

**PRESENCE 2008**

**Proceedings of the 11th Annual International Workshop on Presence  
Padova, 16-18 October 2008**

Printed by  
CLEUP Cooperativa Libreria Universitaria Padova  
Padova 2008

Edited by Anna Spagnolli, Luciano Gamberini

ISBN: 978-88-6129-287-1

*© The copyright for this publication as a whole stands with HTLab. The copyright of each separate paper published within these proceedings remains vested in its author. Authors have assigned to Presence 2008 organizers and ISPR (International Society for Presence Research) the on demand availability rights for their work and the right to create a derivative work from it, including publication on ISPR website and on the conference proceedings.*

## No Vision is Complete without Sound\*: The Sonification of Telepresence

Gary Pettey<sup>1</sup>, Cheryl Campanella Bracken<sup>1</sup>, Bridget Rubenking<sup>2</sup>, Emily Appley<sup>1</sup>, Mycal Brown<sup>1</sup>, Dawnnay Butler<sup>1</sup>, Kelly Cline<sup>1</sup>, Zlatko Coralic<sup>1</sup>, Jennifer Hargiti<sup>1</sup>, Han Liou<sup>1</sup>

<sup>1</sup>Cleveland State University, <sup>2</sup>Indiana University  
{g.pettey@csuohio, c.bracken@csuohio.edu}

### Abstract

*In an experiment exploring the impact of sound on sensations of telepresence, 126 participants watched a video clip using either headphones or speakers. The results reveal that sound is an important factor in stimulating telepresence responses in audiences. Interactions between sound delivery and screen size were also revealed.*

### 1. Introduction

While birds may be able literally to sleep with one eye (and half a brain) open [1], humans appear to monitor their sleeping surroundings using sound. Patient's under anesthesia have demonstrated to learn words played to them through headsets even though they had no explicit memory of hearing them [2]. For humans, the soldier's notion of sleeping with one eye open may actually be more appropriately stated as keeping one *ear open* [3]. Meaningful and unexpected sound appears to be actively processed in active parts of the brain that are "awake" even while we sleep. While we often say vision is our most important sense, in fact, sound and how that sound is reflected within its environment may well give us more information—especially as it can tell us something about things that happen outside our visual range and even in the dark [4].

Engineers have worked hard over the last century to reproduce sound and then to recreate it more realistically in our living rooms and theaters. Stereo split the sound left and right. Surround sound gave the presentation depth. Dolby took the hiss from the tape removing the artificiality of the technology and leaving a cleaner reproduction of the music. Digital recording further recreated sound and a soundscape closer to original sound. Today Dolby 5.1 and surround sound allow home theaters where listeners feel as though they've been dropped into the environment of their movies.

In theaters, Sony Dynamic Digital Sound (SDDS) is an 8 channel reproduction. Digital Theater Systems (DTS) advertise their sound as bringing entertainment to life. Lucas' THX theater system is a balancing of the sound to the environment (soundscape) "At THX we believe that 'sound is fifty percent of the movie experience'," (<http://www.thx.com/products/home/displays/shopping.html>) and by certifying the theater as

THX, "What better way to escape from reality than by going to the cinema? The more immersive the performance, the more you enjoy the feature presentation. This is where a cinema built to THX standards comes into play" [5] (<http://www.thx.com/cinema/index.html>).

Different elements of the aural modality have differential impact on the presence experience. This paper will attempt to do two things: first, it will explore the notion of acoustic architecture and its implications both for mediated experiences from a more conceptual, perhaps less technical perspective; and secondly it will ask the question whether varying elements of the soundscape help us to better understand how the aural modality contributes to dimensions of telepresence?

### 2. Presence

There have been numerous definitions of presence put forth [6-9] The current study defines presence as a perceptual process where the media user somehow looks past or overlooks the technology to experience the medium [10].

#### 2.1. Telepresence

The concept of telepresence has been identified as a multi-dimensional concept [8, 9, 10] with immersion, spatial/physical presence and social/behavioral realism being seen in most telepresence measurement scales [8] Lombard argued for the use of the term "telepresence" instead of "presence" [11] to stress the role of technology. Much of the research is conducted in highly immersive technologies such as virtual reality systems. There is a growing number of researchers within Communication who are investigating telepresence [12-16]. Bracken, Pettey, Rubenking, and Guha offer a definition of Telepresence that makes a clearer break from the virtual reality aspects of Presence and clearly puts it into a psychological context where the media observer uses the technology. Telepresence is an interaction among the user, the content and the technology.

Telepresence requires the use of technology and results in a psychological state in which media users voluntarily suspend the experience of mediation in

order to feel a sense of connection with the mediated content they are using (i.e., a part of the action, connected to characters, involved in the story line). This state is often mediated by the expectation of the technology, the media content, and characteristics of the media user. [15, p. 3]

## 2.2. Telepresence and Television

Several studies have demonstrated that television can evoke sensations of telepresence in viewers [13, 14, 17]. Even prior to these studies addressing telepresence, earlier studies found that under certain conditions viewers can respond to objects and people on the screen as if they were real [18].

A common method to study telepresence and television has been to manipulate screen size. There are consistent findings that larger screens often lead to audience responses that are consistent with telepresence [19, 20, 21] or were classified as telepresence [17, 20, 22].

## 2.3. Telepresence and Sound

Recent studies have examined the impacts of the aural modality on elements of telepresence. While the emphasis in presence research has been on the visual, aural effects alone or their interaction with the visual elements have been shown to be important to dimensions of telepresence.

Congruency between aural and visual modalities has been shown to increase levels of presences and enjoyment [23]. Using a within subjects design, the authors demonstrated that the congruency between the aural and visual channels reinforced subjects experience across a spectrum of telepresence measures [23]. While most generally, their data show that sound vs. no sound showed the largest differences, in several measures they found that subjects perceived differences in acoustical architecture between a conventionally shaped room and a cathedral shaped space about 3x as large.

Nunez (2007) suggests that the expectations of the subject may impact the presence experience [24]. He found that an added sound track (non-diegetic information) congruent with the subject's expectations had impact on how natural the subject found the experience. This congruency, even when controlling for emotional response, predicted to higher reports of the naturalness of the experience. This is interesting given that it reflects a learned response of mediated experience because it has a sound track. The lack of one makes leaves a void in the experience. In an "ear budded" world, perhaps we shouldn't find this tendency surprising. Further, Nunez found relationships between Izard's (1977) DES-II (Differential Emotion Scales) [25] and various levels of presence.

Based on the findings of these studies concerning congruency, the following hypotheses are proposed:

Hypothesis 1a: Participants who have a presentation where the visual and aural modalities are congruent will

report higher levels of *immersion* than the participants who had the incongruent presentation.

Hypothesis 1b: Participants who have a presentation where the visual and aural modalities are congruent will report higher levels of *spatial presence* than the participants who had the incongruent presentation.

Hypothesis 1c: Participants who have a presentation where the visual and aural modalities are congruent will report higher levels of *social realism* than the participants who had the incongruent presentation.

The complexity of sonifying a mixed reality space was explored by Le Groux, Manzolli and Verschure [26]. They argue that sonification is crucial to feelings of immersion. They note the difficulty of sonifying a virtual space in terms of soundscapes, synthetic voices and dynamic emotive music.

Focusing only on the aural modality, other researchers [27] found increased accuracy in identifying sonic events emanating from their rear when synchronized with other vibration. Sonic events originating from the front of the subject were significantly less accurately identified. Information coming from the rear must be indentified and processed largely independently of the visual.

Kallinen and Ravaja (2007) using a within subjects design compared headphones to speakers while listening to news [28]. Subjects reported "liking" the headphones more, even though they reported that "quality" of the sound from the two sources was equal. Subjects reported the speaker condition to be more "realistic" while listening to news from a computer. One interpretation is that that is how they perceive most people listen to the news, or at least news on a computer—a level of social realism. Using physiological measures and self report data, they predicted that headphones would elicit a more intense and immersive listening experience. They found that headphones increased attention and by extension involvement. The suggested this was due to technologies that tend to isolate the listener from her surroundings [29].

Based on the findings of these studies concerning sonification, the following hypotheses are proposed:

Hypothesis 2a: Participants who have a presentation where the aural modality is delivered through headphones will report higher levels of *immersion* than the participants who listened through the television speakers.

Hypothesis 2b: Participants who have a presentation where the aural modality is delivered through headphones will report higher levels of *spatial presence* than the participants who listened through the television speakers.

Hypothesis 2c: Participants who have a presentation where the aural modality is delivered through

headphones will report higher levels of *social realism* than the participants who listened through the television speakers.

### 3. Sonification

One of the problems in getting a handle on the aural modality is that there is no single discipline in which it is studied, and the language used across disciplines is often incompatible [4]. Blesser's work has been at the junction of audio, acoustics, perception and cognitive psychology. Blesser and Salter's 2007 work, *spaces speak, are you listening: experiencing aural architecture* provides a serviceable conceptual framework [4].

#### 3.1. An Aural Language

A *sonic event* is the creation of a sound wave that rises above the background (snapping one's fingers). If we snap our fingers we hear the snap, but we also hear the reflection of that snap from the passive objects in our space (walls, ceiling, floor). Thus the same sonic event is perceived quite differently if done in a small room or if it is done in a large one. The space in which sounds are heard are *soundscapes*. Just as we need light to appreciate a landscape, we need sound to perceive/appreciate a soundscape.

An *aural architecture*, then, "refers to the properties of a space that can be *experienced* by listening" (p.5) [30]. We experience space aurally. We can *taste* a space aurally. We know the difference between the aural space of a bathroom, relatively small physical space with hard reflective surfaces (tiled walls and floor and porcelain fixtures), especially in comparison with a living room, which is normally a larger physical space with softer absorbing surfaces (drapery, carpeting, furniture). We can imagine the different aural experiences of snapping your fingers in one space versus the other. These two soundscapes have different aural architectures. We can imagine/recall the aural architectures of many soundscapes we've experienced. An empty house "sounds" different than an occupied one, a small chapel different from a Gothic cathedral, a fast-food restaurant from an expensive dining room. Our experience of a place is directly related to our experience of the places aural architecture.

That we can imagine the aural architecture of a space assumes that we have a *spatial awareness*. Spatial awareness would include more than the complexity of a reflected sonic event or even series of such events, it also includes the individual's emotional and behavioral experiences with the space. Blesser and Salter conceptualize the complexity of spatial awareness as having "spatial attributes auditory perception, personal history and cultural values" (p. 11). They argue that spatial awareness influences us in at least four ways:

1. **Social Spatiality:** Influences social behaviors. Some spaces may emphasize aural privacy, which may aggravate loneliness or increase social cohesion.
2. **Navigational Spatiality:** Allows us orientation that allows to maneuver. In darkness, spatial awareness actually replaces vision.
3. **Aesthetic Spatiality:** It conveys a sense of the aesthetic. Spaces without low in acoustic features are aurally barren even boring.
4. **Musical Spatiality:** Can enhance complex aural experiences such as music and voice giving them depth and texture (or not).

The compositions of these spatial "hearings" consist of aural elements or packages of aural elements. Consider two phobias that rely heavily on different packages of aural elements: Claustrophobia and Agoraphobia. The sufferers of each phobia suffer similar symptoms, and each has a large component of spatial awareness. A space has an aural architecture. Sonic events within very tight spaces and very large spaces are very different. Think about if you were blindfolded and someone told you to yell your name. If there was a solid object in front of your face you would feel the reflection of that sound. The reflection would be so strong it might startle you—make you duck. Even blindfolded and with no identifiable sonic event, we could tell if we were in a small closet—the close-in feeling is the sense of the aural reflection of the space. Sonic waves travel back to you noticeably quicker than usual. In an open space, say on top of a mountain or a high building, the opposite happens. People's balance is affected in part because the aural reflection does not return in an expected period of time (time is crucial to all aural events). Without that reflection, your inner ear is lacking a normal gyro and it feels disorienting, even nauseating. Blesser and Salter use the bathtub/ocean distinction. In a bathtub the waves of the water return quickly to you, even canceling one another out. In the ocean, no amount of splashing will cause the waves you send out to be reflected to you.

Blesser and Salter describe the range of auditory awareness as: *sensation, perception and affect*. They note that generally awareness is problematic because it implies consciousness, but sensation, they argue, is below the conscious level. Mere detection. We can sense the closeness of the closet without conscious realization of the aural features involved. The lightness of a sunny day when compared with the heaviness of a foggy day is one working at this level. These may well be the important subconscious cues that would be important in elements of immersion and perceived realism, for example.

Perception or recognition is when conscious processes have been engaged but not emotional ones. Cultural influences and personal experiences give meaning to what may have barely risen above sensation. A musician can tell the difference between a violin string and a cello string. Most westerners can tell the difference between a guitar string and a piano (a plucked rather than struck string). Language is a set of sonic events to which we attach cognitions. The more cultural influence and personal experience we have with a language the more meaning we may get from any set of sonic events. At

some point the awareness moves from the cognitive realm to one of emotional meaning.

Finally, affect is emotionally engaged hearing and carries meaning. Hearing a song that reminds you of a party and brings a smile to your face, or hearing a lawnmower that reminds you of long unhappy days mowing as a child constitutes such awareness. Such sounds can create high emotional—even visceral—responses [31]. Sometimes the listener may not even consciously understand the relationship between the sound and his/her reaction. In some instance music and sound has been used to create dream states [32, 33]. Blesser and Salter note that working with affective extremes of awareness are probably too complex to be adequately study without brain scans. As the effect is a complex biological, psychological, social one, and individuals are unlikely to be able to adequately describe their emotional state (describe their feelings), such overt affect effects are probably less fruitful at this point. They suggest that awareness at the other end of the spectrum, they call them subliminal, correspond to more subtle levels of awareness or moods. The task then is to use/find sonic events and acoustical spaces that 1) have some meaningful content for listener and 2) adequately engages the listener. “To use a food metaphor, sonic events are raw ingredients, aural architecture is the cooking style, and, as an inseparable blend, a soundscape is the resulting dish” (p. 15) [4].

### 3.2. Experiencing a Soundscape

Each culture controls what we expect for social distance. Hall divided these spaces into four: the intimate, the personal, the conversational and the public. Acoustic arena interact with these social spaces [34]. The intimate space is highly restricted to very close friends and lovers. Acoustically, this space is very private. Words and sounds are for the intimates only. Silence will greet intruders. The personal space is for friends and acquaintances. Think of a group of students or colleagues talking in the hallway. The conversational space is the comfort space reserved for stranger interactions. Finally, the public space has many varying rules usually based on some social rules for the space. There are few restrictions on a public beach or park, but more restrictions may exist on public indoor spaces such as an airport concourse. How one reacts to a soundscape should differ depending upon the interplay of the acoustical arena and the social distance involved. Headsets would be the most intimate of sound. Listening to a concert in an outdoor venue, a very public one. Watching television in your own living room is pretty private almost intimate, while watching it at someone else’s living room would be much more public.

Based on this explanation of acoustical arena and social distance and considering how this might impact the sense of telepresence when combined with aural and visual congruency of the presentation, we offer the following research questions:

RQ 1a: If headsets create more intimacy (thus more immersion) and congruent presentations are also more

immersive, will they interact to produce effects over and above their individual main effects?

RQ 1b: If headsets create more intimacy and a soundscape that is more separate from the viewing environment and congruent presentations allow for better spatial awareness, will they interact to produce effects over and above their individual main effects?

RQ 1c: If speakers are seen are more real and a congruent presentations allow for higher levels of social realism, will they interact to produce effects over and above their individual main effects?

## 4. Methods

This overall design employed a between-subject 2 x 2 experiment. The 126 subjects watched a video clip where the audio was delivered either through standard headphones or an stereo television speakers. The audio was either the one from the motion picture or one where the audio had alternative non-diegetic sound.

### 4.1. Stimulus

The video shown to the participants was from the movie *Quest for Fire*. This film was made in 1983 and was used because there is no character dialogue. It was also believed that few participants would have previous exposure to the film. The video clip was approximately 15 minutes in length, of which the first a minute and a half was from another source and provided a baseline for the physiological measures (not used in this paper). The movie clip began immediately following the introduction. The natural soundtrack of the film contained sparse classical music that had a primitive sound, which allowed it to match the prehistoric period and theme of the video. The non-diegetic music material was taken from the song Boulevard (Parisian Mix) by Cinemascope from the CD Saint Germain Lounge: Rendez Vous. The non-diegetic music was inserted in the final scene and contradicted the mood of the scene, which showed one prehistoric tribesmen observing another make fire. The incongruent music that was added was electronic and modern in nature. The non-diegetic music was also very upbeat and pleasant sounding, which tended to mitigate the drama of the fire creation which was the climax of the film and the intensity could be seen on one of the men’s’ faces.

### 4.2 Independent Variables

**4.2.1. Soundscape/Delivery** The first independent variable was how the participant experienced the sound. Soundscape was manipulated by having one condition group listen to the video clip through the television speakers (more public) and the other condition group listening through headphones (more intimate).

**4.2.2. Congruence** Congruence had two conditions. The first contained the film's natural congruent audio content. It contained grunts and diegetic sound elements with occasional classically based music soundtrack. The other condition (incongruent) contained the same soundtrack until the final, climatic scene where the music was overlaid. This music was not jarring to viewers in a pretest, but left them confused about what the scene was "trying to say."

### 4.3. Dependent Variables

The amount of presence experienced by the participants was measured using a multidimensional presence scale [8, 13, 14]. The presence dimensions included were immersion, spatial presence and perceptual realism and are detailed below.

**4.3.1. Immersion** Immersion was measured by asking participants to respond to six statements using a scale from not at all (1) to very much (7). The immersion scale was developed to assess the extent to which audience members feel they are absorbed in the media programming. Examples of items include: "How involving was the video," and "How engaging was the story." The scale had a Cronbach's alpha of .90.

**4.3.2. Spatial Presence** Participants responded from very strongly disagree (1) to very strongly agree (7) for six Likert-type statements designed to measure the extent to which television viewers feel a sense of sharing a physical space with the mediated environment (Lombard & Ditton, 1997). Examples of the statements include: "How much did it seem as if the objects and the people you saw/heard had come to the place you were," and "How often when an object seemed to be headed toward you did you want to move out of its ways." The scale had a Cronbach's alpha was .92.

**4.3.3. Social Realism** Participants responded from very not at all (1) to very much (7) for four statements intended to measure the extent to which television viewers feel a sense of realism when viewing television. The four statements were "The events I saw/heard would occur in the real world," "The events I saw/heard could occur in the real world," "How relaxing was the experience?" and "The way the events I saw/heard occurred is a lot like the way they occur in the real world." The scale had a Cronbach's alpha was .82.

### 4.4. Participants

Undergraduate students were recruited from several social science courses to participate. One hundred twenty-six undergraduate students from introductory classes were given credit from their instructors for participation. The group was 57 percent female with an average age of 24.01 and a standard deviation of 7.2.

### 4.5. Procedure

Each participant was met by the experimenter and escorted into a carpeted, 8 x 10 foot room that contained a television, DVD player, and a comfortable chair that faced the television screen. Various other amenities, such as a decorative table lamp and pictures on the wall, provided the verisimilitude of a living room.

The experimenter explained that the participant would be watching a 15-minute video clip and that after the clip they would exit the viewing room. After the participant exited the experiment room, the experimenter escorted the participant to a second room that contained tables and chairs. The experimenter instructed the participant to sit at one of the table and to answer a paper-and-pencil questionnaire. The experimenter emphasized that there were no wrong answers and that the participant should follow the directions in the questionnaire. The entire procedure took about 35 minutes.

### 5. Results

As the design contained multiple independent variables and three dependent variables that previous studies had shown to be moderately correlated, a MANOVA was run. None of the main effects reached significance, but the one-way interaction component was significant (Pillai's Trace,  $F=2.86$ ,  $p<.05$ ). Table 1 presents the ANOVA sets from the MANOVA.

Hypotheses 1a-c predicted higher levels of immersion, spatial presence and social realism in the condition where the sound was congruent with the visual. There is no support for these hypotheses.

Hypotheses 2a-c predicted higher levels of immersion, spatial presence and social realism in the condition where the sound was congruent with the visual. There is no support for these hypotheses.

Research questions 1a-c asked whether congruence would interact with the sound delivery for each of the variables. Previous studies had suggested that sound delivery may interact with other independent variables [15].

Here for the interaction we see that the MANOVA was significant, while (as seen in Table 1), there were no main effects for soundscape/delivery or the congruency of the aural and visual modalities.

RQ1a asked about the effect for soundscape/delivery on immersion. The interaction effect for delivery and congruence for immersion was significant ( $F=7.39$ ;  $p<.01$ ;  $\eta^2=.07$ ). The interaction is transverse with Headphones & Congruent ( $M=23.46$ ;  $SD=8.56$ ) and Speakers & Incongruent ( $M=22.69$ ;  $SD=9.23$ ) in the high cells, and Speakers & Congruent ( $M=18.54$ ;  $SD=6.86$ ) and Headphones & Incongruent ( $M=18.72$ ;  $SD=9.57$ ) in the low cells. These data are presented in Figure 1.

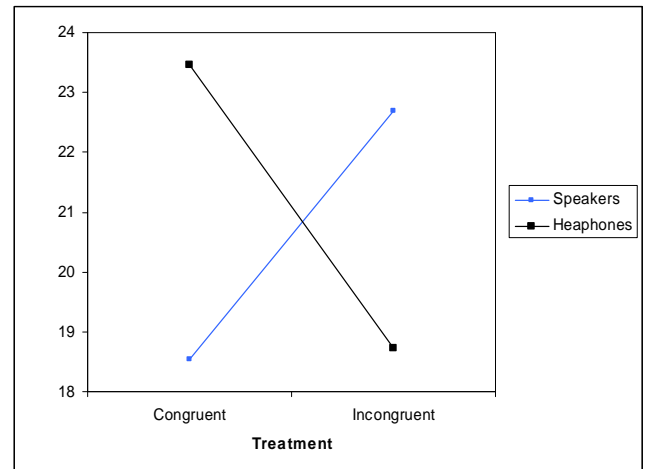
Source of Variation	Sum of Squares	df	Mean Square	F	$\eta^2$
<b>Immersion</b>					
Main Effects					
Congruence	2.36	1	2.36	0.03	.00
Delivery	6.39	1	6.39	0.09	.00
Two-Way Interactions					
<b>Congruence</b>					
<b>x Delivery</b>	<b>548.25</b>	<b>1</b>	<b>548.25</b>	<b>7.39**</b>	<b>.07</b>
<i>Spatial Presence</i>					
Main Effects					
Congruence	25.58	1	25.58	0.41	.00
Delivery	17.91	1	17.91	0.28	.00
Two-Way Interactions					
Congruence					
x Delivery	8.00	1	8.00	0.13	.00
<b>Social Realism</b>					
Main Effects					
Congruence	6.50	1	6.50	0.93	.01
Delivery	2.06	1	2.06	0.09	.00
Two-Way Interactions					
<b>Congruence</b>					
<b>x Delivery</b>	<b>99.33</b>	<b>1</b>	<b>96.33</b>	<b>4.33*</b>	<b>.04</b>

Note: +p<.10; \*p<.05; \*\*p<.01; \*\*\*<.001

**Table 1 Analysis of Variance Table: Congruence and Delivery for Presence Dimensions**

For RQ1b, the soundscape did not interact with the congruency of the presentation for spatial presence.

RQ1c asked about the effect for soundscape/delivery on social reality. The interaction effect for delivery and congruence for social reality was significant ( $F=4.33$ ;  $p<.05$ ;  $\eta^2=.04$ ). The interaction is also transverse and follows the same pattern as with immersion. Headphones & Congruent ( $M=11.61$ ;  $SD=4.80$ ) and Speakers & Incongruent ( $M=10.62$ ;  $SD=4.79$ ) in the high cells, and Speakers & Congruent ( $M=9.31$ ;  $SD=4.74$ ) and Headphones & Incongruent ( $M=9.31$ ;  $SD=4.71$ ) in the low cells. These data are presented in Figure 2.

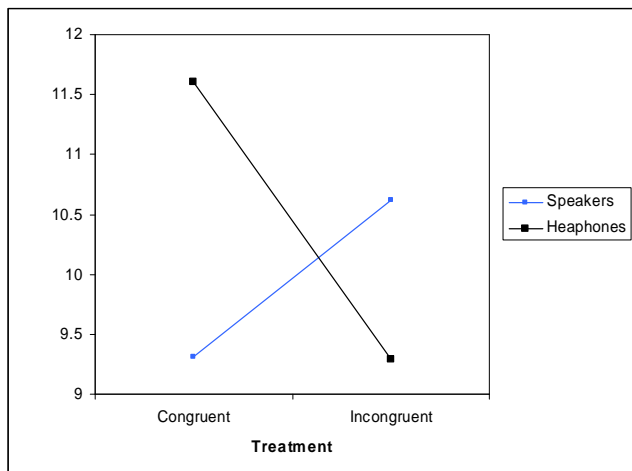


**Figure 1 The Interaction of Sound Delivery and Congruence on Immersion**

## 6. Discussion

This study tried to further explore the conceptualization of sound's role in telepresence. We attempted to provide a theoretical basis for the understanding by using Blesser and Salter (and others) notion of soundscape. Other studies have noted there are differences between perception among listeners according to whether they experienced a presentation through headphones or television speakers. Given that we considered this the best place to begin. Even though other studies found main effect differences between aural and visual congruency in measures of telepresence, we found none. It should be noted that many of those previous studies were working primarily in a virtual reality setting maximizing differences. We worked with in a mediated environment with considerably subtler manipulations. Given that we didn't seek to maximize effects but rather were more concerned with external forms of validity, it may be that increasing the number of subjects would be beneficial in finding main effects for congruency and for soundscapes.

The significant interactions between congruency and soundscapes are noteworthy. That the interactions were transverse, suggests that increasing the number of subjects may not lead to significant main effects. Further that the relationships are the same for both immersion and social realism (but not spatial awareness) suggests we have encountered something about soundscapes that crosses the two types of telepresence.



**Figure 2 The Interaction of Sound Delivery and Congruence on Social Realism**

### 6.1. Immersion

Bracken et al found that sound a main effect found for immersion, but no interaction [15]. They also presented the non-significant interaction. In that study they varied the pace of the content (as well as the screen size). So subjects saw one of two different presentations—both with congruent aural/visual modalities. They found that headsets showed higher immersion than speakers (main effect). This is consistent with our prediction and understanding of the social distance of a soundscape—with headphones being more intimate and thus more immersive. It is also consistent with other findings suggesting that headphones are more immersive. Their non-significant interaction shows that those watching on a personal medium (iPod) showed higher levels of immersion than those who watched on more public one (standard television).

So when the aural/visual are congruent, headphones show higher levels of immersion. This is also the case in this study. Looking only at the congruent, headphones were more immersive than speakers. The question then is what is going on within the incongruent condition. We will further consider this below where social realism shows a similar sonification/congruent relationship.

### 6.2. Spatial Presence

Given the previous work with in this area, finding no impact for spatial presence was surprising. But much of that literature deals with presence and virtual reality. Spatial presence is rarely found in popular media which is less immersive than virtual reality systems or even video conferencing. We believe that our technology may have also been insufficient to produce the effects we expected. Our television sets were new, high quality and stereo, while our headsets were more pedestrian—midranged stereo headphones

that people commonly buy for use with their personal media. We believe that we should try this with top end stereo headsets may be a fairer comparison and worth doing in the future.

### 6.3. Social Realism

Social realism shows a interaction effect with congruency and sonification. Consider first the congruent side. Again here we see that headphones showed higher levels of social realism than television speakers. This is consistent with the findings for immersion and is consistent with the intimacy hypothesis. Kallinen and Ravaja, had suggested that listening to news was more real than headphones, but they didn't use the visual modality [28]. Perhaps when there is discrepancy in the two modalities, people trust (find more real) the aural. Bracken, Pettey, Rubenking, and Guha, may shed some light on this [15]. They report that for their television cells, headphones showed higher levels of social realism than did the speaker cells. That is consistent with the findings here. Looking only at the congruent cells, headphones were higher than speakers. Again it is the incongruent cells that give us pause (as they did with immersion above).

One explanation that is testable is that when the technology leads to expectations of intimacy (and higher levels of telepresence) higher levels are experienced. Congruence in the two modalities would resonate with the intimate expectation and lead to higher levels of telepresence. But when there is incongruence in the modalities, it breaks the intimate expectation—indeed the subjects' expectations are lower, and they find them more consistent with a more public soundscape, which is also less telepresent. The missing variable then is subject expectation of the experience both in terms of content and the technological reproduction of that content. Should this be borne out, it would be an important element as it is not commonly part of the telepresence explication.

Clearly the most important implication of this study is that soundscapes matter. They can be explicated and will have impact on elements of telepresence. Further, the consistent finding with other studies that sound have interactive effects, leads us to conclude that telepresence may need further elaboration possibly including the audience expectations of their mediated experience—both in terms of content and delivery and possibly the interaction between them.

### References

- [1] N. C. Rattenborg, S. L. Lima, C. J. Amlaner. Half-awake to the risk of predation. *Nature*, 397, 397 - 398. 1999.
- [2] C. Deeprose, J. Andrade, S. Varma, N. Edwards. Unconscious learning during surgery with propofol anaesthesia. *British Journal of Anaesthesia*, 92, 171-177. 2004.
- [3] C. Portas, K. Krakow, P. Allen, O. Josephs, J. Armony, C. Frith. Auditory Processing across the Sleep-Wake Cycle Simultaneous EEG and fMRI Monitoring in Humans. *Neuron*, 28, 991 - 999. 2000.



- [4] B. Blesser, L.R. Salter. *Spaces Speak, Are you Listening? Experiencing Aural Architecture*. Cambridge, MA: The MIT Press. 2007.
- [5] URL: <http://www.thx.com/cinema/index.html>.
- [6] T. Kim, F. Biocca. Telepresence via Television: Two Dimensions of Telepresence May Have Different Connections to Memory and Persuasion. *Journal of Computer-Mediated Communication*, 3. 1997.
- [7] K. M. Lee. Presence, explicated. *Communication Theory*, 14, 27-50. 2004.
- [8] M. Lombard, T. Ditton. At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication*, 13, 573-590. 1997. URL: <http://www.jcmc.indiana.edu/vol13/issue2/lombard.html>
- [9] B. G. Witmer, M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7, 225-240. 1998.
- [10] J. Freeman. Implications for the measurement of presence from convergent evidence on the structure of presence. Paper presented to the *Information Systems Division at the annual meeting of the International Communication Association*, New Orleans, LA. May 2004.
- [11] M. Lombard. Presence and Telepresence Scholarship: Challenges Ahead. Paper presented to the *When Media Becomes Reality Workshop*, Berne, Switzerland. 2008.
- [12] J. D. Ivory, S. Kalyanaraman. The effects of technological advancement and violent content in video games on players' feelings of presence, involvement, physiological arousal, and aggression. *Journal of Communication*, 57, 532-555. 2007.
- [13] C. C. Bracken. Presence and image quality: The case of high definition television. *Media Psychology*, 7, 191-205. 2005.
- [14] C. C. Bracken. Perceived source credibility of local television news: The impact of image quality and presence. *Journal of Broadcasting & Electronic Media*, 50, 723-741. 2006.
- [15] C. C. Bracken, , G. Pettey, B. Rubenking, T. Guha. Sounding Out Small Screens and Telepresence: The Impact of Screen Size, Pace & Sound. Presented to the *Information Processing division of the International Communication Association annual convention in Montreal*, Quebec, Canada, May 2008.
- [16] P. Skalski, R. Tamburini. The role of social presence in interactive agent-based persuasion. *Media Psychology*, 10, 385-413. 2007.
- [17] M. Lombard, R. D. Reich, M. E. Grabe, C. C. Bracken, T. B. Ditton. Presence and television: The role of screen size. *Human Communication Research*, 26, 75-98. 2000.
- [18] M. E. Grabe, M. Lombard, R. D. Reich, C. C. Bracken, , T. B. Ditton. Screen size and viewer responses to television: Research findings. *Visual Communication Quarterly*, 6, 4-9. 1999.
- [19] M. Bocker, L. Muhlbach. Communicative presence in video communications. In: *Proceedings of the human factors and ergonomics society 37th annual meeting*, 249-253. October 2003.
- [20] C. Heeter. Being There: The subjective experience of presence. *Presence*, 1, 262-271. 1992.
- [21] B. Reeves, B. Detenber, J. Steuer. *New televisions: The effects of big pictures and big sound on viewer responses to the screen*. Paper presented to the *Information Systems Division of the International Communication Association*. 1993.
- [22] M. Lombard. Direct responses to people on the screen: Television and personal space. *Communication Research*, 22, 288-324. 1995.
- [23] P. Larsson, D. Vastfjall, P. Olsson, M. Kleiner. When What You Hear Is What You See: Presence and Auditory-Visual Integration in Virtual Environments. In: *Proceedings of the 10<sup>th</sup> Annual International Workshop on Presence*, 11-18. Barcelona. 2007.
- [24] D. Nunez. Effects of non-diegetic information on presence: A content manipulation experiment. In *Proceedings of the 10<sup>th</sup> Annual International Workshop on Presence*, 19-25. October 2007.
- [25] C. E. Izard. *Patterns of Emotions: A new analysis of anxiety and depression*, New York: Academic Press. 1972.
- [26] S. Le Groux, J. Manzolli, P. Verschure. Interactive Sonification of the Spatial Behavior of Human and Synthetic Characters in a Mixed-Reality Environment. In: *Proceedings of the 10<sup>th</sup> Annual International Meeting of the Presence Workshop*, 27-34. Barcelona. 2007.
- [27] A. Tajadura-Jimenez, A. Vajamae, J. Kitagawa, H. N. Ho. Whole-Body Vibration Influences Sound Localization in the Median Plane. In *Proceedings of the 10<sup>th</sup> Annual International Meeting of the Presence Workshop*, 35-39. Barcelona. 2007.
- [28] K. Kallinen, N. Ravaja. Comparing speakers versus headphones in listening to news from a computer – individual difference and psychophysiological responses. *Computers in Human Behavior*, 23, 303–317. 2007.
- [29] B.G. Witmer, M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7, 225-240. 1998.
- [30] B. Blesser, L. R. Salter. *Working Vocabulary for the Experience of Aural Architecture*. Cambridge, MA: MIT Press. 2007. URL: <http://www.blesser.net/downloads/Glossary.pdf>
- [31] R. Thayer. *The Biopsychology of Mood and Arousal*. New York: Oxford University Press. 1989.
- [32] F. Besmer. *Horses, musicians, and Gods: The Hausa cult of possession-trance*. South Hadley, MA: Bergen & Garvey. 1983.
- [33] G. Rouget. *Music and trance: A theory of the relations between music and possession*. Chicago: University of Chicago Press. 1985.
- [34] E. Hall, E. *The Hidden Dimension*. New York: Doubleday. 1986.
- \* Samsung's tagline for its S series home theater