

## Scoring Music for *Westworld* Then and Now: A Cognitive Perspective

Annabel Cohen

### **Introduction: Two *Westworlds*, Two Scores**

Since its 2016 debut, *Westworld*, the sci-fi–Western television/HBO series, has engaged audiences in a fictional world that questions the borders between past and future, freedom and determinism. This chapter explores *Westworld*'s musical dimension—the ways music draws a viewer into the show's world and helps explore its questions. The series takes its inspiration from Michael Crichton's 1973 film *Westworld*. Celebrated soundtrack composer and film educator, Fred Karlin, scored this early film, and his soundtrack is an excellent example of prototypical scoring techniques from the 1970s. Award-winning composer Ramin Djawadi scored the 2019 *Westworld* HBO series, as well as the popular *Game of Thrones* 2011-2019 and *Pacific Rim*, 2013. In this chapter, I review scientific studies that demonstrate film music's ability to direct viewers' experiences in powerful and novel ways. I also unpack Karlin's work in the context of early sci-fi–Western, as well as discuss Djawadi's approach to create a more nuanced approach. Later, I turn to my own cognitive model to help explain the composer's power, the listener's susceptibility, as well as contemporary music's ability to convey *Westworld*'s AI. I'll close with a discussion of implications for other cybermedia.

### **Harnessing the Brain's Representations of Musical Meaning and Emotion**

One of music soundtracks' primary goals is to convey a director's vision. Unlike dialogue and sound effects, music resides typically outside the diegesis—it draws viewers into compelling

fictional worlds.<sup>1</sup> Music also engages audiences emotionally in the film.<sup>2</sup> And by employing conventional codes, music establishes contexts of time and place. That this happens confirms the brain's interpretive skills with music across media.

Composers, music supervisors, film directors, and audiences bring tacit knowledge of music to films and media. Psychological research reveals that listeners prefer certain styles of music, can identify a piece's decade of popularity, and recall personal experiences associated with particular musical genres.<sup>3</sup> Participants' abilities to engage with these processes reveal complex knowledge structures established over years of musical exposure. Film composers exploit these listeners' familiarity with music in their scores.

Research shows that musical features convey meanings to listeners. For example, consider a melody's pitch direction. A melodic line that ascends in pitch will be judged by most individuals as conveying a happy meaning, and, conversely, a descending melodic line will be judged as conveying sadness.<sup>4</sup> Going a step further, functional magnetic resonance imaging (fMRI) of the brain during ascending and descending melodies has revealed three locations responsive to pitch

---

### **Acknowledgment**

Co-editors, Selmin Kara, Holly Rogers, Jonathan James Leal, and Carol Vernallis's close readings, constructive critiques, and editorial work are appreciated. Special gratitude is owed to Carol Vernallis for her careful and conscientious labors to both increase the accessibility of and maintain the integrity of the CAM-WN model in the face of word count constraints. Feedback from Noah Fram is also acknowledged. Thank you to University of Wyoming's American Heritage Center for, in the pandemic's midst, granting access to Karlin's score, and Dan Collins, too, for permission to reproduce his score of *Westworld's* opening credits. Gratitude is also expressed for previous support from the Social Sciences and Humanities Research Council of Canada.

<sup>1</sup> Claudia Gorbman, *Unheard Melodies: Narrative Film Music* (Bloomington, IN: Indiana University Press, 1987).

<sup>2</sup> Annabel J. Cohen, "Music as a Source of Emotion in Film," in *The Oxford Handbook of Music and Emotion*, eds. Patrick N. Juslin and John A. Sloboda (Oxford, Oxford University Press, 2010), 879-908.

<sup>3</sup> Carol Lynne Krumhansl, "Listening Niches across a Century of Popular Music," *Frontiers in Psychology*, (March 8, 2017),

<sup>4</sup> William G. Collier and Timothy L Hubbard, "Musical Scales and Evaluations of Happiness and Awkwardness: Effects of Pitch, Direction, and Scale Mode," *American Journal of Psychology* (2001): 355-375.

direction.<sup>5</sup> One contributing factor to this brain activity, suggested by researchers of music, is the alignment of melodic contour with action. An ascending melody may be associated with one kind of motion, such as climbing upward, and a descending melodic line, with another, such as walking downward.<sup>6</sup>

Other neuroimaging research suggests that the same brain areas are activated whether one moves a body part (e.g., leg or tongue) or reads related verbal concepts (e.g., the words *leg*, *face*, respectively). The finding demonstrates that a listener's conceptual understanding is grounded in fundamental interactions with the environment, resulting in *embodied* meaning—be it verbal, musical, or motoric.<sup>7</sup>

Similar brain mechanisms for both perceived and experienced actions have been explained by studies of mirror neurons (in non-human primates' brains) which respond during an activity (such as picking up an object) or watching someone else carry out the activity (such as seeing

---

<sup>5</sup> In an fMRI scanner, participants were tasked with click a button every time they heard a bidirectional, rather than unidirectional sequence. They then participated in a study in which they rated each sequence on a 7-point scale with *sad* and *happy* end points. Yune-Sang Lee et al., "Investigation of Melodic Contour Processing in the Brain Using Multivariate Pattern-Based FMRI," *NeuroImage*, U.S. National Library of Medicine, 57, no. 1 (July 2011): 293-300.

<sup>6</sup> A simple example of this association is provided in an early computer game which played an ascending 3-note pattern each time an avatar picked up an object and conversely a descending pattern when putting it down. Another game similarly paired ascending and descending 3-note patterns with going up and down stairs. See pp. 323-325 in Neil Lerner, "The Origins of Musical Style in Video Games 1977 – 1983," in *The Oxford Handbook of Film Music Studies*, ed. David Neumeyer (Oxford, UK: Oxford University Press, 2014), 319-347.

<sup>7</sup> Olaf Hauk, Ingrid Johnsrude, and Friedemann Pulvermüller, "Somatotopic Representation of Action Words in Human Motor and Premotor Cortex," *Neuron* 41, no. 2 (2004), 301-307, [https://doi.org/10.1016/s0896-6273\(03\)00838-9](https://doi.org/10.1016/s0896-6273(03)00838-9) Hauk, 2004, 2016). Olaf Hauk and L. Small, "Hauk, Olaf, Ingrid Johnsrude, and Friedemann Pulvermüller. 'Somatotopic Representation of Action Words in Human Motor and Premotor Cortex.' *Neuron* 41, No. 2 (2004): 301–7. [https://doi.org/10.1016/s0896-6273\(03\)00838-9](https://doi.org/10.1016/s0896-6273(03)00838-9). Olaf Hauk. "What does it mean? A review of the neuroscientific evidence for embodied lexical semantics" in *Neurobiology of Language*, ed. Gregory Hickok and Steven L. Small (London, UK: Academic Press, 2016), 777-788; George Lakoff and Mark Johnson, *Metaphors We Live by* (Chicago, Ill: University of Chicago Press, 2003). see also Arnie Cox, *Music and Embodied Cognition: Listening, Moving, Feeling, and Thinking* (Bloomington: Indiana University Press, 2017).

someone else pick up the object).<sup>8</sup> An analogous human mirror-neuron system helps describe an audience's experience of performers' activities that unfold through media. In accordance with a mirror system, watching a film produces neural activation similar to that which would occur when carrying out the activity oneself, as if the viewer-listener were the actor on the screen. The concept of mirror system is not limited to a visual depiction of action—listening to music may engage the motor system as if the viewer were the musician performing.<sup>9</sup> Notions of musical embodiment offer further importance to the music presented in a film or TV program.

In addition to meanings conveyed by melodic direction, music conveys pleasantness and unpleasantness through consonance and dissonance. In one experiment, the same melody was harmonized in six different ways to represent six levels of dissonance.<sup>10</sup> Listeners heard these melodies while in a PET brain scanner that measured the amount of blood flow in various areas of the brain. The results showed increased regional cerebral blood flow in the paralimbic and neocortical areas for melodies with increasing or decreasing consonance. These areas are typically engaged in emotional responses in general. In addition, some regions responded only to increasing dissonance, and others to greater consonance.

Many neuroscientific studies reveal that different properties of music activate different parts of the brain; however, few such studies have focused on the brain activity arising when listening to music while watching a film. One rare experiment using fMRI showed activation of the

---

<sup>8</sup> Giacomo Rizzolatti and Laila Craighero, "The Mirror-Neuron System," *Annual Review of Neuroscience* 27, no. 1 (2004): 169-192. .

<sup>9</sup> Katie Overy and Istvan Molnar-Szakacs, "Being Together in Time: Musical Experience and the Mirror Neuron System," *Music Perception* 26, no. 5 (January 2009): 489-504.

<sup>10</sup> Anne J. Blood et al., "Emotional Responses to Pleasant and Unpleasant Music Correlate with Activity in Paralimbic Brain Regions," *Nature Neuroscience* 2, no. 4 (1999): 382-387. .

amygdala (a center of emotional processing) when brief emotionally neutral film were presented with emotional music. No amygdala activity, however, occurred when either the film or music were presented alone.<sup>11</sup> This study suggests that music which elicits emotions causes the brain to search for a source of its meaning.<sup>12</sup> Thus, composers can use emotionally laden music that encourages a viewer to discover more about a film character's motivation, action, or relation to plot. As one example from the *Westworld* film, an unsettling electronic sound with an irregular yet relentless beat accompanies the robot Gunslinger as it chases an innocent human. Although the music causes an unsettled feeling, audiences would tend to attribute the robot to the feeling's source.

### **Film Composers, Resources, and Techniques**

Film and TV composers draw on both external and internal resources, as well as multiple techniques for their craft. External resources, limited by budgets, include support from artisans and current technologies. Internal resources include composers' knowledge of styles and pieces of music, retrievable for matching an emotion or motivation for a filmic action. Internal resources also include imagination and creativity. Djawadi, for example, created the theme for *Westworld* having learned of its context before seeing any of the filming.<sup>13</sup> Composers, music editors, and directors know how to exploit an audience's musical knowledge. At the same time, composers are educators of audiences, teaching new musical meanings through exposure to novel, or sometimes older yet unfamiliar, music in a new context. Decision makers (directors,

---

<sup>11</sup> Eran Eldar, Ori Ganor, Roe Admon, Avraham Bleich, and Talma Hendler, "Feeling the real world: Limbic response to music depends on related content," *Cerebral Cortex*, 17, No. 12 (2007): 2828-2840.

<sup>12</sup> Nicholas Cook, *Analysing Musical Multimedia* (Oxford, UK: Clarendon Press, 1998).

<sup>13</sup> Mike Hilleary et al., "'Westworld' Composer Ramin Djawadi on Why Those Radiohead Covers Keep Coming," *Pitchfork*, November 18, 2016, <https://pitchfork.com/thepitch/1370-westworld-composer-ramin-djawadi-on-why-those-radiohead-covers-keep-coming/> (accessed February 9, 2021).

music editors, and composers) rely on professional experience and training to intuit what music is needed and when. In a sense, they are mind readers, predicting how the audience will hear, interpret, and remember the film and its soundtrack.

Composers also draw on their proficiency with instruments as an internal resource. They may play instruments or be skilled at computer-based digital audio composition and production.

Karlin was a professional trumpet performer. He could play all orchestral instruments and mastered the use of synthesizers and early electronic instruments. Djawadi's primary instrument is guitar, but he studied organ from the age of four to eleven years. His early experience with the organ acquainted him with a wide range of timbres and experiences of musical embodiment.

One technique often employed in creating the music soundtrack is a *temp track*—often composed of swatches of pre-recorded music slotted in to a rough edit of the film—which cues a director where to place scored music, and within which genre and style. This temp track can serve either the director and/or the composer at early postproduction screenings and through final editing.<sup>14</sup> *Westworld's* Jonathan Nolan (co-creator) and Christopher Kaller (supervising music editor) created a temp track for the series' opening and shared it with Djawadi.<sup>15</sup> Often, and in part for budgetary reasons, the supervising music editor and composer begin work only after completion of filming and editing.

---

<sup>14</sup> Fred Karlin and Rayburn Wright, *On the Track: A Guide to Contemporary Film Scoring*, 2nd ed. (New York, NY: Taylor & Francis, 2004), 28.

<sup>15</sup> Kingsley Marshall, "Music as a Source of Narrative in HBO's *Westworld*," in *Reading Westworld*, ed. Alex Goody (Cham, Switzerland: Palgrave Macmillan, 2019), 97-118.

The use of recurring *leitmotifs* (from the German for “leading motif”) is an important scoring technique. In the simplest case, a leitmotiv links a musical theme or motif to a character or plot point. A film score may include numerous leitmotifs, their profiles chosen by the composer, though also sometimes the film director’s and supervising music editor’s input. Filmmakers assume that viewers retain information about the motif that first accompanies an incident or protagonist, and that each reoccurrence will bring to mind those first instances. Such hypothesized one-trial learning should be a mental challenge, but research by Marilyn Boltz provides some support for this.<sup>16</sup> Her study showed better memory for music that was cued by its accompanying visual scene as compared to a condition in which the music had no visual cue. The effect was specific, however, to situations in which the music and scene had similar meanings. Further research by Berthold Hoeckner and Howard Nusbaum showed that music cued visual scenes more effectively than visual scenes cued music, a phenomenon they refer to as the *Casablanca Effect*.<sup>17</sup>

As part of the filmmaking team, the composer employs additional techniques to shape an audience’s attention. Before synchronized film and sound, early film theorist and psychologist Hugo Münsterberg noted the ways music and film controlled an audience’s ordering of concepts and events.<sup>18</sup> He claimed that film was more like music than other art forms, including theatre and photography. Structural and physical aspects of film and TV may, however, change over time. Cognitive psychologist James Cutting and colleagues measured several lower-level

---

<sup>16</sup> Marilyn Boltz, “The Cognitive Processing of Film and Musical Soundtrack,” *Memory & Cognition* 32, no. 7 (2004): 1194-1205.

<sup>17</sup> Berthold Hoeckner and Howard Nusbaum, “Music and Memory in Film and Other Multimedia: The Casablanca Effect,” in *The Psychology of Music in Multimedia*, ed. Siu-Lan Tan et al. (Oxford, UK: Oxford University Press, 2013), 235-263.

<sup>18</sup> Hugo Münsterberg, *The Photoplay: A Psychological Study* (New York: Dover Publications, 1970). Also published as *The Film: a Psychological Study; the Silent Photoplay in 1916* (New York: Dover Publications, 1970).

features of 160 English-language films released between 1935 and 2010 and corroborated earlier findings of a decline in average shot length.<sup>19</sup> They also found an increase in the amount of motion and a decrease in average luminance (amount of light). They offered a single explanation for these linear trends: “Filmmakers have incrementally tried to exercise more control over the attention of filmgoers”.<sup>20</sup> Industry personnel for the *Westworld* series, with its themes of human control and free will, seem especially skilled at directing human attention.

Had the 1973 film *Westworld* been made in the 21<sup>st</sup> century, we might expect novel music which incorporated new styles and technologies. However, as Cutting has pointed out, changes in a film’s lower-level features do not mean changes of style or genre at a higher level.<sup>21</sup> Over time, directors and their teams have simply gotten better at directing audience attention. Thus, we may expect similar functions served by music in the 20<sup>th</sup> century film and the 21<sup>st</sup> century series *Westworld*, even though more external resources, as well as four additional decades of recorded music, were available to Djawadi.

I’ll next discuss the music of the 1973 *Westworld* film as well as the music of the series. My cognitive model of how music works in multimedia will help explain *Westworld*’s two different musical approaches.

### **Musical Challenges in *Westworld*, Then and Now**

---

<sup>19</sup> James E Cutting et al., “Quicker, Faster, Darker: Changes in Hollywood Film over 75 Years,” *i-Perception* 2, no. 6 (2011): 569-576.

<sup>20</sup> Cutting et al., “Quicker, Faster, Darker,” 569.

<sup>21</sup> See also James E. Cutting, “Narrative Theory and the Dynamics of Popular Movies,” *Psychonomic Bulletin & Review* 23, no. 6 (March 2016): 1713-1743. This article reports the analysis of 175 popular movies in 12 studies that show normative aspects of parameters like shot duration and motion, and luminance. These reduced to five stylistic dimensions (music-conversation being one) and were distributed systematically across four acts.



*Westworld* is an example of the sci-fi Western. Both genres, Western and science fiction, have different musical styles, and Karlin and Djawadi both had to deal with these two genres in one score. Musicologist Andrew Granade has traced the development of the two genres.<sup>22</sup> For the Western, he identifies three stages. The music of early Western films was typically "...in a major key with duple meter; featured a bit of syncopation; and stressed simple, chordal accompaniments". Special effects included "... triplet rhythms for galloping horses and tremolo strings for moments of mounting tension."<sup>23</sup> These conventions were influenced by Westerns of the silent era and music for staged melodramas of the late 19<sup>th</sup> and early 20<sup>th</sup> century.<sup>24</sup> This period also included the singing cowboy movies of Roy Rogers and Gene Autry that used guitar, fiddle, banjo, and string bass, and attracted country and western music fans. The second stage began after 1939 with the John Ford films *Stagecoach* and *Drums Along the Mohawk*. The primary influences were American folk songs, hymns, and patriotic songs, scored with a style made popular in Aaron Copland's Western themed ballets *Billy the Kid* (1938) and *Rodeo* (1942), with perfect fourths rising in parallel motion (violating traditional harmony) and "thin orchestration."<sup>25</sup> The final stage is attributed to the influence of the late Ennio Morricone who added electric guitar, non-Western percussion, Mariachi-style trumpet, and pop stylized songs similar to groups like the Beatles. Morricone created a "totality of sound" that blended the sound effects of the West (e.g., horse hooves, gunshots) with the music track. Two further dichotomies added to the semantic cues. First, tonal music was associated with *good*, and atonal music was

---

<sup>22</sup>S. Andrew Granade, "Some People Call Me the Space Cowboy": Sonic Markers of the Science Fiction Western," 2019, [https://www.academia.edu/38629097/\\_Some\\_People\\_Call\\_Me\\_the\\_Space\\_Cowboy\\_Sonic\\_Markers\\_of\\_the\\_Science\\_Fiction\\_Western](https://www.academia.edu/38629097/_Some_People_Call_Me_the_Space_Cowboy_Sonic_Markers_of_the_Science_Fiction_Western) (accessed February 10, 2021) Note: An abbreviated version of this paper appears in *Re-Locating the Sounds of the Western*, ed. Mariana Whitmer and Kendra Preston Leonard (New York, NY: Routledge Press, 2019).

<sup>23</sup> Granade, "Sonic Markers," 7.

<sup>24</sup> Mariana Whitmer, "Melodramatic Music in the Western," *The Journal of Film Music* 5, no.1-2 (2012): 111-112.

<sup>25</sup> Granade, "Sonic Markers," 8.

associated with *bad*. Secondly, acoustic (orchestral) instruments were also associated with good, and electronic instruments with evil.<sup>26</sup> Granade notes the singability of music of Westerns and the use of syncopated ostinatos against a “rhythmically steady though leaping melody”.<sup>27</sup> Granade identifies only two periods for the development of music for science fiction. The first is characterized by electronic or atypical orchestral instruments, such as the theremin, as used by Bernard Herrmann in the 1950s (e.g., *The Day the Earth Stood Still*, Robert Wise 1951). The second is associated with the lush scores and full symphonic orchestra with “heavy brass” as introduced by John Williams in *Star Wars* (George Lucas 1977).<sup>28</sup>

### **Karlin’s *Westworld*, The Movie**

For *Westworld* (1973), Karlin used the elements of the Western and science fiction traditions available to him, without conventions later introduced by Morricone and Williams. Table 1 lists all scenes with music. Gray shading identifies scenes with obvious electronic music; they typically accompany situations involving robots, particularly The Gunslinger. Early in the film, guests Peter and John, having arrived at the park, change into Western clothing, while non-diegetic country hoedown music featuring banjo, guitar, bass, and fiddles confirms the Western context. A trip to the town’s saloon adds harmonica and honky-tonk piano, further confirming the Western diegesis. But the encounter with the robot Gunslinger (see Table 1, 17:12) introduces electronic keyboard and synthesizer consistent with science fiction, which dominates as the film progresses (see Table 1).

---

<sup>26</sup> Quoting Lisa M. Schmidt, “The Popular Avant-Garde: The Paradoxical Tradition of Electronic and Atonal Sounds in Sci-Fi Music Scoring,” in *Sounds of the Future: Essays on Music in Science Fiction Film*, ed. Mathew J. Barthowiak (Jefferson: McFarland & Company, Inc., 2010), 24.

<sup>27</sup> Granade, “Sonic Markers,” 9.

<sup>28</sup> See also, Philip Hayward, “Sci-Fidelity: Music, Sound and Genre History,” in *Off the Planet: Music, Sound and Science Fiction Cinema*, ed. Philip Hayward (Eastleigh, UK: John Libbey Publishing, 2004), 1-29.

< INSERT Table 1 here >

A one-minute excerpt from the written score of *Westworld* reveals how much Karlin accentuates subtle visual motions to advance the narrative. In this case, Peter has just landed in jail, having “killed” the Gunslinger for the second time. John has arranged for a girl to bring Peter a “breakfast” tray. Figure 1 shows, beat by beat, four bars of music that underscore all of the actions, some subtle: Peter gazing downward, eyes open, turning head slightly, blinking, looking down; the tray with a note; covering the tray; and a close-up of the girl. Musically this continues a bluesy, Western, syncopated rhythm orchestrated with fiddle, banjo, harmonica, xylophone, and percussion. The scoring is economical: tremolo on a dissonant chord before it resolves (adding suspense) with a shot on Peter’s open eyes; a pizzicato xylophone note accompanying the blink; a hollow percussion triplet underscoring the focus on the tray and note. As timbres, rhythms, and harmonies change, so too do contexts and meanings. In the next four bars, the girl looks up, while the sheriff sleeps on-duty in the background. Peter looks at the sheriff, takes the tray, and the girl then turns. During this sequence, the melodic motif gives way to a steady 4/4 beat of hollow percussion atop unsettled seventh chords, furthering the plot without words. The melody and rhythm say: there’s a joke here, but on whom? Peter? The audience? The sheriff? Is Peter in trouble, or is this just the fun of *Westworld*? This multi-element yet spare musical score continues to add context and suspense until the denouement: an explosion in a jail cell, enabling Peter’s jailbreak.

<INSERT Figure 1 here>

Karlin's skill can further be appreciated in an 8-second segment in which each of the three dormant worlds appear successively at the end of their overnight reset. While a different composer might think to score three short motifs uniquely representative of the West, Rome, and Medieval times, Karlin chose a single two-note motif: the opening notes of the traditional military bugle reveille (*sol-do* in the *do-re-mi-fa-sol-la-ti-do* scale, a musical interval of an ascending perfect fourth). He repeats each *sol-do* four times with decreasing loudness almost echo-like as the camera leaves one world for the next (see Figure 2a, showing this decreasing loudness pattern three times over the 8 seconds). Karlin selected an appropriate timbre for each of the worlds: banjo/Western world, flute/Roman world, and trumpet/Medieval world. The subtle difference in the sounds is shown through a spectrograph (Figure 2b). These musical choices can occupy both diegetic and non-diegetic domains; while the sounds reoccur in their respective worlds, they also signal that each of these runs on its own unique program.

<INSERT Figure 2 here>

Karlin also manipulated acoustical instruments electronically to reflect the partially-human characteristics of the robots. As he says: "I had a single violin—I played all this music incidentally to capture this horse galloping—it didn't sound like any violin, because I wanted that shrieking, primitive quality. So it's a little electronic, but manipulated from an acoustic instrument."<sup>29</sup>

Karlin's music serves its traditional roles. The piano in the saloon contributes to the diegesis, as does music on period instruments in the Medieval banquet hall scenes. The Western-style

---

<sup>29</sup> From a 1994 interview published in *Soundtrack Magazine* 13, No. 52, republished at [https://westworld.fandom.com/wiki/Westworld\\_\(1973\\_film\\_soundtrack\)](https://westworld.fandom.com/wiki/Westworld_(1973_film_soundtrack)) (accessed February 10, 2021).

country and banjo music provides context for the town of Sweetwater. Atonal music and dissonant, electronically controlled sounds code two spaces as sci-fi: the laboratory where robots are built and reconditioned as well as the master control room of *Westworld*'s three domains. The regular duple (1 2 3 4) and triple (1 2 3) beat elements combine into an unfamiliar 10/4 time pattern (1 2 3 1 2 3 1 2 3 4) to create an irregularity (in Western cultures) due to an unexpected "extra" last beat of the 4/4 (i.e., 1 2 3 4) when only the repeated 3/4 (1 2 3) is expected. This irregular halting rhythm is well suited to the chase by the evil Gunslinger robot.

### ***Westworld*, The HBO Series**

In interviews with director Nolan, supervising music editor Kaller, and producer Stephen Semel, film theorist, composer and director Kingsley Marshall reveals Djawadi had enormous resources for the TV series' score—a stark contrast to Karlin's one-man band.<sup>30</sup> Kaller was hired full-time for 18 months during the creation of the series, and composer Djawadi was also employed from the start. The producers' and practitioners' commitment to the series' score is atypical.

I'll focus on Season 1's episodes, which are most closely aligned to the *Westworld* film. One difference between Karlin's and Djawadi's approaches arises from the treatment of the piano. In Karlin's *Westworld*, the piano is a bar-room instrument requiring a human pianist. In Djawadi's *Westworld*, the piano is now a *player piano* that makes music either by a human pianist or by internal mechanics, including its piano roll. While the instrument is chronologically accurate, the player piano's repertoire of songs draws anachronistically from the late 20<sup>th</sup> century. The film's directors decided on this use of instrument and source music and relied on Djawadi for their

---

<sup>30</sup> Marshall, *Music as Source of Narrative*.

successful execution. Granade identifies the music's three purposes: "...first, it activates the audience's personal memories of the song, increasing engagement; second, it builds metadiegetic thematic connections when the music recurs as orchestrated underscore; and third, it works diegetically by controlling the Hosts."<sup>31</sup>

From a cognitive psychological perspective, a viewer's associations to music are often unique and based on experience, though these associations are also shaped by demographics and culture. Individual knowledge of music can be personal: it may be hard for a younger or older person to appreciate that a song, such as Amy Winehouse' "Back to Black" (2006), is so revered by a middle-aged demographic. In the *Westworld* TV series, sources as diverse as early 20<sup>th</sup> century classical composer Claude Debussy's *Rêverie* (1890) or *Clair de Lune* (1913), Italian opera, and Radioheads' "Fake Plastic Trees" (1995), will most likely not be familiar to all audiences, and prior familiarity won't much hinder an emotionally rich filmic experience.<sup>32</sup> Familiarity with music, however, can occasionally detract from its effectiveness, as older associations complicate the film's new context. Films that receive awards for best song seldom win awards for best film or best director. It seems a song that stands out on its own may disengage the audience from the story.<sup>33</sup> A positive relation is found, however, between awards for film *score* and film direction or best film, suggesting that In the *Westworld* series, pop songs like "Back to Black" that lack words and are reworked for player piano, with or without full orchestration, have different resonance. Here, prior association to known songs may add to emotional impact, via an

---

<sup>31</sup> Granade, "Sonic Markers," 26-27.

<sup>32</sup> Ann-Kristin Herget, "Well-Known and Unknown Music as an Emotionalizing Carrier of Meaning in Film," *Media Psychology* (2020):1-28.

<sup>33</sup> Dean Keith Simonton, "Film Music: Are Award-Winning Scores and Songs Heard in Successful Motion Pictures?," *Psychology of Aesthetics, Creativity, and the Arts* 1, no. 2 (2007): 53-60.

unconscious level. So believes director Nolan, who is credited with choosing the hits Djawadi rendered for player piano.

Granade also states that the player piano establishes context and supports the narrative through repetition. When the plot moves in unexpected directions, the music provides a familiar foundation. This music may also provide clues about Westworld inhabitants and its puzzles. For example, Marshall notes that unheard lyrics of “Back to Black” are “relevant to Maeve’s awakening”.<sup>34</sup> The song reflects that the Hosts have “‘died a hundred times’ over the course of their narrative, but lack access to these experiences.” Marshall notes the *Clair de Lune* cue brings a further extra-textual connection to the Paul Verlaine poem that inspired the piece, which refers to Italian masked theatre.<sup>35</sup> The piece makes its first appearance in episode 5, Season 1, in a guitar version when Dolores and William are walking through Pariah, and then performed by Dr. Ford in the bar visited by William and Teddy. “Back to Black”’s unvoiced lyrics are probably more available to most viewers than Verlaine’s poem and may be more effective, but research on the roles of unheard text and lyrics would be needed to determine these relations.

Granade’s third use is unique to the *Westworld* TV series: music’s control over the Hosts. Music’s control of android behavior might strike the audience as science-fictional, but in truth, music controls human behavior in the quotidian. Music in relation to humans’ daily life is paradoxically one of both control and freedom. On the freedom side, music enables social activities that may otherwise be difficult. Love songs enable the expression of affection for someone without having

---

<sup>34</sup> Marshall, *Music as Source of Narrative*, 110.

<sup>35</sup> With (translated) lines “Peopled with maskers delicate and dim, that play on lutes and dance and have an air, of being sad in their fantastic trim,” Marshall, *Music as Source of Narrative*, 111.

to use words. Dance music enables people to come in closer proximity than may otherwise accord with social norms. Music may activate a part of the brain engaged in emotional rather than logical interpretations of the world. Music plays a role in emotional regulation<sup>36</sup> and coping in adolescence.<sup>37</sup> Music can provide energy and change mood. Music's centrality can be discerned in patients with Parkinson's disease: in the presence of rhythmic music, many regain a more normal walking stride.<sup>38</sup> Music is used by athletes to reduce perceived exertion and increase output.<sup>39</sup>

Given music's control over human behavior, the *Westworld* soundtrack's control of the Hosts by music might be regarded as a caricature of the power of music on humans. Viewing an everyday dance party from above would show that activity starts and stops in accordance with the starts and stops of the music. Yet, when music controls the androids in *Westworld*, it seems 'out of this world.' *Westworld*'s piano music additionally controls the robots and viewers on two levels of activity—the robots in the diegesis and the audience's attention.

Whereas Djawadi never disappoints with his original, catchy melodies (e.g., the infectious "Sweetwater" theme) and periodic lush orchestration, it is the piano that attracts our attention, beginning with the opening credits and extending through most of Season 1. When the hosts

---

<sup>36</sup> Margarida Baltazar and Suvi Saarikallio, "Toward a Better Understanding and Conceptualization of Affect Self-Regulation through Music: A Critical, Integrative Literature Review," *Psychology of Music* 44, no. 6 (2016): 1500-1521.

<sup>37</sup> Dave Miranda, "A Review of Research on Music and Coping in Adolescence," *Psychomusicology: Music, Mind, and Brain* 29, no. 1 (2019): 1-9.

<sup>38</sup> Miek J de Dreu, Gert Kwakkel, and Erwin E. H. van Wegen, "Rhythmic Auditory Stimulation (RAS) in Gait Rehabilitation for Patients with Parkinson's Disease: a Research Perspective," in *The Oxford Handbook of Neurologic Music Therapy*, ed. Michael H. Thaut and Volker Hoemberg (Oxford: Oxford University Press, 2014), 69-93.

<sup>39</sup> Costas I. Karageorghis and David-Lee Priest, "Music in the Exercise Domain: A Review and Synthesis (Part I)," *International Review of Sport and Exercise Psychology* 5, no. 1 (2012): 44-66.



become skilled at manipulating the music, they gain agency over their world. Interestingly, this dynamic extends to film directors and the popular music industry at large: power is expressed and retained through musical control.

As an example of cybermedia, *Westworld* uniquely provides a musical opportunity to explore free will, consciousness, and reality. The guests at Westworld must distinguish humans from android Hosts in a world where the consequences of human actions have neither moral nor permanent physical implications. As the Hosts acquire consciousness and free will, audiences may consider their own relation to free will, consciousness, and reality.

The music helps to convey these concepts and relationships by transporting the audience member on a non-diegetic vehicle into the diegesis, but then suddenly halting, changing direction, or metaphorically crashing that vehicle. Marshall notes that “slowing or distorting familiar cues is a device used within the series to signal an imminent disruption of the park’s narrative”.<sup>40</sup> The music generally is either diegetic (a pianist playing in a saloon) or non-diegetic, like when, after a day of bank holdups and brawls in Sweetwater, electronic, atonal music provides a background to the nightly pick-up of “dead and wounded” Hosts. Why and how non-diegetic music is understood by audiences is a topic of much interest to film-music scholars, including Claudia Gorbman<sup>41</sup>, Jeff Smith<sup>42</sup>; and Robyn Stillwell,<sup>43</sup> as well as scientists such as Siu-Lan Tan and

---

<sup>40</sup> Marshall, *Music as a Source of Narrative*, 209.

<sup>41</sup> Gorbman, *Unheard Melodies*.

<sup>42</sup> Jeff Smith, “Once More into the Breach: Interrogating Ben Winters’s Nondiegetic Fallacy,” in *Voicing the Cinema: Film Music and the Integrated Soundtrack*, ed. James Buhler and Hannah Lewis (Urbana, Illinois: University of Illinois Press, 2020), 260-277.

<sup>43</sup> Robyn Stilwell, “The Fantastical Gap Between Diegetic and Non-Diegetic,” in *Beyond the Soundtrack*, ed. Daniel Goldmark, Lawrence Kramer, and Richard Leppert (Berkeley and Los Angeles: University of California Press, 2007), 184-202.

colleagues, who have conducted related studies.<sup>44</sup> It's notable that a film director would take pains to make a scene hyperrealistic, and then add music within which it didn't belong.<sup>45</sup>

Djawadi explores the diegetic/non-diegetic divide, disrupting the audience's absorption in Westworld through music that intimates, "Something is not right." Characters and audiences also aren't hearing the same music, after all. I've described Karlin's two-note series as straddling diegetic and non-diegetic worlds. Djawadi's player piano has a similar double role. In the saloon, the pianist is part of the diegetic action; the bar room's characters are aware of the piano music only to the extent that anyone would be aware of hired musicians. Yes, as a non-diegetic element, the visual image of the player piano roll also signals scene changes—or, as Marshall describes, a restarting of the programmed behavior loop.<sup>46</sup> But the player piano signals something deeper, as a mechanical system capable of executing one of the most human of tasks, the performance of complex piano music. Music, like language, distinguishes humans from other living creatures. And yet a mechanical device like the player piano can reproduce and sometimes even exceed human musical virtuosity. Moreover, the piano, and the classical music it affords, has always been a symbol of high society and sophistication. This is represented in the TV series by the music of Debussy. Yet the piano also has served as "cultural go-between, as a medium through which social spheres that stood in opposition to each other could nonetheless nourish

---

<sup>44</sup> Siu-Lan Tan, Matthew P. Spackman, and Elizabeth M. Wakefield, "The Effects of Diegetic and Nondiegetic Music on Viewers' Interpretations of a Film Scene," *Music Perception* 34, no. 5 (January 2017): 605-623.

<sup>45</sup> *Blazing Saddle's* appearance of the Count Basie orchestra on the desert has become a classic example. For further discussion see Jeff Smith, "Bridging the Gap: Reconsidering the Boundary between Diegetic and Non-Diegetic Music," *Music and the Moving Image* 2, no. 1 (2009): 1-25; and Annabel J. Cohen, "Resolving the Paradox of Film Music Through a Cognitive Narrative Approach to Film Comprehension," in *The Social Sciences of Cinema*, ed. James C. Kaufman and Dean Keith Simonton (Oxford, UK: Oxford University Press, 2014), 57 – 83.

<sup>46</sup> Marshall, "Music as a Source of Narrative," 105, paraphrasing music editor Kaller.

each other”.<sup>47</sup> Pianos in the early 20<sup>th</sup> century brought together people of many walks of life, for example, piano teachers going into homes of aristocrats. In the *Westworld* series the piano provides the foundation for traversing between diegetic and non-diegetic worlds, and between the experience of attention and inattention, between consciousness and unconsciousness, and free-will and determinism.

The musical elements of the opening credits (see Figure 3) present surprising shifts between non-diegetic and diegetic roles of the piano, startling the audience, and over the course of the episodes allowing for contemplation of the film’s themes such as that of freedom and control. The computer video production company, Elastic, worked with the film’s directors and composers to create “something that you’ll enjoy watching the first time, and the more you watch the show, the more you’ll see how these images reflect the story and its significance”.<sup>48</sup> Lasting one minute and 45 seconds, it first shows a giant automated device (Bar 4) producing a fiber that becomes a piano string. Then, more fiber creates the skeleton and muscles of a horse, and finally an android. The cello theme begins an ascending motif—A<sub>3</sub> C<sub>4</sub> D<sub>4</sub> E<sub>4</sub> (Bars 3-4). The piano (Bars 5-6) echoes the motif in reverse (descending order—A<sub>3</sub> E<sub>4</sub> D<sub>4</sub> C<sub>4</sub>). The cello “call” is echoed by the piano “response,” teaching the piano, in a sense. The piano, however, reverses the tones, getting it wrong. Further, piano-key hammer action does not allow continuous control over the string. Once a piano note is hit, there is no turning back—the program must run its course.

---

<sup>47</sup> James Parakilas, *Piano Roles: Three Hundred Years of Life with the Piano* (New Haven, CT: Yale University Press, 1999), 4.

<sup>48</sup> Quoting Patrick Clair the director the *Westworld* titles for all three season in an interview with Jennifer Vineyard, “The Secrets Behind *Westworld*’s Opening Title Sequence,” *Vulture*, October 5, 2016, <https://www.vulture.com/2016/10/westworld-title-sequence-secrets.html> (accessed February 9, 2021); see also Staff “Patrick Clair on ‘*Westworld*’ Title Sequences, Creativity & Robotics,” *ARTpublika Magazine*, June 1, 2020, <https://www.artpublikamag.com/post/the-beautiful-the-grotesque-emmy-winning-motion-graphics-designer-patrick-clair-talks-westworld-titl> (accessed February 9, 2021).

The bowing motion afforded by the violin family, by contrast, provides continuous control over the tone.

<Insert Figure 3 around here>

The ascending cello line repeats (Bar 7). The piano produces a jarring note ( $B_{\flat 3}$ , Bar 9) outside the key, creating a dissonant tritone (6 semitones) with the previous final note ( $E_4$ ) of the cello motif. The piano then offers a distorted reversed variation of the cello theme (Bar 9-10). With the introduction of a new surprising jaunty phrase (Bar 11), astonishingly, a skeletal right hand at a piano keyboard appears performing the music, transporting the piano from the non-diegetic world into the diegesis. A shot with *both* skeletal hands follows (Bar 13) repeating the jaunty theme a scale-step lower. Following an interlude based on the first theme but in double time, the jaunty theme returns with more surprises—the two hands raise (or are raised) from the keyboard (Bar 21), but the keys continue to move consistent with continuing piano music. It has been stunning to see a skeletal hand created and then play a piano, but now, even more astounding, the piano plays by itself! The dual roles of the piano in the diegesis (automatic moving keys) and non-diegesis (providing emotional associations with eery, jaunty, melodic music) are clarified with a further surprise, a shot of a player-piano roll (Bar 22) with its perforated paper controlling the specific piano strings, whose location coincides with the aligned moving black holes in the paper. A shot from a longer vantage point includes the entire player piano (Bar 23).

The player piano reminds us that 21<sup>st</sup> century electronics are not needed to control a complex musical instrument that produces a beautiful musical theme. Player pianos became commercially

available at the turn of the 20<sup>th</sup> century.<sup>49</sup> The piano rolls resemble the early punched paper tape cards that operated the first digital computers. A parallel can thus be drawn between the player piano and artificial intelligence (AI), where AI refers to the carrying out of human tasks by nonhumans.<sup>50</sup>

In an interview prior to Season 3's release, Jonathan Nolan and Lisa Joy identified fashion as the most challenging aspect of representing what the forthcoming series' world would be like in 50 years. Although music wasn't what came to his mind, it is at least equally challenging to think how music, particularly popular music, would change over the next decades. The half-hour interview included background music from *Westworld* and also showed Joy seated at an electric piano performing Debussy's technically and conceptually challenging *Clair de Lune*. The use of the player piano in the *Westworld* TV series draws attention to the question of human creativity and free will. The 1973 *Westworld* film missed this opportunity, and not because the player piano had not yet been invented. The idea was cleverly picked up by the creators of the 21<sup>st</sup> century series, and added to the depth of cognitive processing in which audiences may engage, and consequently to the meaningfulness of the experience of this modern sci-fi Western.

### ***A Cognitive Model: CAM-WN***

I propose the Congruence Associationist Model with Working Narrative (CAM-WN, see Figure 4) to account for several aspects of how music contributes to a viewer's experience of film and

---

<sup>49</sup> Timothy Dean Taylor, *Music and Capitalism: A History of the Present* (Chicago: The University of Chicago Press, 2016).

<sup>50</sup> Kurt Vonnegut's (1952) book *Player Piano* is a story about workers displaced by technological advance. In one chapter, two persons who have lost their jobs due to automation converse in a bar, that has a player piano, and they disparage the fact that this device has taken a job away from a human pianist. Jonathan Nolan claims that the use of the player piano in *Westworld* pays homage to Vonnegut and this idea.

video.<sup>51</sup> It offers a perspective on the diegetic/non-diegetic distinction and emphasizes commonalities in the mental processing of music and other types of media (i.e., visual, text, speech, sound effects).

The model first assumes that a film audience member wishes to become absorbed in a story, and in a way similar to reading a novel or newspaper. Here, the viewer relies on two sources of information. One is the moment-by-moment unfolding of audiovisual content—sights and sounds that impinge on the human sensorium. At Level A, in the model (see Figure 4), these packages of physical energy are represented as “surfaces.” Sound is divided into three types of acoustical information (speech, music, and sound effects) and light is divided into scenes and text. These categories are in accordance with five channels identified by film theorist Christian Metz.<sup>52</sup> Kinesthetic information is included as an additional surface that accommodates stimulation arising from recently automated theater seats (e.g., D-Box, MX4D), which immerse viewers through movements forward, backward, and sideways. Kinesthetics also account for more subtle embodied aspects of perception, such as sensations associated with motion and pressure that might correlate with the viewer’s perception of filmic action.<sup>53</sup>

At the top of the diagram, Level E describes the viewer's higher-level understanding of sensory input. This level can detect when notes are in or out of tune, what sentences mean, and events, like what falling down stairs might sound like. Supporting this is knowledge of the story’s,

---

<sup>51</sup> Annabel J. Cohen, “Film Music and the Unfolding Narrative,” in *Language, Music, and the Brain: a Mysterious Relationship*, ed. Michael A. Arbib (Cambridge, MA: The MIT Press, 2013), 173-201.

<sup>52</sup> Christian Metz and Michael Taylor, *Film Language: A Semiotics of the Cinema* (Chicago, Ill: Univ. of Chicago Press, 2007). Originally translated from the French by Michael Taylor in 1974.

<sup>53</sup> David A. Rosenbaum, *Human Motor Control*, 2nd ed. (Amsterdam, NE: Academic Press is an imprint of Elsevier, 2016), 129-130.

music's, and language's grammar,<sup>54</sup> as well as its sequential action.<sup>55</sup> The model (which necessarily leaves out content) describes only story grammar, but can be taken for representing grammars in general.

It is assumed that a viewer's coding of Level A's information is carried out early on to Level B for all sensory surfaces. This coding at Level B is of two types. The first is structural (e.g., patterning in time, or configural such as smooth of jagged visual edges) and the second is semantic or meaning (e.g., positive/negative valence, or item recognition). The model divides Level B's "Structure" and "Meaning" through diamond-shapes representing these two analyses of Level A's sensory surfaces. It's assumed that text, speech, visual scenes and actions, music, sound effects, and kinesthetic information all have both meaning and configural patterns.

As recently revealed by Gao et al., a brain network determines the congruence of meaning coming from different sources of information.<sup>56</sup> There may also be such a network for determining the degree of multimodal configural congruence. As Bregman has identified in his notion of Auditory Scene Analysis,<sup>57</sup> similar motion patterns from two independent streams suggest a single source of the two streams (e.g., increasing pitch and decreasing loudness in the auditory realm, deepening color and increasing size in the visual realm, moving lips and changing speech sounds as a multimodal example). CAM-WN assumes that in the face of limited

---

<sup>54</sup>Siu-Lan Tan, "Scene and Heard: The Role of Music in Shaping Interpretation of Film," in *The Routledge Companion to Music Cognition*, ed. Richard Ashley and Renee Timmers (London: Routledge, Taylor & Francis Group, 2019), 364-376, 372.

<sup>55</sup> Jeffrey M. Zacks, "Event Perception and Memory," *Annual Review of Psychology* 71, no. 1 (April 2020): 165-191.

<sup>56</sup> Chuanji Gao et al., "An fMRI Study of Affective Congruence across Visual and Auditory Modalities," *Journal of Cognitive Neuroscience* 32, no. 7 (2020): 1251-1262.

<sup>57</sup> Albert S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound* (Cambridge, MA: MIT Press, 2006).

attention, the brain focuses on congruencies within and across different surface origins as the source of meaningful actions. At Level B, an arrow points from Music structure to Visual structure, reflecting heightened significance attached to cross-modal (here music- and visual structural) congruencies. For example, a galloping rhythm will direct attention to a galloping horse rather than a horse that is walking or trotting.

Thus, at Level B, decoding of the multisensory surfaces from Level A takes place along two lines: configural structure and semantic or associationist meaning. Configural structure refers to such things as temporal patterning (e.g., the motion pattern of a galloping horse, the intensity or timing patterns characterizing musical rhythms; or on a longer time span, the boundary between the beginning of one event and the start of another). It also can identify spatial relations (e.g., a crash heard from the left-hand side of the room). Analysis at Level B produces a pool of semantic and structural information. An assumption of the CAM-WN models is that the brain does not care where the information is coming from, at least for some aspects, and instead is interested in redundancies. If auditory and visual temporal patterns align, then this visual information will have priority over visual information that is not aligned with the auditory pattern.

A further feature of CAM-WN is that throughout the system of levels, some coding is quick and coarse-grained, and other coding is slower and refined. The fast and coarse/ slow and refined concept is consistent with various theories of perception, memory, and attention<sup>58</sup>. In CAM-WN, the output of quick coding ascends to the higher brain centers (dotted vertical arrows beginning

---

<sup>58</sup> Christoph Herrmann, Mattias H. J. Munk, and Andreas K. Engle, "Cognitive Functions of Gamma-Band Activity: Memory Match and Utilization," *Trends in Cognitive Sciences* 8, no. 8 (2004):347-55.



at A), to cue the cognitive processes (at E) in their task of deciphering a story underlying the multisensory barrage of cues from A and B. Here the brain can be understood as an inference machine, sometimes referred to as a Helmholtz machine (Clark, 2013)<sup>59</sup> in recognition of Helmholtz' concept of 'unconscious inference' that is fundamental to everyday perception. The idea is that we see what we expect to based on experience.<sup>60</sup>

Two types of processes now co-occur. One is the continuing bottom-up analysis of the multisensory barrage (Levels A to B) and the other is the top-down inferencing (Level D) based on weak cues that have leaked quickly through from bottom-up fast encoding. This inferencing is based on experience and story, music, language, and sequential-behavior grammars. To achieve the creation of the best story, the brain makes the best match between the bottom-up fine-grained analysis of the multisensory information and the top-down inferences. The matching process takes place at Level C. Because of the dynamic nature of this ongoing process, from the opening to the final credits of the film, this best match is referred to as the *working narrative*. In other words, the experience of *Westworld* is the audience member's working narrative.

How can this model be helpful? As an example, consider a scene with the player piano projecting the famous popular *Back to Black* by Amy Winehouse. If the audience member has never heard the song before, through processing at Level E, the piece may sound like slightly modernized traditional honkytonk music that might have been played in an early 20<sup>th</sup> century

---

<sup>59</sup> Andy Clark, "Whatever next? Predictive Brains, Situated Agents, and the Future of Cognitive Science: Behavioral and Brain Sciences," 36, no. 3, (June 2013):181 – 204.

<sup>60</sup> The role of experience in the interpretation of media is also highlighted in Chapter X by Noah Fram, "Expecting the twist: How media navigate the intersections among multiple sources of prior knowledge," in *Cybermedia: Explorations in science, sound, and vision*, ed. Selmin Kara, Jonathan Leal, Holly Rogers, and Carol Vernallis (London: Bloomsbury Press, 2021).

Western setting. If the audience member is a fan of the song, they may feel connected to the scene, and open to overlooking the mismatch of a modern piece played in a 20<sup>th</sup> century saloon. They may even remember unheard lyrics that reinforce *Westworld*'s themes. These individual differences will lead to slightly different working narratives or experiences of the episode (Level C). One's accuracy of sensory encoding and prior knowledge will limit or enhance one's working narrative.

As another example of the use of CAM-WN, consider how diegetic awareness is explained by the top-down matching process. The best match between sensory analysis (outcomes of Levels A, B) and top-down inference regarding the story (Levels E, D) wins at Level C. Once the story diegesis is established, some information will enter consciousness (Level C) and other information will not. In the opening credits, attention is directed to the creations of the piano, horse, and android. The non-diegetic music offers acoustical, music-grammatical, and emotional information. The music begins with a relatively low pitch, establishing an A minor tonality; the B<sub>♭</sub> at bar 9 violates the tonality. Its underlying ostinato tempo is a relentless four sets of triplets per bar, with a melody at bar 12 that doubles the meter to 24 16<sup>th</sup> beats per bar. The unsettling emotion arising from this combination of features finds its explanation (Eldar et al., 2007) in the visual objects depicted. Initially, the acoustical information (Level B) finds no match from Level D. There is no reason to be visualizing a cello or a piano. The acoustical information therefore does not form part of the initial diegesis. It is outside the working narrative.... that is, until we see the skeletal hand at the piano. Now the top-down inferences (Level D), based on the image of a hand playing the keyboard, can match the acoustical information arising from Level A to B and the match between the image of the piano and the sound of the piano leads to consciousness of

the piano sound, at Level C, the working narrative. At level B, temporal congruency of the visual hand motion and fingers moving on the keys and the auditory pattern of the melody direct attention to the hand, making it prominent in consciousness. The meaning of the music (relentless, unsettling, ominous) picked up from the semantic analysis at B and matched by meanings inferred from the higher level at D also match, bringing into consciousness an emotionally charged engagement for this *Westworld*'s scene. The sudden switching from the non-diegesis to diegesis in the opening credits is but one example of how the music is used throughout the film, and this use, assisted by the exploitation of 21<sup>st</sup> century computer-based editing studios coupled with the imaginative creative team offer more sophisticated and nuanced play with the audience's mind. The audience experiences a moving in and out of the diegesis that's more striking than what's offered from the *Westworld* film, produced over four decades earlier.

The CAM-WN model reminds us that the brain can engage with an analysis that counters our intuitions. For example, it might seem obvious that music has an acoustic element, but the brain might separately 'perceive' an acoustic element and emotional element and encode music as such, using only what it needs to make sense of the story. The brain can discard the acoustic element if there is no reason for a musical instrument to be in the scene, while at the same time preserving the emotion carried by the music, which belongs there. This seems consistent with Smith's (2020) discussion of the diegetic/non-diegetic distinction from the perspective of the "...spectator's narrative comprehension. The crux of the issue is not the kinds of departures from everyday experience of which film-worlds are capable, but rather how viewers could ever

process any story information without some tacit application of real-world knowledge.”<sup>61</sup> Here I suggest that “story information” is Level A of the CAM-WN model and “real-world knowledge” is Level E. Smith continues, noting film perception’s reliance on normal cognitive processing, including heuristics and then proposes that the film score should be considered alongside film’s other formal systems, where functions carried out by music can be carried out by other means (and vice versa). His idea and the CAM-WN model both concur about the common coding principle (for meaning and structure) of all filmic categories and prediction processes. Finally, Smith recognizes that norms associated with sound design enable viewers to make inferences about the sound’s relation to the narrative and, as well, that departures from these norms can disrupt “normal conceptual schemas.”<sup>62</sup> These features are compatible with the CAM-WN model, as exemplified by the admission of the piano sound into the Westworld working narrative.

Scientists are currently delineating the neural networks underlying various sensory, perceptual, and cognitive processes. Consistent with my approach, research on event perception identifies the human sensitivity to common event structures. One line of research examines human segmentation of events, often using video representations, as reviewed by Zacks (2020)<sup>63</sup> although boundary formation in melody has also been explored in segmentation tasks.<sup>64</sup> Further studies with eye-tracking, neuroimaging, and computational modeling may soon tell us more. Interdisciplinary research teams should also help us determine what brain regions are involved

---

<sup>61</sup> Smith, *Once More into the Breach*, 274.

<sup>62</sup> Smith, *Once More into the Breach*, 276.

<sup>63</sup> Zacks, *Event Perception*.

<sup>64</sup> Bradley Frankland and Annabel J. Cohen, “Parsing of Melody: Quantification and Testing of the Local Grouping Rules of Lerdahl and Jackendoff’s A Generative Theory of Tonal Music,” *Music Perception* 21, no.4 (2004): 499-543.

with processing complex multimedia.<sup>65</sup> The imaginative as well as the traditional uses of music in *Westworld* offer fertile ground for empirical research.

## **Summary and Conclusion**

Drawing from cognitive psychology and neuroscience, this chapter explores two major film composers scoring of similar cybermedia themes. Common functions of music within science fiction and western genres were evident, despite passages of time and different formats.

However, psychological research suggested that in the 21<sup>st</sup> century, directors and the film-maker teams may be getting smarter about holding audience attention. Newer music technologies, such as better computer music systems, and clever new uses of traditional instruments (such as *Westworld*'s player piano) have expanded narrative depth and links across non-linear plots. The Congruence-Association Model with Working Narrative (CAM-WN) was discussed as an aid to understanding the brain's multiple tasks of integrating audio and visual tracks of a film, with the conclusion that the experience of a film or TV drama might be understood as a working narrative created from the match between bottom-up multisensory analytical processes and the top-down inferences based on remembered experience and the grammars of story, music, and language. The bottom-up/top-down matching process explains why the audience typically does not as much 'hear' but nevertheless feels the music of film soundtracks, as this acoustical information, unlike the emotion of the music, is not matched by the top-down inferences.

The 21<sup>st</sup> century demands more from film makers than ever before, given the many kinds of media

---

<sup>65</sup> Siu-Lan Tan et al., "Future Directions for Music and Sound in Multimedia," in *The Psychology of Music in Multimedia*, ed. Siu-Lan Tan et al. (Oxford, United Kingdom: Oxford University Press, 2013), 391-405. See also, Annabel J. Cohen, "Congruence-Association Model and Experiments in Film Music: Toward Interdisciplinary Collaboration," *Music and the Moving Image* 8, no. 2 (2015): 5-24.

that compete for audiences. In winning audience attention to contemporary cybermedia, Jonathan Nolan and Lisa Joy, with the help of their creative team, have exploited the player piano as a vehicle to transport the audience in and out of the narrative, simultaneously heightening their consciousness of their own consciousness. Both the *Westworld* movie and series use music in traditional ways to transport viewers into a fictional world and establish its context; however, the *Westworld* TV series uniquely exploits music as a means of exploring the concept of control of the androids, by the androids, and of the audience. The chapter lays a foundation for further studies of these musical phenomena in cybermedia.