



**UNIVERSITÀ  
DI PARMA**

**DIPARTIMENTO DI MEDICINA E CHIRURGIA**

**CORSO DI LAURE MAGISTRALE IN PSICOBIOLOGIA E  
NEUROSCIENZE COGNITIVE**

**L'ESPERIENZA DI "PRESENZA" MODULATA DAL  
SETTAGGIO AUDIO DURANTE L'ASCOLTO DI SEQUENZE  
SONORE CINEMATOGRAFICHE: UNO STUDIO  
COMPORTAMENTALE**

**Relatore:**

***Prof. VITTORIO GALLESE***

**Controrelatore:**

***Dott. NUNZIO LANGIULLI***

**Laureando:**

***VINCENZO FAZIO***

**ANNO ACCADEMICO 2020-2021**



# INDICE

<b>ABSTRACT</b> .....	5
<b>CAPITOLO 1 – INTRODUZIONE</b> .....	9
1.1 Il Film: Componenti Audio-Visive .....	9
1.2 Immersività e Presenza .....	13
1.3 Simulazione Incarnata: Coinvolgimento Fisico ed Emotivo .....	15
1.4 Obiettivi Sperimentali .....	19
<b>CAPITOLO 2 – MATERIALI E METODI</b> .....	22
2.1 Esperimento di Validazione degli Stimoli Sperimentali .....	22
2.1.1 Campione Sperimentale .....	22
2.1.2 Stimoli Sperimentali .....	22
2.1.3 Procedura .....	23
2.1.4 Analisi e Selezione degli Stimoli .....	24
2.2 Esperimento sulle Modalità di Riproduzione Acustica .....	27
2.2.1 Setup .....	27
2.2.2 Campione Sperimentale .....	30
2.2.2.a Questionari .....	31
2.2.2.b Test Audiometrico .....	32
2.2.2.c Test di Screening 1 – Loudness .....	33

2.2.2.d Test di Screening 2 –	
Localizzazione della Sorgente Sonora . . . . .	35
2.2.2.e Analisi del Campione Sperimentale . . . . .	36
2.2.3 Stimoli Sperimentali . . . . .	38
2.2.4 Paradigma Sperimentale . . . . .	39
<b>CAPITOLO 3 – ANALISI E RISULTATI . . . . .</b>	<b>46</b>
3.1 Analisi Statistiche . . . . .	46
<b>CAPITOLO 4 – DISCUSSIONE E CONCLUSIONI . . . . .</b>	<b>54</b>
4.1 Discussione . . . . .	54
4.2 Conclusioni e Futuri Sviluppi . . . . .	56
<b>BIBLIOGRAFIA . . . . .</b>	<b>59</b>

## ABSTRACT

Nella nostra esperienza quotidiana veniamo frequentemente in contatto con dispositivi di riproduzione audio che possono essere in grado di generare virtualmente spazi acustici più o meno fedeli a quelli reali: la scena sonora di un film può essere infatti percepita e “vissuta” differentemente in base a che la si ascolti al cinema o da uno *smartphone*.

L'utente che abbia accesso ad indizi di posizione degli oggetti riprodotti e che si senta localizzato nello spazio percepito può esperire la sensazione di presenza nell'ambiente virtualmente riprodotto in modo analogo a come si sentirebbe presente nel mondo reale, attivando quei meccanismi preriflessivi e automatici descritti nella teoria della *simulazione incarnata*.

In questo studio è stata investigata la sensazione di presenza modulata dalla modalità di presentazione di stimoli acustici estratti da sequenze cinematografiche. Gli stimoli sono stati suddivisi in tre categorie in base alla spazializzazione delle sorgenti sonore: riproduzione monofonica (una sorgente sonora, frontale centrale); riproduzione stereofonica (due sorgenti sonore, frontali laterali); riproduzione *surround* (sei sorgenti sonore, distribuite intorno al partecipante). L'esperienza di presenza è stata investigata indirettamente chiedendo ai partecipanti di giudicare gli stimoli in base al realismo e gradimento della riproduzione acustica e alla propria percezione di coinvolgimento fisico ed emotivo.

I risultati hanno chiaramente dimostrato che la riproduzione in modalità *surround*, rispetto alle altre, ottiene valutazioni migliori in termini di coinvolgimento fisico ed emotivo e percezione di realismo e piacevolezza, suggerendo che quei meccanismi di simulazione incarnata, che sappiamo essere attivati nel mondo reale, possano essere

elicitati anche in presenza di ambienti virtuali (*simulazione liberata*) e che possano essere modulati in base agli indizi di spazializzazione acustica presentati dalla riproduzione sonora.

# ABSTRACT

In our daily experience we frequently interact with audio devices that may be able to virtually develop an acoustic space similar to the real one: the sound scene of a movie can in fact be perceived and "lived" differently based on whether it is reproduced by a cinema theatre speaker system or by a smartphone speaker.

The user who can recognize the position clues of virtual objects and who feels localized in the perceived space can experience the feeling of presence in a virtual environment. This could happen in a similar way to how we would feel present in the real world: activating those pre-reflexive and automatic mechanisms described in the *embodied simulation* theory.

In this study we investigated the feeling of presence modulated by the playback mode of acoustic stimuli extracted from cinematic sequences. The stimuli were divided into three categories based on the spatialization of the sound sources: monophonic reproduction (one sound source, central front); stereo reproduction (two sound sources, frontal sides); surround reproduction (six sound sources, distributed around the participant). The experience of presence was investigated indirectly by asking participants to judge the stimuli based on the realism and liking of the acoustic playback and their perception of physical and emotional involvement.

The results clearly demonstrated that surround reproduction, compared to the others, reaches higher ratings in terms of physical and emotional involvement, perception of realism and liking. A possible explanation is that embodied simulation mechanisms, which are known to be activated in the real world, may be elicited even in virtual

environments and can be modulated by the acoustic spatialization cues played by the sound system.



# CAPITOLO 1

## INTRODUZIONE

*“Alcune azioni sono accompagnate da sonorità caratteristiche e di conseguenza”,  
certi suoni “ci fanno comprendere, quasi vedere, il comportamento di qualcuno che non  
vediamo”*

(Gallese & Guerra, 2015)

### **1.1 Il Film: Componenti Audio-Visive**

Nella nostra esperienza quotidiana è facilmente riconoscibile e individuabile la presenza di supporti tecnologici che permettono agli utilizzatori di visualizzare ed ascoltare persone, oggetti e luoghi attraverso una mediazione digitale: schermi visivi e sistemi di riproduzione sonora di ogni dimensione, dallo *smartphone* al cinema, sono ormai ovunque e sono strumenti indispensabili allo svolgimento di molte attività, compresa quella ricreativa. Gli individui hanno quindi imparato a interagire in modo naturale con rappresentazioni mediate digitalmente di cose e di persone, trasformando di fatto l'ambiente virtualmente proposto in una realtà effettivamente vissuta. Anche quando l'individuo interagisce con un ambiente ricreato appositamente al fine dell'intrattenimento, si instaurano una serie di relazioni tra il soggetto reale e gli oggetti virtuali che sono rappresentazioni di cose, luoghi e persone e che sono caratterizzati da

azioni, emozioni e sensazioni; queste relazioni sono ben riconoscibili durante l'esperienza percettiva del *medium* per eccellenza capace di portarci e di trasportarci in una dimensione alternativa della realtà: il film.

La visione di un film permette allo spettatore di essere testimone di una realtà alternativa ricreata innanzi ai suoi occhi che può essere più o meno convincente e coinvolgente fino a stimolare un'ulteriore ricostruzione mentale dell'ambiente circostante diverso da quello fisicamente presente. Anticipando la definizione dei costrutti di "immersività" e di "presenza" (che verranno definiti in maniera più dettagliata nel Paragrafo 1.2), ci si chiede quale possa essere la relazione tra la partecipazione ad esperienze immersive diversamente modulate in termini di riproduzione audio-visiva e la conseguente possibile esperienza di presenza del soggetto nell'ambiente virtuale proposto.

L'ambiente virtuale, dipendentemente dalle caratteristiche oggettive del sistema di riproduzione in termini di realismo ("immersività") e dalle caratteristiche soggettive dello spettatore in termini di coinvolgimento fisico ed emotivo ("presenza"), può infatti elicitarne risposte fisiologiche, psicologiche e comportamentali in qualche modo paragonabili alle stesse risposte elicitate da un ambiente reale: un film ci può far piangere o ci può far ridere, ci può sorprendere o spaventare a tal punto da farci scattare sulla sedia, possiamo talvolta arrivare a provare disagio fisico o eccitazione osservando le esperienze di qualcuno che non è presente con noi ma è riprodotto virtualmente per noi.

Affinché questa connessione tra reale e virtuale si verifichi è necessario che la riproduzione virtuale possa essere in qualche modo interpretata dallo spettatore come verosimile: le immagini e i suoni possono avere definizioni qualitative e spaziali differenti

comportando diversi livelli di fedeltà alla scena reale riprodotta. Lo sviluppo tecnologico è sicuramente un elemento fondamentale per la costruzione di un'esperienza partecipativa di un individuo reale in un ambiente non-reale e mediato.

Le origini esclusivamente visive del film, la cultura occidentale che attribuisce più importanza alla visione piuttosto che all'ascolto (Sterne, 2003) e soprattutto la posizione privilegiata del senso della vista rispetto all'udito nell'essere umano, potrebbero avere indirizzato gli sviluppi tecnologici e le ricerche verso l'analisi e il perfezionamento della sola componente visiva del film. Molte ricerche hanno investigato l'effetto della modulazione della presentazione visiva su diverse caratteristiche psicologiche, fisiologiche e comportamentali dello spettatore: come riportato in numerosi studi, uno schermo più grande comporterebbe un senso di presenza maggiore (Böcker & Mühlbach, 1993; Heeter, 1992; Reeves et al., 1993) oltre che un incremento della memoria (Kim & Biocca, 1997), dell'attenzione e dell'*arousal* (Reeves et al., 1999). Tuttavia, anche la modalità di distribuzione spaziale del suono gioca un ruolo fondamentale nella formazione dell'esperienza di presenza. Arnheim, riflettendo sulla dimensione spaziale, afferma che il suono ha la capacità di trasformare l'immagine bidimensionale in un'apertura su una realtà tridimensionale (Elsaesser & Hagener, 2009), affermazione affiancabile al pensiero di Di Donato il quale sottolinea come il suono (in formato 5.1) agisca su uno spazio a 360° e non a 180° come di solito avviene con le immagini (Di Donato, 2006).

Sebbene le ricerche sulla componente acustica abbiano suscitato minore interesse rispetto alla controparte visiva, alcuni studi hanno comunque investigato le risposte psicologiche e comportamentali di partecipanti ai quali venivano somministrati suoni in modalità di riproduzione differenti:

- Bracken e colleghi hanno rilevato come la riproduzione del suono tramite cuffie comporti un senso di immersione nell'ambiente maggiore rispetto alla riproduzione tramite altoparlanti televisivi (Bracken et al., 2010), in quanto le cuffie contribuirebbero a creare la sensazione che il suono provenga anche dalle direzioni laterali e posteriori oltre che dalla sola direzione frontale come invece avviene nella riproduzione televisiva;

- Lipscomb e colleghi hanno riscontrato assegnazioni di attributi verbali migliori a contenuti sonori presentati in modalità *surround* rispetto alla modalità stereofonica (Lipscomb & Kerins, 2004);

- Västfjäll ha mostrato che riproduzioni audio a sei canali ricevono stime di presenza significativamente più elevate rispetto alle riproduzioni stereo (due canali) e mono (un canale) (Västfjäll, 2003);

- Kobayashi e colleghi hanno trovato maggiori punteggi di presenza assegnati in condizioni di riproduzione sonora spazializzata rispetto alla condizione non-spazializzata, inoltre hanno riscontrato una maggiore attivazione del sistema nervoso simpatico in condizioni di suono spazializzato (Kobayashi et al., 2015);

- Lessiter e colleghi hanno evidenziato che l'aggiunta al sistema di riproduzione di un *subwoofer* per l'emissione di suoni a bassa frequenza influenza positivamente la sensazione di presenza del partecipante definita in termini di percezione dello spazio, realismo e immersione (Lessiter et al., 2001).

Considerando queste ricerche, sembrerebbe quindi che una maggiore caratterizzazione spaziale del suono riprodotto permetta all'ascoltatore di immergersi più efficacemente nello spazio sonoro virtuale. Come infatti suggerisce anche Di Donato, la direzionalità

del suono riprodotto in modalità *surround* può essere sfruttata per produrre diversi effetti: oltre ad aumentare lo spazio della scena osservata tramite la riproduzione di rumori ambientali, si può attivare “*una dimensione altra, che si aggiunge o addirittura oltrepassa quella del visibile*” (Di Donato, 2006).

## **1.2 Immersività e Presenza**

In un lavoro del 1997 Slater e Wilbur hanno, per la prima volta ripreso in esame criticamente i concetti di “immersività” e “presenza”, spesso confusi, disambiguandone il significato. I due autori hanno definito l’“immersività” come una proprietà oggettiva del sistema tecnologico di riproduzione (ad esempio di un sistema di realtà virtuale che è più immersivo di un sistema 3D) e la “presenza” come l’esperienza psicologica soggettiva di sentirsi situati in un ambiente mediato – di “essere lì” (Slater & Wilbur, 1997).

Secondo Slater e Wilbur un sistema immersivo è in grado di offrire simulazioni della realtà ad alta fedeltà attraverso molteplici canali sensoriali; ne deriva che l’esperienza di presenza correla positivamente con il livello di immersività del sistema (Cummings & Bailenson, 2016; Slater & Wilbur, 1997).

Dagli anni '40 del secolo scorso in poi, abbiamo assistito ad un intensivo sviluppo tecnologico nel settore dell'intrattenimento volto ad incrementare la sensazione di presenza, in termini di coinvolgimento fisico ed emotivo dell'utente, tramite il perfezionamento delle riproduzioni audio e video, sempre più realistiche. Con lo scopo di elevare il livello di immersività dei sistemi di riproduzione si è anche sperimentato il coinvolgimento di altri canali sensoriali quali olfatto, tatto e sistema vestibolare, come

avvenuto nel 1962 quando fu realizzato da M. Heilig il dispositivo di riproduzione *Sensorama*, in grado di permettere allo spettatore la visione stereoscopica di un film implementata con stimoli quali movimenti del sedile, vibrazioni, suoni stereofonici, correnti d'aria e aromi che venivano somministrati in modo appropriato durante la riproduzione nel cinematografo appositamente costruito (García-Pereira et al., 2020).

Seguendo questa linea di pensiero, i cinema, luoghi di intrattenimento dove il coinvolgimento dello spettatore è fondamentale per la riuscita dell'esperienza ricreativa, si sono dotati nel tempo di sistemi di riproduzione sonora sempre più sofisticati: inizialmente i primi riproduttori monofonici semplicemente duplicavano la stessa traccia acustica per incrementare l'intensità dei suoni e l'avvolgimento senza però ottenere direzionalità (Beck, 2003). In seguito, con l'avvento della registrazione stereofonica in più tracce il suono veniva distribuito su più canali tramite altoparlanti posizionati intorno allo spettatore. Quest'ultima fase ha visto molte sperimentazioni sia a livello del numero di canali utilizzati (7 nel *Cinerama*, 9 nel *Kinopanorama*, 4 nel *CinemaScope*, fino all'attuale *standard 5.1 del Dolby Stereo*) sia a livello del numero e del posizionamento spaziale degli altoparlanti (fino a 96 installati nel *Broadway Theatre* di New York), e ha portato alla realizzazione degli attuali sistemi *surround*, per cinema e per utilizzo domestico, che circondano lo spettatore ricreando uno spazio sonoro tridimensionale intorno ad esso (Sbravatti, 2019; Wright, 2005).

Recentemente sono emerse nuove tecnologie di riproduzione che non solo puntano al totale coinvolgimento dell'utente, ma offrono esperienze sensoriali alternative al mondo reale: la *virtual reality* sostituisce il mondo reale con uno digitale/virtuale, la *augmented reality* arricchisce il mondo reale con oggetti virtuali, la *mixed reality* connette il mondo

reale a mondi virtuali in una percepita continuità spazio-temporale (Maas & Hughes, 2020).

Tutti questi sviluppi tecnologici sono stati pensati espressamente per fornire al “percipiente” (termine introdotto da Sbravatti ad indicare la più completa esperienza percettiva rispetto al termine più utilizzato di “spettatore”), situato in un ambiente mediato, l’illusione di un ambiente non-mediato: la “presenza”, riscontrabile in termini di coinvolgimento fisico ed emotivo, eccitamento e senso di partecipazione (Lombard et al., 2000; Sbravatti, 2019).

Wirth e colleghi propongono un modello che definisce la formazione dell’esperienza di presenza con due premesse:

1. l’utente deve avere accesso a indizi di posizione che rendano l’ambiente mediato come uno spazio possibile nella realtà;
2. l’utente deve sentirsi localizzato nello spazio percepito.

Quando l’esperienza di presenza viene elicitata, l’utente si sente localizzato nello spazio riprodotto e percepisce l’illusione di un ambiente non-mediato, le sue capacità mentali e di azione si connettono quindi allo spazio mediato (Wirth et al., 2007).

### **1.3 Simulazione Incarnata: Coinvolgimento Fisico ed Emotivo**

Il costrutto di “presenza” si potrebbe inserire come parte integrante del meccanismo sul quale si basa la teoria della “simulazione incarnata” proposta da Gallese. Secondo questa teoria, alla base della capacità di comprendere il comportamento intenzionale altrui vi è infatti un meccanismo di base, automatico e preriflessivo, che sfrutta l’organizzazione

funzionale intrinseca del sistema motorio: la *simulazione incarnata* (Gallese, 2007). Un atto motorio osservato o ascoltato, ma anche solo immaginato, genererebbe quindi nel soggetto percipiente una riproduzione interna (una simulazione incarnata) di quello stesso atto motorio al fine di innescare nel soggetto la comprensione di esso senza necessariamente comportare l'effettiva riproduzione a livello di comportamento osservabile esternamente.

Nel modello neuro-cognitivo di Gallese e Guerra, il *sistema cervello-corpo*, l'azione, la percezione e la cognizione sono strettamente integrate: “*comprendiamo il senso di molti dei comportamenti e delle esperienze altrui grazie al riuso degli stessi circuiti neurali su cui si fondano le nostre esperienze agentive, emozionali e sensoriali in prima persona*” (Gallese & Guerra, 2015). La cognizione umana sarebbe quindi strettamente collegata alla corporeità situata degli individui che si realizza pienamente con l'esperienza percepita e che prende parte ad un processo di fondamentale importanza per l'essere umano: l'intersoggettività, definita come condivisione dell'evidenza naturale del mondo degli altri, ed esperita tramite la consonanza intenzionale, ovvero la sintonizzazione di intenti con altri individui (Gallese, 2010).

La simulazione incarnata come meccanismo di base del sistema cervello-corpo dei primati permetterebbe la relazione diretta con lo spazio, gli oggetti, le azioni, le emozioni e le sensazioni altrui tramite l'attivazione di rappresentazioni sensori-motorie nel cervello dell'osservatore; inoltre, sempre secondo Gallese, la simulazione incarnata avrebbe un ruolo fondamentale anche nella generazione delle immagini mentali, affermazione che trova riscontro positivo nei recenti studi sulle *vitality forms* (Di Cesare et al., 2015), ed è l'ipotesi fondamentale su cui si basano le terapie di riabilitazione tramite immaginazione motoria (Dodakian et al., 2014).



L'ambiente in cui il soggetto trova la propria collocazione e con cui può interagire fornisce gli stimoli (visivi, uditivi e tattili) percepiti dall'individuo (Rizzolatti et al., 1997); la visione di un oggetto o anche il solo suono generato da un'azione in una possibile localizzazione nello spazio peri-personale evoca la simulazione di azioni appropriate verso la localizzazione spaziale di provenienza dello stimolo (Gallese & Guerra, 2015). Graziano e Gross hanno infatti dimostrato che alcuni neuroni premotori, detti trimodali, oltre a rispondere a stimoli tattili e visivi, rispondono al buio anche a suoni diretti verso la ristretta porzione di spazio tridimensionale ancorata al corpo (Graziano et al., 1999).

Il meccanismo della simulazione incarnata fin qui presentato troverebbe il substrato neurale su cui svilupparsi in un particolare gruppo funzionale di neuroni denominati "neuroni specchio". Gli studi originali di Rizzolatti sui macachi definivano "neuroni specchio" quella classe di neuroni motori che si attivano sia nel momento in cui il soggetto esegue un'azione finalizzata, sia nel momento in cui il soggetto osserva l'esecuzione della stessa azione finalizzata da parte di altri individui (Rizzolatti & Destro, 2008). Ulteriori studi condotti su primati non-umani hanno dimostrato che anche solo l'ascolto di un suono, come quello prodotto dalla rottura di un guscio di una nocciolina americana, attiva questi neuroni che risultano avere quindi caratteristiche sia motorie che audio-visive (Kohler et al., 2002). Keysers e colleghi suggeriscono che questi neuroni codificano le azioni in modo astratto indipendentemente dalla sorgente dell'informazione (acustica o visiva), pertanto l'integrazione multimodale generata da stimoli di diversa natura può essere usata per fornire rappresentazioni significative e riconoscere azioni rilevanti nell'ambiente (Keysers et al., 2003).

Tramite l'utilizzo di tecniche quali la stimolazione magnetica transcranica (TMS), la risonanza magnetica funzionale (fMRI) e l'elettroencefalografia (EEG) è stata dimostrata anche nel cervello umano l'esistenza di circuiti analoghi a quelli utilizzati dal macaco per "rispecchiare" internamente le azioni osservate (Fadiga et al., 1995; Muthukumaraswamy & Johnson, 2004; Rizzolatti et al., 1996). Come dimostra uno studio di Buccino e colleghi, a differenza del macaco nell'uomo l'omuncolo motorio può essere attivato anche durante l'osservazione della pantomima di varie azioni, in assenza quindi dell'oggetto su cui sarebbe invece diretta l'azione transitiva (Buccino et al., 2001).

Ovviamente ogni individuo porta con sé una storia personale unica che lo contraddistingue e lo identifica; quindi, quando si utilizza il termine "specchio" si sta in realtà esprimendo una metafora, in quanto la percezione delle azioni, delle emozioni e delle sensazioni presenti nell'ambiente esterno non viene riprodotta in maniera oggettiva e immutabile (come farebbe un comune specchio) ma viene metabolizzata e restituita in base all'esperienza vissuta del percipiente.

Grazie a questi meccanismi di simulazione che condividono gli stessi circuiti neuronali sia in esecuzione che in osservazione comprendiamo quindi le esperienze degli altri individui che popolano il mondo che ci circonda. Nel caso dell'esperienza cinematografica, il soggetto si trova immerso in un mondo popolato dalle rappresentazioni di altri individui; l'altro osservabile non è più un individuo fisicamente presente ma è un oggetto con caratteristiche tali (la figura presentata sullo schermo e i suoni emessi dagli altoparlanti) da creare nel percipiente l'illusione di potersi relazionare con un individuo con cui condivide similarità, in un mondo illusoriamente tridimensionale nel quale è in grado di orientarsi, attivando di conseguenza le simulazioni precedentemente descritte. La differenza tra la percezione del mondo reale e la percezione

del mondo mediato non sarebbe quindi di tipo categoriale ma più prettamente di tipo dimensionale: la dimensione reale e la dimensione virtuale innescano comportamenti impliciti di imitazione, simulazioni, che nel primo caso possiamo definire come *simulazioni incarnate*, mentre nel secondo potremmo definirle più propriamente come *simulazioni liberate*, in quanto libere dall'intrusività del mondo reale: assistere a uno spettacolo di finzione ci libera dai coinvolgimenti personali diretti con la realtà quotidiana, siamo quindi liberi di esperire emozioni e sensazioni “a distanza di sicurezza” (Gallese & Guerra, 2013).

#### **1.4 Obiettivi Sperimentali**

Una completa comprensione dei meccanismi che regolano l'esperienza audio-visiva non è ancora stata raggiunta, quindi, al fine di descrivere in maniera sempre più approfondita la fenomenologia dell'esperienza di presenza, è importante chiedersi come le diverse modalità di presentazione di un ambiente mediato possano influenzare le percezioni e i conseguenti processi cognitivi e fisiologici del percipiente.

Lo studio comportamentale che viene qui presentato è inserito in un progetto più ampio, attualmente in corso, che ha l'obiettivo di investigare, tramite anche l'utilizzo di tecniche di *neuroimaging* funzionale come l'EEG, l'effetto della modulazione del setting audio-visivo, considerando le sue componenti singole e la combinazione delle due, sull'esperienza di “presenza” dello spettatore durante la visione ed ascolto di sequenze sonore cinematografiche.

Il presente studio si è concentrato sull'investigazione della diversa modalità di riproduzione della sola componente acustica del film: diverse sequenze filmiche sono

state scomposte nelle componenti audio e video e le tracce sonore estrapolate sono state somministrate in tre diverse condizioni di riproduzione che corrispondono al livello crescente di immersività del sistema: riproduzione monofonica, riproduzione stereofonica, riproduzione *surround*.

Per non denaturare l'esperienza di fruizione multimediale sono state utilizzate sequenze filmiche opportunamente estratte e validate. Come riportato da Sonkusare e colleghi, prove emergenti e considerazioni teoriche suggeriscono che il cervello potrebbe essere "sintonizzato" più su stimoli naturalistici che su stimoli artificiali; definendo euristicamente "naturalistici" o "ecologici" quegli stimoli presenti nella nostra esperienza vissuta quotidianamente, possiamo includere in questa categoria anche i film (Sonkusare et al., 2019). Nella meta-analisi condotta da Fernández-Aguilar e colleghi viene riportato che la tecnica di laboratorio più semplice ed efficace per l'induzione di stati emotivi nei partecipanti prevede l'utilizzo di materiale audio-visivo estratto da film (Fernández-Aguilar et al., 2019).

Le tracce sonore nelle tre condizioni sperimentali sono state valutate in termini di coinvolgimento fisico ed emotivo, livello di realismo percepito e piacevolezza, in quanto, come fatto notare da Lombard e colleghi, la modalità di misurazione del costrutto di presenza nei diversi studi risulta essere a volte inappropriata (Lombard et al., 2000): la domanda diretta "**Quanto ti sei sentito presente?**" potrebbe non essere interpretata in modo chiaro dal partecipante. Sulla base di precedenti studi che hanno ricercato il costrutto di presenza, si è quindi scelto di porre domande meno dirette volte a misurare le componenti che formano l'esperienza di presenza: l'immersività, in termini di realismo dello spazio percepito, il coinvolgimento fisico e il coinvolgimento emotivo del percipiente (Lessiter et al., 2001; Lombard et al., 2000).

Ci si è chiesti se una maggiore caratterizzazione dello spazio sonoro fornita da un sistema di riproduzione *surround* potesse attivare quei meccanismi di simulazione liberata precedentemente visti che in questo studio comportamentale verrebbero stimati in relazione alla sensazione di presenza percepita dal partecipante. Come descritto in precedenza, in un partecipante che abbia accesso a indizi di posizione degli oggetti nello spazio può essere evocata l'esperienza di presenza tramite il sentirsi localizzato nella scena ambientale riprodotta; ipotizziamo quindi un maggiore coinvolgimento sia fisico che emotivo durante la somministrazione degli stimoli nella condizione di riproduzione *surround*, la quale fornisce un maggior numero di indizi di posizione degli oggetti all'interno dello spazio acustico virtualmente generato.

Ai partecipanti è stato inoltre chiesto di stimare il livello di realismo e gradimento (piacevolezza) delle scene riprodotte nelle tre diverse condizioni. Ci aspettiamo che il realismo indicato sia maggiore nella condizione di riproduzione *surround*, in quanto la dinamicità degli oggetti localizzati nello spazio risulta essere più fedele ai movimenti degli stessi all'interno di uno spazio sonoro reale. Inoltre, pur mancando in letteratura riferimenti specifici all'esperienza estetica nelle diverse condizioni di spazializzazione acustica, ipotizziamo punteggi di piacevolezza maggiori nella condizione *surround* in quanto più un suono viene riprodotto fedelmente rispetto al suono reale e più potrebbe essere apprezzato per le sue qualità sonore.

## CAPITOLO 2

### MATERIALI E METODI

#### 2.1 Esperimento di Validazione degli Stimoli Sperimentali

##### 2.1.1 Campione Sperimentale

Il campione sperimentale era composto da 100 partecipanti (67 maschi e 33 femmine, con età media  $M$  di 23.6 e deviazione standard  $DS$  di  $\pm 4.9$  anni, compresa in un *range* di 18 - 38 anni).

Il reclutamento è avvenuto *online* tramite la piattaforma Prolific (Palan & Schitter, 2018). La selezione si è basata su tre caratteristiche: i partecipanti dovevano avere un'età compresa tra i 18 e i 40 anni, dovevano parlare fluentemente l'inglese, non dovevano avere una storia pregressa di disturbi psichiatrici. Tutti i partecipanti sono stati ricompensati tramite la piattaforma Prolific con una tariffa oraria di circa 7.5 £/ora.

##### 2.1.2 Stimoli Sperimentali

Per la validazione degli stimoli sono state utilizzate 185 sequenze filmiche della durata di 10 secondi estratte da uno sperimentatore dopo la visione di 31 film candidati all'Oscar

per le categorie “miglior sonoro” e “miglior montaggio sonoro” nelle edizioni comprese tra il 1979 e il 2019. I criteri di individuazione delle scene filmiche erano i seguenti:

- nella scena non doveva essere presente musica;
- nella scena non dovevano essere presenti dialoghi;
- la scena doveva essere caratterizzata con un buon impatto sonoro e buona spazializzazione acustica.

L'estrazione degli stimoli audio-visivi in formato mp4 è stata effettuata partendo da materiale originale in formato PCM 5.1. Nessun'altra modifica è stata effettuata.

### **2.1.3 Procedura**

Il paradigma sperimentale per la validazione degli stimoli è stato scritto in PsychoPy (v3.0) e ospitato dal servizio Pavlovia per l'esecuzione *online* del programma e la gestione dei dati raccolti (Peirce, 2007).

Il campione sperimentale è stato suddiviso in modo casuale in 10 gruppi da 10 partecipanti ciascuno. In maniera casuale sono state assegnati 19 differenti stimoli a nove gruppi e 14 differenti stimoli a un gruppo, per un totale di 185 stimoli. Il primo gruppo, al quale sono stati assegnati solo 14 stimoli, aveva la funzione di gruppo pilota per testare che il paradigma sperimentale procedesse senza intoppi.

I partecipanti dovevano visionare, esclusivamente sullo schermo di un PC e tramite l'utilizzo di cuffie o *speakers* esterni, uno stimolo, presentato in maniera casuale, e successivamente dovevano rispondere a una delle seguenti tre domande che apparivano sullo schermo in modo casuale:

1. “How would you judge the **Dynamicity** of the scene?”  
(Come giudicheresti la dinamicità della scena?)
2. “How would you judge the **Emotional Valence** of the scene?”  
(Come giudicheresti la valenza emotiva della scena?)
3. “How would you judge the **Emotional Intensity** of the scene?”  
(Come giudicheresti l'intensità emotiva della scena?)

Ogni stimolo era ripetuto tre volte durante la sessione sperimentale e dopo ogni presentazione era seguito da una domanda differente. In totale quindi il partecipante si sottoponeva a 57 prove (42 solo nel caso del primo gruppo) per una durata totale della sessione sperimentale di circa 15 minuti.

Le domande richiedevano al partecipante di esprimere un giudizio personale sulla Dinamicità, Valenza Emotiva e Intensità Emotiva degli stimoli audio-visivi tramite una scala VAS (Visual Analogue Scale) che andava da 0 a 100 per le domande sulla Dinamicità ed Intensità Emotiva, e da -50 a +50 per la domanda sulla Valenza Emotiva. Le domande erano volte all'individuazione di stimoli con contenuti emotivi, misurati secondo l'approccio dimensionale delle emozioni di Russell (Russell, 1980), ad alta attivazione e valenza negativa.

Oltre ai punteggi assegnati sono stati registrati i tempi di risposta.

#### **2.1.4 Analisi e Selezione degli Stimoli**

Essendo questo uno studio effettuato *online* senza la supervisione degli sperimentatori, le condizioni di esecuzione non potevano essere totalmente controllate portando ad una



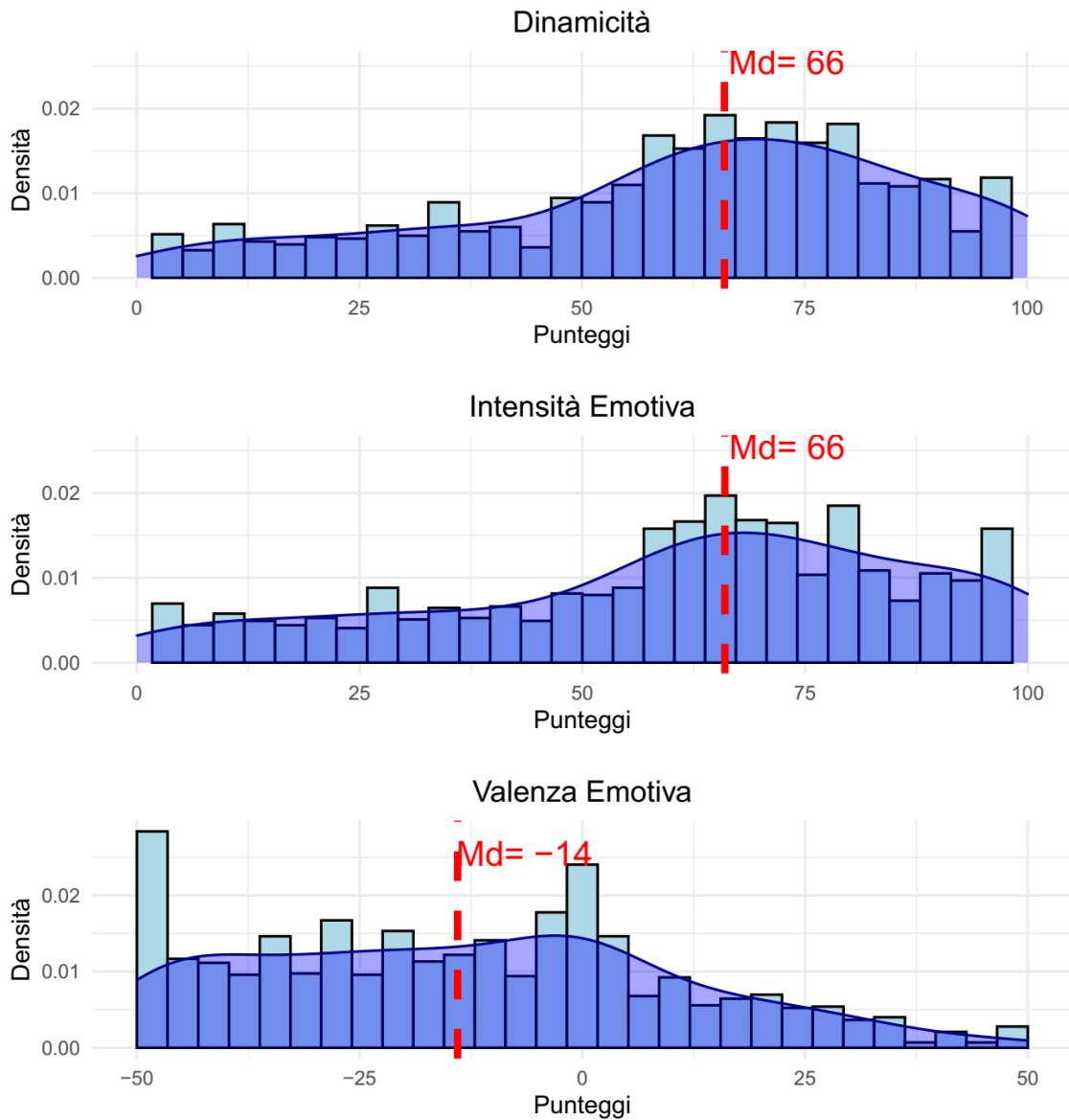
qualità dei dati inferiore rispetto a quella di dati raccolti in laboratorio (Sauter et al., 2020). Per evitare quindi di introdurre nell'analisi dati registrati in seguito ad errori occasionali (dovuti all'uso accidentale improprio del mouse o a distrazioni del partecipante), sono stati esclusi i punteggi assegnati con tempo di risposta inferiore o uguale a 500 millisecondi o superiore a 10 secondi (303 *data point* esclusi per tempi di risposta superiori ai 10 secondi; nessun *data points* al di sotto di 500 millisecondi).

È stata inoltre effettuata la *outlier detection* (rilevamento di valori anomali rispetto agli altri) sui punteggi assegnati ad ogni stimolo per ogni domanda e sui tempi di risposta, escludendo ulteriori 183 *data points* per i tempi di risposta (nessuno per i punteggi).

Partendo dal presupposto che vi sia correlazione tra immersività e intensità di emozioni percepite come rilevato da Vish (Visch et al., 2010), che le emozioni negative abbiano priorità di processazione in quanto innescano meccanismi di difesa tramite l'attivazione precoce dell'amigdala (Figueiredo et al., 2003; Leppänen & Nelson, 2009) e che stimoli audio-visivi a valenza negativa estratti da film inducono maggiore *arousal* rispetto a quelli a valenza positiva (Fernández-Aguilar et al., 2019), tra gli stimoli proposti ai partecipanti sono stati ricercati quelli con alta Dinamicità, alta Intensità Emotiva e con Valenza Emotiva negativa. Sono quindi stati selezionati quegli stimoli che avevano ricevuto punteggi tali da rispettare le seguenti condizioni:

- I punteggi medi di Dinamicità e Intensità Emotiva dovevano essere superiori alle mediane delle corrispondenti categorie (*Mdn* Dinamicità = 66, *Mdn* Intensità Emotiva = 66);
- il punteggio medio della Valenza Emotiva doveva essere inferiore alla mediana della categoria corrispondente (*Mdn* Valenza Emotiva = -14).

L'analisi ha portato alla selezione di 50 stimoli che presentavano le caratteristiche desiderate; in Figura 2.1 sono riportate graficamente le densità dei punteggi assegnati alle tre domande e le mediane.



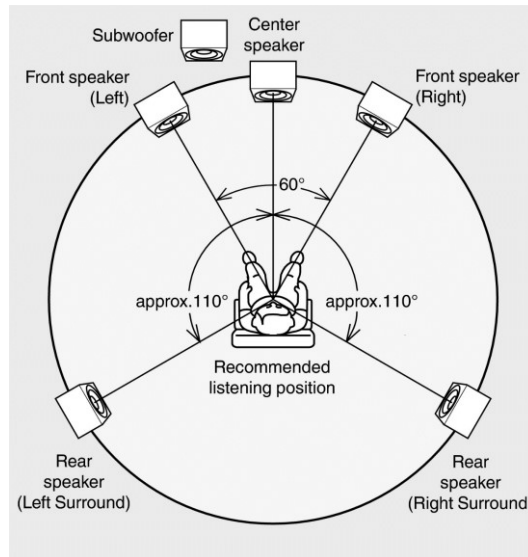
**Figura 2.1** *Densità dei Punteggi e Mediane*

*Densità dei punteggi assegnati alle domande e Mediane (linea rossa tratteggiata)*

## 2.2 Esperimento sulle Modalità di Riproduzione Acustica

### 2.2.1 Setup

Una cabina audiometrica silente (IAC-Acoustics) alta 2 metri, larga 2,5 metri e profonda 2,1 metri è stata allestita con un sistema *surround* di riproduzione audio composto da cinque altoparlanti APART MASK4C (impedenza 8 Ohm) e un *sub-woofer* APART SUBA165 (impedenza 4 Ohm), un televisore (LG UHD, 42 inch, 16/9) e una sedia con tavolino mobile su cui poter utilizzare un mouse collegato ad un computer esterno alla cabina. Il partecipante sedeva al centro della cabina audiometrica silente, mentre i cinque altoparlanti sono stati posizionati e orientati come mostrato in Figura 2.2 seguendo la raccomandazione ITU-R BS.1116-1, in modo da direzionare il suono in un punto centrale della cabina che identificava la corretta posizione d'ascolto (ITU-R, 1997). Nello specifico, su un carrello TV da pavimento, disposto frontalmente rispetto al partecipante, è stata realizzata un'intelaiatura alla quale è stata fissata su supporto VESA, a 95 cm di altezza, la televisione e tre altoparlanti così disposti: due ai lati della televisione, (canale sinistro denominato "L", *Left* e destro denominato "R", *Right*) e uno al di sopra di essa (canale centrale denominato "C", *Central*). Agli angoli posteriori, rispetto alla posizione del partecipante, sono stati posizionati due supporti da pavimento (uno per angolo) sui quali sono stati installati gli ultimi due altoparlanti (canale sinistro posteriore denominato "Ls", *Left Surround* e destro posteriore denominato "Rs", *Right Surround*); il *sub-woofer* (canale denominato "LFE", *Low Frequency Effects*) è stato posizionato a terra, all'incirca sotto il canale "L".



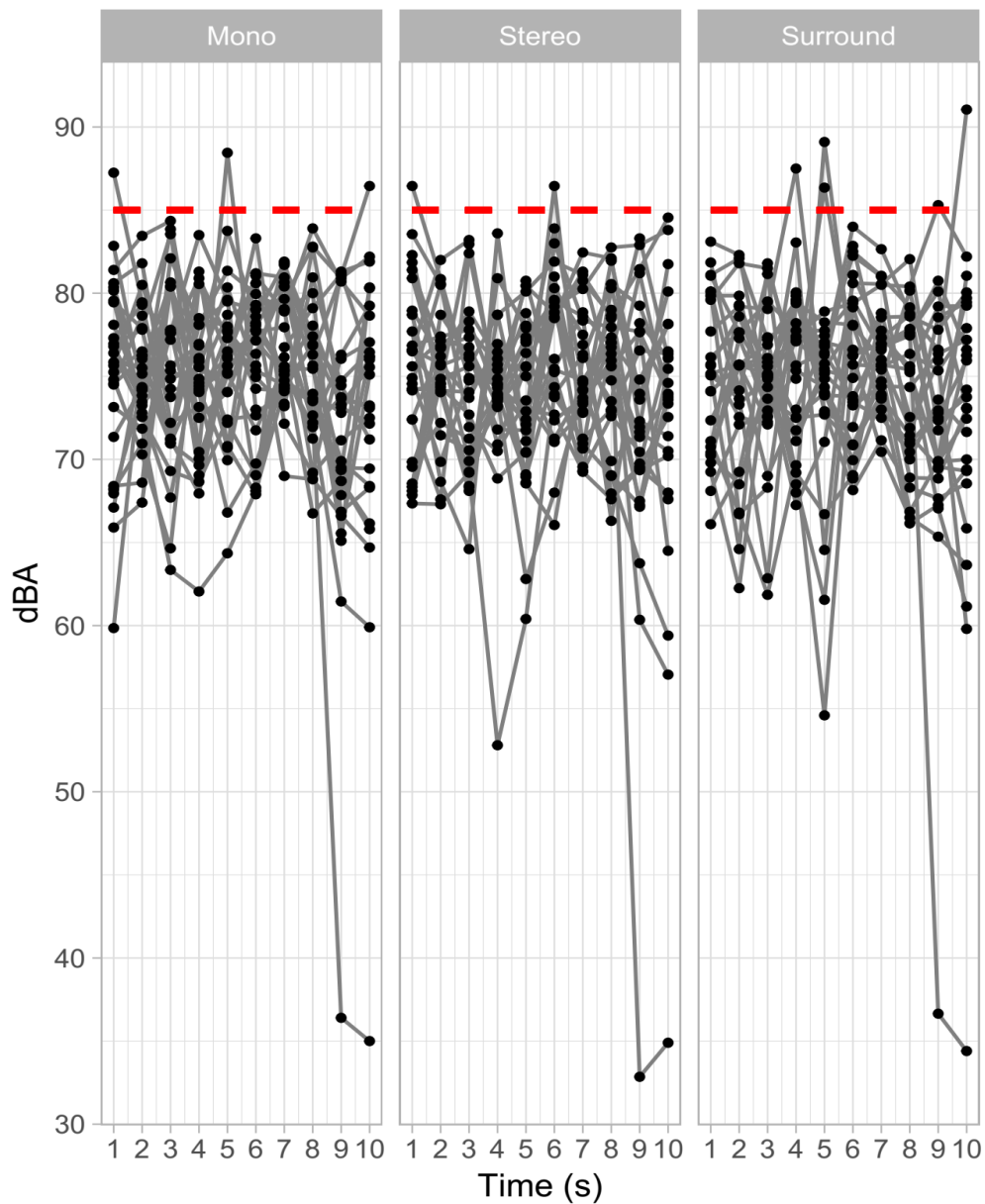
**Figura 2.2** *Disposizione Sistema Surround – Raccomandazione ITU-R BS.1116-1*

*Immagine tratta da <https://manuals.marantz.com>*

Il sistema era alimentato da un amplificatore DENON AVR-X1600H, posizionato all'esterno della cabina audiometrica silente, che amplificava il segnale audio proveniente dal computer dal quale venivano somministrati gli stimoli. In base al posizionamento effettuato e al riverbero acustico, misurato nella cabina utilizzando il microfono fornito nella dotazione dell'amplificatore e posizionato in corrispondenza del punto d'ascolto, il sistema è stato calibrato per correggere la risposta in frequenza dei diversi canali audio effettuando una *room correction* tramite l'utilizzo del software Audyssey (Paul, 2009).

Il livello di *loudness* di tutti gli stimoli (che saranno presentati nei Paragrafi 2.2.2.c/d e 2.2.3) è stato registrato con un fonometro (Gain Express, standard applicato IEC651 tipo 2, tipo ANSI 2 SI 0.4) posizionato in corrispondenza del punto d'ascolto; l'analisi delle registrazioni, riportata graficamente in Figura 2.3, ha permesso di individuare il livello di amplificazione da utilizzare durante l'esperimento, e ha verificato che i livelli di riproduzione acustica fossero al di sotto della soglia ritenuta pericolosa per l'udito (85

dB per 8 ore consecutive) definita e standardizzata dal National Institute for Occupational Health and Safety (NIOSH) nella raccomandazione ONE (Occupational Noise Exposure) (Murphy & Franks, 2002).



**Figura 2.3** *Sound Pressure Level*

*Sound Pressure Level (dBA) degli Stimoli nel Tempo e soglia di pericolosità (linea rossa tratteggiata)*

## 2.2.2 Campione Sperimentale

Al fine di ottenere dati affidabili è stato opportuno selezionare i partecipanti in modo da rispettare alcuni criteri fondamentali:

- i partecipanti non dovevano presentare problemi di udito;
- i partecipanti dovevano essere in grado di percepire differenze tra diversi stimoli acustici e tra diverse modalità di riproduzione sonora;
- i partecipanti dovevano essere in grado di valutare una serie di stimoli acustici riprodotti in successione;
- i partecipanti dovevano essere “*un-trained/naives subjects*” come descritto nella raccomandazione ITU-T P.800, non dovevano quindi avere abilità specifiche nella valutazione di sistemi di riproduzione acustica; questo criterio, se rispettato, permette la generalizzazione alla popolazione dei risultati ottenuti che non sarebbe possibile se il campione di partecipanti fosse composto da “*experienced/expert subjects*” descritti nella raccomandazione ITU-T P.831, P832 (Rec, 1996, 2000; TSG, 2004).

Per controllare la presenza di tutti i requisiti elencati, i partecipanti sono stati sottoposti ad un adattamento della Generalised Listener Selection (GLS) descritta da Zacharov e colleghi (Bech & Zacharov, 2007; Mattila & Zacharov, 2001) che comprende la somministrazione di:

- 6 questionari,
- un test audiometrico,
- due test di *screening* per valutare le capacità discriminative di *loudness* e localizzazione della sorgente sonora presentati in sequenza casuale.

Le somministrazioni del test audiometrico e dei test di *screening* sono state eseguite tramite il *software* MATLAB (Matlab, 2012).

### 2.2.2.a Questionari

Ogni partecipante doveva compilare, in maniera telematica tramite Moduli di Google (Google Inc.), una batteria di questionari, per ottenere informazioni su diversi aspetti dell'esperienza acustica individuale, che potesse accertarne la corretta inclusione nel campione, così composta:

1. GLS-1 (Generalised Listener Selection): questionario composto da 13 *items* che indagano aspetti generali di interesse: il partecipante non doveva essere esperto o avere familiarità nell'utilizzo di sistemi *surround*, non doveva occuparsi professionalmente di audio e/o acustica, non doveva aver partecipato precedentemente a test formali di ascolto (Bech & Zacharov, 2007).

2. QEAV (Questionario sull'Esperienza Audio-Visiva): questionario composto da 11 *items* che misura il grado di conoscenze in ambito cinematografico e di montaggio audio-visivo.

3. B-MEQ (Brief Music Experience Questionnaire): questionario composto da 53 *items* che misura il grado di conoscenze in ambito musicale sia come ascoltatore sia come musicista, suddiviso nelle sottoscale *Commitment to music*, *Innovative musical aptitude*, *Social uplift*, *Affective reactions*, *Positive psychotropic effects*, *Reactive musical behavior* (Werner et al., 2006).

4. IRI (adattamento italiano all'Interpersonal Reactivity Index): questionario composto da 28 *items* che quantifica la responsività empatica, definita come capacità individuale di vicariare gli stati emotivi di altri individui durante

l'osservazione di esperienze altrui, attraverso la misura integrata di componenti affettive e cognitive. Il questionario si articola in quattro sottoscale: *Fantasia*, indaga la propensione a identificarsi con personaggi fittizi della letteratura, del cinema o del teatro; *Perspective Taking*, sonda la capacità di adottare il punto di vista altrui; *Considerazione Empatica*, valuta la tendenza dei soggetti a provare compassione, preoccupazione e calore nei confronti di altre persone che vivono esperienze spiacevoli; *Disagio Personale*, fa riferimento ai casi in cui l'essere spettatore di esperienze spiacevoli che accadono ad altri provocano un senso di sconforto e di ansietà nei soggetti stessi (Albiero et al., 2006).

5. ITQ (Questionario sulle Tendenze Immersive): questionario composto da 34 *items* che misura la capacità o tendenza dei partecipanti ad immergersi in diverse situazioni ambientali (Witmer & Singer, 1998).

6. VMIQ-2 (Vividness of Movement Imagery Questionnaire): questionario composto da 36 *items* sulla capacità di immaginazione motoria del proprio sé corporeo. Il questionario si articola in tre sottoscale: *External self Visual Imagery* (EVI), misura l'immaginazione visiva in terza persona; *Internal first-person Visual Imagery* (IVI), misura l'immaginazione visiva in prima persona; *Kinesthetic Imagery* (KIN), misura l'immaginazione delle sensazioni cinestesiche legate al movimento immaginato (Roberts et al., 2008).

### **2.2.2.b Test Audiometrico**

Ogni partecipante è stato sottoposto, all'interno di una cabina audiometrica silente, ad un breve test audiometrico tonale per misurarne la soglia uditiva attraverso l'applicazione "Frequency Response of the Ear , Hearing Test" sviluppata in MATLAB (Rawashdeh, 2021). Il partecipante indossava cuffie *over ear* (Audio-technica ATH-MSR7) ed il



programma somministrava contemporaneamente ad entrambe le orecchie un tono puro, partendo da una frequenza di 1 KHz fino ad arrivare a una frequenza di 16 KHz con intervalli di 1 KHz, a livelli di *loudness* decrescenti. Il compito del partecipante era quello di indicare se aveva sentito il tono somministrato utilizzando il mouse per premere un pulsante di conferma. Quando il partecipante non era più in grado di percepire il tono puro (e quindi non premeva il pulsante) il programma somministrava il tono puro con la frequenza successiva. Le curve audiometriche risultanti sono state successivamente confrontate con le curve audiometriche per persone otologicamente normali differenziate per età secondo lo standard ISO - International Organization for Standardization (7029, 1984) e i partecipanti il cui udito non sembrava essere nella norma sono stati esclusi dal campione sperimentale.

### **2.2.2.c Test di Screening 1 - Loudness**

Ogni partecipante è stato sottoposto ad un compito a scelta forzata per valutare la capacità di discriminare suoni riprodotti con livelli di *loudness* differenti. Il partecipante sedeva all'interno della cabina audiometrica silente, posizionato centralmente rispetto alla disposizione dei cinque altoparlanti del sistema *surround*, come descritto nel Paragrafo 2.2.1.

Per la realizzazione di questo test di *screening* sono stati creati cinque differenti stimoli acustici della durata di 500 millisecondi che venivano riprodotti contemporaneamente su cinque canali del sistema *surround* (il canale LFE era escluso). Gli stimoli consistevano in un rumore bianco campionato a 48000 Hz con livelli di *loudness* differenti indicati in Loudness Units relative to Full Scale (LUFS), unità di misura descritta nello standard EBU (European Broadcasting Union) R 128 per il quale 1 LU (Loudness Unit)

corrisponde alla misurazione relativa di 1 dB su scala digitale e  $0 \text{ LU} = -23 \text{ LUFS}$  (EBU–Recommendation, 2011; ITU-R, 2007). Il livello più alto di *loudness* è stato fissato a -23 LUFS, in base allo standard EBU R 128, mentre il livello di *loudness* degli altri stimoli è stato fissato a -3 LUFS dal precedente in base alla minima differenza udibile descritta da Larsen e colleghi (EBU–Recommendation, 2011; Larsen et al., 2008). Gli stimoli presentavano quindi i seguenti livelli di *loudness*:

1. -23 LUFS,
2. -26 LUFS,
3. -29 LUFS,
4. -32 LUFS,
5. -35 LUFS.

Il partecipante doveva ascoltare ad occhi aperti una coppia di stimoli acustici riprodotti in successione e contemporaneamente dai cinque altoparlanti e doveva ogni volta indicare, in base alla propria percezione ed entro un tempo massimo di 5 secondi, quale stimolo era stato riprodotto col livello di *loudness* maggiore o se entrambi gli stimoli erano stati riprodotti con lo stesso livello di *loudness*.

Gli stimoli venivano presentati in sequenza casuale nei diversi possibili accoppiamenti (compresi gli accoppiamenti tra stimoli uguali) per un totale di 25 prove. Ogni coppia di stimoli veniva accompagnata da un *cue* visivo: il primo stimolo della coppia era accompagnato dalla scritta sullo schermo “Primo”, mentre il secondo stimolo riprodotto era accompagnato dalla scritta “Secondo”. Alla fine di ogni riproduzione appariva sullo schermo la domanda “**Quale suono aveva il volume più alto?**” e il partecipante doveva rispondere con il mouse scegliendo tra le opzioni “Primo”, “Uguali” e “Secondo”.

Essendo gli accoppiamenti tra stimoli con livelli di *loudness* differenti tra loro per 6 o più LUFS facilmente discriminabili in quanto la differenza è ampiamente al di sopra della soglia di discriminazione tra due stimoli uditivi, ed essendo gli accoppiamenti tra stimoli con livelli di *loudness* differenti tra loro per 3 LUFS difficilmente distinguibili dagli accoppiamenti tra stimoli con livelli di *loudness* uguali tra loro, il criterio di accettabilità per la prestazione del partecipante è stato stabilito al minimo di 80% di risposte corrette (20 su 25) come indicato nella descrizione dei *matching* test sulle abilità uditive eseguiti da Bech e Zacharov (Bech & Zacharov, 2007; Larsen et al., 2008).

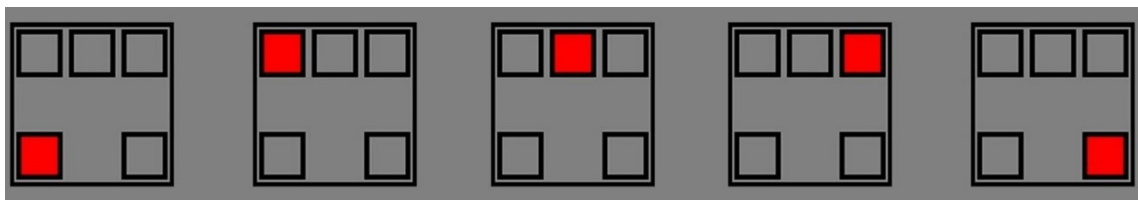
#### **2.2.2.d Test di Screening 2 – Localizzazione della Sorgente Sonora**

Ogni partecipante è stato anche sottoposto ad un compito a scelta forzata per valutare la capacità di riconoscere la direzione della sorgente sonora (localizzazione del suono).

Per la realizzazione di questo test di screening sono stati prodotti cinque stimoli acustici differenti aventi tutti una durata di 500 millisecondi e campionati a 48000 Hz. Ogni stimolo era costituito da rumore bianco di sottofondo proveniente da quattro canali e da un tono puro a 1000 Hz riprodotto da un solo canale (anche in questo caso il canale LFE era escluso).

La costruzione degli stimoli acustici è stata eseguita seguendo le indicazioni riportate da Lopatka e colleghi, i quali hanno registrato risposte corrette sulla localizzazione della sorgente sonora al 79.8% per stimoli che presentavano una differenza di 25 LUFS tra i livelli di *loudness* del segnale e del rumore (Lopatka et al., 2016). Conseguentemente, il livello di *loudness* del tono puro è stato fissato a -23 LUFS in base allo standard EBU R 128 (EBU–Recommendation, 2011), mentre il livello di *loudness* del rumore bianco di sottofondo è stato quindi impostato a -48 LUFS.

Il compito del partecipante era quello di ascoltare lo stimolo ed indicare la direzione di provenienza del tono puro riprodotto casualmente da uno dei cinque altoparlanti. Alla fine di ogni riproduzione appariva sullo schermo la domanda “**Da quale altoparlante proveniva il beep?**” e il partecipante doveva rispondere con il mouse, entro un tempo massimo di 5 secondi, scegliendo una tra le mappe di localizzazione proposte a video come mostrato in Figura 2.4. Ogni stimolo veniva ripetuto quattro volte, per un totale di 20 prove.



**Figura 2.4** *Mappe di Localizzazione*

*Le mappe rappresentavano la possibilità di risposta per i canali (da sinistra verso destra) “Ls”, “L”, “C”, “R”, “Rs”*

Il criterio di accettabilità per la prestazione del partecipante è stato stabilito al minimo di 80% di risposte corrette (16 su 20) per quanto riguarda l’individuazione della direzione della sorgente sonora, basandosi sui risultati degli esperimenti sulla localizzazione della sorgente sonora riportati da Lopatka e colleghi (Lopatka et al., 2016).

### **2.2.2.e Analisi del Campione Sperimentale**

La somministrazione dei questionari e dei test audiometrico e di *screening* (visti nei Paragrafi 2.2.2.a/b/c) hanno portato alla selezione di 17 partecipanti su 20 esaminati (9 maschi e 8 femmine, con età media  $M$  di 31.8 e deviazione standard  $DS$  di  $\pm 7.0$  anni,

compresa in un *range* di 24 - 42 anni). Il campione selezionato era bilanciato per quanto riguarda il genere e presentava un alto livello di scolarizzazione (*M* di 15.5 e *DS* di  $\pm 2.3$  anni); i punteggi minimi e massimi riscontrati nelle diverse scale e sottoscale dei questionari somministrati sono riportati in Tabella 2.1 insieme alle medie e alle deviazioni standard delle risposte medie dei partecipanti ad ogni scala e sottoscala dei questionari.

<i>Questionario</i>	<i>Min.</i>	<i>Max.</i>	<i>M</i>	<i>SD</i>
BMEQ C	13	28	19.3	4.8
BMEQ I	7	33	20.0	7.1
BMEQ S	8	18	12.2	2.8
BMEQ A	36	50	43.3	3.5
BMEQ P	34	71	58.5	11.1
BMEQ R	19	45	35.2	6.3
IRI CE	3	13	7.4	3.6
IRI DP	11	26	17.7	4.8
IRI PT	0	25	8.6	5.6
IRI FS	0	18	8.1	4.5
QTI	46	94	73.8	13.4
VMIQ-2 EVI	12	48	26.8	10.3
VMIQ-2 IVI	12	55	27.8	13.2
VMIQ-2 KIN	12	55	26.8	13.0

**Tabella 2.1** *Statistiche Descrittive per ogni scala e sottoscala dei Questionari*

*BMEQ = Brief Music Experience Questionnaire (C = Commitment to music, I = Innovative musical aptitude, S = Social uplift, A = Affective reactions, P = Positive psychotropic effects, R = Reactive musical behavior)*

*IRI = Interpersonal Reactivity Index (CE = Considerazione Empatica, DP = Disagio Personale, PT = Perspective Taking, FS = Fantasia)*

*QTI = Questionario sulle Tendenze Immersive*

*VMIQ-2 = Vividness of Movement Imagery Questionnaire-2 (EVI = External self Visual Imagery, IVI = Internal first-person Visual Imagery, KIN = Kinesthetic Imagery)*

### 2.2.3 Stimoli Sperimentali

Gli stimoli sperimentali erano costituiti dalle tracce sonore estratte dai 50 stimoli audiovisivi selezionati tramite la validazione *online* (si veda il Paragrafo 2.1.4). Le tracce audio sono state estratte dagli stimoli dapprima come tracce a sei canali (Surround) in formato wav (canali = 5.1; *sample rate* = 48000 Hz; *sample size* = 16 bit) e successivamente come tracce Stereofoniche a due canali (da ora in poi Stereo) in formato wav (canali = 2; *sample rate* = 48000 Hz; *sample size* = 16 bit). Per ottenere la terza condizione Monofonica (da ora in poi Mono) le tracce dei due canali Stereo sono state mediate in un unico canale in formato wav. Nella condizione Mono la traccia è stata registrata sul canale C mantenendo gli altri cinque canali vuoti; nella condizione Stereo le tracce sono state registrate separatamente sui canali L e R mantenendo gli altri quattro canali vuoti.

Per ogni traccia audio sono stati calcolati il *range* di *loudness* (indice di dinamicità della sequenza audio), il valore di *true-peak* in dB (il livello massimo in valore assoluto della forma d'onda sonora digitale) e i valori di *momentary*, *short-term* e *integrated loudness* in LUFS. I valori calcolati di *loudness* si differenziano per l'ampiezza della finestra temporale analizzata all'interno della traccia audio: 400 millisecondi per il valore di *momentary loudness*, 3 secondi per il valore di *short-term loudness* e l'intera durata della traccia per il valore di *integrated loudness*. Il valore di *integrated loudness* delle tracce audio ottenute nelle tre differenti condizioni è stato normalizzato al *target level* di -23 LUFS secondo lo standard EBU R 128 che determina l'allineamento *loudness* e il livello massimo permesso dei segnali audio (EBU–Recommendation, 2011).

La *outlier detection* degli stimoli acustici è stata effettuata sui valori di *range* di *loudness* e *true-peak* considerati in ognuna delle tre condizioni di riproduzione, ottenendo 12 *outliers* (5 per il *range* di *loudness* e 7 per il *true-peak*).

Dei 50 stimoli audio-visivi validati ne sono stati esclusi 21, quelli indicati dalle analisi statistiche come *outliers* (6 per la luminosità\*, 3 per la dinamicità\*, 5 per il *range* di *loudness* e 7 per il *true-peak*), riducendo il numero degli stimoli a 29. Un ulteriore stimolo è stato escluso in quanto non presentava riproduzione sonora nei canali posteriori (Ls, Rs) nella condizione Surround; un ultimo stimolo è stato escluso in quanto ritenuto dallo sperimentatore non idoneo alla somministrazione perché presentava contenuti visivi\* sensibili. Sono quindi state selezionate 27 tracce audio in tre diverse condizioni di riproduzione (27 stimoli Mono, 27 stimoli Stereo e 27 stimoli Surround) per un totale di 81 stimoli sperimentali.

Tutte le estrazioni ed elaborazioni audio-video sono state effettuate con il *software* Adobe Premiere Pro CC (v12.0); tutte le analisi sono state effettuate tramite i *software* MATLAB (R2020a), PIVlab (v2.41) e R (v1.2.5001) (Matlab, 2012; Team, 2019; Thielicke & Stamhuis, 2014).

## **2.2.4 Paradigma Sperimentale**

Il paradigma sperimentale prevedeva lo svolgimento in successione delle seguenti fasi.

1. Compilazione, prima dell'arrivo, dei sei questionari per la selezione dei partecipanti (si veda Paragrafo 2.2.2.a) tramite Google Forms (Google Inc.).

---

\* La procedura di analisi della componente visiva degli stimoli non viene riportata essendo trascurabile ai fini di questo lavoro sperimentale.

2. Esecuzione del test audiometrico (si veda Paragrafo 2.2.2.b).
3. Esecuzione dei due test di *screening* per la selezione dei partecipanti (*Loudness* e Localizzazione della Sorgente Sonora) somministrati in ordine casuale (si vedano Paragrafi 2.2.2.c/d).
4. Training.
5. Svolgimento dell'Esperimento sulle modalità di riproduzione acustica.
6. Compilazione di due questionari di controllo.
7. *Debriefing*.

Dopo l'arrivo in laboratorio, ai partecipanti veniva chiesto di leggere e firmare il consenso informato per la partecipazione all'esperimento.

Successivamente, il partecipante veniva fatto accomodare sulla sedia posizionata all'interno della cabina audiometrica silente e veniva istruito sull'importanza di mantenere fisse le posizioni della sedia e della propria postura per tutta la durata dell'esperimento al fine di non alterare la percezione degli stimoli acustici, dopodiché veniva lasciato solo nella cabina con la porta chiusa e la comunicazione con l'esterno avveniva tramite un interfono.

Venivano quindi effettuati i test per la selezione dei partecipanti introdotti nei Paragrafi 2.2.2.b/c/d.

Per determinare il livello di *loudness* col quale avrebbe eseguito il test audiometrico, il partecipante indossava le cuffie *over ear* dalle quali veniva somministrato un tono puro a 3.5 KHz con un basso livello di *loudness*; il partecipante doveva indicare verbalmente se riusciva a percepire il suono. Il suono veniva somministrato più volte a livelli di *loudness* differenti fino ad individuare il livello minimo di percezione del partecipante



che veniva usato durante tutto il test audiometrico. Il partecipante veniva quindi istruito verbalmente su come effettuare il test audiometrico che svolgeva in autonomia come descritto precedentemente (Paragrafo 2.2.2.b).

Terminato il test audiometrico al partecipante venivano fatte togliere le cuffie e la somministrazione di tutti i successivi stimoli acustici avveniva utilizzando il sistema *surround*.

Venivano quindi illustrate verbalmente le procedure di esecuzione dei test di *screening* (*Loudness* e Localizzazione della Sorgente Sonora) e venivano somministrati due stimoli di riferimento che sarebbero stati utilizzati successivamente nei test, in modo che il partecipante potesse comprendere che tipo di suoni avrebbe dovuto valutare. I due test venivano presentati in ordine casuale e il partecipante aveva modo di leggere le istruzioni e fare domande prima dell'esecuzione di ogni test. Il partecipante svolgeva quindi ad occhi aperti i due test come precedentemente descritto nei Paragrafi 2.2.2.c/d.

Terminati i test di selezione lo sperimentatore verificava che il partecipante avesse risposto accuratamente ad un minimo di 80% degli stimoli somministrati come precedentemente indicato nei Paragrafi 2.2.2.c/d; se il criterio di selezione non veniva rispettato, il partecipante veniva escluso prima di procedere oltre. Il partecipante idoneo veniva quindi istruito verbalmente su come svolgere il compito di valutazione degli stimoli sperimentali e su come esprimere la propria valutazione.

Ogni partecipante doveva ascoltare ad occhi aperti gli 81 stimoli riprodotti in sequenza casuale per due ripetizioni per un totale di 162 prove suddivise in tre blocchi da 54 stimoli intervallati da una pausa di durata variabile a discrezione del partecipante; ogni blocco

durava circa 15 minuti. Prima della fase sperimentale ogni partecipante doveva inoltre svolgere un training composto da tre prove.

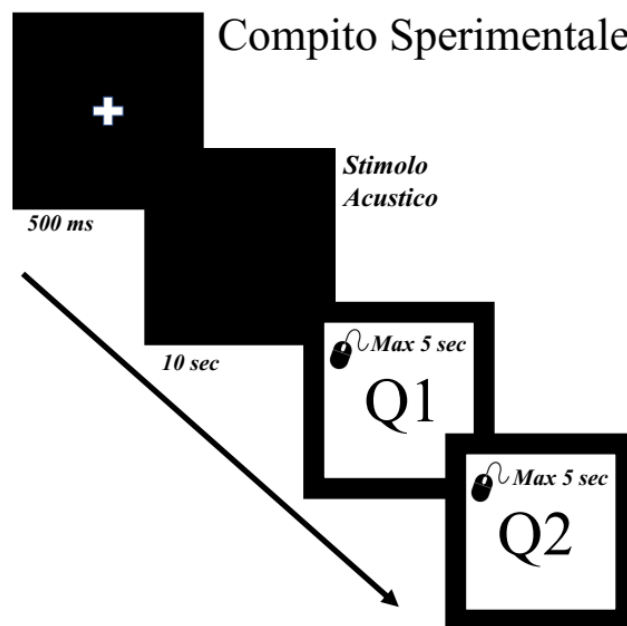
Dopo la prima riproduzione di uno stimolo, il partecipante doveva rispondere a due di quattro domande proposte in ordine casuale; dopo la seconda ripetizione dello stesso stimolo, il partecipante rispondeva alle restanti due delle quattro domande proposte in ordine casuale. Ogni stimolo riceveva quindi un *rating* per ognuna delle domande sperimentali. Le domande di valutazione della riproduzione sonora che apparivano sullo schermo erano le seguenti:

- “**Quanto ti è sembrata REALISTICA?**”
- “**Quanto ti sei sentita/o COINVOLTA/O EMOTIVAMENTE?**”
- “**Quanto ti è PIACIUTA?**”
- “**Quanto ti sei sentita/o IMMERSA/O FISICAMENTE?**”

Ogni prova, quindi, prevedeva la presentazione di una croce di fissazione per la durata di 500 ms, la presentazione di uno stimolo sperimentale seguito da due domande ad ognuna delle quali il partecipante doveva rispondere il più velocemente ed accuratamente possibile (e comunque entro un limite temporale prestabilito di 5 secondi) selezionando col mouse una qualsiasi posizione all'interno di una scala VAS i cui estremi erano contrassegnati dalle etichette “**Poco**” e “**Molto**” (Figura 2.3).

I significati dei termini presenti nelle domande venivano chiariti durante l'esposizione delle istruzioni; sullo schermo veniva visualizzata una pagina che riportava le quattro domande accompagnate dalle seguenti definizioni:

- “Per **REALISMO** si intende il livello di fedeltà di riproduzione di una scena sonora”
- “Per **COINVOLGIMENTO EMOTIVO** si intende quanto ti sei emozionato”
- “Per **PIACEVOLEZZA** si intende quanto hai apprezzato la sequenza audio”
- “Per **IMMERSIVITÀ FISICA** si intende l’impressione di essere fisicamente presenti all’interno della scena sonora riprodotta”



**Figura 2.3** Schema Temporale del Compito Sperimentale

*Q1 e Q2: Domande Sperimentali*

Al termine della sessione sperimentale il partecipante doveva compilare i seguenti due questionari:

1. F-IEQ (Film Immersive Experience Questionnaire): questionario composto da 31 *items* per la misurazione della capacità immersiva dello spettatore durante la visione di film (adattato alla sola esperienza acustica), suddiviso nelle sottoscale *Captivation*, *Real-world dissociation*, *Comprehension*, *Transportation* (Rigby et al., 2019);

2. Post-Q: questionario composto da nove *items* costruito *ad hoc* dallo sperimentatore per ottenere maggiori informazioni sulla familiarità del partecipante agli stimoli e al sistema di riproduzione utilizzati durante la sessione sperimentale.

I punteggi minimi e massimi, le medie e le deviazioni standard delle risposte medie dei partecipanti ad ogni sottoscala del questionario F-IEQ sono riportati in Tabella 2.2.

<i>Questionario</i>	<i>Min.</i>	<i>Max.</i>	<i>M</i>	<i>SD</i>
F-IEQ CAP	44	77	62.8	10.6
F-IEQ DIS	10	20	15.4	2.6
F-IEQ COM	9	27	15.5	4.5
F-IEQ TRA	16	27	23.3	3.3

**Tabella 2.2** *Statistiche Descrittive per le sottoscale del Questionario F-IEQ*

*F-IEQ = Film Immersive Experience Questionnaire (CAP = Captivation, DIS = Real-world dissociation, COM = Comprehension, TRA = Transportation)*

Terminata la compilazione dei questionari veniva effettuato un breve *debriefing* durante il quale veniva esplicitato lo scopo dell'esperimento e il partecipante aveva la possibilità di fare domande agli sperimentatori.

A tutti i partecipanti sono stati forniti in modalità elettronica un Foglio Informativo e una Nota Informativa contenenti indicazioni sullo scopo dell'esperimento, sulle modalità di partecipazione all'esperimento, sul trattamento dei dati personali e sulla copertura assicurativa. Lo studio è stato svolto in accordo con il comitato etico locale "Comitato etico Area Vasta Emilia Nord" ed è stato condotto in accordo con la Dichiarazione di Helsinki del 2013 (WMA, 2013). Tutte le sessioni sperimentali si sono svolte rispettando il Protocollo condiviso sulle misure per il contrasto e il contenimento della diffusione del virus SARS-CoV-2/COVID-19 negli ambienti di lavoro (Ministero della Salute, 2021).

## CAPITOLO 3

### ANALISI E RISULTATI

#### 3.1 Analisi Statistiche

I partecipanti hanno assegnato, utilizzando una VAS, punteggi compresi nell'intervallo tra 0 e 100 ad ognuna delle quattro domande di ricerca (si veda Paragrafo 2.2.4) in ognuno dei tre livelli della condizione sperimentale.

Per verificare l'esistenza di differenze significative tra le risposte date dai partecipanti nei tre livelli della condizione sperimentale in interazione con le domande sperimentali è stata condotta un'ANOVA a misure ripetute. Le risposte dei partecipanti sono state considerate come variabile dipendente continua, mentre la condizione sperimentale è stata utilizzata come variabile indipendente categoriale a tre livelli (Mono, Stereo e Surround) e la domanda sperimentale è stata considerata come seconda variabile indipendente a quattro livelli (Coinvolgimento Emotivo, Immersività Fisica, Realismo, Piacevolezza). Tutti i prerequisiti per l'ANOVA sono stati controllati: assenza di *outlier* significativi, normalità delle distribuzioni, sfericità delle covarianze tra i diversi gruppi (condizione sperimentale e domanda di ricerca).

Dopo aver verificato l'assenza di *outliers* univariati e multivariati, per tutte le distribuzioni è stato effettuato il test di Shapiro-Wilk sulla normalità: nessuna distribuzione è risultata normalmente distribuita, e un'ulteriore ispezione visiva dei relativi *QQ plot* ha confermato i risultati ottenuti. Il numero di osservazioni, le medie e le

deviazioni standard delle risposte medie dei partecipanti alla domanda di ricerca nei diversi livelli della condizione sperimentale sono mostrate in Tabella 3.1.

<i>Condizione</i>	<i>Domanda</i>	<i>N</i>	<i>M</i>	<i>SD</i>
Mono	Coinvolgimento Emotivo	453	36.3	24.4
Mono	Piacevolezza	456	38.4	23.5
Mono	Immersività Fisica	453	36.0	26.3
Mono	Realismo	457	45.3	24.6
Stereo	Coinvolgimento Emotivo	452	45.2	24.9
Stereo	Piacevolezza	456	44.8	25.3
Stereo	Immersività Fisica	453	50.5	24.8
Stereo	Realismo	455	53.7	23.6
Surround	Coinvolgimento Emotivo	457	49.9	24.9
Surround	Piacevolezza	457	51.5	27.0
Surround	Immersività Fisica	455	60.5	25.2
Surround	Realismo	451	60.4	24.9

**Tabella 3.1** *Statistiche Descrittive*

*N: numerosità dei dati raccolti; M: media dei punteggi assegnati; SD: deviazione standard*

La sfericità delle covarianze è stata controllata effettuando il test di Mauchly. I risultati hanno indicato la violazione del prerequisito di sfericità delle covarianze tra i livelli della condizione sperimentale ( $W = .280, p < .001, \varepsilon = .58$ ), tra i livelli della domanda sperimentale ( $W = .287, p = .003, \varepsilon = .66$ ) e nell'interazione tra i livelli della condizione e della domanda sperimentale ( $W = .030, p < .001, \varepsilon = .45$ ).

In seguito alla violazione dei prerequisiti di normalità e di sfericità si è optato per l'utilizzo di una *two-way* ANOVA fattoriale robusta, generalizzazione del metodo di Welch (20% *trimming level*) (Mair & Wilcox, 2018). I risultati dell'ANOVA sono

mostrati in Tabella 3.2; non vengono riportati i gradi di libertà in quanto la computazione della funzione analitica utilizza un valore critico corretto (Mair & Wilcox, 2018).

<i>Predictor</i>	<i>F</i>	<i>p</i>
Condizione	419.3	.001
Domanda	101.8	.001
Condizione x Domanda	32.8	.001

**Tabella 3.2** *Two-way ANOVA*

*Effetti principali (Condizione e Domanda) e effetto di interazione (Condizione x Domanda)*

Il primo effetto principale (Condizione), il secondo effetto principale (Domanda) e l'effetto d'interazione (Condizione x Domanda) sono risultati significativi, si è proseguito quindi con l'analisi *post-hoc* sui confronti a coppie delle medie *trimmed*. I risultati sono riportati nelle Tabelle 3.3, 3.4 e 3.5, gli intervalli di confidenza al 95% e i *p-value* sono stati corretti per la computazione di test multipli (Mair & Wilcox, 2018); le medie dei punteggi assegnati alle quattro domande di ricerca nei tre livelli di condizione sperimentale sono rappresentati graficamente in Figura 3.1.

<i>Contrast</i>	$\Psi$	<i>CI Lower</i>	<i>CI Upper</i>	<i>p</i>
Mono vs Stereo	-11.70	-14.08	-9.33	0.000
Mono vs Surround	-20.03	-22.42	-17.64	0.000
Stereo vs Surround	-8.33	-10.70	-5.95	0.000

**Tabella 3.3** *Confronti a Coppie (effetto principale: Condizione)*

$\Psi$  : differenza tra medie *trimmed*; *CI Lower* : limite inferiore intervallo di confidenza della media; *CI Upper* : limite superiore intervallo di confidenza della media



Come mostrato in Tabella 3.3 tutti i confronti a coppie tra i livelli della condizione sperimentale sono risultati significativi: considerando complessivamente le quattro domande di ricerca, Surround ha ricevuto punteggi maggiori ( $M = 55.5$ ,  $DS = 26.0$ ), seguito da Stereo ( $M = 48.5$ ,  $DS = 25.0$ ), mentre Mono ha ricevuto punteggi minori ( $M = 39.0$ ,  $DS = 25.0$ ); la differenza maggiore è stata riscontrata tra Mono e Surround.

<i>Contrast</i>	$\Psi$	<i>CI Lower</i>	<i>CI Upper</i>	<i>p</i>
CE vs P	-0.77	-3.93	2.40	0.525
CE vs IF	-5.80	-9.09	-2.51	0.000
CE vs R	-9.92	-12.98	-6.87	0.000
P vs IF	-5.04	-8.38	-1.70	0.000
P vs R	-9.16	-12.26	-6.05	0.000
IF vs R	-4.19	-7.35	-0.89	0.002

**Tabella 3.4** *Confronti a Coppie (effetto principale: Domanda)*

*CE: Coinvolgimento Emotivo; P: Piacevolezza; IF: Immersività Fisica; R: Realismo;  $\Psi$ : differenza tra medie trimmed; CI Lower: limite inferiore intervallo di confidenza della media; CI Upper: limite superiore intervallo di confidenza della media*

La Tabella 3.4 evidenzia che anche i confronti a coppie tra i livelli della domanda sperimentale sono risultati significativi ad esclusione del confronto tra Coinvolgimento Emotivo e Piacevolezza. Realismo ha ottenuto i punteggi maggiori ( $M = 53.1$ ,  $DS = 25.1$ ), seguita da Immersività Fisica ( $M = 49.0$ ,  $DS = 27.3$ ), Piacevolezza ( $M = 44.9$ ,  $DS = 25.8$ ) e Coinvolgimento Emotivo ( $M = 43.8$ ,  $DS = 25.3$ ).

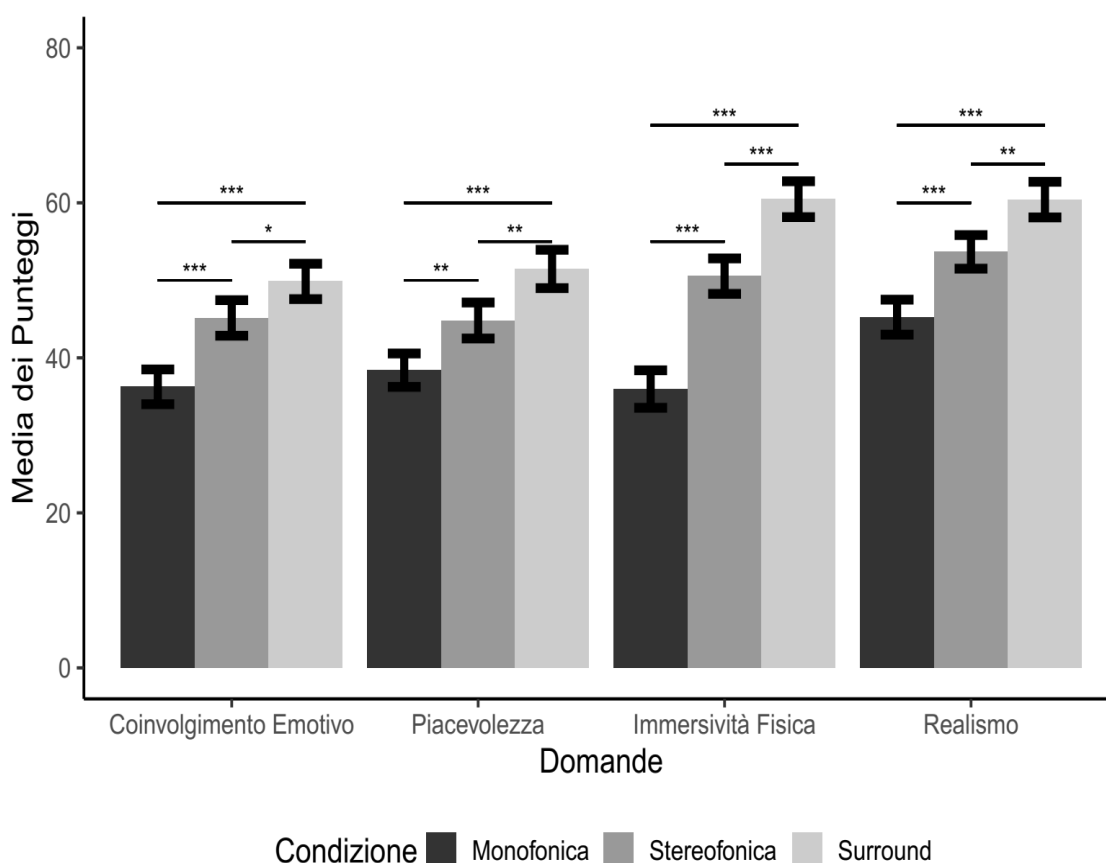
<i>Condizione</i>	<i>Domanda</i>	$\Psi$	<i>CI Lower</i>	<i>CI Upper</i>	<i>p</i>
Mono	CE vs P	-2.28	-8.23	3.68	0.940
Mono	CE vs IF	1.20	-5.34	7.74	0.940
Mono	CE vs R	-10.05	-16.16	-3.94	0.000
Mono	P vs IF	3.47	-2.89	9.84	0.940
Mono	P vs R	-7.77	-13.70	-1.84	0.001
Mono	IF vs R	-11.25	-17.76	-4.73	0.000
Stereo	CE vs P	0.43	-5.99	6.84	0.940
Stereo	CE vs IF	-6.35	-12.35	-0.34	0.015
Stereo	CE vs R	-9.30	-15.32	-3.28	0.000
Stereo	P vs IF	-6.77	-13.02	-0.52	0.012
Stereo	P vs R	-9.73	-15.99	-3.46	0.000
Stereo	IF vs R	-2.96	-8.80	2.88	0.940
Surround	CE vs P	-1.52	-8.02	4.98	0.940
Surround	CE vs IF	-11.85	-17.85	-5.84	0.000
Surround	CE vs R	-10.92	-16.99	-4.86	0.000
Surround	P vs IF	-10.33	-16.84	-3.83	0.000
Surround	P vs R	-9.41	-15.97	-2.85	0.000
Surround	IF vs R	0.92	-5.14	6.99	0.940
Mono vs Stereo	CE	-10.75	-16.91	-4.59	0.000
Mono vs Surround	CE	-16.95	-23.02	-10.89	0.000
Stereo vs Surround	CE	-6.20	-12.30	-0.11	0.020
Mono vs Stereo	P	-8.047	-14.27	-1.82	0.001
Mono vs Surround	P	-16.19	-22.58	-9.81	0.000
Stereo vs Surround	P	-8.15	-14.95	-1.34	0.004
Mono vs Stereo	IF	-18.29	-24.68	-11.91	0.000
Mono vs Surround	IF	-30.00	-36.48	-23.52	0.000
Stereo vs Surround	IF	-11.71	-17.62	-5.79	0.000
Mono vs Stereo	R	-10.00	-15.97	-4.03	0.000
Mono vs Surround	R	-17.83	-23.94	-11.73	0.000
Stereo vs Surround	R	-7.83	-13.83	-1.83	0.001

**Tabella 3.5** *Confronti a Coppie (effetto interazione: Condizione x Domanda)*

*CE: Coinvolgimento Emotivo; P: Piacevolezza; IF: Immersività Fisica; R: Realismo;  $\Psi$ : differenza tra medie trimmed; CI Lower: limite inferiore intervallo di confidenza della media; CI Upper: limite superiore intervallo di confidenza della media*

Come si può vedere dalla Tabella 3.5, la differenza maggiore in Surround è stata riscontrata tra le domande Coinvolgimento Emotivo e Immersività Fisica, in Stereo tra Piacevolezza e Realismo, in Mono tra Immersività Fisica e Realismo. Si può inoltre notare che in tutti i livelli della condizione non è stata riscontrata alcuna differenza

significativa tra le domande Coinvolgimento Emotivo e Piacevolezza. In Surround e in Stereo la differenza tra le domande Immersività Fisica e Realismo non è risultata significativa, inoltre in Mono la domanda Immersività Fisica non è stata evidenziata come significativamente differente dalle domande Coinvolgimento Emotivo e Piacevolezza (per le medie dei punteggi e deviazioni standard si veda la Tabella 3.1). Tutti i confronti tra i livelli della condizione sperimentale all'interno della stessa domanda di ricerca sono risultati significativi: la differenza maggiore si è sempre ottenuta tra Mono e Surround, mentre la differenza minore è stata riscontrata tra Stereo e Surround (ad esclusione della domanda Piacevolezza dove la differenza minore è stata ottenuta nel confronto tra Mono e Stereo).



**Figura 3.1** Grafico di Confronto fra le Medie dei Punteggi

Medie dei punteggi assegnati alle quattro domande di ricerca nei tre livelli di condizione sperimentale con barre di intervallo di confidenza al 95% e confronti post-hoc (\*\* $p < .001$ ; \* $p < .01$ ; \* $p < .05$ )

Come schematizzato graficamente in Figura 3.1, in generale (considerando tutti i livelli della domanda sperimentale insieme), e in particolare per ogni livello della domanda sperimentale:

- la differenza maggiore è stata ottenuta tra le condizioni Mono e Surround;
- i punteggi assegnati nella condizione Surround erano maggiori di quelli assegnati nelle condizioni Mono e Stereo;
- i punteggi assegnati nella condizione Stereo erano maggiori di quelli assegnati nella condizione Mono.

Considerando invece ogni livello distinto della condizione, i punteggi più alti si sono registrati:

- in Mono per la domanda Realismo ( $M = 45.3$ ,  $DS = 24.6$ );
- in Stereo per la domanda Realismo ( $M = 53.7$ ,  $DS = 23.6$ );
- in Surround per la domanda Immersività Fisica ( $M = 60.5$ ,  $DS = 25.2$ );

mentre i punteggi più bassi sono stati ottenuti:

- in Mono per la domanda Immersività Fisica ( $M = 36.0$ ,  $DS = 26.3$ );
- in Stereo per la domanda Piacevolezza ( $M = 44.8$ ,  $DS = 25.3$ );
- in Surround per la domanda Coinvolgimento Emotivo ( $M = 49.9$ ,  $DS = 24.9$ ).

Per le analisi statistiche è stato utilizzato il software R (Team, 2019). Nello specifico sono stati utilizzati i seguenti pacchetti: “rstatix” per la determinazione delle statistiche di confronto delle distribuzioni di dati (Alboukadel Kassambara, 2021); “ggplot2” per la costruzione dei grafici (A Kassambara, 2020); “ez” per il controllo della sfericità delle

covarianze (Lawrence, 2016); “WRS2” per le analisi robuste ANOVA e post-hoc (Mair & Wilcox, 2018).

# CAPITOLO 4

## DISCUSSIONE E CONCLUSIONI

### 4.1 Discussione

L'obiettivo del presente studio era quello di investigare l'effetto della spazializzazione della riproduzione del suono sulla sensazione di "presenza" del partecipante. A questo scopo sono stati selezionati stimoli acustici ecologici (si veda Paragrafo 2.1.2) e presentati ai partecipanti in tre modalità di riproduzione contraddistinte per i diversi livelli di caratterizzazione spaziale (mono < stereo < *surround*); i partecipanti hanno quindi valutato ciascuno stimolo in base al realismo e gradimento della riproduzione acustica e alla propria percezione di coinvolgimento fisico ed emotivo, permettendo di ricavare una stima indiretta della sensazione di presenza (punteggi più alti portano a maggiore presenza; per la definizione del costrutto si veda il Paragrafo 1.2).

Il campione sperimentale è stato selezionato in base alle capacità di valutazione di stimoli acustici: sono stati selezionati partecipanti "*un-trained/naives*" in modo da poter generalizzare i risultati ad una popolazione senza abilità specifiche nella valutazione di sistemi di riproduzione acustica (si veda il Paragrafo 2.2.2). I questionari somministrati per stimare la capacità immersiva dei partecipanti durante l'ascolto di tracce sonore (Tabella 2.2) hanno evidenziato punteggi nella media, confermando la pertinenza del campione selezionato ai fini della generalizzazione dei risultati di questo studio alla popolazione.

Per la valutazione delle riproduzioni sonore sono state proposte quattro diverse domande: l'analisi dei dati ha evidenziato che le domande sul coinvolgimento emotivo e sulla piacevolezza non sono risultate differenti in nessuna modalità di riproduzione, suggerendo che i due costrutti potrebbero essere valutati con simile criterio soggettivo.

In modalità *surround* i punteggi maggiori si sono ottenuti con la domanda sull'immersività fisica, mentre invece per la stessa domanda si sono riscontrati i punteggi minori in modalità di riproduzione monofonica. Inoltre, nei livelli più alti di immersività del sistema (riproduzione stereofonica e *surround*), le domande sull'immersività fisica e sul realismo non sono risultate significativamente differenti. Infine, i punteggi sul realismo della riproduzione sonora sono risultati minori per la somministrazione in modalità monofonica, intermedi per la stereofonica e maggiori per il *surround*. Questi dati confermerebbero la corretta percezione del livello di immersività del sistema di riproduzione da parte dei partecipanti (si veda il Paragrafo 1.2) e sono in linea con la meta-analisi di Cummings e colleghi, i quali riportano che l'esperienza di presenza correla positivamente con il livello di immersività del sistema (Cummings & Bailenson, 2016).

Secondo quanto emerso dai risultati di questo studio, la riproduzione acustica in modalità *surround* sembrerebbe quindi essere considerata più realistica e gradevole rispetto alle altre due modalità (mono e stereo), comportando maggiore coinvolgimento fisico ed emotivo, mentre la riproduzione acustica monofonica sembrerebbe quella meno in grado di elicitare la sensazione di presenza. Una possibile interpretazione di questi risultati potrebbe essere data dalla maggiore disponibilità di indizi di posizione degli oggetti localizzati nello spazio acustico virtuale, che porterebbe quindi il partecipante a sentirsi maggiormente localizzato nella scena riprodotta e conseguentemente a "vivere" una più intensa esperienza di presenza.

I risultati di questo studio confermano pertanto quelli ottenuti nei precedenti studi che hanno investigato la sensazione di presenza modulata dalla spazializzazione della riproduzione del suono (Bracken et al., 2010; Kobayashi et al., 2015; Lessiter et al., 2001; Västfjäll, 2003) e offrono un ulteriore supporto all'ipotesi che una maggiore spazializzazione comporti una maggiore sensazione di presenza in quanto, a differenza degli studi sopracitati che prevalentemente utilizzavano un solo stimolo presentato in condizioni differenti, in questo lavoro sono stati somministrati stimoli eterogenei ed ecologici che permettono una maggiore generalizzazione dei risultati.

## **4.2 Conclusioni e Futuri Sviluppi**

Questo studio fornisce nuovi dati su come il livello di dettaglio spaziale di una scena presentata nella sola modalità acustica possa influenzare le percezioni e le sensazioni del percipiente dando origine all'esperienza psicologica di presenza.

Come ipotizzato nel Paragrafo 1.4, la riproduzione sonora in modalità *surround*, rispetto alle altre condizioni:

- induce un maggiore coinvolgimento fisico ed emotivo;
- viene interpretata come più realistica;
- è generalmente più apprezzata per le qualità del suono.

Si potrebbe ipotizzare quindi che una migliore caratterizzazione dello spazio sonoro riprodotto possa attivare maggiormente i meccanismi interni di simulazione (di azioni ed emozioni) che ci farebbero interpretare l'ambiente virtuale similmente all'esperienza di presenza nel mondo reale.



In ottica di futuri sviluppi tecnologici di sistemi di riproduzione portatili o per uso domestico volti a coinvolgere e ad intrattenere piacevolmente l'utente, sarebbe opportuna una particolare attenzione alla componente acustica, in particolare alla modalità di distribuzione del suono. L'obiettivo sarebbe quindi fornire indizi acustici verosimili sulla localizzazione degli oggetti virtuali nella scena riprodotta tramite l'utilizzo di più canali diversificati per l'emissione delle tracce sonore come avviene comunemente nelle sale cinematografiche.

Questo studio rappresenta la prima fase di un progetto di ricerca più ampio il cui scopo ultimo è la più completa caratterizzazione dell'esperienza di presenza modulata dal settaggio audio-visivo. Il paradigma sperimentale e gli stimoli acustici utilizzati in questo lavoro verranno riproposti in un secondo esperimento durante il quale verrà registrata l'attività cerebrale dei partecipanti tramite EEG ad alta densità (128 elettrodi, sampling rate = 500 Hz) con l'obiettivo di rilevare desincronizzazioni del ritmo Mu (indice di coinvolgimento fisico in prima persona o di simulazione incarnata, più propriamente "liberata" in questo contesto) e asimmetrie nel ritmo frontale Alfa (indice di coinvolgimento emotivo).

Queste due fasi verranno poi riproposte in maniera analoga per la componente visiva; gli stimoli selezionati saranno somministrati nella sola modalità video tramite tre dispositivi con dimensioni differenti: uno *smartphone*, un *laptop* e un televisore. Per i limiti dovuti al paradigma sperimentale che prevede l'immobilità del partecipante durante le registrazioni EEG, non sarà possibile randomizzare le riproduzioni video sui tre schermi, ma ogni partecipante visualizzerà tutti gli stimoli in un'unica modalità di riproduzione, confrontando successivamente i risultati *between subjects*.

Come ultima fase del progetto le condizioni emerse come più attivanti nella sola modalità acustica e nella sola modalità visiva verranno combinate somministrando quindi ai partecipanti stimoli audio-visivi propriamente composti e confrontando i nuovi risultati delle registrazioni EEG con quelli ottenuti da un campione di controllo che eseguirà lo stesso compito in condizioni di riproduzione non spazializzata (es: riproduzione monofonica su *laptop*).

## BIBLIOGRAFIA

- 7029, I. S. O. (1984). *Acoustics-Threshold of hearing by air conduction as a function of age and sex for otologically normal persons*. International Organization for Standardization Geneva.
- Albiero, P., Ingoglia, S., & Lo Coco, A. (2006). Contributo all'adattamento italiano dell'Interpersonal Reactivity Index. *Testing Psicometria Metodologia*, 13(2), 107–125.
- Bech, S., & Zacharov, N. (2007). *Perceptual audio evaluation-Theory, method and application*. John Wiley & Sons.
- Beck, J. S. (2003). *A quiet revolution: Changes in American film sound practices, 1967–1979*. The University of Iowa.
- Böcker, M., & Mühlbach, L. (1993). Communicative presence in videocommunications. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 37(3), 249–253.
- Bracken, C. C., Pettey, G., Guha, T., & Rubenking, B. E. (2010). Sounding out small screens and telepresence. *Journal of Media Psychology*.
- Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., & Freund, H. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience*, 13(2), 400–404.

- Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology, 19*(2), 272–309.
- Di Cesare, G., Di Dio, C., Marchi, M., & Rizzolatti, G. (2015). Expressing our internal states and understanding those of others. *Proceedings of the National Academy of Sciences, 112*(33), 10331–10335.
- Di Donato, M. (2006). *La spazializzazione acustica nel cinema contemporaneo. Tecnica, linguaggio, modelli di analisi.* (Onyx).
- Dodakian, L., Stewart, J. C., & Cramer, S. C. (2014). Motor imagery during movement activates the brain more than movement alone after stroke: a pilot study. *Journal of Rehabilitation Medicine, 46*(9), 843–848.
- EBU–Recommendation, R. (2011). *Loudness normalisation and permitted maximum level of audio signals.*
- Elsaesser, T., & Hagener, M. (2009). Teoria del film: un'introduzione. *Piccola Biblioteca Einaudi. Nuova Serie, 445.*
- Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology, 73*(6), 2608–2611.
- Fernández-Aguilar, L., Navarro-Bravo, B., Ricarte, J., Ros, L., & Latorre, J. M. (2019). How effective are films in inducing positive and negative emotional states? A meta-analysis. *PloS One, 14*(11), e0225040.
- Figueiredo, H. F., Bodie, B. L., Tauchi, M., Dolgas, C. M., & Herman, J. P. (2003).

- Stress integration after acute and chronic predator stress: differential activation of central stress circuitry and sensitization of the hypothalamo-pituitary-adrenocortical axis. *Endocrinology*, 144(12), 5249–5258.
- Gallese. (2007). Before and below ‘theory of mind’: embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480), 659–669.
- Gallese. (2010). Le basi neurofisiologiche dell’intersoggettività. *SOCIETÀ DEGLI INDIVIDUI (LA)*.
- Gallese, V., & Guerra, M. (2013). Forme di simulazione e sti (mo) li cinematografici. *Reti, Saperi, Linguaggi*, 2.
- Gallese, Vittorio, & Guerra, M. (2015). *Lo schermo empatico: cinema e neuroscienze*. Raffaello Cortina Editore.
- García-Pereira, I., Vera, L., Aixendri, M. P., Portalés, C., & Casas, S. (2020). Multisensory experiences in virtual reality and augmented reality interaction paradigms. In *Smart Systems Design, Applications, and Challenges* (pp. 276–298). IGI Global.
- Graziano, M. S. A., Reiss, L. A. J., & Gross, C. G. (1999). A neuronal representation of the location of nearby sounds. *Nature*, 397(6718), 428–430.
- Heeter, C. (1992). Being there: The subjective experience of presence. *Presence: Teleoperators & Virtual Environments*, 1(2), 262–271.
- ITU-R. (1997). Recommendation BS.1116-1, Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems. In

*International Telecommunications Union Radiocommunication Assembly.*

- ITU-R. (2007). *Rec. ITU-R BS.1770-1, Algorithms to measure audio programme loudness and true-peak audio level.* International Telecommunications Union.
- Kassambara, A. (2020). *ggplot2. Based Publication Ready Plots [R Package Ggpubr Version 0.2. 5].*
- Kassambara, Alboukadel. (2021). Pipe-Friendly Framework for Basic Statistical Tests [R Package Rstatix Version 0.7. 0]. *February. [https://Mran. Microsoft. Com/Web/Packages/Rstatix/Index. Html](https://Mran.Microsoft.Com/Web/Packages/Rstatix/Index.Html).*
- Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., & Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research, 153*(4), 628–636.
- Kim, T., & Biocca, F. (1997). Telepresence via television: Two dimensions of telepresence may have different connections to memory and persuasion. *Journal of Computer-Mediated Communication, 3*(2), JCMC325.
- Kobayashi, M., Ueno, K., & Ise, S. (2015). The effects of spatialized sounds on the sense of presence in auditory virtual environments: a psychological and physiological study. *Presence: Teleoperators and Virtual Environments, 24*(2), 163–174.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science, 297*(5582), 846–848.
- Larsen, E., Iyer, N., Lansing, C. R., & Feng, A. S. (2008). On the minimum audible

- difference in direct-to-reverberant energy ratio. *The Journal of the Acoustical Society of America*, 124(1), 450–461.
- Lawrence, M. A. (2016). ez: Easy analysis and visualization of factorial experiments. *R Package Version*, 4(2).
- Leppänen, J. M., & Nelson, C. A. (2009). Tuning the developing brain to social signals of emotions. *Nature Reviews Neuroscience*, 10(1), 37–47.
- Lessiter, J., Freeman, J., & Davidoff, J. (2001). Really hear? The effects of audio quality on presence. *Proceedings of the Fourth Annual International Workshop on Presence*, 288–324.
- Lipscomb, S. D., & Kerins, M. (2004). An empirical investigation into the effect of presentation mode in the cinematic and music listening experience. *Proceedings of the 8th International Conference on Music Perception & Cognition, Evanston, IL*.
- Lombard, M., Reich, R. D., Grabe, M. E., Bracken, C. C., & Ditton, T. B. (2000). Presence and television. The role of screen size. *Human Communication Research*, 26(1), 75–98.
- Lopatka, K., Kotus, J., & Czyzewski, A. (2016). Detection, classification and localization of acoustic events in the presence of background noise for acoustic surveillance of hazardous situations. *Multimedia Tools and Applications*, 75(17), 10407–10439.
- Maas, M. J., & Hughes, J. M. (2020). Virtual, augmented and mixed reality in K–12 education: A review of the literature. *Technology, Pedagogy and Education*, 29(2), 231–249.

- Mair, P., & Wilcox, R. (2018). *Robust statistical methods using WRS2. The WRS2 Package.*
- Matlab, S. (2012). Matlab. *The MathWorks, Natick, MA.*
- Mattila, V.-V., & Zacharov, N. (2001). Generalized listener selection (GLS) procedure. *PREPRINTS-AUDIO ENGINEERING SOCIETY.*
- Ministero della Salute. (2021). *Protocollo condiviso di aggiornamento delle misure per il contrasto e il contenimento della diffusione del virus SARS-CoV-2/COVID-19 negli ambienti di lavoro.*  
[https://www.salute.gov.it/imgs/C\\_17\\_pagineAree\\_5383\\_0\\_file.PDF](https://www.salute.gov.it/imgs/C_17_pagineAree_5383_0_file.PDF)
- Murphy, W. J., & Franks, J. R. (2002). Revisiting the NIOSH criteria for a recommended standard: Occupational noise exposure. *International Congress and Exposition on Noise Control Engineering*, 19–21.
- Muthukumaraswamy, S. D., & Johnson, B. W. (2004). Primary motor cortex activation during action observation revealed by wavelet analysis of the EEG. *Clinical Neurophysiology*, 115(8), 1760–1766.
- Palan, S., & Schitter, C. (2018). Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17, 22–27.
- Paul, A. (2009). *Audyssey DSX 10.2 Surround Sound Overview.*
- Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13.
- Rawashdeh, S. (2021). *Frequency Response of the Ear , Hearing Test.* MATLAB Central File Exchange.



<https://www.mathworks.com/matlabcentral/fileexchange/16101-frequency-response-of-the-ear-hearing-test>

Rec, I. (1996). P. 800: Methods for subjective determination of transmission quality.

*International Telecommunication Union, Geneva, 22.*

Rec, I. (2000). P. 832, Subjective performance evaluation of hands-free terminals.

*International Telecommunication Union.*

Reeves, B., Detenber, B., & Steuer, J. (1993). *New television: The effects of big pictures and big sound on viewer responses to the screen.*

Reeves, B., Lang, A., Kim, E. Y., & Tatar, D. (1999). The effects of screen size and message content on attention and arousal. *Media Psychology, 1*(1), 49–67.

Rigby, J. M., Brumby, D. P., Gould, S. J. J., & Cox, A. L. (2019). Development of a questionnaire to measure immersion in video media: The Film IEQ. *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*, 35–46.

Rizzolatti, G., & Destro, M. F. (2008). Mirror neurons. *Scholarpedia, 3*(1), 2055.

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1997). The space around us. *Science, 277*(5323), 190–191.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., & Fazio, F. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Experimental Brain Research, 111*(2), 246–252.

Roberts, R., Callow, N., Hardy, L., Markland, D., & Bringer, J. (2008). Movement imagery ability: development and assessment of a revised version of the vividness

- of movement imagery questionnaire. *Journal of Sport and Exercise Psychology*, 30(2), 200–221.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161.
- Sauter, M., Draschkow, D., & Mack, W. (2020). Building, hosting and recruiting: A brief introduction to running behavioral experiments online. *Brain Sciences*, 10(4), 251.
- Sbravatti, V. (2019). *La cognizione dello spazio sonoro filmico: un approccio neurofilmologico*. Bulzoni editore.  
[https://books.google.it/books?id=0\\_pqyQEACAAJ](https://books.google.it/books?id=0_pqyQEACAAJ)
- Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 6(6), 603–616.
- Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic stimuli in neuroscience: critically acclaimed. *Trends in Cognitive Sciences*, 23(8), 699–714.
- Sterne, J. (2003). *The audible past*. Duke University Press.
- Team, R. D. C. (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2011. URL: <https://www.R-project.org>. [Google Scholar].
- Thielicke, W., & Stamhuis, E. (2014). PIVlab—towards user-friendly, affordable and accurate digital particle image velocimetry in MATLAB. *Journal of Open Research Software*, 2(1).

- TSG, I. (2004). *15, Digital network echo cancellers (recommendation)*. tech. rep., Tech. Rep. G. 168, ITU-T.
- Västfjäll, D. (2003). The subjective sense of presence, emotion recognition, and experienced emotions in auditory virtual environments. *CyberPsychology & Behavior, 6*(2), 181–188.
- Visch, V. T., Tan, E. S., & Molenaar, D. (2010). The emotional and cognitive effect of immersion in film viewing. *Cognition and Emotion, 24*(8), 1439–1445.
- Werner, P. D., Swope, A. J., & Heide, F. J. (2006). The music experience questionnaire: Development and correlates. *The Journal of Psychology, 140*(4), 329–345.
- Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., Saari, T., Laarni, J., Ravaja, N., & Gouveia, F. R. (2007). A process model of the formation of spatial presence experiences. *Media Psychology, 9*(3), 493–525.
- Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence, 7*(3), 225–240.
- WMA. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *Jama, 310*(20), 2191–2194.
- Wright, B. A. (2005). *Making films sound better: the transition to Dolby sound in Hollywood cinema*. Carleton University.