about the effect that its widespread dissemination may have on listening habits. Abeles et al. (1984) point out that the pervasiveness of recorded music means that the role of listening has become much more important in music education as students now listen to recorded music far more frequently because it is so widely available. These authors maintain that the majority of music listening now takes place outside the music classroom and express grave concern about the superficial approach to music listening that they believe most students adopt. Contentiously perhaps, they propose that educators should help their students to develop an understanding of two distinct approaches to music listening: one approach consists of "keen attention to and contemplation of sounds", while the other is one of "casual attention to the sounds". They report that "the basis of the two approaches to music listening is that some music must be listened to carefully if the listener is to appreciate its intricacies, while other music is structured simply enough to be grasped without much effort" (p. 124). Most contemporary popular music is learned "without much effort", they maintain, and would thus not demand the detailed listening approach that is required for other, more complex, music. They concede that some observers may find fault with these assertions, but argue that in the absence of the listening strategies they suggest, students will continue to listen
to all music in a superficial way, never suspecting that some types of music offer other levels of listening experience. In comparing popular music with similar forms of contemporary artistic expression, Goodson (1992) summarises the position aptly:

Pop art is designed for instantaneous entertainment. To have only experienced light, melodramatic productions and then unsuspectingly witness a serious play can be a jolt to the senses. One is likely to miss its subtleties, its sophisticated tragedy or humor, and find it altogether unsatisfying. (p. 19)

Further concerns have been raised about the function of music in computer-mediated communications where music is often presented in combination with visual information such as moving images and text. Elliott (1995) believes that music listening is a "context dependent" process in which the events which surround auditory sensation influence the ways in which it is perceived and understood. From a music education perspective, this observation raises questions about the effectiveness of music listening when music is combined with stimuli presented in various sensory modalities. In 1970, Reimer warned that

unfortunately, new technological capabilities have produced some listening devices that are gimmicky, non-musical and non-educative... no amount of non-musical picturizations, moving or still, of color wheels or amusing drawings or charming narrations will fill the bill. Musical aids should focus on musically expressive events and characteristics using simple, descriptive, non-emotional and non-interpretive terms. (p. 120)

While Reimer's comments were made well before computer-mediated communication had attained its current pervasive influence, the implications of his assertions continue to resonate. Present understanding of the effects of the simultaneous presentation of auditory stimuli with visual stimuli in computerised music instruction is limited. If Reimer's contention that some forms of visual stimuli are inappropriate additions to
music listening remains tenable, it is important that music educators and researchers develop a better understanding of the effects of combined media presentation on music listening so that they can minimise negative outcomes that may arise as a result of multimedia information presentation. A first step in the process might be to examine the relationships between the audio components of computer-mediated instruction and participant music listening habits.

The extensive body of literature relating to listening in musical experience encompasses a range of issues that are of vital concern in the current investigation. While confirming the widely held belief that listening underpins and permeates all musical activity, most commentators attest to the inherent complexity of the music listening process and the difficulties encountered in its investigation. These authors agree that music listening occurs at different levels and that instructional interventions can be targeted specifically to develop music listening skill. However, there is little consensus about precisely what the various levels of music listening entail, nor the most appropriate instructional interventions for the development of music listening skills. Central to the present study are the discussions in the literature that pertain to the use of educational technology in the development of music listening skills. The nexus between technology and music listening is explored further in the ensuing analysis of the influence of audio technology on music listening.

The Influence of Audio Technology on Music Listening

Thomas Edison's experimentation with sound recording in the 1870s provided the foundation for a transformation of human musical experience. Jourdain (1997) likens Edison's invention of sound recording to Gutenberg's development of the printing press and notes that by providing a means for the widespread dissemination of recorded music, Edison had initiated a process that was fundamentally to alter our relationship with music. While music had once been the domain of only those wealthy enough to afford
"live" music making, recorded music provided a means by which people with more modest resources could access music as an integral part of their daily lives.

As the quality of sound reproduction improved, and the affordability and portability of playback devices increased, the importance of recorded music began to be recognised. Seashore (1938) was one of the first observers to acknowledge the influence that sound recording and broadcasting were beginning to exert on musicians and their audiences. He suggested that

> when listeners from every part of our nation can hear the same musician at the same time in actual performance, and when the musician may stand up before us to be seen and heard in the same song for generations to come, a new type of responsibility is thrown upon the performer and his instrument. (p. 360)

The widespread availability of recorded sound has had a substantial influence on music listening habits, particularly for people living in industrialised societies who now regularly find themselves exposed to music as they engage in other activities (Radocy and Boyle: 1979, Durant: 1990, Bontinck: 1991, Alten: 1994, Meyer: 1994, Stevens: 1994). The influence of recorded music in contemporary society is further emphasised by Abeles et al. (1984) who maintain that "no other technological changes have had such an impact on the world of music as electro-acoustic reproduction. Not only did it put many performers out of work...it affected greatly the way people listen, what they listen to, their need to perform music, and the way music is marketed. It should also affect the way music is taught" (p. 122).

The extensive broadcasting of music on radio and television and the increasing use of portable audio devices such as the Walkman have led to a situation where music is present in almost every facet of human activity. The dominant position of recorded music is emphasised by Jourdain (1997) who observes that "we dance to music, shop to music,
clean house to music, exercise to music, make love to music. Yet only occasionally do we sit down and intently listen to music" (p. 238).

As the quantity of recorded and broadcast music has increased, observers such as Abeles et al. (1984), Storr (1992) and Meyer (1994) have warned that the quality of music listening has been diminished as a result of its ubiquitous presence. Meyer (1994) reports that one of the consequences of being "otherwise occupied" as we listen to music is that it encourages a fragmented and casual approach to listening. A similar sentiment is expressed by Storr (1992) who believes that "a perpetual background of good music to which little serious attention is paid diminishes both the music and the listener" (p. 111). While some individuals claim that background music enhances their ability to study, in concurring with Zuckerkandl (1973), Storr maintains that background music "interferes with intense concentration" (p. 111).

Researchers have long held that in traditional classroom settings, students spend most of their instructional time listening. In one of the first studies of classroom listening habits Wilt (1950) observed that elementary school students spend almost 60 percent of their class time listening. Recent reports by Heinich et al. (1993) and Russell and Molenda (1996) suggest that school students spend about 50 percent of their class time listening, while university students spend up to 90 percent of their instructional time listening. Importantly however, Russell and Molenda point out that "without someone standing over them or speaking with them face-to-face, some students do not pay attention to audio instruction. They may hear the message but not listen to and comprehend it" (p. 436). Nummela and Rosenberg (1986) maintain that some form of active feedback is the key to maintaining student concentration and report that "unless the student is provided with immediate access to further information through questioning or some other means, he or she will cease to listen or actively focus on events in the classroom. Thus an interactive teaching process is more appropriate than a purely didactic one..." (p. 100).
Goodson (1992) suggests that without guidance and sufficient encouragement to interact with music examples, many music students may not profit from their listening experiences because they do not concentrate on the aural stimuli. Without adequate time on the listening task, participants will not develop an appropriate level of listening skill. The problem of maintaining student attention is further detailed by Heinich et al. (1993) who maintain that

the receiver [listener] must be able to direct and sustain concentration on a given series of sounds (the message). He or she must have the skill to think ahead as the message is being received (we think faster than we hear, just as we think faster than we read or write) and use this time differential to organize and internalize the information so that it can be comprehended. (p. 162)

The relationship between the level of concentration and the quality of listening has likewise been addressed by Alten (1994) who concludes that "listening is not reading while playing music. It is not talking or shouting during a concert. It is not paying attention only to the picture in a film or TV program. It is not riding a bicycle while listening to a Walkman...if you are not listening, sound remains part of the environment; it does not become a part of your consciousness" (p. 7).

The use of music to create an environment, and in particular the role of music in commercial settings, has recently received considerable research attention. Recorded music is now habitually used to manipulate moods and behaviours for commercial advantage (Rosenfeld: 1985, Haack: 1992, Lanza: 1994, Heller and Campbell: 1997). In his investigation of the effects of the widespread occurrence of programmed music in commercial spaces in north America, Sterne (1997) observes that music increasingly functions as "part of a fabricated environment that is aimed at getting visitors to stay and buy more" (p. 25). Music listening now more frequently occurs in public settings such as shopping malls and supermarkets than in traditional locations such as concert halls.
This change of music listening venue concerns Meyer (1994) who suggests that it has resulted in "a decline in the intensity involved in listening" (p. 319). He reflects a grave disquiet about the so-called *canned music* that is often heard in public settings as he asserts it encourages a habit of casual listening. Confirmation of Meyer's concern about casual listening is provided by Sterne who reports that "the music in question is not meant to be listened to, but to be heard" (p. 30).

In spite of the many benefits of recorded music, its widespread presence in contemporary society has led to what many commentators view as an overexposure which promotes a superficial approach to music listening that tends to be habituated. *Hearing* music in a superficial manner encourages people to ignore musical detail rather than take the time to *listen* attentively. The resulting habit of hearing music in a fragmentary manner prevails over comprehensive listening. Such habits are perpetuated by the need to filter the often intrusive auditory stimuli that frequently surround us. Cherry (1953) was the first to identify the so-called "cocktail party" effect whereby listening attention can focus exclusively on one of several speakers. Jones and Yee (1993) report that "we often effortlessly 'tune out' distracting sound patterns which arise from various sources (eg. speakers, instruments, vibrating objects, etc.), and follow the output from just one source" (p. 69). Likewise, Heinich et al. (1993) emphasise the human brain's remarkable capacity for filtering unwanted sound and point out that people often gradually lose awareness of sounds that they found intrusive when first encountered. However, Alten (1994) suggests that attempts to ignore unwanted music may "desensitize aural acuity and make it difficult to enjoy worthwhile sound" (p. 8). Furthermore, as Storr (1992) and Meyer (1994) have suggested, when people hear music while engaged in other activities simultaneously, they may enjoy a superficial encounter with sensuous aspects of the music, but they rarely develop a conscious awareness of musical content. Schoenberg (1951) characterised the superficial appeal of music elegantly by noting an aspect of
musical experience he describes as "the perfume of a work, that narcotic emanation of music which affects the senses without involving the mind" (p. 147).

From an educational perspective, superficial music listening habits present problems as the process of developing a rich understanding of music is one which requires attentive mental involvement. In his seminal treatise on experience in education, Dewey (1938) maintained that

the belief that all genuine education comes about through experience does not mean that all experiences are genuinely or equally educative. Experience and education cannot be directly equated to each other. For some experiences are mis-educative. Any experience is mis-educative that has the effect of arresting or distorting the growth of further experience... an experience may be immediately enjoyable and yet promote the formation of a slack and careless attitude; this attitude then operates to modify the quality of subsequent experiences so as to prevent a person from getting out of them what they have to give. (p. 25)

Dewey's assertions support the notion that a casual or inattentive approach to music listening, whilst perhaps "immediately enjoyable", may result in the formation of inappropriate listening habits. Indeed, a careless attitude to music listening may hinder the development of competent listening skills which commentators such as Schoenberg (1951) have suggested require a focused and attentive approach. Schoenberg aptly characterised the importance of attentive music listening when he pointed out that "...one cannot do justice to a work of art while allowing one's imagination to wander to other subjects related or not. In the face of works of art one must not dream, but one must try hard to grasp their meaning" (p. 146-147).

Overexposure to background music tends to exacerbate the concentration problems caused by casual music listening with the result that, as Goodson (1992) reports, "much of the content of music goes unperceived because of undeveloped or desensitized
listening skills" (p. 14). Further concern arises from the findings of an investigation by Salame and Baddeley (1989) which indicate that background music, and in particular vocal music, may have a disrupting effect on short term memory. While the study involved memory tasks which required the recall of numbers, the researchers contend that other cognitive tasks such as "reading, counting, calculating, and reasoning" may likewise be effected by background music.

Casual music listening habits are being cultivated further as a result of an escalation in the use of recorded music into computer software and other related forms of mediated communication. For example, recorded music is now routinely employed in interactive multimedia software where it frequently serves as a background to visual information in a similar manner to the way in which music is used in film and television. In IMMI, where recorded music is usually intended to be the primary focus of attention, participant interactions with music examples are often accompanied by simultaneous interactions with information presented in a variety of other media formats such as graphics, moving images, and text (Berz and Bowman: 1994). The intricate web of multimedia information that arises as a result of these interactions creates another layer of complexity which may at times distract participants from attentive music listening. As Hayward and Orrock (1995) observe,

CD-ROM will undoubtedly change (one of) the ways we experience retailed, pre-recorded music. Any new medium or form changes the music experience and has implications for how we relate to music... CD-ROM will change the experience as well - not necessarily the musical aspect itself, but the context that enfolds, 'colours' and contextualises that experience. (p. 74)

Scepticism about the capacity of instructional technology to promote attentive music listening has existed for many years. As previously noted, Reimer (1970) has criticised the use of "gimmicky" listening devices and other "non-musical" prompts in the
development of music listening skill. Likewise, Radocy and Boyle (1979) reported that the "elegant presentation of a musical concept or principle is not learning. No matter how attractive and expensive the instruments, books, or other instructional paraphernalia, they are not learning. Learning requires action on the part of the learner" (p. 285). These observations sound a warning about the distractions that may arise as a result of the presentation of music instruction in multimedia formats. In spite of these early concerns, in recent times music educators have enthusiastically embraced interactive multimedia music instruction, apparently convinced that the benefits of its use outweigh the problems which may arise as a result of conflicting media stimuli (Higgins: 1992, Berz 1995, Bray: 1997). Commentators such as Elliott (1995) and Jourdain (1997) continue to caution about the effects of sound recording and mediated communication on music listening. Elliott, for example, argues that "all forms of media tend to distort the information they convey" (p. 102). However, Boody (1992) takes a more positive view of the potential of digital audio technology than Elliott by observing that

we know that music is an aural art, and we know that performance musicians have always learned by listening. We have always tried to bring that listening experience to others but have seldom been fully successful. That is part of why we recognize the value of methodologies such as those of Carl Orff and Zoltan Kodaly: they help us to focus the students on becoming active listeners. This new technology enables us to examine music by listening to it, rather than using notation or other abstraction of the sounds. (p. 27)

The current survey of literature relating to the effect of audio reproduction technology on music listening habits indicates that the ubiquitous presence of music in contemporary society often encourages casual listening habits which when transferred to a music instruction setting may result in a decrease in levels of attentive music listening. A second concern relates to music listening in mediated instruction where the complex interaction of media elements can contribute to an overloading of human information processing
capabilities with the result that participants may not pay adequate attention to music examples. As efficient music learning depends on a high level of active and attentive music listening, IMMI participants who exhibit a low level of interaction with music examples are unlikely to gain an appropriate level of instructional benefit from the interactive multimedia programs they use.

Audio Technology and Computerised Music Instruction

The pivotal role of audio technology in the evolution of computerised music instruction is reflected in reports by Allvin (1971), Kostka (1974), Boody (1975), Watanabe (1980), Upitis (1983), Hofstetter (1985), Bales (1986), Higgins (1992), Peters (1992), Berz and Bowman (1994, 1995), Stevens (1994), Mobley (1996), Rudolph (1996). The history of computerised music instruction has been well documented in such reports and thus will not be reiterated in detail in the present analysis. Instead, the current survey will focus on the facilitating role of audio technology in the evolution of computerised music instruction from the perspective of a four-stage model of the development of computerised music instruction suggested by Peters (1992) and refined by Berz and Bowman (1994, 1995). The model identifies four stages in the evolution of computerised music instruction, which began in the late 1950s with a research and development phase that lasted until approximately 1965. Since 1965, progress has occurred in three further development cycles, each of which lasted approximately ten years. The present survey identifies milestones in the evolution of audio technology and analyses their influence on the development of computerised music instruction.

The longstanding influence of technology in musical endeavour has been characterised by Bales (1986) who claims that "music has always been a 'high tech' field...[and that] the machines we call instruments are, in most cases, highly sophisticated. They would not have been possible without advances in modern technology" (p. 2). Likewise, Durant (1990) maintains that musicians have always relied on technological development and
that "all forms of music making are dependent on some kind of deliberately designed and specialised equipment or technology" (p. 178). He suggests that rather than simply being led by technological development, musicians have tended to seize upon an innovation and adapt it to suit their circumstances. The widespread influence of technology in musical activity is confirmed by Stevens (1994) who points out that "throughout history, musicians and other artists have embraced new technology wherever it has offered new and tangible benefits for their creative endeavours" (p. 26).

It is perhaps partly due to the emphasis that musicians have historically placed on technological innovation that music educators have been actively involved in the development of computerised music instruction since computers were first employed for educational purposes in the late 1950s (Higgins: 1992, Placek: 1992). The four-stage model of the evolution of computerised music instruction that provides a structural framework for this discussion identifies the first phase in the evolution of computerised instruction as a "developmental period" that began in the late 1950s and continued until about 1965. Research at this time was primarily directed towards the development of computer hardware for educational purposes, but the potential application of computers in music instruction was given scant attention. Berz and Bowman (1995) maintain that the hardware development of this period is best viewed as a preparatory phase which preceded the practical implementation of computers in music instruction.

Computer sound reproduction during this early developmental period was extremely limited and sound activity was typically confined to the use of "beeps" which were used to draw attention to system failures and matters that required an immediate response from the operator (Aartzen: 1993, Barron and Kysilka: 1993, Blattner: 1993, Hereford and Winn: 1994, Daniels: 1995). As the audio reproduction capability of early computers was poor, research into computerised music instruction initially focused on the integration of various analogue audio devices, such as reel-to-reel tape recorders and
audio cassette players. Although the first such studies by Kuhn and Allvin (1967) and Deihl (1969) confirmed the feasibility of linking audio reproduction devices to computer hardware, the poor sound fidelity of early analogue audio equipment and the lack of random access to sound were identified as critical limiting factors that were of particular concern in computerised music instruction (Watanabe: 1980).

A second, "mainframe period" in the development of computerised instruction occurred between approximately 1965 and 1978 when most activity took place at American universities such as Florida State University, Stanford University, University of Delaware, and University of Illinois. Research continued to focus on feasibility studies which described the evolution of specific computerised music instruction programs based on learning theories derived from programmed instruction. Much effort was devoted to the development of instructional software use on the Programed Logic for Automatic Teaching Operations (PLATO) system (Hofstetter: 1981a, 1981b, 1985, Taylor: 1982). Later research began to assess the effectiveness of these instructional programs, often in comparison to traditional classroom teaching (Higgins: 1992).

Goodson (1992) reports that even though feasibility studies and trials of computerised music instruction had been conducted since the late 1960s, practical implementation and investigation of the effectiveness of computerised music instruction in educational settings did not commence until approximately 1973. Studies by Placek (1972), Peters (1974), and Eddins (1978) assessed the suitability of linking various enhanced forms of analogue audio technology to mainframe computer systems (such as PLATO) in an attempt to address poor sound quality and the lack of random sound access. However, while improvements were reported, progress was incremental, in spite of early attempts to integrate technological advances such as audio cards that were specifically designed to augment computer sound reproduction. The consensus view at the time was that the inadequacy of computer audio hardware was the most pressing issue constraining the
evolution of computerised music instruction (Watanabe: 1980). While the viability of computer-assisted instruction (CAI) as a tool for music teaching and learning had been established in early feasibility studies, the poor quality of computer audio reproduction meant that many music educators were understandably reluctant to consider the pedagogical ramifications of a technology that they perceived as producing inferior sound (Kemp: 1986).

A third stage in the development of computerised music instruction, which Berz and Bowman (1994) label as the "microcomputer / traditional computer-assisted instruction (CAI) period", occurred between approximately 1978 and 1989. The most salient feature of this period was the introduction of personal computers such as the Apple II in 1978, the IBM-PC in 1981, and the Macintosh in 1984. The widespread availability of these affordable machines helped to establish a critical mass of computing power in education which facilitated the practical implementation of CAI in schools and universities. Berz and Bowman (1994, 1995) report that a new phase of feasibility and effectiveness studies ensued (Hofstetter: 1978, 1979, 1981a, Hultberg, Hultberg and Tenny: 1979, Deal: 1985). These studies often focused on the transfer of instructional software originally developed for use on mainframe computers such as PLATO, to the new generation of microcomputers. Most of these computerised music instruction programs employed drill-and-practice instructional models derived from behaviourist learning theories (Peters: 1979).

During this third development period, longstanding problems of poor computer sound quality and limitations of access to sound were successfully addressed with the introduction of digital audio recording in 1982 (Barron and Kysilka: 1993). The subsequent introduction of the Compact Disc (CD) in 1983 dramatically improved the quality of computer audio reproduction and facilitated almost instantaneous random recall of specific music cues (Boody: 1992). When allied to a computer system, the CD
could be controlled automatically by commands written into instructional software which facilitated precise control over searching and playback of music excerpts. The technological capacities of the CD considerably increased not only sound quality and access to sound, but also the type and amount of information that could be handled. Various CD formats were developed that were capable of storing large amounts of any information that could translated into a digital format. Thus, graphics, moving images, text, and sounds could be stored with equal ease on CD and could be retrieved and presented by a computer with speed and accuracy (Boumans: 1996).

Sound implementation in computerised music instruction during this third developmental phase was also facilitated by the rapid improvement of computer hardware that increased the flexibility and audio capabilities of the personal computer. For example, the Macintosh computer had sound recording and playback facilities as a standard feature, and while relatively primitive by current standards, the modest audio capabilities of the early Macintosh foreshadowed an increasing emphasis on the use of sound in computing activities. Another significant development was the introduction of Musical Instrument Digital Interface (MIDI) in 1981, which offered an innovative protocol for the communication of musical information between formerly incompatible computer musical instruments (Peters: 1992). MIDI became an industry standard and led to the development of sequencing and music notation programs such as Finale (1987), Performer (1987), and Cubase (1988) which provided yet another means by which computers could be used to control the delivery and manipulation of sound (Hereford and Winn: 1994, Stevens: 1994).

A fourth evolutionary stage in the development of computerised music instruction began in 1989 and continues to the present. This phase is identified by Berz and Bowman (1994) as the "emerging technologies period". Characteristic of computerised music instruction in this phase is the convergence of digital media technologies and a shift
towards flexible models of instruction that arise from cognitive learning theories. Facilitated by technological progress, new instructional designs are being developed in which interactive multimedia, artificial intelligence, and virtual reality models encourage users to adopt exploratory and divergent approaches to learning (Moore: 1992, Berz and Bowman: 1994, Stevens: 1994). Interactive engagement with the components of mediated instruction is one of the key paradigms in these new models of computerised instruction in which flexible access and a high level of user control of instructional sequencing are considered fundamental (Schwier and Misanchuk: 1993, Berz and Bowman: 1994).

The implementation of sound in computerised music instruction continues to be driven by substantial improvements in the power of personal computers and by the linking of computers to external data delivery mechanisms such as CD-ROM. Technological progress in the processing speed and storage capacity of the recent generation of personal computers has facilitated the transfer of the large digital data files that are required to effectively deliver new models of computerised instruction that employ interactive multimedia. Alten (1996) confirms that new personal computers are now almost universally equipped with CD technology capable of the simultaneous delivery of not only high quality digital sound with fast, flexible, random access, but also visual data in the form of animations, graphics, moving images, and text.

Improvements in computer hardware have also facilitated a trend toward network computing in which emerging digital technologies are being integrated to form complex multimedia information delivery systems that are likely to play an increasingly important role in education. For example, the recent substantial increase in the use of computer networks such as the Internet has been driven by the introduction of the World Wide Web (WWW) in 1993. Internet computer software such as Netscape Navigator (1994) and Internet Explorer (1995) employ an interactive multimedia computing environment for information presentation that has become the principal means of data transfer on the
Internet as users access graphics, moving images, text, and sounds by clicking on hypertext links within a WWW document.

Sound reproduction on the Internet was initially limited by so-called "bandwidth" constraints which meant that the large digital data files required to transfer high quality sound across the Internet could not be handled with sufficient speed. However, technological progress in data compression and the continuing evolution of software such as QuickTime (1989) and RealPlayer (1995) mean that the delivery of high quality sound over the Internet is increasingly efficient. In a recent guide to the design and implementation of audio on the Web, Patterson and Melcher (1998) report that "just as graphics-enhanced Web sites now outnumber text-only sites, pretty soon Web pages incorporating audio – be it music clips, sound effects, spoken word, or ambient sound – will be the norm, and sites without sound will seem strange and primitive" (p. xvi).

The delivery of computerised music instruction over the Internet is at an embryonic stage of development, however the potential of the Web for music instruction is being continually enhanced by the ongoing improvements in Internet technology (Hayward and Orrock: 1995, Czencz and Hayward: 1997). The rapid proliferation and widespread use of music-related Internet sites is confirmed in a recent survey of Internet music resources by Bryant (1996) who reported that well-known music sites such as the Internet Underground Music Archive (IUMA) (1998) are accessed by more than 100,000 users per day. However, while there is extensive and increasing interest in the Internet as a tool for musical communication, the potential of the WWW for the delivery of IMMI has yet to be widely explored. CD-ROM currently remains the most practical delivery mechanism for IMMI, although hybrid formats that integrate CD-ROM with Internet technologies are beginning to emerge (Czencz and Hayward: 1997).

Current trends in computer usage indicate that interactive multimedia is one of the most far reaching of the "emerging technologies" (Higgins: 1992, Berz and Bowman: 1994,
While the available evidence suggests that interactive multimedia may be well suited to music instruction, its use in music teaching and learning is presently at a rudimentary stage. Only a very limited amount of research has been conducted into the complex interactions that occur between participants and the media components of computerised instruction and the pedagogical implications of using interactive multimedia in music instruction have yet to be comprehensively investigated (Stevens: 1994, Berz and Bowman: 1994).

This survey of the facilitating role of audio technology in the development of computerised music instruction is now succeeded by a review of pedagogical concerns that arise from the combined presentation of sound with the other media components of IMMI. A summary of the conceptual origins of interactive multimedia and its early use in IMMI leads to a discussion of the perceived benefits and potential problems in the design, development, and implementation of sound in IMMI, which provides a context for an investigation of additional issues that emerge from IMMI research conducted to date. Studies by Lee (1989), Hughes (1991), Sigurjonsson (1991), Goodson (1992), Duitman (1993), Berz (1995), and Bray (1997), which represent the comparatively sparse body of literature pertaining to the use of interactive multimedia in music instruction, are examined with reference to the principal concern of the current investigation, namely, participant interactions with the audio components of IMMI.

Interactive Multimedia and Music Instruction

The concept of interactive multimedia instruction is not new; educators have been presenting their lessons in a "multimedia format" for hundreds, if not thousands, of years. Reiser (1987) suggests that the roots of multimedia instruction are foreshadowed in the work of Comenius in the 1600s and Pestalozzi in the 1800s, who proposed that "we learn things through our senses and therefore real objects and illustrations should be used to supplement oral and written instruction" (p. 13). Pestalozzi’s (1894) educational
philosophy was based on the premise that learning starts with active experience. As Jacobs (1992) points out, Pestalozzi identified

three mental skills: making sounds, forming images, and imagining concepts – what he calls 'three elementary processes'. In essence, they amount to interactive multimedia, and Pestalozzi’s whole approach - that of incidental, intuitive learning – is the one to which modern hypermedia technology is most suited. (p. 116)

Current forms of interactive multimedia have their roots in audiovisual instruction which evolved primarily in the period between 1950 and 1980 (Reiser: 1987, Niemiec and Walberg: 1989, Heinich et al.: 1993). In the 1960s, the term "multimedia" was used to describe instructional packages that consisted of print materials, films, slides, audio-tapes and so forth (Barker and Tucker: 1990). Two principal differences exist between audiovisual instruction and current interactive multimedia. First, interactive multimedia is delivered by computer, and second, interactive multimedia programs are generally intended for use by an individual user who is often free to peruse the content of a program in a very flexible, non-linear manner (Goodson: 1992, Gayeski: 1996). Computer delivery allows the presentation of multimedia instruction to be precisely and dynamically structured to meet individual learning styles (Carlson: 1991). It is the speed of access, the flexibility, and the comprehensive database searching capabilities that further distinguish interactive multimedia instruction from traditional audiovisual instruction (Russell and Molenda: 1996).

While a "multimedia" approach to instruction has been adopted by educators for hundreds of years, the use of computer technology in the delivery of multimedia instruction is a relatively recent phenomena. The first commercial publication of an IMMI program was Winter’s (1989) release of the inaugural disc in his Voyager Companion CD-ROM series. Entitled Ludwig Van Beethoven – Symphony No. 9 , Winter’s CD-ROM is typical of many IMMI programs that focus on a specific
composer or musical composition and provide a musicological analysis of the selected work which is viewed from an historical and social perspective. In these programs, which aim to develop musical knowledge, the musical sounds themselves represent a critical "mainstream" of information (Aarntzen: 1993) that situates music listening at the centre of the instructional interaction. Companies such as Voyager and Warner New Media have continued to publish a wide range of CD-ROM music programs that are intended for both educational and entertainment purposes. A recent survey of IMMI resources (Bray: 1997) identified more than one hundred and fifty CD-ROM music programs and revealed widespread interest in the use of interactive multimedia for music instruction.

Renwick and Walker (1992) point out that the study of music is by its nature a multimedia activity that involves diverse tasks such as listening, performing, reading, and analysing printed scores. Music learning frequently requires the combined application of the auditory and visual senses, as students listen to musical sounds while viewing a score or related text. Interactive multimedia programs, which are capable of the simultaneous presentation of visual and aural stimuli, appear to offer considerable potential for music instruction and there is a consensus amongst commentators that one of the fundamental advantages of using interactive multimedia in music instruction is its capacity to link sound to visual symbols (Winter: 1989, Peters: 1992, Lehrman: 1993, Stevens: 1994, Berz and Bowman: 1994, Jaworski, 1996, Parakilas: 1996). Although it pre-dates the widespread introduction of IMMI, Kemp’s (1986) report on the potential of microtechnology in music education points out that "developing the sound-symbol relationship in children's thinking has always been somewhat problematic in music teaching and this may be due in part to the failure of teachers to ensure that an internalised experience of a sound is firmly embedded within the knowledge of its symbol" (p. 40). Kemp feels that the capacity of computers to link sound with its graphical representation offers a major breakthrough in fostering the "sound-symbol relationship". The ability of computerised music instruction to cross-reference music
examples with printed notation and text is also particularly appealing to Stevens (1994),
who notes that "it is only when the connection is made between what is seen in notated
form and what is played and/or heard that true musicianship and musical understanding
can be developed in our Western musical tradition" (p. 18).

A practical illustration of the advantage of using interactive multimedia to link sound and
visual elements is provided by Lehrman (1993) who contends that "a computer that lets
the student hear the music while simultaneously pointing out and correcting her (sic)
mistakes not only provides far more valuable feedback than do corrections on paper
returned the next day but also makes the process a far more rewarding experience,
making her eager to continue" (p. 41). The benefits of linking sounds to symbols in
IMMI programs are confirmed by Jaworski (1996) who suggests that

to the learner they [IMMI programs] offer the advantage of pursuing
an individual path through a rich collection of appropriate media,
for example, listening at one instant to an audible rehearsal of Beethoven's
music, followed at the user's choice by a pictorial summary of the environment
that Beethoven worked in, supplemented by a textual analysis of his
compositions. Thus, the student's ears can be addressed by the sound of
Beethoven's Fifth Symphony, while the computer screen is filled with the
conventional musical representation of the score. (p. 216)

Consequently, the ability to link sounds with visual stimuli represents a critical attribute
of interactive multimedia when it is used in music instruction. Digital audio recordings of
musical works, which can be activated either automatically using pre-programmed
computer commands or at the discretion of the user, are an essential element in the
design of most IMMI program as they facilitate rapid and precise control of IMMI audio
components.

Sound in Computerised Instruction
While the advantages of using sound in conjunction with visual stimuli in computerised music instruction have been identified, and substantial improvements in computer sound reproduction capabilities have been achieved, the audio component of computerised instruction has often been neglected by instructional designers and researchers (Aarntzen: 1993, Barron and Kysilka: 1993, Hereford and Winn: 1994, Daniels: 1995). The lack of research-based understanding about how audio components can be effectively implemented in computerised instruction can perhaps best be understood by considering the sequence of events in the technological evolution of the computer. Initially, interactions between computer and user were predominantly reliant upon the visual mode of communication. Information was relayed from the machine to its operator primarily through a visual display unit, while input to the computer was achieved mostly by using keyboard commands which when typed, appeared on the visual display. Eventually, with the invention of the pointing and selection device known as the "mouse", communication between user and computer was conducted with the assistance of graphical user interfaces (GUI) such as those of the Macintosh and Windows operating systems, which likewise continue to reinforce the dominance of visual modes of communication in human-computer interaction.

Partly as a result of this focus on the visual modality in computing, and also because of the relatively poor quality of sound in many early computer systems, research in the field of computerised instruction has tended to concentrate on matters relating to its visual elements rather than upon its audio components (Hativa and Reingold: 1987, Mountford and Gaver: 1990, Jasper: 1991, Aarntzen: 1993, Barron and Kysilka: 1993, Barron and Atkins: 1994). The scarcity of audio-related research has meant that comprehensive research-based guidelines for the design and implementation of sound in computerised instruction have yet to evolve and instructional designers have consequently had little choice but to adopt an intuitive approach to sound design. In the absence of proven development guidelines, sound design in computerised instruction often appears to have
been arbitrary. Mann (1995) asserts that poor sound design, which stems from inadequate planning, has resulted in "an ambivalence about using audio in computerised instruction, and has been blamed for inconsistent effects" (p. 16). An intuitive approach to sound design in computerised instruction has not yielded optimal results and poor sound design frequently exacerbated problems of mediocre sound quality that hampered many early computerised music instruction programs. Poor sound design and limitations in the audio quality of many early computer music instruction programs led to a reluctance from music educators about using computers in music instruction that persisted for some time (Kemp: 1986).

In an attempt to improve upon intuitive models of sound design, Mann (1995) proposes a Structured Sound Function (SSF) model for the design and implementation of sound in computerised instruction. His "case-based method, unlike the intuitive approach, requires the researcher to assign structural and functional attributes from the SSF model to specific software" (p. 18). Mann maintains that the SSF model is "conceptual" rather than "procedural" and as such does not provide "specific guidelines" for the combination of sound with vision, but rather aims to "facilitate an understanding of those attributes which impinge on the design and implementation of sound in computerised instruction" (p.18). The SSF model presents a series of strategies and sub-strategies that relate to the structure and function of sound in computerised instruction. For example, structural attributes relate to pacing or rhythm in the relationship between sound and vision, while functional attributes include the use of sound to define character, place, or time. Mann concludes that while the SSF model is not ideal, his recent research suggests that it has potential which needs additional exploration.

Further illustration of problems in sound design for computerised instruction is supplied by Daniels (1995) who points out that software designers too often "utilize 'stock' sounds that are 'thrown' in without contemplating or understanding the relationship
between audio and visuals..." (p. 2). Sound is frequently used as an effect to gain the participant’s attention, but is rarely fully integrated into the multimedia environment. Nonetheless, Daniels concludes that "the key to finding the best audio for a particular visual is to experiment and listen carefully. As with any other aspect of multimedia design, the audio element is somewhat intuitive and changes with each new situation" (p. 7). He suggests that good practice in the design of the audio components of computerised instruction is to model sound design upon successful aspects of existing multimedia programs.

Although the intuitive modelling of sound design on successful interactive multimedia programs may prove beneficial, the importance of sound in music instruction indicates that IMMI sound design requires a more structured research-based approach. Audio components currently enjoy a similar level of availability as other media resources such as graphics, text, and moving images in computer-mediated communications, and sound offers a variety of unique and complementary capabilities when used on its own or in combination with other media components. With the expanding use of interactive multimedia in instructional settings, there is an increasing need for software designers to extend the role of sound in computerised instruction. The provision of effective guidelines for the design and implementation of sound in IMMI means that research in the area remains a matter of urgency.

While knowledge of the characteristics and functions of sound in computerised music instruction is presently limited, and the effects of interactions between visual and aural stimuli in IMMI have yet to be comprehensively investigated, it is possible to draw upon research conducted in non-music domains that may assist in the cultivation of research-based knowledge about the role of sound in IMMI. The recent growth in computer-mediated communication has stimulated a modest amount of research into the role of the audio components of computerised instruction (Barton and Dwyer: 1987, Hativa and
In one of the most comprehensive overviews of design issues relating to the use of sound in computerised instruction, Aarntzen (1993) points out that audio can draw and hold the attention, complement the visual information on the screen, support the student reading the text on the screen, minimize the amount of information that it is necessary to present on the screen, announce some event, and motivate the student. (p. 355)

Aarntzen’s comments reflect a consensus that the audio channel of communication has an outstanding potential in computerised instruction which remains largely unexplored. In assessing the role of sound in the multimedia interface, Blattner (1993) confirms that the audio channel of communication is one of the most compelling modes of instructional presentation. She maintains that sound has a powerful capacity to create rich sensory environments that encourage listener involvement, enhance concentration, and engage the mind.

The essential characteristics of sound are identified by Gaver (1989) who reflects Fleming and Levie’s (1978) assertion that sound is transient and sequential and that it exists in time and over space, whereas visual objects exist over time and in space. While sound is transitory, it is nonetheless omnidirectional, whereas visual objects are persistent, but directional. It is not necessary to look at sound in order to perceive it and therefore it is possible to hear things that cannot be seen. Sound and vision therefore provide complementary sensory input mechanisms and as Hereford and Winn (1994) suggest, one of the principal benefits of sound is that it is "very good at telling the eyes where to look" (p. 212).
Observers commonly identify four audio components that occur in computerised instruction, namely (1) music, (2) speech, (3) sound effects, and (4) silence (Aarntzen: 1993, Daniels: 1995, Alten: 1996). These audio components generally appear in combined forms, but can occur individually and may be presented as either the "mainstream" of information or used to complement visual information, perhaps by displaying data that is difficult to convey visually (Aarntzen: 1993). When an audio component functions as the mainstream of information it is the principal focus of the presentation. Where an audio component complements visual information it appears either "redundantly" or "non-redundantly" in support of the visual components of a presentation (Hereford and Winn: 1994). Redundant manifestation of an audio component refers to the simultaneous occurrence of "semantically identical information that will be processed by the receiver through multiple sensory channels" (Aarntzen: 1993:355). For example, redundant presentation occurs where a printed text appears on the computer screen while a simultaneous vocalisation of that precise text emanates from the computer's audio channel. Importantly however, Edwards (1992) concludes that "the definition of redundancy has to be subjective and context-dependent. Information which is redundant to one person may not be so to another" (p. 149). Grimes (1990:15) maintains that redundancy, which he defines as "the semantic match or mismatch between the audio and video messages in the presentation" is one of the most "troublesome" variables encountered in the investigation of audiovisual instruction as it is difficult to define and control.

In non-music domains, the audio component of music rarely functions either as the mainstream of information or in a redundant capacity as it is usually employed for affective purposes to underscore the mood or emphasise an important concept. According to Blattner (1993), music is frequently used in a complementary role where it "serves as a kind of cohesive, filling empty spaces in the action or dialogue… music can
complete the total picture and produce a kind of dramatic truth, which the visual element is not always capable of doing" (p. 78). Daniels (1995) suggests that "no other sound is as effective as music in communicating complex emotions and moods" (p. 6). He endorses the belief that music can define a locale, establish a period in history, identify characters and events with recurring themes, and provide transitions from one idea to the next. Furthermore, musical elements such as tempo and rhythm can be varied to influence the pace of a presentation and heighten the intensity of its visual components. Literature reviewed here suggests that music usually performs a complementary role in computer-mediated communication. However, a notable exception occurs in computerised music instruction where music commonly represents the mainstream of information presentation and where music notation is often viewed simultaneously (redundantly) while a recording of the relevant musical work is heard.

Aarntzen (1993) maintains that speech is generally the most appropriate audio component to use where the audio channel is the principal means of information presentation. Speech can occur as either narration or dialogue, and narration is frequently employed to deliver information or instructions which can be illustrated with on-screen images. According to Daniels (1995), narration can be effective where computer screen space is limited, as it reduces the need to display text on screen. Where narration is presented in conjunction with on-screen text, the narration should be highly redundant to that text since both text and speech are perceived as verbal communication and interference may occur in processing where text and speech are non-redundant (Aarntzen: 1993). Narration is therefore most useful as a replacement for text rather than as an addition to it. Narration is likewise preferred for attracting attention to visual details, while the pacing of narration can intensify the visual experience by complementing the mood of the visual presentation (Mann: 1995).
Sound effects only occasionally function as the mainstream of information in computerised instruction (for example, where recorded sounds of naval vessels might be used to instruct trainee sonar operators). They are however frequently used to draw and maintain attention or in a complementary role to add subtle enhancements to visual information (Aarntzen:1993, Daniels:1995). Perhaps the most common manifestation of sound effects in computerised instruction occurs with messages and alerts that keep the user informed about the current status of the computing environment. The use of sound effects to draw attention to matters which require urgent attention, such as system failures, considerably increases the bandwidth of communication between the computer and the user (Mountford and Gaver: 1990). Furthermore, the ongoing development of earcons (auditory messages that augment the system status information that computers supply to users) is an example of continuing interest in extending the structured use of sound effects in the human-computer interface (Gaver: 1987, Blattner et al.: 1989, Blattner and Greenberg: 1992, Brewster: 1996a, 1996b).

The audio component of silence is rarely mentioned in discussions of the audio elements of computerised instruction. According to Daniels (1995), silence is probably the most "underrated" of audio elements. He points out that silence can be used effectively to build tension and that it can have a strong affective influence, especially when its occurrence is unexpected. In instructional settings, silence can facilitate perceptual and cognitive processing by allowing participants to concentrate on information presented in other sensory modalities.

**Human Information Processing of Multimedia Stimuli**

As noted previously, there is a considerable amount of enthusiasm about the potential of interactive multimedia in instructional settings, which often arises out of a strong expectation that the use of multiple media stimuli will enhance learning. The enthusiasm for multimedia stimuli in computerised music instruction has meant that "in virtually all
of the commercially available hypermedia applications in the field of music, learners are directed to read computer displays while music is playing" (Berz and Bowman: 1994:42). However, researchers have detected a number of potential problems that may result from the simultaneous presentation of multiple stimuli. Daniels (1995) maintains that combined audio-visual components take on dynamics and meanings that are different from those that occur when either medium is presented alone. In terms of the present investigation, an issue of major concern is the interference that may occur to IMMI participants’ attention and concentration as a result of the simultaneous presentation of aural and visual stimuli. When IMMI users encounter multiple stimuli, it is possible that perceptual and cognitive processing may be hampered by competing processing needs that cause interference that leads to lapses in attention and concentration (Salame and Baddeley: 1989). As effective instruction depends upon a learner attending to instructional stimuli and processing them appropriately (Fleming: 1980), lapses in attention and concentration, as a result of competing media stimuli, may result in a reduction of instructional effectiveness.

The enthusiasm about the use of multiple media stimuli in instructional settings rests largely on research that has been conducted in "non-music" domains. Meta-analyses of such research by investigators including Kulik and Kulik (1991), Khalili and Shashaani (1994) and Fletcher-Flinn and Gravatt (1995) suggest that computerised multimedia instruction often produces superior instructional outcomes when compared to traditional instruction. However, there is robust debate and a substantial amount of opposing evidence which stems from contradictory theories and conflicting research evidence. For example, commentators such as Schramm (1973, 1977), Salomon (1979a, 1979b), Clark (1983, 1985), and Kozma (1991) have been involved in an ongoing debate about the contentious issue of whether various forms of media differentially influence learning. In a provocative critique of related literature, Clark (1983), argued that "media do not influence learning under any conditions… media are mere vehicles that deliver
instruction, but do not influence student achievement any more than the truck that
delivers our groceries causes changes in our nutrition" (p. 445). Clark challenged media
researchers to establish a stronger theoretical framework that might improve upon the
existing paradigm of comparative research, which has produced equivocal results in
many investigations of the effects of media on learning. Kozma (1991) responded to the
challenge by developing a theoretical framework in which

media are distinguished by cognitively relevant characteristics of their
technologies, symbol systems, and processing capabilities...these
characteristics, and the instructional designs that employ them, interact
with learner and task characteristics to influence the structure of mental
representations and cognitive process. (p. 179)

Essentially, Kozma views mediated instruction as an active and constructive process in
which learners select characteristics of media (such as symbol systems) with personal
relevance according to an individual’s existing mental models: the selected media
characteristics facilitate learning for that individual. Consequently, Kozma argues that
media do influence learning, for some learners, in certain tasks. It is reasonable to
conclude that Clark and Kozma represent the diametrically opposing viewpoints of a
debate that is likely to continue for some time (Clark: 1994a and 1994b, Jonassen,
1997). Notwithstanding these controversies in media research, the present review
proceeds with a summary of theories of attention and human information processing in
preparation for a discussion of research relating to the use of multiple media stimuli in
"non-music" domains and its possible application in the IMMI setting.

Investigators have been actively studying human attention and the perception of visual
and auditory stimuli for more than 40 years. According to Fleming (1987) "attention,
perception, and learning are not discrete processes. They are richly intertwined – both
practically and theoretically" (p. 236). While a substantial body of related literature
exists, it is reasonable to conclude that conflicting theories prevail about how humans process multiple media stimuli and how such processing may affect attention and concentration in instructional settings (Broadbent: 1958, Travers: 1964, Treisman: 1964, 1969, Severin: 1967a, 1967b, Hsia: 1968, Shiffrin: 1976, Baggett: 1984, Robinson: 1985, Salame and Baddeley: 1989, Edwards: 1992, Jones and Yee: 1993, Shih and Alessi: 1996). Theories of attention often employ diverse terminology and describe differing processes with the result that commentators have yet to agree on a precise definition of attention. Johnston and Dark (1986) point to a "widespread reluctance to define attention", while Jones and Yee (1993) stress that as attention is an "inferred construct", attempts to define the concept are inevitably "theory bound" and that "ultimately, definitions of attention become theories of attention" (p. 70). As a consequence perhaps, Glover, Ronning, and Bruning (1990) have defined attention broadly as "a person's allocation of cognitive resources to the tasks at hand" (p. 51) and report that in spite of some individual differences, human beings can generally pay attention to only a very limited number of tasks simultaneously. Such limitations of attention are of critical importance in education settings as instructional stimuli that are not attended to will not be processed or learned.

The study of human perception and attention first gained an impetus in the late 1950s when Broadbent (1958) proposed a model of attention which suggested that stimuli were "selected" for processing very early in the analysis of information. Broadbent theorised that humans possess limited information processing capabilities and that such limitation necessitates the "early selection" of stimuli for processing. According to Broadbent, selected stimuli would subsequently be channelled for further processing, while information that was not selected remained unprocessed and would thus be ignored. A later model of attention developed by Treisman (1964, 1969) suggested that Broadbent's theory of "early selection" was incomplete. Treisman proposed a model of "attenuated processing" in which she argued that different "channels" of information could be
processed simultaneously. According to Treisman however, limitations in human processing capabilities meant that some channels of information received a higher level of processing, while others were processed at a reduced or "attenuated" level.

In an apparent contradiction of the work of Treisman, concurrent research into the effects of the simultaneous presentation of visual and aural stimuli by Travers (1964) suggested that due to limited processing capabilities, learners could generally only use information from one source at a time. Travers reported that

so far, there does not seem to be a single contemporary scientist who takes the position that the human can receive more information if exposed to two or more sources simultaneously than if exposed to one, nor if the information is transmitted through two sense modalities rather than one. (p. 375)

According to Travers, not only did additional stimuli offer no learning advantage, their presence contributed to the risk of "overloading" the human processing system. In an attempt to reconcile the conflicting theories of human information processing, Hsia (1968) hypothesised that the use of multiple channels in the presentation of information would enhance communication provided that the central nervous system was not overloaded. He reported that "the problem concerning communication efficiency lies in the reinforcing of one channel with the other while concurrently keeping down between-channel interference effects" (p. 250). According to Hsia, maximising the effectiveness of combined audio and visual information requires the optimisation of information presentation, that is, the combined amount of auditory and visual information should approach the maximum information processing capacity of the central nervous system. Where information processing capacity is exceeded, overload will result, but where the amount of information is substantially less than the capacity for processing, resources will remain under utilised.
Subsequent research by Severin (1967b) provided further supporting evidence of the advantages of using multiple media stimuli when communicating. Severin’s "cue summation" theory asserted that learning would be increased when the number of stimuli is increased, but only in cases where such stimuli reinforce one another. Severin concluded that

> multiple-channel [media] communications appear to be superior to single-channel communications when relevant cues are summated across channels, neither is superior when redundant between channels, and are inferior when irrelevant cues are combined (presumably because irrelevant cues cause interference between them). (p. 397)

While the summation of cues presented in multi-channel formats may result in learning gains, Severin cautioned that "unexpected interference" can occur between media channels. He noted that adding more channels will not necessarily increase learning as "often the material which a communicator believes is related in two channels may not be related in the mind of the receiver because of his (sic) own prior associations or lack of any prior associations" (p. 398). Severin’s study suggests that where "unrelated information" is presented in different media, learning is less efficient.

In his 1970 investigation of multichannel processing in perception, Lindsay found that it was not the amount of information presented by a stimulus, but rather "the difficulty in discriminating the signals in a given channel that causes performance to break down under multichannel conditions" (p. 154). His findings indicate that it is a problem in distinguishing which information to attend to, rather than stimulus overload, that causes lapses in information processing. In surveying the current understanding of mechanisms of perception and attention, Lindsay identified three models of human information processing, namely, single channel processing, parallel processing, and multiplexing, but concluded that each of the possible models was "consistent with existing data and it
seems unlikely that current experimental techniques will allow us to distinguish unequivocally between them" (p. 170).

Further evidence to support the effectiveness of multiple-channel communication is provided in Paivio’s (1971, 1986, 1991) "dual coding theory" which proposes that human information processing is served by two separate symbolic systems, one specialised for dealing with verbal information (both printed text and speech) and another "non-verbal" system that is targeted toward spatial information and mental imagery. The processing of verbal information is assumed to be "sequential" while the processing of visual information is "synchronous" and relatively free from sequential constraints, which means that the two symbol systems can operate in a complementary manner. Although processing of information from these symbol systems can function independently, most information processing involves connections between the two symbol systems which supplement one another and may have an additive effect on information recall. For example, a person looking at a picture of a flower might engage in covert verbalisation of the word "flower". As the content of the picture is consequently registered in both memory systems it is said to be dual coded. Paivio argues that learning is superior with dual coding, as "the learner creates more cognitive paths that can be followed to retrieve the information" (Najjar: 1996:134). Moreover, Levie (1987) reports that the dual coding mechanism accounts for the recurrent research finding that recall for pictures is superior to the recall of words.

In an attempt to explain human information processing which directly contradicted the early selection model of Broadbent, Shiffrin (1976) proposed a "full processing" model in which the selection of stimuli does not occur until after stimulus recognition processes have occurred. According to this model, stimulus recognition occurs automatically without conscious effort and the selection of which stimuli to attend to occurs in memory rather than as a part of the perceptual process. Shiffrin held that memory limitations
account for the inability of humans to retain all the perceptual analysis that occurs automatically.


Broadly speaking, these research tasks can be classified into two categories: resource-limited tasks and data-limited tasks (Norman and Bobrow: 1976, Nusbaum and Schwab: 1986). Resource-limited tasks are those for which performance would improve if more processing resources were devoted to them. For example, where multiple stimuli are being encountered, focusing resources on a particular stimulus (at the expense of other stimuli) would improve performance in processing the selected stimulus. Data-limited tasks are those tasks where the quality of the stimulus constrains processing performance. For example, poor audio quality of a music recording will limit stimulus processing. Importantly, with data-limited tasks, the allocation of further resources to the processing of the task is unlikely to substantially improve processing as performance is limited by the poor quality of the stimulus rather than a lack of processing resources.
In the current study, where IMMI participants encounter high quality digital sound at the same time as viewing high resolution graphics and textual information, the quality of the data resources is unlikely to impose substantial limitations on the processing of media stimuli. More concern arises out of possible limitations in human information processing capabilities which may place marked restrictions on the capacity of IMMI participants to assimilate information presented simultaneously with multimedia stimuli. Research into the effects of simultaneous multimedia presentation of information in computerised music instruction is extremely limited, while simultaneous multimedia presentation in computerised instruction in non-music domains has attracted considerable interest. Although the investigation of human processing of multimedia stimuli has frequently produced contradictory findings, research from non-music domains yields a valuable body of information that is relevant to the current investigation. The pertinent research relates not only to the perceptual and cognitive effects of simultaneous multimedia information presentation, but also to the influence of redundant (and non-redundant) presentation of information on human information processing.

Tergan (1997) points out that it has been widely assumed by educational theorists, cognitive psychologists and instructional designers that the addition of media stimuli and the use of multiple modes of representation in computerised instruction will increase learning. The advantages of combining sensory inputs are reported by Oz and White (1993) who concur with the often cited contention that humans retain approximately twenty percent of what they hear, forty percent of what they hear and see, and seventy five percent of what they hear, see and do (Wilt: 1950, Treichler: 1967). Evidence to support this assertion can be found in studies by Haber and Myers (1982), Fleming (1987), and Levie (1987) who report, for example, that the recall of words is enhanced by associating pictures with the words being memorised. In summarising the advantages, Fleming (1987) points out that "pictures and words can be reciprocally beneficial; words can delimit and interpret pictures and pictures can help define, exemplify, and make
memorable words (Mills, 1980). Captions and labels can have a determinative effect on how or whether pictures are studied” (p. 242). Fleming maintains that the use of pictures to reinforce verbally presented information has repeatedly been shown to significantly increase learning when compared to the use of verbal cues alone. Importantly however, he notes that while the combined use of the visual modality (pictures) and the aural modality (speech) increases capacity above that achieved with either modality individually, "this effect occurs when information in the two modalities is related rather than redundant or discrepant. Also, adding more information to both modalities can overload the system so much that the learner must choose to attend to one or the other" (p. 243-244). Further studies by Drew and Grimes (1985) and Hicks (1987) found that the presence of auditory stimuli in combination with visual stimuli has a positive effect on the long-term retention of information.

However, there is much conflicting evidence. A meta-analysis of related research by Robinson (1985) suggests that visual stimuli tend to dominate other media in both perceptual and memory tasks. Robinson reports that "when visual and auditory signals are presented simultaneously, subjects generally respond to the visual input and are often unaware that an auditory signal has occurred" (p. 2). The dominance of the visual over the aural modality is further confirmed in reports by Colavita (1974), Posner, Nissen, and Klein (1976), Hayes and Kelly (1984), Pezdek and Stevens (1984), and Moore and Dwyer (1994). Moreover, Beckwith (1992) suggests that increasing the number and complexity of media stimuli does not necessarily lead to increased cognitive performance. In reflecting on his early use of the computer programming language LOGO in computer-assisted multimedia music programs, he observes that "too rich a sensory stimulation actually seemed to impede the abstract thinking needed to make headway here. The output medium almost had to be relatively impoverished - simple straight-line drawings, square-wave beeps, and 'staticky' scratches in lieu of drumbeats - in order to get the full benefit of this kind of integration" (p. 25). Conversely, a study by
Chadwick (1992) of the role of audio in an interactive multimedia museum exhibit found that reducing the number of media components in the program by removing the audio channel resulted in a substantial lowering of the user completion rate when compared to the completion rate for the program which contained audio. While Chadwick’s investigation did not measure the instructional effects of the presence or absence of audio, the lowered completion rate resulting from the absence of audio is a concern as participants who do not complete interactive multimedia sessions are unlikely to accomplish the instructional objectives of the session.

Another issue which has assumed substantial importance in the related research literature is that of the symbol systems that occur in media (Levie: 1987). In his examination of the relationships between media and learning, Salomon (1979b) provides an explanation of the symbol systems which he asserts are the fundamental characteristic of media. He reports that "media of communication and instruction are most essentially distinguishable from each other in terms of the symbol systems one can best use with them" (p. 133). Reflecting the essential tenets of Goodman’s (1976) influential theory of symbols, Salomon maintains that symbol systems are "a set of elements such as words, numbers, shapes, or musical scores that are interrelated within each system by syntactic rules or conventions and are used in specifiable ways in correlation to fields of reference" (p. 132). For example, Salomon argues that music notation is a symbol system that is "well differentiated" and correlates closely with the sounds it represents. Other symbol systems, such as pictures, are "non-notational" in the sense that they are "syntactically and semantically dense", which means that they can represent a wide variety of ideas or events depending of the interpretation of the viewer. Salomon believes that media are not "invariable, discrete entities" and that "no medium is exclusively associated with one symbol system, nor does one symbol system serve only a single medium" (p. 133). He concludes that "not all media can instruct equally effectively. [However,] if their unique symbolic capabilities are capitalized upon, each medium
addresses itself to different constellations of mental skills, thus benefits learners of different aptitudes, and serves different educational ends” (p. 144).

Salomon’s observations underscore the widely held belief that the symbolic diversity of interactive multimedia facilitates human information processing as users can choose, from a variety of symbolic representations of the instructional content, those representations that best suit their learning needs. For example, in a study of symbolic representation and its effect on learning, Nugent (1982) found that "the combination of symbol systems evidently exploits the potential of each and promotes a synergistic interaction. It appears that the two [symbol systems] provide alternate means of processing information and are complementary. Students are able to alternate between modes and use each to its best advantage" (p. 168). Likewise, Shavelson and Salomon (1985) contend that where information is presented in more than one symbol system, learners are free to shift from one mode of representation to another and thus can receive information in their preferred symbol system. Similar studies by Levie and Lentz (1982) and Mayer and Anderson (1991, 1992) provide further supportive evidence that the redundant presentation of information improves instructional outcomes when compared to single media presentation.

Investigators have also examined the differential effects of the provision of media in varying symbol systems on the learning of low and high aptitude participants. In a meta-analysis of such studies, Najjar (1996) found that

multimedia is most effective for people with low prior knowledge or aptitude in the domain being learned. This may be because experts already have a cognitive model and large amounts of information for new knowledge to connect to, but novices do not. Alternatively, novices may not know which information is important and on which information they should focus their attention. (p. 137)
Likewise, Hativa and Reingold (1987) hypothesised that the provision of information in differing symbol systems may improve instructional outcomes for low aptitude students, however, the results of their investigation indicate that the effects were inconsistent. The researchers examined the differential effects of color, animation, and non-verbal sounds in the computerised instruction of geometric concepts, and concluded that "it is possible that the concurrent use of three different symbol systems for stimulating student’s attention exceeds the amount of information that some students are able to assimilate and to process, or that it interferes with the encoding processes of the different symbol systems which, in turn, causes a decrease in learning efficiency" (p. 302).

Evidence that conflicts with the assumed instructional efficacy of using multiple media stimuli also arises out of Barton and Dwyer’s (1987) investigation of the influence of audio redundancy on the learning of groups of participants with differing intelligence levels. Their findings indicated that while audio redundancy produced positive results for high IQ participants, performance for low IQ participants was significantly impeded. The result was a direct rebuttal of their hypothesis that audio redundancy would assist low IQ participants. In an attempt to explain their findings, the authors postulate that low IQ participants may have been pre-occupied with switching from one sensory mode to another and thus had little time "for reflective thought and cognitive integration" (p. 96). Low IQ participants, they suggest, may lack the "cognitive schemata" to profit from complex multi-modal instruction.

A subsequent investigation of the relationship between visual and auditory information by Barron and Atkins (1994) examined textual redundancy (visually presented text with narration) in computerised instruction. While the assumed benefit is that such redundancy will result in increased learning, these investigators found that the redundant presentation of text and narration had no significant effect on learning outcomes or upon the time it took participants to complete the study task. A possible explanation for these findings is suggested by Daniels (1995) who maintains that text and narration are coded
within the same (verbal) symbol system and that the processing of sensory stimuli is not optimal when two stimuli are coded within the same symbol mechanism. Accordingly, he recommends that "when information is presented across channels, it should be highly correlated to improve learning and avoid inter-channel interference. Additionally, the multimedia author should ensure that images or verbals presented across channels are not conflicting" (p. 3). Further warning about the interference that can occur when redundant stimuli are presented is provided by Heller and Martin (1995) who claim that such duplication of information "greatly diminishes" the overall impact of the presentation.

Baggett (1984) examined the role of temporal contiguity in the formation of associations between visual and spoken stimuli in dual-media presentations such as film. In settings where speech was used to describe and explain visual images, she found that "the temporal order in which visual and auditory elements were presented differentially influenced the formation of dual media associations" (p. 408). The results of Baggett’s investigation suggest that where visual and auditory components reinforce each other semantically, the visual component of instruction is best presented either early or simultaneously with the auditory component. The findings indicate further that where the audio component preceded the visual component, much of the information from the audio component was lost. However, Baggett’s investigation was confined to an assessment of the effects of temporal shifts between visual and non-musical stimuli and did not examine the effect of temporal shifts in the presentation of music in combined media presentations.

Shih and Alessi (1996) analysed the relative advantages of text and voice and their effect on learning outcomes from interactive multimedia courseware. While eighty-two percent of participants in their study preferred combined text and voice presentation over either text or voice alone, the results indicate that learning performance in the combined media condition was inferior to that in the text only condition. The authors postulate that when
combined with visual stimuli, narration may interfere with attention and cognitive processes. However, as the vast majority of participant’s preferred the presentation of both text and narration simultaneously with the visual content (in spite of the apparent decrement in the learning performance), the authors observed that "what you like is not necessarily good for you" (p. 216).

One of the few studies that specifically examines the use of music in combined media presentations is Salame and Baddeley’s (1989) investigation of the effects of background music on short term memory. The investigators report that when visually presented sequences of digits were displayed simultaneously with background music, disruptions occurred in the immediate serial recall of the numbers. The study revealed that silence and (artificially generated "pink") noise were less disruptive to short term recall than instrumental music, which was less disruptive than vocal music. The investigators found that "background music can certainly disrupt immediate verbal memory, particularly in the case of vocal music" and theorised that background music may cause similar disruptions to occur with other cognitive tasks such as "reading, counting, calculating, and reasoning" (p. 119). They conclude that the results of their investigation may have important ramifications in view of the widespread prevalence of background music in contemporary society. When considered from the perspective of the current investigation, these findings, which indicate that background music interferes with memory and other cognitive tasks, suggest that the simultaneous presentation of music with visual stimuli in IMMI has the potential to hamper information processing and the assimilation of instructional messages.

Despite the conflicting research evidence, there remains a high level of enthusiasm about interactive multimedia and its capacity to link visual and auditory stimuli in IMMI. For instance, Berz and Bowman (1994) maintain that the association of pictures with sounds "would seem to be of considerable value in teaching the elements of a given piece of
music" (p. 43). The presentation of music in conjunction with other sensory stimuli has a unique importance in music instruction, especially in IMMI where audible musical stimuli are frequently the "mainstream" of information presentation. As music listening assumes a far more crucial role in IMMI than it does in other forms of computerised instruction, the development of understanding about the effects of combined media presentation on human information processing remains a vital research issue.

While the related literature provides support for the widely held belief that multimedia information presentation may enhance learning in the IMMI setting, the literature cited in the current review reflects a body of evidence which indicates that interference can occur as a result of the simultaneous or redundant presentation of information in multiple sensory modes. Whereas the combined presentation of information in complementary sensory modalities appears to enhance learning in situations where the multimedia presentation modality matches the cognitive processing style and other characteristics of the learner, for participants whose learning strategies favour single media presentation, such redundancy is likely to interfere with their assimilation of the instructional content of computerised instruction programs (Nugent:1982, Barton and Dwyer: 1987). There is also evidence to suggest that interactive multimedia users may not always choose the most instructionally effective media combinations (Shih and Alessi: 1996). Perhaps the most salient issue to emerge from this analysis is the disruption to short term memory that can result from the simultaneous occurrence of background music and visual stimuli. The results of Salame and Baddeley’s (1989) study raise serious concerns about interference to perceptual and cognitive processing that may result from the simultaneous presentation of music and visual stimuli in IMMI.

As previously reported, participants’ interactions with the media components of IMMI are highly individualistic and are likely to vary widely between users (Burkman: 1987, Higgins: 1992, Schwier: 1995). Accordingly, consideration needs to be given not only to
the perceptual and cognitive effects of combined media presentation, but also to the high level of instructional flexibility that is inherent in most IMMI programs. The complex interaction between instructional flexibility, media component selection, and participant navigation in IMMI is reflected by Berz and Bowman (1994) who point out that "people are not able to attend very efficiently to streams of simultaneous multiple information… learners might not be able to pay attention to both music and the textual explanations, let alone consider making navigational choices" (p. 42). The current review proceeds with an examination of literature relating to instructional flexibility in interactive multimedia and its effect on participants’ navigation, orientation, and instructional outcomes.

The Effects of Instructional Flexibility

Flexibility is recognised as one of the most valuable attributes of interactive multimedia and learner control is the cornerstone of that flexibility (Higgins: 1992, Stevens: 1994). As Berz and Bowman (1994) observe "current audio-CD tutorials allow the student almost total control in terms of content, sequence, and mode of presentation (text, verbal presentation, music, animation, or pictures)" (p. 8). Users can redefine both the structure and content of the learning environment and, as Nelson and Palumbo (1992) report, interactive multimedia programs

eliminate the linear, arbitrary sequencing of traditional text, allowing users to freely browse through a knowledge base. Emphasis is placed on a representational architecture which allows authors to link semantically and logically related information in conceptual webs that mirror some of the associational power of human memory. (p. 287)

Observers such as Kearsley (1988) andPhillipo (1989) confirm the longstanding assumption that the structural flexibility of interactive multimedia matches and perhaps simulates the essential characteristics of the semantic networks of human information processing. A commonly held belief is that the apparently parallel relationship between
human cognitive processing and the flexibility of interactive multimedia facilitates learning (Bush: 1945, Jonassen: 1988c, 1990, Gygi: 1990, Najjar: 1996). Music educators, such as Stevens (1994), also attest to the assumption that interactive multimedia "builds on the premise that users should be able to organise information using a computer in the same way as they organise it in their own minds; that is the free association of thoughts in the mind should be able to be emulated in a computer environment by the author making multi-dimensional connections between 'objects' (text, visual images, sound images etc.)" (p. 57).

However, concerns have been voiced about the validity of the assumption that a correspondence between the structural flexibility of interactive multimedia and the structural organisation of human cognitive processing facilitates learning (Whalley: 1990, Tergan: 1997). For example, critics argue that even the most elaborate hypertext networks are vastly overshadowed by the structural complexity of human cognitive processing and also that "arbitrary webs" of hypertext information often have little semantic relevance to users (Jonassen: 1990). Likewise, commentators such as Gygi (1990:281) question the widely held view that interactive multimedia instruction facilitates learning as it is somehow "inherently motivating", and that consequently it offers a persuasive solution that will encourage students to learn. As with many emerging educational technologies, novelty effects have been observed in the application of interactive multimedia in instructional settings. However, users rapidly become accustomed to the features of interactive multimedia and novelty effects, such as increased motivation, quickly subside (Berz and Bowman: 1995). As Schwier (1995) reports, giving control to the learner can enhance motivation, but will not necessarily improve learning outcomes and may increase the time spent on instruction. The belief that interactive multimedia represents a kind of technological panacea that is capable of solving longstanding educational difficulties is firmly refuted by Tergan (1997). He emphasises that learning cannot be attributed solely to the influence of a particular
instructional technology, but that it arises from the complex interaction of variables such as learner characteristics, instructional methods, characteristics of the instructional content, as well as the delivery mechanism.

A great deal of conflicting evidence has emerged in the literature relating to instructional flexibility and learner control. Reeves (1993) argues that the widely studied issue of learner control in computerised instruction (e.g., Steinberg: 1977, 1989, Williams: 1993) has been plagued by definitional, theoretical, methodological, and analytical problems which have led to a perception that much of the literature in the field is flawed. Since a substantial body of contradictory evidence exists, the “intuitive appeal” of learner control has never been empirically confirmed and a consensus has failed to emerge about the efficacy of allowing learner control in computerised instruction (Goforth: 1994).

Despite the existence of conflicting research evidence, instructional flexibility is still regarded as highly desirable by many commentators, even though it is acknowledged that the freedom it generates can cause problems. Interactive multimedia programs frequently provide far less instructional guidance than earlier drill-and-practice software and often contrast sharply with traditional classroom instruction models where the teacher largely controls instructional sequencing. The high level of learner control inherent in interactive multimedia instruction allows participants to make inappropriate instructional choices which may constrain learning (Schwier: 1995). Moreover, Cates (1991) warns that "many students may not have the skills and abilities necessary to learn efficiently using computer-assisted instruction" (p. 130). He maintains that the high level of learner control in computerised instruction means that students are often faced with a bewildering array of choices which may prove confusing. Garhart and Hannafin (1986) likewise found that students often felt overwhelmed by the number of choices available, and that many were poor judges of their own level of understanding of the material covered in a lesson. These investigators report a low correlation between the students’
perceived understanding of the instructional content and their performance on comprehension tests. Tergan (1997) confirms that many users cannot cope adequately with the complexity and flexibility of interactive multimedia structures and need "modelling and scaffolding support as well as more experience in using hypertext-based technologies for constructive learning" (p.267). Moreover, Nelson and Palumbo (1992) advise that

it is not sufficient to simply present information on a computer screen and assume that the information will be accurately, and completely transferred to the knowledge base of the user. Even multiple modes of presentation (a current theme of hypermedia) do not assure such transfer. Designers of such systems need to provide specific learner activities focusing on the relationships between the information in the database. (p. 293)

Related investigations by Trip and Roby (1990), McGrath (1992), and Shin, Schallert and Savenye (1994) suggest that students who have poor meta-cognitive skills may exit instructional sessions before they have adequately surveyed the program content or completed its instructional objectives.

Observers have identified two problems that often arise out of instructional flexibility, namely, disorientation and cognitive overload, which may account for the difficulties that interactive multimedia users frequently encounter. For example, disorientation occurs when users become "lost in hyperspace" (Conklin: 1987) while attempting to navigate through complex interactive multimedia databases (Canter, Rivers & Storrs: 1985, Canter, Powell, Wishart & Roderick: 1986, Jonassen: 1988a, Marchionini: 1988).

Likewise, Gygi (1990) reports that interactive multimedia "users are prone to develop hypertext-related symptoms – spatial and conceptual disorientation, for instance …" (p. 279). She maintains that "the major problem facing [interactive multimedia] systems has to do with managing complexity – how not to overwhelm users with vast amounts of
information… the effort and concentration needed to maintain several tasks or trails at the same time can be overwhelming" (p. 284). In another examination of disorientation in an interactive multimedia setting, Edwards and Hardman (1989) identified "three distinct situations when users feel lost: (a) not knowing where to go next, (b) knowing where to go, but not knowing how to get there, and (c) not knowing the current position relative to the overall hypermedia structure" (cited in Dias and Sousa: 1997:175).

Conversely, Jacobs (1992) notes that while the majority of observers assume that being "lost in hyperspace" is a fundamental obstacle to learning, there is increasing support for the notion that "getting lost may be regarded as a desirable or even necessary part of the process of structuring" (p. 120). Support for the latter view can be traced to an early report by Berlyne (1960) who suggested that curiosity about what comes next contributes to the attractiveness of information. Likewise, Fleming (1987) concurs with Mouly’s (1973) assertion that "moderate uncertainty may induce careful attention" (p. 237). More recently, Schaumberg (1996) found that participants who reported feeling disorientated in an interactive multimedia program rated the program as more attractive than those who did not feel disorientated. The notion of allowing uncertainty to stimulate instruction is consistent with models of discovery-based learning espoused by Bruner (1960, 1966), Gardner (1983, 1985, 1989), Papert (1980, 1993) and Webster (1987, 1988, 1995) which encourage self-guided exploration, divergent thinking, and an intuitive approach to the construction of knowledge. However, critics argue that a lack of structure and direction in discovery-based learning can lead to misconceptions and disorientation that constrains learning.

A second problem to arise out of the instructional flexibility of interactive multimedia is that of cognitive overload. Cognitive overload frequently occurs as a result of interactive multimedia participants needing to devote a large proportion of their cognitive resources to orientation and navigation decisions and consequently, the remaining information
processing resources may be insufficient to adequately attend to the instructional content of the program. For example, in a study of orientation and disorientation in an interactive multimedia language instruction program, Trip and Roby (1990) found "some students reported becoming completely disorientated and viewing only a small portion of the lexicon" (p. 123). When assessing the conspicuous orientation problems encountered by some participants in their study, Trip and Roby point out that disorientation can lead to "increased cognitive load which will reduce mental resources available for learning… if mental resources are engaged by navigational tasks, and if those same resources are needed for learning, it would be logical that achievement should suffer to the extent that navigation is demanding" (p. 120).

While disorientation and cognitive overload in interactive multimedia are frequently viewed from the perspective of the design features of the programs themselves, the difficulties also relate to user characteristics such as the level of prior domain knowledge of and the extent of previous experience with interactive multimedia systems. Gygi (1990) reports that

the major problem for first time users of any complex computer system, especially if they are entering new knowledge domains, is lack of guidance…how are they to decide which questions to ask, which associations to make, which hypermedia links to follow? Whereas experts may require total freedom in navigating through information spaces, novices may get quickly lost or disorientated. Their only strategy may be random choice. (p. 284)

Berz and Bowman (1994) concur with Gygi’s observations by suggesting that "either because of unfamiliarity with the environment or because of a lack of adequate prior knowledge about the subject being studied, students may choose incorrect or inappropriate search strategies given the teacher’s instructional goal" (p. 42). In an investigation of the navigational pathways that users follow through hypertext
documents, Horney (1993b) found that users "with different skills and purposes will navigate hypertext documents using markedly different patterns…" (p. 81). He concludes that interactive multimedia documents "can be rich in structure, a richness that readers can exploit or ignore" (p. 82).

Whether IMMI participants engage with the structural "richness" of the interactive multimedia program as described by Horney, or whether they ignore it, is a principal concern of the present investigation. Moreover, an examination of the factors that influence participant interactions with IMMI music examples forms a central concern of this work. A critical issue is that disorientation and cognitive overload which can occur as a result of a high level of instructional flexibility may encourage IMMI participants to minimise or to avoid listening to music examples. As Berz and Bowman (1994) have observed, "students are not limited to a linear presentation of the music, but may jump from section to section - a process that may disrupt the continuity of the music" (p. 41).

Berz and Bowman (1994) suggest that IMMI participants may need to adopt new strategies in order to learn efficiently from interactive multimedia. They point out that "these strategies are possibly quite different from those employed in other, more conventional approaches, such as with books (text and pictures) and lectures. There is a tendency for students to move off task when using CAI, such as getting too involved in the gaming aspect of instructional games or moving to unrelated material in hypermedia programs" (p. 57). Without appropriate instructional strategies and guidance, IMMI participants may not engage at an optimal level with IMMI music examples and thus will have restricted opportunities to develop an understanding of the musical concepts that the program is attempting to teach. Disorientation, and the cognitive overload that frequently results from attempts to navigate in complex unfamiliar instructional settings, may contribute to the tendency for IMMI users to move "off task" and ignore music listening. The available evidence suggests that without adequate instructional guidance, some
IMMI participants may fail to conceptualise their multimedia interactions in an instructionally effective manner, particularly where there is a high degree of learner control and where participants are new to an instructional domain or have limited prior experience of the characteristics of interactive multimedia systems (Gygi: 1990).

**Previous Investigations of IMMI**

Researchers have only recently begun to investigate the use of interactive multimedia in music instruction and very few studies have been conducted that specifically target IMMI (Berz and Bowman: 1995). As Higgins (1992) has observed, research has been driven by continuing rapid technological evolution and as a result, literature in the field is dominated by developmental and feasibility studies. Only a relatively small number of effectiveness studies, which typically seek to compare IMMI with traditional instruction, have been conducted. In one of the first studies of the application of interactive multimedia in music instruction, Lee (1989) developed an IMMI "listening station" designed to assist participants in the aural recognition of core music literature. When compared to the results of participants who used traditional "repeated listening" strategies, subjects using the IMMI listening station demonstrated not only superior performance in recognition tests, but achieved the improved results in shorter listening times. Lee attributes the superiority of the IMMI listening station to the high level of participant control.

A similar study by Hughes (1991) compared the effectiveness of an interactive multimedia music "listening station" with that of traditional "repeated listening" strategies in the development of recognition ability of standard music repertoire. The most salient finding was that the performance of subjects who used the listening station was significantly higher on post-treatment recognition tasks than subjects who used traditional listening strategies. Hughes argues that the observed differences can be attributed largely to the design of his program which, in contrast to previous
computerised music listening programs that simply automated earlier programmed instruction models, gave listeners in the experimental group a high level of precise control over music selection and audition. Hughes reports from "informal discussions" with members of the experimental group that "considerably less time was spent listening to those musical works included in the study than to those required listenings in the general class schedule" (p. 50). The implication of this observation is that the interactive multimedia listening station reduced the amount of listening time required to attain mastery of the music listening repertoire. The results of Hughes’ study appear to support this contention, even though the total time that participants in each treatment group spent listening to music was not recorded. Hughes judiciously acknowledges that the amount of time participants spend listening to music may have important ramifications when comparisons are being made between computer-mediated music listening and traditional music listening strategies. He recommends that future studies of interactive multimedia music listening should scrutinise participant listening times. The current investigation of participant interactions with audio components of IMMI embodies Hughes’ recommendation.

In a case study of student interactions with twentieth century art music, Sigurjonsson (1991) employed a researcher-designed IMMI program which facilitated "non-directed" music listening. Participants listened to a short composition by Webern, then rated the piece on four criteria, namely, liking of the piece, performance quality, complexity, and their level of understanding of the piece. In response to a hypothetical question designed to stimulate contemplation of the composition, participants provided a written discourse describing the essential characteristics of the piece and their reactions to it. Following a second audition of the piece, participants again rated the piece according to the criteria used in the first rating. As Sigurjonsson had predicted, participants tended to rate the piece higher on each of the four criteria in the second rating. The researcher suggests that increases in the second rating may have been due to increased participant attention and
interaction with the composition as a result of their having to prepare the written discourse.

Sigurjonsson found that both the total amount of time that participants spent in the IMMI program and the level of participant interaction with the program varied widely. Overall, graduate students spent more time using the IMMI program, and had a higher level of interaction with program components than undergraduate students. The investigation also revealed positive relationships between the amount of time participants spent in the program, their level of interaction with the program, and the comprehensiveness of their response in the written discourse about the composition.

Goodson (1992) designed an introductory module of an IMMI program in which she aimed to teach fundamental musical concepts such as pitch and dynamics. In testing the instructional effectiveness of the module she found that participants using the IMMI program required 66 percent less instructional time to produce equal or higher achievement than participants who received traditional instruction. Goodson emphasises the importance of music listening in IMMI and calls for further research to analyse the learning processes participants employ when they use IMMI programs. The present study examines IMMI learning processes from the perspective of participant interactions with IMMI audio components.

While the studies of Lee, Hughes, Sigurjonsson, and Goodson were concerned with the effectiveness of researcher-designed IMMI programs, Duitman (1993) investigated the instructional effectiveness of commercially-available IMMI programs which, he maintained, differ from researcher-designed programs in terms of their complexity and the scope of the information they contain. Such differences, he argued, indicate that commercial IMMI programs warrant specific investigation. Like the earlier investigators, Duitman compared the learning of students who used IMMI programs as a music listening resource with that of students who used traditional listening strategies. While he
found statistically significant learning improvements in both treatment conditions, no significant learning differences were found between the treatments. In an apparent contradiction of the earlier findings of Lee, Hughes, and Goodson, the results of Duitman’s study suggest that subjects in the IMMI group spent significantly more time using the programs than subjects in the traditional listening strategies group had spent using the listening laboratory. Crucially however, the finding that the IMMI participants spent more time "directly interacting with music information" (p. 49) is based on participant self-reports of usage of listening laboratory resources which are notoriously unreliable (Hargreaves: 1986, Tuckman: 1994, Gall et al. 1996, Burns: 1997). Nonetheless, Duitman prudently acknowledges the importance of the "time factor" in IMMI, and calls for research "to determine how much of the CD-ROM time of use is spent listening to the music only, how much time is spent listening while reading about the music, and, finally, how much time is spent reading alone" (p. 49). Duitman’s call for further research into the "time factor" in IMMI participants’ music listening points explicitly to the principal concern of the current investigation which, likewise, employs commercially–available IMMI programs.

In the first substantial investigation of participant navigation behaviours in IMMI, Berz (1995) found that participants accessed the components of his researcher–designed IMMI program in widely divergent ways. Berz reports considerable variation in the types of navigation decision (linear browsing, hypertext leaps, jumps to menu), the amount of time participants spent studying the content of individual program screens, and in participants’ choice of media presentation mode (eg., video, sound, and text). From the perspective of the current investigation, perhaps the most striking finding was that while participants frequently accessed IMMI music examples, they rarely listened to the music excerpts in their entirety. For example, Berz found that 26 percent of participants interrupted every music example that they selected before it had concluded, while 68 percent of participants interrupted at least half of the music examples that they had
chosen. Only one participant listened to each selected music example in its entirety. Berz expresses grave concern about the educational implications of such a high level of participant interruption of IMMI music examples. He notes that the interruption of music examples has the potential to undermine music learning, as students fail to gain an overall conception of the aesthetic and structural properties of a musical work because they may never hear a complete performance of it. While Berz points out that the brevity of the music examples employed in his IMMI program and the limited instructional objectives of the study may have contributed to the high level of interruption of music examples, the interruptions to musical flow were akin to "stopping reading in mid-sentence when studying literature and poetry…” (p. 181). He nominates the interruption of musical flow in IMMI as a critical research issue, and the results of his study of navigation behaviours in IMMI provide robust support for the ongoing investigation of that issue in the current study.

In addition to the investigations outlined above, reviews of the application of technology in music instruction provide an invaluable overview of emerging trends in IMMI. Reports by Higgins (1992), Berz and Bowman (1994, 1995), Stevens (1994), and Bray (1997) synthesise research findings and identify critical research issues, evaluate currently available IMMI resources, and make recommendations for future research. For example, Higgins points out that researchers have found it difficult to keep pace with advancing technology which has tended to evolve so quickly that by the time an investigator has conducted a study using current technology, that technology has been superseded. The rapid evolution of instructional technology, and a desire to remain at the "cutting edge" by continually adopting innovative technologies, has meant that music educators have not sufficiently tested "in place" technologies, and have yet to develop a solid research foundation for the use of technology in music instruction. Likewise, Berz and Bowman reflect a growing concern that research has been overly concerned with the evaluative comparison of computerised music instruction with traditional teaching
methods. These critics argue that researchers have focused too heavily on attributes of IMMI technology itself without adequately addressing learner characteristics and the interactions that occur between learners and the instructional technology. Higgins maintains that investigators have too often proceeded on the assumption that it is the technology that influences instructional outcomes without acknowledging that factors such as learner ability, prior knowledge, and motivation play a critical role in the instructional interaction. He points out that "a unique feature of technology instruction is the effect of the manner in which each individual learner approaches the media and the choices made in the interaction" (p. 491). As a result of the diversity of choices and interactions available to the participant, Higgins reasons that the effects of computerised instruction will differ for each individual, even though they may have used the same computer program. Consequently, he recommends that "formative, open-ended observational research" be undertaken urgently to establish the relationships between individual interaction styles of learners and the attributes of instructional technology. The present investigation carries out Higgins’ recommendation by examining participant interactions with IMMI audio components.

This examination of studies that specifically investigate the application of interactive multimedia in computerised music instruction indicates that the current body of research is very limited. Nonetheless, each of the investigations reviewed here has identified critical research issues that warrant further scrutiny. A consensus to emerge from these studies is that there is a lack of understanding of how individual participants interact with the components of mediated instruction. The present investigation addresses this lack of research-based understanding by addressing a specific subset of the problem, namely, participant interactions with the audio components of IMMI.