Sound & Vibration 2026, 60(6), 4715.
doi: https://doi.org/10.59400/sv4715

Review

Acoustic homogenization of background sound effects and user auditory fatigue on short video platforms: A conceptual review of mechanisms and evidence

Qiming Wang

School of Fine Arts, Anshan Normal University, Anshan 114000, China; 13804223445@163.com

CITATION

Wang Q. Acoustic homogenization of background sound effects and user auditory fatigue on short video platforms: A conceptual review of mechanisms and evidence. Sound & Vibration. 2026; 60(6): 4715. https://doi.org/10.59400/sv4715

ARTICLE INFO

Received: 24 July 2026
Revised: 21 August 2026
Accepted: 26 August 2026
Available online: 7 September 2026

Abstract: Short video platforms have quickly emerged as a source of digital media consumption, which is overwhelmingly researched in terms of visual saturation, dependence on algorithms, and performance response. The auditory aspect, and, in particular, the homogenization of the background sound effects, is not researched thoroughly. This review will examine how user auditory fatigue can be caused by acoustic homogenization on short video platforms. It states that the repetition of similar background music structure, transition signals, loudness, timbral textural profiles, and emotionally coded sound patterns produces a platform-mediated auditory scenery with a lack of diversity and continued sensory demand. The article builds a conceptual model based on the insights of acoustics, psychoacoustics, media studies, human-computer interaction, and platform studies to construct a link between acoustic homogenization and auditory fatigue by proposing mechanisms, which include perceptual saturation, augmented listening effort, dynamic compression, enhanced temporal turnover, algorithmic reinforcement, and cross-modal overload. It also examines existing methods of measurement, such as acoustic feature analysis, subjective fatigue evaluation, behavioral measures, and physiological measures, and points to the weaknesses of the existing empirical findings. This review makes three key contributions: first, it introduces acoustic homogenization as a novel analytical concept in short video research; second, it provides an interdisciplinary synthesis linking platform sound ecology to auditory fatigue through multiple mechanisms; and third, it offers practical implications for creators, platform designers, and digital well-being governance.

Keywords: short video platforms; acoustic homogenization; background sound effects; auditory fatigue; digital well-being

1. Introduction

Short video platforms are one of the most influential types of consumption of digital media in the world in recent years. Their fast emergence has transformed the trends of information access, entertainment, socialization, and consumerism. In comparison with old-fashioned length-based audiovisual media, short video content is defined by discontinuous viewing time, quick switching of scenes, intensive rate of recommendation, and highly optimized multimodal stimulation [1]. Previous studies on the issue of short video platforms have thus focused primarily on such questions as engagement through algorithms, capturing visual attention, behavioral addiction, emotional contagion, and information overload. Nevertheless, although there is a rich literature on the visual and behavioral aspects of short video viewing, the soundscape of these platforms has been discussed in a significantly less systematic way. Specifically, the background sound effects as a constant and frequently standardized platform content have not been properly interpreted in terms of user well-being [2–5].

Audio is not just something short video content adds on. Background music, ambient sound layers, rhythmic beats, transition effects, interface feedback sounds, and sound cues that have been encoded as emotional have an influential role in determining the immersive experience of users. These acoustics play a role in the perceived excitement, tone of emotion, memory, and attention distribution, as well as the consumption choice [6]. The ecology of platforms makes sound even more important on platforms that feature short videos. First, videos are short and require quick emotional and attentional attention, which drives the creators to use the most salient and recognizable sound resources. Second, the content formats that have already been demonstrated as appealing are rewarded by the recommendation system and, thus, encourage people to use trending audio clips, popular background music templates, etc.sound design logic over and over again. Third, the convergence of acoustic style is further accelerated by integrated editing tools, commercial sound libraries, and viral imitation mechanisms. Consequently, users are being exposed more to a highly repetitive and standardized aural environment, regardless of the visual content seemingly being diverse. This effect can be called the acoustical homogenization of the background sound effects [7–9].

Acoustic homogenization is the growing similarity and repetition of the sound structures in the content of the platform, such as similarity of rhythm, timbre, loudness profile, melodic contour, spectral density, emotional coding, and editing patterns. On short video platforms, homogenization commonly takes the form of repetition of the same or very similar background music, repeated viral audio memes, the patterns of compressing loudness, overuse of transition effects, and predictable sound design that would maximize click-through and retention. As much as this homogenization can enhance the recognizability, efficiencies of production, and platform-level circulation of content, it can also produce unintended expenses for users. Auditory fatigue, which is a condition of burden associated with listening, is one of the possible costs, and it entails perceptual exhaustion, lesser auditory comfort, greater annoyance, mental effort, and lesser attentional tolerance following extended or frequent exposure to audible sound [10–12].

The fact that acoustic homogenization may lead to auditory fatigue is all the more important in the context of the short video. Compared to the usual audiovisual media, short video viewing is marked by endless scrolling, disrupted exposure, and quick rotation between videos [13]. Every video can only last a few seconds, but the users can have long cumulative sessions during which acoustically related stimuli appear with minimum recovery. The recurrence of alike beats, maximum loudness, and sound emotion may decrease auditory novelty and maintain sensory demand. Furthermore, in short video settings, audio is infrequently perceived alone, but it is in contact with dynamic visual content, text overlays, speech fragments, and platform notifications, all of which increase the cognitive load. In such circumstances, it is not just discrete sound events that are being processed by the auditory system of the user, but instead a highly compressed, repetitive, and algorithmically supported soundscape is being experienced.

The environment can contribute to subjective discomfort, effort in listening, irritability, and mental fatigue, although the damaging sound pressure levels in the traditional clinical sense are absent [3, 14–16].

Although this issue is relevant, existing literature is still in pieces. The nature of noise annoyance, listening effort, sound quality, and fatigue has been studied in the past in the fields of acoustics and psychoacoustics, both in a work or laboratory setting [17]. The emotional impact of music, audiovisual synchronization, and the persuasion potential of sonic branding are discussed in media studies. The problem of notification fatigue, interface sound design, and sensory overload in digital systems has been discussed in human-computer interaction research. In the meantime, the homogenizing impact of algorithms, imitation cultures, and production of content based on templates has been examined in communication and platform studies on short video ecosystems. Nevertheless, such bodies of literature seldom overlap in such a way that they can explicitly describe the impact of acoustic standardization of platform-mediated environments on user auditory fatigue. To put it another way, it does not yet have a combined framework which would be able to connect the production logic of short video audio and the structural phenomenon of acoustic homogenization, on the one hand, and the outcome of the experience of auditory fatigue on the other [18].

This paper fills that gap by suggesting a review-based study of the effects of acoustic homogenization of the background sound effects on the user's auditory fatigue on short video platforms. Instead of considering platform audio as a kind of secondary aesthetic feature, this paper makes the audio environment a central focus of investigation and considers it as a problematic constituent of digital media health and user experience. The intended purpose of the review is to explain the ways platform ecology can lead to homogenized audio, what processes can mediate the relationship between repetitive exposure to acoustics and fatigue, how fatigue can be measured, and what remedies can be applied to make auditory platforms healthier for users. Emphasizing that the background sound effects take precedence over the foreground speech or music alone, the article also indicates the fact that the sonic layer is frequently overlooked, and it does influence perception at all times, being a continuous element under the surface of overt content [19–21].

This article is innovative in a number of ways. To begin with, it presents the concept of acoustic homogenization as a new analytical term in studies of short videos. The available literature tends to treat content homogenization either through visual style, recommendation results, replication of memes, or cultural convergence, but seldom systematically isolates the auditory aspect. This paper follows the debate on the homogenization of platforms into sonic space and claims that background sound effects represent one of the most significant locations where the process of standardization can be measured and experienced in an actual way [22].

Second, this review provides a clear relationship between platform sound ecology and auditory fatigue, a relationship that has yet to be developed in communication research, as well as in the study of auditory. Auditory fatigue is usually studied in connection with occupational noises, extended listening, or speech intelligibility activities, but short video situations are a qualitatively different environment that involves intermittent and continuous micro-exposures, repetition of emotional cues, and cross-modal overload. Placing auditory fatigue into the context of using the media of short video platforms, the article expands the conventional range of studies on fatigue and indicates that the digital entertainment setting may also serve as a significant source of auditory stress.

Third, the article uses an interdisciplinary synthesis, encompassing acoustics, psychology, media studies, platform studies, and human-computer interaction. It is required to integrate it due to the inability of a particular discipline to be able to explain the phenomenon completely. Acoustic science assists in the determination of measurable characteristics, i.e., loudness compression, spectral concentration, and repetition rate. Habituation, depletion of resources, and the development of irritations are explained by cognitive psychology. The media and platform studies enlighten the industrial and algorithmic conditions, which promote sound convergence. Human-computer interaction provides insights into user comfort, interface design, and sustainable digital experience. The review provides a wider conceptualization than is present in the literature because it pulls these strands into one another [23].

Fourth, the article highlights the application of fatigue-conscious audio design to creators, platforms, and policymakers. Although acoustic standardization can be commercially effective, it can also compromise user comfort in the long run and platform sustainability. Greater knowledge of auditory fatigue can, therefore, be used to inform sound recommendation diversity, creator toolkit development, control strategies in loudness control, and digital well-being regulation. This changes the focus of the platform engagement metrics towards a more user-oriented and health-based assessment of platform audio environments [24].

Overall, the article addresses a growing yet under-researched issue in the digital media context: the likelihood that the homogenization of background music-like sound effects on short video websites is a source of auditory fatigue in users. Since the short video is still one of the focal points in daily life, it becomes more and more critical to understand what users watch, as well as what they listen to over and over again. This review seeks to establish a platform on which future studies can be conducted to understand the healthier auditory ecologies in media platforms through the definition of the phenomenon, synthesizing relevant mechanisms, identifying the empirical approaches, and emphasizing its novelty in the literature.

This review employed a systematic search strategy to identify relevant literature across multiple disciplines. The following databases were searched: Web of Science, Scopus, PubMed, Google Scholar, ACM Digital Library, and Communication & Mass Media Complete. The search was conducted between January and March 2026 using the following search terms and Boolean combinations: ('acoustic homogenization' OR ’sound standardization' OR 'audio repetition' OR ’sonic convergence') AND (’short video' OR ’short-form video' OR 'TikTok' OR 'Reels' OR 'Shorts') AND ('auditory fatigue' OR 'listening effort' OR ’sensory overload' OR 'auditory discomfort' OR 'digital fatigue') AND ('background sound' OR 'background music' OR ’sound effect' OR ’sonic environment' OR ’soundscape'). Reference lists of identified articles were hand-searched for additional relevant sources. Studies were included if they: (a) addressed sound characteristics in digital media contexts; (b) examined auditory fatigue, listening effort, or sensory overload; (c) discussed platform-native audio design; or (d) provided empirical or theoretical insights into repetitive sound exposure effects. Exclusion criteria included studies focused solely on clinical hearing loss, occupational noise exposure exceeding safety limits, or musical performance without media context. From an initial pool of 847 articles, 94 were selected for full-text review based on title and abstract screening, with 48 ultimately included in the synthesis based on relevance and quality.

2. Conceptual foundations and theoretical framework

2.1. Acoustic homogenization as a platform-mediated auditory phenomenon

In short video platforms, acoustic homogenization is the tendency of the background sound effects to gradually converge into incredibly close sonic structures, affective functions, and production logics [25]. Homogenization in this context does not imply the reiteration of the same sounds that are reused. Instead, it refers to a more general decrease in acoustic diversity, in which various videos slowly turn into using similar rhythmic patterns, profiles of timbre, loudness curves, transition indicators, and templates of sound coded with emotions. The user can also have a restricted and repetitive auditory ecology even in cases where the audio files themselves are not the same at all, since the sonic grammar of the underlying grammar is quite uniform [10,12].

This process cannot be imagined without the industrial and algorithmic arrangement of the production of short videos. Most sites have in-built editing software, sound libraries, trend-following software, and circulation driven by imitation, and are therefore hastening the creation of content. These affordances celebrate audio choices that are instantly recognizable, emotionally efficient, and confirmed to be able to maintain engagement. This leads to the sound design being tailored to the environment of quick attention capture in a very competitive feed environment, rather than being acoustically rich and well-balanced. It is the constant repetition of similar background music, the punchy transitions, the more powerful beats, and the textures of sound that have been compressed and are not incidental to its aesthetic choice, but rather reinforced by structural logic on platforms [26].

One of the critical points is that homogenization cannot be perceived as an issue of the creator’s behavior only. It is also a recommended product of the infrastructures that magnify the already successful forms of sonics. In situations where the algorithmic systems favor content related to popular sound templates, there is an incentive to encourage creators to recreate the same acoustic patterns. In this regard, platform design, creator strategy, and user feedback loops co-produce homogenization. The sound world that users live in is not thus an impartial totality of personal decisions as much as a platform-conditioned soundscape that is mediated by the efficiency of circulation and data-based visibility [27].

Conceptually, acoustic homogenization is different in two significant respects compared to audio repetition, in general. To begin with, it is systemic and not discrete. A user cannot hear the same sound repeated many times in immediate succession, and yet he or she remains in a sonic field, which consists of structurally identical cues. Second, it has experiential consequences which are more than familiarity. Background effect perceptions may be less perceptually new, less emotionally varied, and cumulative listening loads may be greater with repeated exposure to acoustically similar effects. Therefore, acoustic homogenization is not merely a descriptive characteristic of the platform content but a possibly consequential phenomenon of user sensorial experience.

2.2. Background sound effects in short video platforms: functions and transformations

The background sound effects take a unique place in short video media since they are used as both an aesthetic enricher, emotional controller, attentional keystone, and platform-native style marker. The background sound effects do not always work as strongly as the foreground sound effects, unlike foreground speech, which mainly expresses semantic content, or foreground music, which may take over the expressive frame. They are used to influence mood, pace, suspense, transition, and perceived intensity that need not necessarily get into conscious focus. This is because of its low-threshold but continuous presence, which renders them particularly significant in auditory exposure of short video environments [2,28].

Background sound can be used in classic forms of audio-visual to either aid in the continuity of a narrative or the realism of the environment [29]. In short video platforms, the role changes with the compressed time structure of content. Videos shorter than several seconds have to be emotionally oriented within a short time. Background sounds are therefore usually meant to be rich in salience, instantaneously readable, and closely coupled with visual information. Rhythmic pulses are used to indicate excitement, soft ambient beds are used to indicate intimacy, exaggerated clicks or swooshes provide momentum, and brief loops of music make otherwise disjointed material stable. These elements of sound can be used to balance out the limited length of the medium by speeding up the process of emotional inference [30].

But the efficiency of this design is the cause of convergence. Because only specific patterns of sounds are always useful in severe time compression conditions, producers and media are more inclined to use a relatively small collection of tested audio structures. This, in the long run, makes the scope of background sonic expression smaller. What seems to be the artistic profusion at the stage of content types might thus conceal a still smaller acoustic arrangement at the stage of aural composition. That is, the short video platforms may appear heterogeneous but sound more and more alike.

This brings up a key problem of analysis: it is at this point that the background sound effects are neglected simply because they are below the scope of direct attention. A catchy tune or a verbal utterance might be apparent to users; however, repetitive sounds of transition, compression of ambience, and formulaic expressions of emotion are more likely to be unconscious in the feed. However, these induced layers build up during extended viewing periods and can have a stronger influence on fatigue than single foreground audio events. They exert their influence not in a spectacular but in a persistent manner, and this persistence makes them the key point of any serious study of auditory burden in short video setups [31].

2.3. Auditory fatigue: From physiological concept to digital media experience

Auditory fatigue has been studied in the past in the context of long-term listening load, work in noisy environments, speech perception challenges, and unfavorable acoustic conditions. In traditional auditory studies, the notion is often used in reference to temporarily decreased auditory sensitivity , added effort in listening activities, or perception of fatigue, pain, and irritation following prolonged sound exposure. Nonetheless, auditory fatigue should be conceptualized in a broader way when it is applied to short video platforms. It cannot be reduced to changes in hearing thresholds or medical sound damage. Instead, it is a multidimensional condition consisting of perceptual exhaustion, cognitive exhaustion, emotional irritation, and a lesser ability to hear more [15].

This larger concept is required due to the fact that the digital media environment offers a different form of listening condition. Consumers on short video websites do not have to be subjected to long continuous amounts of extreme sound. Rather, they are exposed to a series of brief and highly stimulating clips that are acoustically similar most of the time. The tiring effect of this type of environment is less one of isolated intensity than a micro-aggregation. Repeats, sharp turns, compression of the dynamics, and affective overcoming can all enhance the load of auditory processing, even though individual sounds are not considered harmful in isolation.

There should therefore be a critical difference between acoustic harm and auditory fatigue. The latter is usually estimated in objective terms of sound pressure excess or physiological harm. The latter consists of subjective and cognitive aspects that can manifest themselves far below detrimental volume levels. The user can complain of irritability, mental fatigue or the need to cease listening despite the moderately low volume of playback. This means that auditory fatigue in a platform environment should be addressed as both an auditory problem in the limited physiological definition, but also as an experience of sensory design [32].

It is important to note that sound perception in digital media contexts is inherently linked to vibration phenomena, particularly in mobile device usage where speaker-induced vibrations contribute to tactile and auditory experiences. Research on vibration characteristics of structural systems, thermal contact mechanics , and surface roughness interactions. These considerations are relevant to short video platforms where device-based audio delivery involves complex vibration transmission pathways that may influence user comfort and fatigue.

2.4. Theoretical perspectives for explaining the relationship between homogenization and fatigue

Several theoretical approaches can be used to understand why acoustic homogenization can augment auditory fatigue in short video channels [3]. An applicable framework is that of auditory load theory, which implies that listening is more effortful in thick, monotonous, or constantly attention-demanding sound environments. The situation with the short video is that they will be required to decode fast audio alterations repeatedly, although they will be in the presence of structurally similar sonic stimuli. This produces an odd mixture of newness at the clip succession and staleness at the acoustic patterning. The outcome is no restful familiarity but continued low-level processing demand [33].

Habituation theory also provides insightful information, though it has to be applied with caution. With repeated exposure to the same stimulus, there may be decreased responsiveness over time, and this is an indication that the users may just get used to the effects of the backgrounds that occur repeatedly. Nevertheless, platform audio is not commonly repeated in an entirely stable form. Rather, it reappears with minor variations in the ever-varying videos. This creates a paradoxical state where the auditory system acknowledges the familiar structures and is unable to disengage completely due to the fact that every clip has small variations and new audiovisual surroundings. Habituation can therefore be partial, and attentional strain can be sustained in the presence of attentional adaptation [34–36].

Another dimension is cognitive resource theory, which focuses on the fact that attention and processing have a limited capacity. Short video settings are subject to repetitive actions of orienting, filtering, and emotional valuation. Background sound effects, particularly when they are exaggerated or formulaic, can use up cognitive resources to indicate a shift, enhance affect, or oppose speech and visual text. This build-up could lead to mental exhaustion over long sessions. Notably, homogenization can aggravate the situation by rendering the stimuli meaningfully less informative and keeping their demand on attention intact. When this happens, the users waste resources on sounds whose perceived or emotional value reduces as they are played.

The soundscape theory also extends the analysis by considering the auditory experience to be more of a condition of the environment and not a series of detached sound objects. In this sense, it is not simply whether certain sonic effects are repetitive or not, but whether the sonic background (in general) promotes comfort, variance, and perceptual harmony. Soundscapes created by short video platforms are often discontinuous, salient, emotionally overdetermined, and with low diversity in auditory texture. These soundscapes can destabilize the conditions of restorative listening and encourage chronic sensory irritation. The theoretical shift can be useful as it brings the platform audio study close to the broader discussions of environmental quality and human health [37–39].

Lastly, platform studies introduce a structural dimension that is required. Phantomization and algorithmic governance theories propose that repetition and standardization are not the by-products of an optimization system but instead the result of an optimization system. Acoustic homogenization in this perspective is an expected side effect of those infrastructures that are intended to optimize retention, memetic circulation, and the efficiency of the creators. The importance of this point of view is that, in this way, fatigue cannot be simply described as a psychological response on the user level. Rather, it brings out the manner in which auditory load can be integrated into the very structure of platforms.

2.5. Toward an integrated framework for this review

When combined, these conceptual and theoretical views point to the fact that the acoustical homogenization-auditory fatigue relationship should be viewed as a multilevel one. **Figure 1** illustrates the integrated conceptual framework that has been proposed in this review. On the level of content, homogenization minimizes sound diversity and strengthens repetitive sound patterns. This can be reduced at the perceptual level and in the maintenance of the processing requirement. The user can get annoyed, depleted, and have less tolerance for listening at the cognitive and affective levels. On the platform level, recommendation systems and creation tools repeat the circumstances that predispose users to such fatigue [40,41].

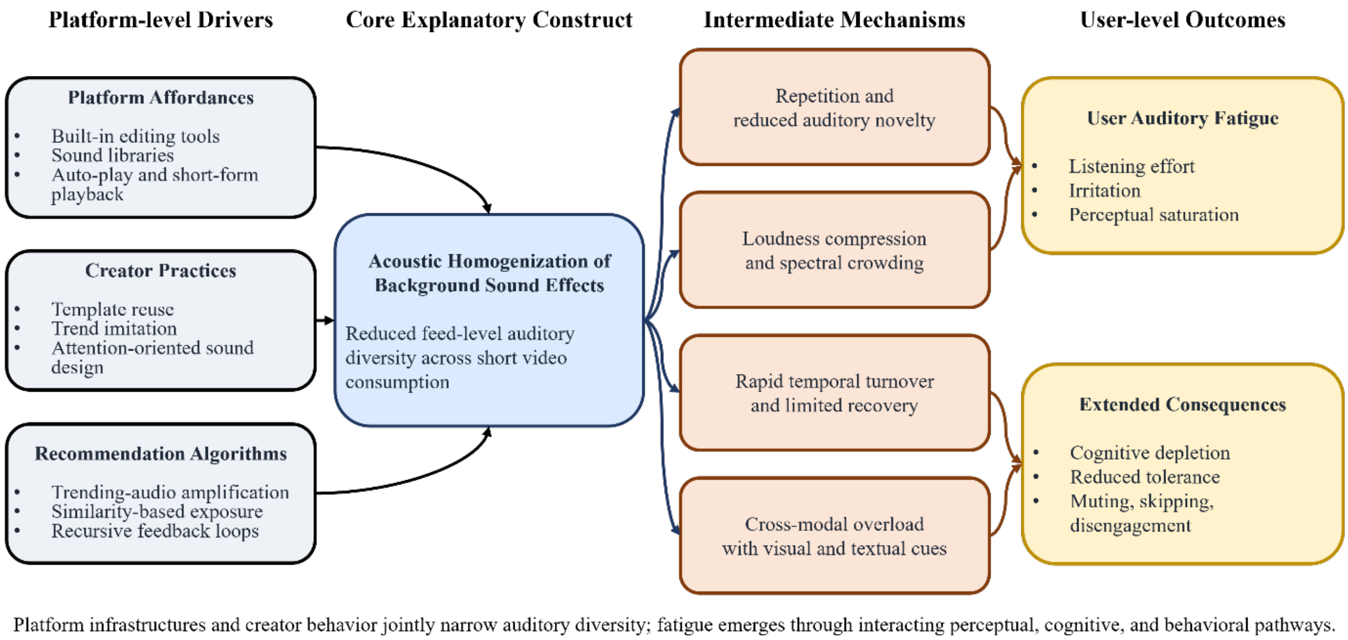

Figure 1. Proposed conceptual framework illustrating the hypothesized relationship between acoustic homogenization and auditory fatigue in short video platforms.
Note: This framework synthesizes mechanisms identified from the literature and is presented as a theoretical model requiring empirical validation.

The usefulness of this combined framework is that it unites micro-level auditory characteristics with macro-level platform frameworks. Too strong an emphasis on the parameters of sounds would be the risk of overlooking the algorithmic and cultural processes that generate homogenization. Being too broad in platform logic may lead to the overlook of the auditory fatigue sensory specificity. These subjects need both of these levels together to be analyzed using an SCI-standard. It is based on that assumption that the current review provides the view of user auditory fatigue on short video platforms, not as a simple matter of high-volume audio, but as a symptom of the new, platform-optimized, and homogenized auditory ecology. **Table 1** presents the concepts and definitions of analysis embraced in this review [42].

Table 1. Core concepts and analytical definitions in this review.

Concept	Definition in this article	Analytical significance
Acoustic homogenization	The convergence of background sound effects toward similar rhythmic, timbral, loudness, spectral, and affective patterns across short video content	Serves as the central explanatory variable for understanding repetitive auditory exposure on short video platforms
Background sound effects	Non-foreground sonic elements such as ambient pads, transition sounds, rhythmic fillers, emotional cues, and supportive audio layers accompanying short videos	Represent the primary sonic layer through which cumulative auditory exposure occurs
Auditory fatigue	A multidimensional state of perceptual weariness, listening effort, annoyance, cognitive depletion, and reduced tolerance for continued auditory stimulation	Functions as the core user outcome examined in this review
Platform auditory ecology	The total sound environment generated by platform design, creator practices, recommendation systems, and user listening conditions	Provides an ecological framework for linking sound design to user experience
Feed-level acoustic diversity	The extent of variation in sonic properties across a sequence of recommended videos during user scrolling	Helps explain cumulative exposure effects beyond single-video analysis
Algorithmic reinforcement	The repeated circulation of acoustically similar content through recommendation systems that privilege familiar and high-performing sound formats	Explains why homogenization is not random, but structurally reproduced
Cross-modal overload	The amplification of sensory burden through the interaction of repetitive sound with rapid visuals, captions, speech, and emotional cues	Clarifies why auditory fatigue on short video platforms is rarely caused by sound alone

3. Mechanisms by which acoustic homogenization influences auditory fatigue

3.1. Repetition, perceptual saturation, and the erosion of auditory novelty

Repeated exposure to sound structures that are very similar to each other in a succession of short videos is one of the most direct ways in which acoustic homogenization can generate auditory fatigue. Ideally, repetition is not necessarily negative. Sonic motifs repeated in most media situations contribute to familiarity, emotional integrity, and mnemonic remembrance. But, on short video media, repetition is seen under conditions of enhanced content turnover, discontinuous attention, and extended session time. In this case, repetitive background sound effects no longer work as integrating aesthetic devices and become perceptual saturation devices [43].

The effect of perceptual saturation occurs when the auditory system is repeatedly exposed to a type of stimulus that is similar enough to be considered familiar, but is persistent enough to be intrusive. It is not silence or a significant difference between clips that the listener is experiencing, but a tight range of repeating sound forms. These can be parallel rhythmic loops, identical transition clues, similar timbral brightness, and similar emotional sound indicators. Consequently, auditory novelty is reduced, although the need for perceptual processing does not vanish. This imbalance is essential. It is not that the users learn to be accustomed to the sound, but that they proceed to process repetitive acoustic information that is generating less and less experiential value [44].

This process applies particularly to short video ecosystems since novelty is frequently present in visual form instead of in the auditory. The user scrolls through seemingly varied content categories, artists, and stories, but the background sound patterns are extremely consistent. The discrepancy between visual change and auditory consistency can enhance fatigue because it does not enrich the listening experience. The ear is, in effect, repetitious, and the eye is changeable. The result of such divergence may be a kind of overt monotony of the ear, hidden yet growing in a supposedly vibrant media space [45].

Most importantly, repetition here must not be diminished to the literal re-use of the same soundtrack. Acoustic homogenization usually does not work through direct copying, but through structural similarity. Sounds can vary in a superficial way and still have the same functions, pacing, spectral emphasis, or emotional tone. This complicates the fatigue-producing nature of repetition since the user might not realize how certain sounds are being repeated. However, the overall outcome can also be identical, resulting in the loss of auditory tolerance, the development of greater irritation, and the longing to stop listening entirely.

3.2. Loudness compression, spectral crowding, and increased listening effort

A second key mechanism is the acoustic intensity and spectral organization of platform sound. Video audio in the form of short videos is often created to be immediately conspicuous within very short periods of time. To create this effect, platform-native editing conventions and creators tend to utilize compressed dynamics, high-average loudness, and compact spectral content. These characteristics make content more salient, but they can also increase the listening effort and reduce the listening comfort over time [46].

This is especially noteworthy for dynamic range compression. Compression enhances more evenly intense listening by making the quiet and loud passages less distinct. It is beneficial when creators want to get attention immediately, particularly on mobile devices in a noisy environment. However, an audio stream where the average intensity is continuously high does not leave a lot of room for perceiving relief. Although absolute volume may not be hazardous, the lack of dynamic contrast can create an impression of permanent listening pressure. Moments of softness, distance, and acoustic spaciousness needed to aid recovery when engaging in long-term media use are denied to listeners.

This is worsened by spectral crowding. Background sound effects on short video sites have many frequency regions that are linked to clarity, brightness, and immediacy, especially in the mid and high frequencies. This aids the sounds to break through device speakers and compete with other sound elements like speech and interface indications. Nevertheless, continued exposure to spectrally crowded audio may cause listening to be more difficult. In the case when a great number of sounds can be heard over the same frequency range, hearing discrimination is challenged, and the total sound environment can be perceived as harsh, congested, or overwhelming. This does not necessarily make one consciously uncomfortable initially, but it is a gradual wearing out of the tolerance to noise [47].

Another significant contradiction in the design of platform audio is also exposed in this mechanism. Aspects that enhance short-term attentional capture can impair the long-term sustainability of senses. Impact is optimized using loudness compression and spectral concentration, which helps compress stimulation in front of the waveform and removes acoustic nuance. This can change background sound, which is a supportive aspect, into an overextending aspect, especially after repetition. Therefore, auditory exhaustion in the short video setting can be caused not only by the amount of exposure but also by the microstructure of the sound.

This observation has a critical implication. The studies that concentrate on the acoustically homogenized design might not identify the role of the sound pressure level in the development of fatigue. Clinically, the sound is not too loud and therefore makes the users feel tired. Compression, density, and spectral distribution also determine listening effort, in addition to intensity. A more sophisticated description of auditory fatigue should therefore take into consideration how the platform's native sound design standardizes these features in a manner that exaggerates cumulative load [48,49].

3.3. Accelerated temporal turnover and the absence of auditory recovery

The use of short video platforms exposes the user to a unique time regime, with content chunks being short, with the transition between them fast, and the ability to keep scrolling and to provide the sense of uninterrupted audiovisual stimuli. This time sequence is the key to comprehending how the process of acoustic homogenization is transformed into fatigue. The users of traditional media tend to stay longer in the same coherent aural setting, which gives the sound environment time to stabilize. With short video consumption, by comparison, the user is exposed to a sequence of sonic micro-events, each of which is meant to capture attention in a brief period, although it is commonly constructed on extremely similar acoustic models. This discontinuity and similarity form a rather exhausting listening experience [50].

It is not that users hear lots of sounds over a brief duration, but that they do not have sufficient time between sounds for proper auditory rest. There is no need to be silent to recover listening, but some form of reduced stimulation, dynamic contrast, or continuity in the environment is required. Quite the contrary, this is what is usually offered by short video platforms. The last note of one clip is followed by the next note of another one, very often with rapid changes in beat, cue timing, and affective coding. The continuous introduction of new acoustic phenomena, even in cases of homogenization of the sound design in clips, requires a new orientation. The users are forced to reset their perceptual expectations repeatedly and are not given adequate time to undergo auditory decompression.

According to this mechanism, fatigue can be attributed to the time packaging of sound in addition to the content of sound. When homogenized audio is inserted into the feed paradigm of chopping listening into an infinite number of micro-exposures, it can even be more tiring. These exposures might appear inconsequential on an individual basis, but when they are organized over time, they cause a long-term state of auditory alertness. The user is always ready for the next sound signal, emotional signal, or rhythmic trigger. In the long run, this can lead to low-grade tension and listening fatigue, though individual clips can be short and seem lightweight [50].

Another lesson learned is that any temporal acceleration undermines the occurrence of reflective auditory participation. Sound is instrumentally processed as opposed to pensive. It gets involved in a fast process of attraction, evaluation, and substitution. Recurring sound effects in such a cycle might no longer add to the sense but might become the means of repetition in a system of perpetual sensory mobilization. It is, perhaps, one reason why sound exhaustion on short video platforms can be less directly the feeling of negative response towards a single sound annoyance and more of a general weariness due to constant sonic stimulation.

3.4. Algorithmic reinforcement and recursive exposure to acoustically similar content

Creator imitation or production convenience is not the only cause of acoustic homogenization; it is supported by recommendation systems, which reproduce and amplify acoustically successful content in a continuous process. This algorithmic aspect is important since it will convert sonic familiarity into an aesthetic disposition into a structural arrangement of user exposure. Where the recommendation systems find some sounds, musical styles, or emotion-based audio stimuli effective in retention and interaction, they will tend to spread similar material further. The user is then not only exposed to popular videos but also to repetitive auditory logics contained in the videos [51,52].

This forms a recursive exposure cycle. Material with familiar and interesting patterns of sounds works. Its performance gives value to the algorithm. Adoption of similar content is then suggested, and creators are encouraged to use similar audio approaches. The recommendation feed is filled with similar-sounding content by degrees, and so is the listening experience of the user. In this respect, algorithmic personalization can become a paradoxical decrease in auditory diversity in spite of the fact that it can seem more adaptable in delivering content to individual tastes.

The first point that is critical in this case is that the preference of the user should not be misunderstood with the well-being of the user. In the short run, systems can come to know that a user will respond positively to some of the sound styles, but the optimization could produce more fatigue with repetitions in the long run. What may be a delight in the present may be tiring in the past. There is thus an algorithmic contradiction of temporality between engagement efficiency and sensory sustainability. The site is optimized to attract repeatedly, whereas the user can experience progressive auditory loss [53].

Fatigue's purely psychological interpretations are also made complicated by this mechanism. When exposure to homogenized sound is structurally scheduled according to the recommendation systems, then fatigue cannot be discussed only in the framework of overuse or poor self-regulation by individuals. The platform is engaged in shaping the circumstances of repetitive listening. That does not nullify user agency, but it does imply that auditory fatigue is supposed to be thought of as an infrastructural design phenomenon to some extent. This revelation can be relevant to both research and regulation, as it shifts the burden off the individual users onto the general structure of platform-mediated listening [54,55].

3.5. Cross-modal overload and the amplification of fatigue through audiovisual coupling

Sound alone can hardly generate auditory fatigue on short video sites. Background sound effects go hand in hand with visual movement, captions, rhythms of editing, facial expressions, and emotionally charged images. This cross-modal synthesis augments the convincing and engaging power of content, but can also boost exhaustion in the case of homogenization of the auditory aspect. In the repetitive use of sound effects, they do not work within an empty sensory space, but are continuously matched with visual novelty and behavioral cues, further overwhelming attention and affective processes [56].

The reason behind this is that audiovisual congruence has the potential of enhancing the salience of otherwise insignificant sounds. A transition signal corresponding to a cut, zoom, or gesture is more perceptually strong than the identical signal that is heard on its own. The tendency to reinforce the fatigue-inducing effect of homogenized background sound with repeated exposure to such couplings may thus be viewed as a way of creating a larger framework of multimodal stimuli. The listener is not simply hearing the repetition of cues, but he is sensing them as a component of a combined attentional experience. This increases the chances that sound is adding up to cumulative overload even in the background of conscious awareness [57].

Meanwhile, auditory repetition can be covered with visual novelty, so the users will find it more difficult to detect the fatigue or describe it. Without realizing the contribution of repetitive sonic environments, users can explain their exhaustion as a result of too much scrolling or the overall tiredness of the screen. The analytical significance of perceptual masking is significant since it can be used to understand why auditory fatigue is not well-researched in platform research. The auditory load is actual, yet it is shared among multimodal experiences and thus cannot be identified with ease using random observation [58,59].

Cross-modal overload also indicates that homogenized background sound could increase switching costs in the mind. The visuals on the screen are new with every new clip, but the same audio effects can be used to create the effect of emotion, speed, or tension in the same manner. New semantic and visual data and repetitive sound structures of an affective nature have to be integrated in the brain. The long-term result of this may be not only sensory fatigue but also decreased flexibility of attention and emotional numbing. What appeared to be stimulating at first might turn out to be exhausting as the same sonic stimuli are constantly added to ever-changing pieces of content.

3.6. A cumulative model of fatigue formation in homogenized auditory environments

Collectively, these processes suggest that auditory fatigue on short video platforms is both cumulative and multicausal and strongly influenced by platform conditions. Repetition decreases novelty, but does not decrease processing requirements. Spectral crowding and loudness compression enhance the effort of listening. Quick time turnover inhibits sound recovery. The recommendation systems expose users to more similar sounds in a recursive way. Background audio is enhanced in terms of its salience and persistence by cross-modal coupling. None of these mechanisms completely explains fatigue, but their interplay forms a consistent route in which acoustic homogenization can be experientially salient [60].

The importance of this cumulative model lies in the fact that it criticizes the single-variable explanations of the narrow type, in terms of volume, length of content, or sensitivity of the user. It seems that the combination of micro-acoustic design, temporal organization, algorithmic reinforcement, and multimodal stimulation leads to auditory fatigue in short video use. This implies that the interventions should also be multilevel. It can be that merely reducing volume or screen time is not enough when the root cause of the problem in the background is that the auditory ecology has been homogenized and optimized to produce instantaneous effects.

Critically, the most significant observation is that fatigue is not a side effect that occurs at the periphery of platform use. It can be embedded in the very systems which can make short video content interesting, expandable, and commercially effective. The processes that optimize sonic recognizability and user retention can also destroy auditory comfort as time goes by. Any serious platform audio design assessment needs to take this tension into account. **Table 2** is a synthesis of the principal mechanisms by which acoustic homogenization can be a cause of auditory fatigue. It redefines auditory fatigue not as a small nuisance but as an important measure of the sensory cost of contemporary short video ecologies [60,61].

Table 2. Proposed mechanisms linking acoustic homogenization to user auditory fatigue.

Mechanism	Description	Expected effect on users
Repetition and perceptual saturation	Recurrent exposure to structurally similar sound effects reduces auditory novelty while maintaining sensory demand	Monotony, irritation, reduced listening tolerance
Loudness compression	Narrow dynamic range produces persistently intense auditory stimulation with limited relief	Increased listening effort, subjective pressure, faster fatigue accumulation
Spectral crowding	Dense concentration of sound energy, especially in attention-grabbing frequency ranges, makes auditory processing more effortful	Perceived sharpness, clutter, overstimulation
Accelerated temporal turnover	Continuous scrolling exposes users to many short sonic events with little recovery time between clips	Cumulative fatigue, sensory vigilance, reduced restoration
Algorithmic recursive exposure	Recommendation systems repeatedly present acoustically similar content based on prior engagement patterns	Reinforced sameness, prolonged exposure to narrow sound repertoires
Cross-modal coupling	Repetitive background sound is amplified by synchronization with visual novelty, captions, and affective editing	Higher attentional burden, masked but intensified fatigue
Reduced feed-level diversity	Even when individual clips differ visually, the auditory feed remains within a narrow stylistic range	Diffuse exhaustion across extended viewing sessions

4. Evidence base: Measurement approaches and empirical findings

4.1. The current evidence landscape and the challenge of indirect inference

The empirical basis of cognition of the relation between acoustic homogenization and auditory fatigue in short video platforms is not well-rounded, haphazard, and staged. Currently, no developed literature exists that directly and comprehensively quantifies the effects of homogenized background sound effects on the short video feed on user auditory fatigue in realistic contexts of use. Rather, the evidence at hand will have to be assembled out of a few domains, which overlap to a certain degree, such as psychoacoustics, environmental sound, media psychology, digital platforms, music information retrieval, and human-computer interaction. This renders the evidence base to be promising and problematic. It has a promising future since the surrounding areas provide ideas, techniques, and discoveries that are highly indicative of plausible mechanisms. It is troubling since the transfer of these disciplines to the short video setting is not often direct [62].

One of the major challenges is the discrepancy between platform realism and laboratory control. Auditory research studies frequently manipulate individual variables, including loudness, frequency distribution, repetition, or listening effort under very controlled conditions. In comparison, the use of the short video is a multimodal, algorithmically curated, and dynamic experience. Sound is integrated into changing images and emotional scenes at a high rate, and exposure patterns are determined by scrolling behavior as opposed to experimental sequences. Consequently, any evidence that is obtained through traditional listening experiments can determine the applicability of auditory burdens without necessarily measuring the extent to which the burdens are accumulated when using a platform in real life. On the other hand, platform studies can be realistic; however, they will not provide the acoustic accuracy necessary to determine the exact characteristics of sounds that cause fatigue. The present evidence situation is then typified more by concurrent evidence rather than evidence per se.

This does not render the inquiry speculative. Instead, it emphasizes the necessity of synthesis to establish what is already empirically based and what is merely inferential. To be cautious, however, a review has to evaluate three levels of evidence: first, quantification tools of acoustic homogenization; second, tools of auditory fatigue or other burdens; and third, empirical studies, the results of which, either directly or indirectly, inform the connection between repetitive sonic conditions and user strain. This method can also demonstrate the weaknesses of the available literature and identify the spheres where future activities should be made more specific [63]. For better understanding, a summary of such evidence is given below in **Table 3**.

Table 3. Categorization of evidence supporting the acoustic homogenization-auditory fatigue relationship.

Evidence type	Examples from literature	Contribution to argument	Limitations
Empirical (Direct)	Studies measuring auditory fatigue in repetitive sound exposure; listening effort studies with compressed audio; platform behavior studies on skipping/muting patterns	Directly tests mechanisms linking homogenization to fatigue outcomes	Limited direct research on short video platforms specifically; most studies use simplified stimuli
Empirical (Indirect)	Research on notification fatigue; interface sound irritation; music overexposure effects; noise annoyance studies	Provides analogous findings from similar phenomena that support plausibility	May not fully transfer to short video context; differences in exposure patterns and multimodal integration
Theoretical	Auditory load theory; habituation theory; cognitive resource theory; soundscape theory; platform studies	Offers explanatory frameworks for why homogenization should produce fatigue	Requires empirical testing of proposed mechanisms; not sufficient alone to establish effect
Methodological	Studies on feature extraction; subjective fatigue questionnaires; behavioral logging; physiological monitoring	Provides tools and approaches for investigating the relationship	Need integration across methods; ecological validity concerns

4.2. Measuring acoustic homogenization in short video environments

Any empirical study of acoustic homogenization must have a method of defining and measuring similarity among background sound effects. It is not a simple methodological exercise. Exact copying of audio tracks is not the only way of homogenization; it can also arise in acoustic structure repetitions, parallel affective cues, convergent aesthetics of production, or repetitions of rhythm, timbre, and loudness combinations. Measurement, therefore, should not be confined to mere identification of reused sounds, but should be the expression of the reduction of meaningful acoustic diversity in a segment of the feed or platform.

Acoustic feature extraction is one of the methods. Background sound effects are measured and processed in this framework based on measurable parameters that include tempo, density of onsets, spectral centroid, bandwidth, harmonicity, dynamic range, contours of loudness, timbral texture, and repetition rate. The distance between audio samples of a multidimensional feature space can then be used to estimate similarity. When there is a large number of short videos concentrated over a small number of feature profiles, then it can signify acoustic homogenization [64]. The usefulness of the method is its objectivity and scalability. Platform content can be computed in large volumes, which enables researchers to study the question of whether recommendation-based feeds tend to systematically decrease sonic diversity.

Nevertheless, there are limitations of feature-based methods as well. First, extracted acoustic similarity is not perceptually significant in all cases. Two sounds can sound very similar in terms of computation but be perceived differently by the human listener. Second, most research on short video audio is devoted to foreground music instead of background sound effects, and it might neglect the more subtle sonic layers that are most likely to be sensitive to cumulative exposure. Third, feature extraction tends to separate sound from its location in time and role in audiovisual representation. However, a sound which is associated with a dramatic shift in the visual image can have a vastly different effect to the same sound in an environment that is not so stimulating. Computational similarity, therefore, ought to be considered as an essential point of origin rather than an overall explanation of homogenization [65].

A second method is template and trend analysis. In this case, researchers can find recurring platform-native sound conventions, including shared transition noises, emotional stingers, suspense signals, ambient pads, or viral sound formats. This method is particularly helpful for recording cultural and functional homogenization instead of acoustic similarity. It may show how makers are reiterating to use sound in accordance with platform expectations, how editing software directs some default sonic tendencies, and how recommendation algorithms enhance already recognizable sound. It is more advantageous in that it has interpretive richness. It explains why, as well as that, sounds are similar, and how their repetition comes to be a normalized aspect of platform culture [66].

However, this approach is not without its problems. Coding recurring sound templates can prove to be a tedious process and might rely on subjective analysis. Researchers have to choose whether similarity is based on the properties of the waveforms, affective functionality, editing location, or user identification. An example of a transition cue is that it can differ acoustically and still be functionally the same between clips. This ambiguity is analytically fruitful, but difficult to provide a methodology. It postulates that acoustic homogenization is most effectively quantified by using the summation of both computational features analysis and interpretive platform analysis as opposed to either.

The third and more crucial approach is the feed level diversity indexing. Rather than quantifying individual audio objects, researchers are able to quantify the extent of acoustic variation perceived by a user over time in a series of suggested clips. These indices can be a combination of track reuse, similarity of features, dynamic range change, affective range, and change of timbre between scrolling sessions. This method is especially useful since fatigue may also be conditioned not only by the personal sound qualities, but also by the time exposure to the recurrence of a similarity of repeated exposures. That is, it is not merely the fact that a single clip is acoustically constricted but, over an extended time, the entire feed continues to reposition the user in the same auditory range [67].

4.3. Measuring auditory fatigue: Subjective, behavioral, and physiological indicators

In case acoustic homogenization proves difficult to measure, auditory fatigue is even more difficult to measure. Fatigue may not manifest itself as one apparent symptom in the context of short videos [68]. Instead, it can include a complex of subjective discomfort, listening discomfort, irritability, decreased attention capacity, mental fatigue, and withdrawal of behavior. This multidimensional state cannot be appropriately measured by one metric. The combination of physiological, subjective, and behavioral indicators is the most informative evidence, therefore. **Table 4** sums up the key methodological techniques for measuring acoustic homogenization and auditory fatigue, along with their key strengths and weaknesses [69, 70].

Table 4. Measurement approaches for acoustic homogenization and auditory fatigue.

Measurement domain	Main indicators	Typical methods	Strengths	Limitations
Acoustic homogenization	Tempo similarity, spectral centroid, loudness contour, timbral resemblance, repetition rate, dynamic range	Audio feature extraction, clustering, similarity analysis, corpus-based platform audio mining	Scalable and objective; suitable for large datasets	May miss perceptual meaning and contextual audiovisual function
Functional sound template convergence	Recurring suspense cues, transition sounds, emotional pads, rhythmic fillers	Content coding, platform trend analysis, creator tool analysis	Captures platform-native audio conventions and cultural repetition	Labor-intensive and partly interpretive
Feed-level acoustic diversity	Variation in sound features across a user's scrolling session	Sequential feed analysis, recommendation exposure mapping	Reflects cumulative real-use exposure more accurately	Requires realistic session reconstruction
Subjective auditory fatigue	Listening effort, annoyance, discomfort, overstimulation, desire to stop listening	Questionnaires, rating scales, post-exposure interviews	Directly captures user experience	Vulnerable to reporting bias and limited self-awareness
Behavioral fatigue indicators	Skipping, muting, lowering volume, shortening session duration, content abandonment	Platform usage logs, controlled behavioral observation	Ecologically relevant and scalable	Behavior may reflect multiple motives beyond fatigue
Physiological burden	Heart rate variability, electrodermal activity, pupillometry, EEG attention/load markers	Laboratory experiments, multimodal monitoring	Can detect non-conscious strain and arousal	Difficult to isolate sound effects from visual and emotional confounds

Note: EEG: Electroencephalography.

The most direct means of reaching the lived experience of auditory fatigue is still subjective self-report measures. The participants can be requested to measure the listening effort, annoyance, auditory discomfort, irritability, overstimulation, or wish to terminate exposure to various sound conditions following interaction. These measures are unavoidable since fatigue is in some way experiential and can even be present without the presence of externally manifested manifestations. They come in handy, especially to differentiate familiarity and repetition that are a burden. A user may be familiar with background sounds that he or she does not dislike, or may complain of increasing fatigue, although not consciously aware of the repetition. This distinction can only be found through self-report [71].

Meanwhile, self-report data may be susceptible to bias, low introspection, and retrospective distortion. Auditory discomfort is not always easily separable from general digital fatigue, particularly in highly multimodal settings. They can even credit exhaustion to screen time, content volume, or visual stimuli and underreport the contribution of repetitive sound. Background sound effects are a particular issue that is particularly acute since they tend to influence experience without their direct subject of attention. Consequently, there is a need but not a sufficiency of subjective measures.

Another significant supplement is offered by behavioral indicators. When in platform environments, auditory fatigue can be observed in an increased rate of skipping, muting, reducing volume, reduced persistence of session, moving to text-based or silent content, or abandonment of videos with specific sound signatures. These habits may be indirectly used to prove that certain auditory surroundings cause user intolerance. Behavioral data can be recorded in an unobtrusive and large-scale way, unlike self-reports. They prove to be of great value in investigating cumulative effects during real usage sessions, and not just on immediate verbal judgments.

But the interpretation of behavior should be careful. It is possible that a user can skip content due to a variety of reasons that are not connected to fatigue, such as boredom, lack of interest in the subject, or negative opinion regarding visuals. Sensory burden may not be the reason why people mute, but it reflects the situation. This is why the behavioral information is more persuasive to be triangulated with the acoustic analysis and subjective reporting. When the exposure reduction is systematically applied by the users in highly homogenized sound conditions, with the additional report of increased strain during listening, the argument of fatigue effect becomes much stronger [72].

A third level of evidence is provided by physiological measures. They can be heart rate variability, electrodermal activity, pupillometry, EEG indicators of attentional effort, or other arousal and cognitive load correlates. This kind of measure can, in principle, allow determining whether repeated and compressed sound environments are making stress or processing harder than users can easily express. They come in handy, especially in identifying low-level strain that accumulates over time and may not necessarily be noticed as a painful experience on the conscious level.

Nonetheless, auditory fatigue does not have direct physiological proxies. The increase in arousal could be a sign of involvement as opposed to load, and alterations in the physiological response could be promoted by visual or emotional stimuli as opposed to sound. The difficulty, then, is the design of the experiment. The conditions are to be compared in studies, and the influence of acoustic similarity, loudness patterning, and audiovisual coupling should be separated. The most informative physiological measures are not intended to be applied alone as sole evidence, but rather applied in a multimethod design that can differentiate fatigue and general stimulation [73].

4.4. What existing empirical findings suggest

Even though there are only limited direct studies on acoustic homogenization and auditory fatigue in short video platforms, there are a number of directions that seriously indicate the possibility of the relationship [3]. The studies in the repetitive sound exposure field have always demonstrated that repetition may augment irritation and perceived load when change is minimal and recovery is inadequate. This should not mean that repetition is always bad, but it does indicate that closed acoustic space may be exhausting in the long run, especially when stimuli are salient and functional. Such results are quite applicable to short video websites, where there are repeated stimuli in quick sequences during long viewing periods [2].

Listening effort studies can also be of significance. Psychoacoustic research has demonstrated that sound conditions due to compression, spectral density, or informational clutter have the potential to increase the perceived effort to listen despite moderately high levels of volume. It serves the purpose of arguing that auditory fatigue on short video platforms can be caused by the design characteristics of compressed loudness and crowded frequency material as opposed to an explicitly dangerous sound level. In other words, fatigue might be used as an indicator of stimulation quality and structure, rather than its intensity.

Similar studies in digital and interactive media are also heading the same way. These may include notification fatigue, interface sound irritation, and sensory overload in highly mediated environments, which imply that repetitive short-duration sounds may become annoying when repeated, attention-seeking, and hard to ignore. Short video background sound effects are similar to notifications in their purpose and structure, but they have one important thing in common: they are programmed to elicit instant orientation. Repeated exposure to these cues during a scrolling session is thus most likely to cause cumulative burden, despite any one of the cues seeming inconsequential.

Music psychological evidence, as well as streaming behavior, provides partial similarities as well. Familiar audio can at first increase preference, and overexposure tends to cause satiation, boredom, or avoidance. Within the short video frame, though, it is not the problem of musical boredom as much as it is the problem of sensory weariness through structurally similar audio being attached to high-speed content. Such a difference is important since the effect of fatigue can manifest itself in advance, even when the users do not actively dislike the sound. Sameness and intensity can be imposed on the auditory system as the user keeps interacting behaviorally with the platform [74].

The only thing lacking is evidence that links these findings directly to real-life situations of using short videos. There is substantial evidence to suggest that repetitive, compressed, and high-density auditory experiences may amplify load, but comparatively little research has been conducted to measure this on the basis of recommendation-based feeds and actual platform dynamics. In this way, the existing empirical image is indicative as opposed to authoritative. It raises a cause of concern and warrants additional research, although it is not yet a causal map that is settled in its entirety.

4.5. Methodological limitations and critical gaps

The primary flaw of the existing evidence base is discontinuity. The acoustic analysis, fatigue measurement, and platform behavior are rather frequently researched in isolation. Similarity mapping can be a study in computational audio that does not require understanding if it imposes a burden on the user. User studies can identify frustration or exhaustion, but not the sound characteristics that cause the effect. Platform analytics can show the patterns of skipping or muting, but not the cause, which is sonic repetition. Such division constrains the depth of explanation and comes up with sturdy conclusions [75].

The second limitation is the ecological validity. Most experiments involve simplified audio stimuli or brief lab exposures, which do not scale up the cumulative structure of scrolling. However, fatigue on the short video platform will most probably be time-dependent, sequence effects, and time-repetition of the algorithm. In the absence of feed-level exposure designs, studies will be susceptible to underestimating the accumulation of small burdens in dozens or hundreds of clips.

The third issue is the lack of background sound effects attention as a separate issue, opposed to music or speech. Most audio studies of short videos focus on the songs, voiceovers, or overt sound memes, but the more subtle but constant layer of background effects is possibly more of a central interest to fatigue as it is active throughout, and across diverse content. This discontinuity has a wider reflection of biased salient sonic objects at the cost of ambient or supportive auditory design [76].

The other important problem is confounding. As the use of short video is multimodal in nature, auditory fatigue cannot be separated from either visual overstimulation, emotional arousal, or general cognitive overload. This fact, however, is no reason to be conceptually vague. Quite the contrary, it highlights the necessity of experimental and mixed-method designs, which involve comparing matched visual content in varying auditory conditions, following behavioral changes in real-life sessions, and combining self-report and physiological and acoustic measures.

4.6. Toward a stronger empirical agenda

The available evidence base, albeit not exhaustive, is already adequate to prove that the suggested connection between acoustic homogenization and auditory fatigue is a theoretically plausible and empirically reachable fact. Methodological integration is what is now required. The analysis of acoustic diversity at the feed level in the future needs to be integrated with realistic exposure tests and multimodal fatigue analysis. **Figure 2** suggests a multimethod study design that could be used in the future to conduct empirical research on the topic of acoustic homogenization and auditory fatigue. They are to contrast not only loud and quiet, liked and disliked sounds, but also acoustically different and acoustically homogenized series. They are expected to look at cumulative exposure in scrolling but not isolated clips. The most important thing is that they must consider background sound effects as an object of study instead of some minor decorations [61].

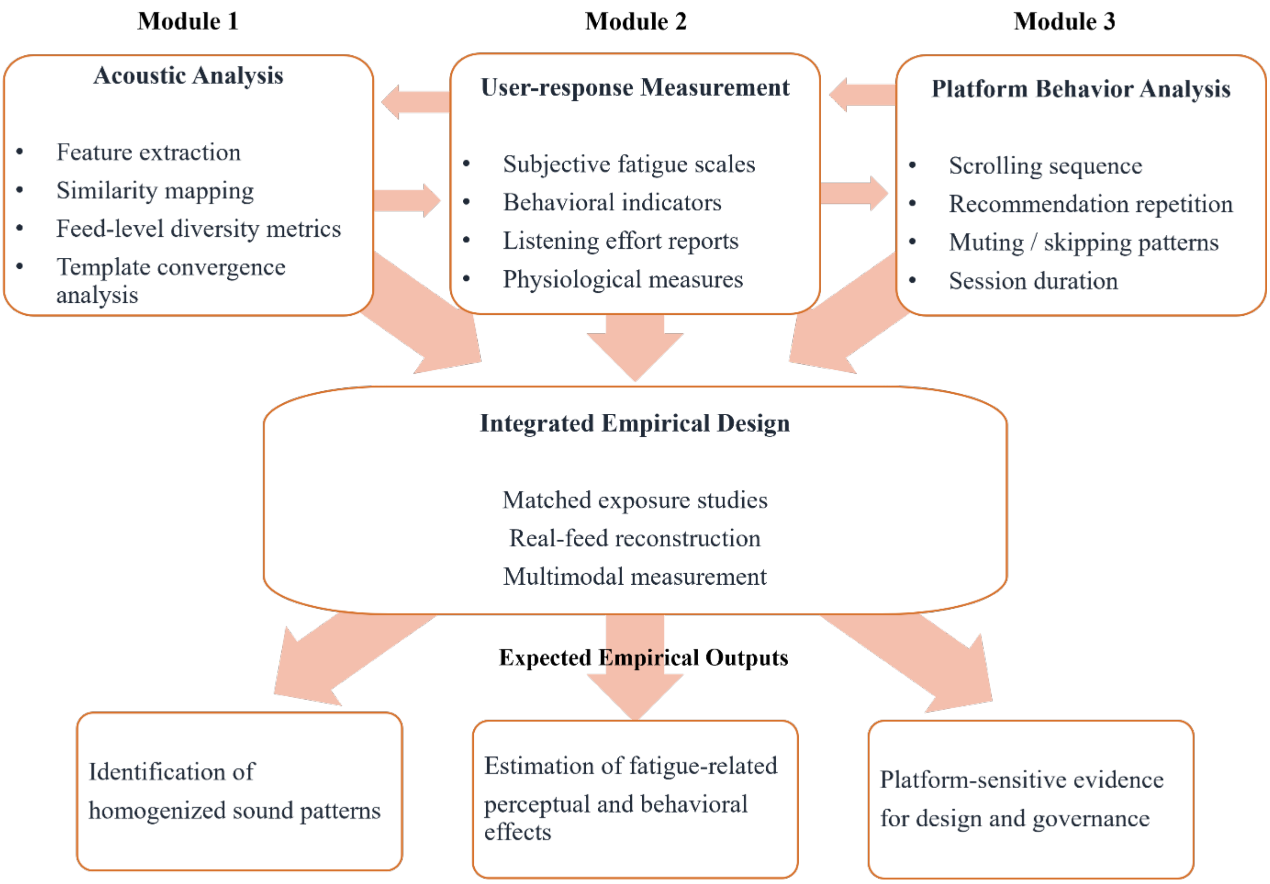

Figure 2. Methodological framework for studying acoustic homogenization and auditory fatigue.

This kind of agenda would go a long way in empowering the field. It would enable researchers to go beyond the general arguments of digital overload and towards a more specific approach to how platform-specific auditory design influences user well-being. It would also aid in making clear whether or not auditory fatigue is to be viewed as a marginal nuisance or a structural implication of platform-optimized sound ecologies. Currently, there is limited evidence to support the latter case, although it is heavily leaning that way, as definitive evidence has yet to be compiled. That is why it is not a premature, but an urgent issue [77].

5. Moderating factors and practical implications

5.1. User-level moderating factors in the formation of auditory fatigue

There is not much likelihood that the effect of acoustic homogenization has a consistent effect on auditory fatigue in all users. Although the exposure of large populations to common background sound patterns may be triggered by short video platforms, the extent to which the pattern will be perceived as a burden will be determined by a set of individual-level moderating factors. They are auditory sensitivity, age, media use habits, emotional regulation style, attentional control, ability to maintain cognitive load, and the listening context of the situation. Any serious description of fatigue must thus not be purely deterministic, so that homogenized sound causes the same effects for all listeners [78].

Baseline auditory sensitivity is one of the most significant user-level variables. There are those users who are more responsive to repetitive or spectrally dense sounds, even when the volume is moderate. To these users, background sound effects of compressed dynamics, high-frequency focus, or rhythmic regularity could hasten discomfort and perceived listening effort. Other people might find these sonic environments easier to deal with, particularly when they are accustomed to using high-intensity media or when they view short video material regularly using personal listening devices. The habituation, however, should not be interpreted as an easy way out. Users who constantly subject themselves to homogenized audio are less consciously aware of irritation despite cumulative fatigue in the form of a lower limit of listening, loss of attention, or a flattened effect.

There are also complex effects of age on fatigue formation. The younger users who tend to have more commitment to platform-native audiovisual styles might experience increased exposure to the fast and repetitive sound formats and may not rate the immediate annoyance as much. Simultaneously, they could be at higher risk of permanent sensory overload due to the increased length of cumulative exposure and the more frequent scrolling behavior. Homogenized short video audio can be considered more obtrusive or overstimulating to older users since it is more discontinuous to the established audiovisual timing. It is not that one of the age groups is more vulnerable in an absolute sense, but it is more probable that vulnerability is the product of the interplay between sensory expectations, the intensity of use, and exposure to platform rhythms [60].

It is also important in terms of attentional style and tolerance to cognitive load. When multitasking, commuting, or resting following other types of mentally demanding activities, users who watch short videos when performing these tasks are likely to have fewer resources to deal with repetitive and fast-changing audio. Even moderate sonic repetitions in such states can turn out to be disproportionately fatigative. In comparison, a more intentional and limited consumption of content by users can make the same audio sound stimulating and not wearying. This implies that auditory fatigue cannot be comprehensible in isolation from the wider ecology of daily life. Platform sound design is engaged in relation to already existing states of mental fatigue, environmental sound, and simultaneous task needs.

These observations have a critical implication. The issue of fatigue responsibility should not be personalized at the user level to exempt platforms of design responsibility. Structural effects do not rule out the presence of moderators. Instead, it implies that homogenized soundscapes will tend to place disproportionate pressures on populations, and more susceptible users may get exposed to those more severely. This is the reason why this issue is relevant within the digital well-being considerations. Once platform audio systems cause fatigue in some groups of users in a systemic manner, then it is not subjective variability but rather an inquiry of unequal exposure to avoidable sensory stress [79].

5.2. Content- and platform-level moderating factors

In addition to the user factors, the impacts of acoustic homogenization are highly influenced by the content type, platform framework, and the listening environment. Not every short video is based on the background sound in a similar manner, and not every platform presents the exposure to audio with the same intensity. These content and platform-level variables intervene between the provision of repetitive sound, its perception, and its accumulation over the use sessions [2, 80, 81].

One of the most important moderators is content genre. Clips with an entertainment focus, promotional material, dance videos, and humorous edits frequently include background sound effects that are very salient and are intended to enhance rhythm, humor, suspense, or emotional immediacy. Sound, in such genres, is highly bound to visual action and is optimized to get a quick hit, which means that homogenization is particularly probable and fatigue is particularly realistic. Short videos of the educational, informational, or documentary type might be less liberal with the use of background sound, or they might be more functional, so the repetitive load on the senses may be less, even in the case of acoustic similarity. Nevertheless, the difference is not to be exaggerated. Even information content is becoming influenced by platform-native editing norms, and the educational producer tends to use the same suspense indicators or transition sounds, or inspirational ambient beds, to stay competitive in the feed. Genre differences are therefore important, although they are typically reduced by platform norms [82, 83].

The recommendation system is another important moderator. Social networks with a high volume of recycled trending audio, that encourage the use of recognizable sound templates or put users into limited loops of preferences, might encourage homogenization to a higher extent than networks with more extensive feed diversity. Recommendation systems are not merely a neutral means of delivery; they establish the time arrangement of the similarity. The user might not feel bothered by a single familiar sound pattern in some instances, but in the case of repetition, cues similar to them, during a short period of time, can produce a significantly greater burden of fatigue. In that regard, feed level diversity can be more significant than clip level novelty. Visual differences can seem to indefinitely refresh a platform but restrict the auditory repertoire of the users [84].

Infrastructure and creator tools editing also work as moderators. In-service music libraries, automatic transition effects, single-button audio enhancement, and trend-based sound suggestions are all less frictional in the content creation process, but they lead creators to a limited set of acoustically standardized choices. This not only has an impact on what can be sounded but also on what turns out to be the norm, readable, and rewarded. The greater the number of platforms that simplify content creation using reusable sound templates, the more auditory homogenization is likely to become a part of daily production. Fatigue is therefore not solely an issue of user choice—what they want to watch—but rather what creators are guided to produce.

Fatigue is additionally mediated by listening mode. Sound playable on headphones can be more immersive and more emotive, but also repeats more often by making the environment less contrasting and the compressed sound or spectrally dense sounds more salient. Audio playback on the smartphone can reduce certain acoustic details, but maintain the aggressive middle-range emphasis of platform audio. Auto-play settings may also serve to increase the load by depriving the user of the chance to reload between clips. These differences indicate that auditory fatigue is not determined just by the content. It occurs as an outcome of the platform delivery, device affordances, and user listening conditions [85].

The bigger critical observation is that content and platform-level moderators manifest the distribution of fatigue across media systems and not just individuals. Sound design decisions do not work in a vacuum; they are either magnified or diminished by advice structures, template structures, and playback situations. This supports the position that auditory fatigue is an ecological phenomenon that is based on platform experience organization.

5.3. Implications for creators and platform designers

In the event that acoustic homogenization is a factor in auditory fatigue, then the problem has direct consequences for content creators and platform designers. The implications do not only refer to the comfort of users, in a small sense; they also refer to the quality of long-term engagement, trust towards platform environments, and the sustainability of digital attention economies. In the short term, a system that optimizes immediate capture of the senses at the expense of increasing fatigue over time may seem effective, but it will degrade the user experience over the long term [10].

To creators, the main connotation is that sound cannot be viewed as a mere means of quickening attention or heightening emotion. Background audio also determines the level and length of audience tolerance. The repetition of the same suspense signals, dramatic escalations, comfortable background emotion, high-stakes transitions, and their combination in the thousands of clips can render a single video platform-native and commercially competitive, but when duplicated thousands of times, add to an overall wearying aural effect. This implies that innovative distinction in sound design is not only potentially of value as an artistic decision but also as an audience-care method. The increased diversity of pacing, timbral texture, silence, and dynamic contrast can probably be useful in maintaining the receptivity of listeners.

To platform designers, it has more structural implications. Acoustic diversity may be an important design parameter in the recommendation system instead of merely optimizing the system to achieve only click-through, retention, or trend conformity. Currently, the success of algorithms tends to strengthen repetitive formats of sounds, as they are simple to identify and statistically sound. However, a platform that is sensitive to fatigue would not only inquire about whether a sound is doing you well at a specific time, but also would, over time, ask whether to continue exposing the user to similar sounds as they find that it reduces comfort. This would need to change the current approach to engagement-maximizing metrics towards models that take into account cumulative sensory effects [38].

Sound libraries and editing tools used on a platform should also be reconsidered.

Existing creator ecologies usually favor efficiency, virality, and frictionless replication. Costs are associated with this design logic, as far as fatigue is concerned. When libraries default recurrently predict the identical emotional stimuli and transition impacts, auditory homogenization is effectively embedded within the artistic infrastructural provisions of the platform. One viable solution would be to increase sonic diversity on the tool level by encouraging wider sound palettes and fatigue-sensitive mixing presets, and discouraging the extensive use of highly compressed or high-salience templates.

The argument for reconsidering the normalization of loudness and dynamic shaping of short video platforms also holds a lot of weight. Consistently optimized audio that is maximally salient might prove useful when it comes to engaging the listener in the short term, and harmful when it comes to listening comfort in the long term. More adaptive systems would be able to maintain clarity without making all contents equal in terms of their intensity profiles. It would not necessitate giving up platform-native aesthetics all the way, but it would entail being conscious of the fact that auditory impact and auditory sustainability are not synonymous [86].

These implications are important since they expose a very fundamental belief in platform design, which is that whatever performs best in terms of gaining attention is good design. That assumption can no longer be held as far as auditory well-being is concerned. Effective capture of attention might be accompanied by poor sensory care. In a more adult design approach, the capacity of users to listen would be regarded as a limited resource that should be preserved instead of harvested.

5.4. Public health, digital well-being, and regulatory implications

The realistic implications of acoustic homogenization are not confined to the creators and platform engineers but to the larger issues of population health and online well-being. Video platforms of a short format have gained a large proportion of daily life, particularly among the younger demographics, and their sensory space is becoming an influential determinant of the normal patterns of relaxation, distraction, learning and socializing. In such circumstances, auditory fatigue cannot be ruled out as a trifle inconvenience. It can be a minor yet common media-related stress that has some consequences in terms of attentional and mood control and quality of life [87].

In the digital well-being perspective, existing discussions have been extremely concentrated on screen time, addiction-like behavior, misinformation, and mental health impacts associated with being exposed to the contents. The sound aspect has been relatively overlooked. Nevertheless, users do not passively view the short video visually but live in a multisensory space where sound constantly affects arousal, affect, and fatigue. Being aware of this fact extends the notion of digital well-being to the issues of information and behavioral aspects, as well as the quality of mediated sensory experience. Platforms could meet traditional usability standards but create an unhealthy sound ecology.

Regulatory relevance is also applicable to this recognition. Current discourses of media sound tend to focus on the highest volume limits or advertising sound levels. Even though such measures are critical, they cannot be used to deal with the type of fatigue to be discussed here. Even when the sound is not too loud relative to absolute sound, the effect of homogenization may be a burden. Repeat, compression, density, forced continuity through auto-play feeds all can cause cumulative fatigue without being out of conventional acoustical safety limits. This implies that a wider regulatory vocabulary should be used that is capable of responding to sensory sustainability and not just acute harm [38].

The future governance system must be done with caution. The excessively strict control over the sound design of the platform may limit creative design or have simplistic criteria not applicable in various cultural conditions. However, soft governance mechanisms are practicable and preferable. These might consist of voluntary practices in fatigue-conscious audio design, disclosure of sound recommendation behaviors, industry conventions on loudness and dynamic range in short-form media, and design audits which measure sensory diversity as well as engagement consequences. These would not do away with homogenization based on platforms, but they might initiate a shift towards making auditory well-being a valid area of responsibility.

Equity is another aspect to be taken into consideration. Homogenized auditory environments may be more impactful to users with sensory sensitivities, neurodivergent profiles, and hearing vulnerabilities or users with high cumulative media exposure, than others. When platforms do not pay attention to these differential burdens, they are likely to promote a normalization of an exclusionary sensory norm that is best adapted to the most flexible or commercially useful consumers. It is specifically due to the fact that public health thinking re-contextualizes such issues as ones of individual preference rather than issues of environmental fairness and preventable burden that it matters [88].

5.5. Toward fatigue-aware auditory ecology in short video platforms

The broader implications of these moderating factors and their practical meanings are that they bring us out of the discussion of isolated symptom management to ecological redesign. Auditory fatigue cannot be considered only by the recommendation to turn the volume down, spend less time on screens, or become more disciplined, but such solutions can be effective. These suggestions are not complete as they offer to leave the responsibility of adaptation fully to the individual and do not touch the homogenized structure of platform sound.

A more fruitful approach would be to cognitively think about fatigue-conscious auditory ecology. This entails creating a brief video space where variation, recovery, perceptual balance, and sonic diversity will be considered as major components of user experience. In **Figure 3**, a multilevel system of intervention among creators, platforms, and governance is provided. It implies the understanding that background sound effects are not passive elements of sensory infrastructure. It also implies that the maximization of sound with respect to virality and retention is limited to certain points when the efficiency of the platform turns into a depletion of users [89].

The theory of auditory ecology comes in handy as it focuses on inter-relational harmony and not individual factors intervention. It would not be a healthier platform or sound environment, as it would not necessarily remove repetition, trending audio, or emotionally charged sound design. Instead, it would decrease the over-convergence, recapture the significant variation, and create an environment where the users are not subject to constant sonic torture. This can be compared to larger trends in responsible technology design that have been moving towards placing more importance on sustainability, user autonomy, and long-term well-being over performance-based metrics. **Table 5** sums up the main moderating variables identified in this review and the practical implications that go with them [90].

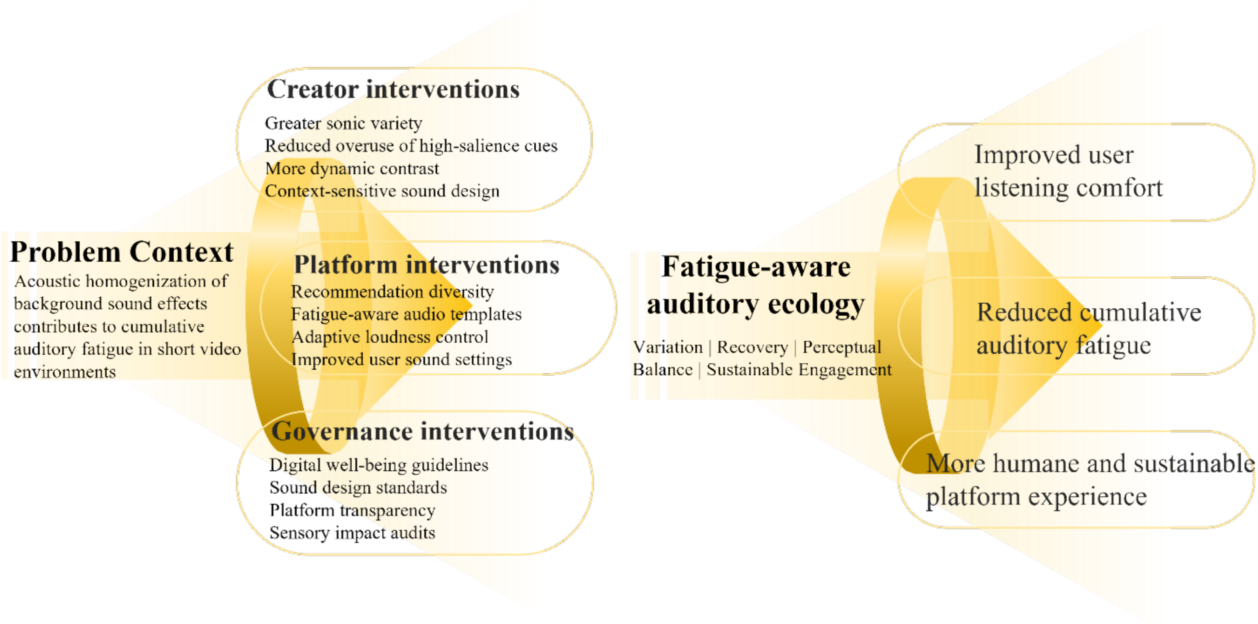

Figure 3. Fatigue-aware intervention framework for creators, platforms, and governance.

Table 5. Moderating factors and practical responses for fatigue-aware platform design.

Dimension	Moderating factor	Potential influence on fatigue	Practical implication
User-level	Auditory sensitivity	More reactive users may experience repetitive or compressed sounds as tiring more quickly	Encourage optional sound personalization and user control
User-level	Age and media habits	Different familiarity with platform-native audio may shape tolerance and cumulative exposure patterns	Consider age-sensitive platform well-being strategies
User-level	Cognitive load and situational context	Multitasking, stress, or environmental noise may reduce tolerance for repetitive audio	Support easier muting, pausing, and sound customization
Content-level	Genre	Entertainment and promotional clips often use more intense and repetitive background sound than informational content	Promote genre-specific guidance for healthy sound design
Platform-level	Recommendation repetition	Frequent recycling of similar audio content may intensify feed-level homogenization	Introduce acoustic diversity metrics into recommendation systems
Platform-level	Editing tools and default templates	Built-in sound libraries may channel creators toward standardized sound choices	Expand sound palettes and reduce default overreliance on high-salience cues
Device/listening mode	Headphones, speakers, autoplay	Different playback conditions alter immersion, salience, and recovery opportunities	Optimize playback settings for comfort, not only impact
Governance level	Digital well-being standards	Lack of standards leaves sensory burden largely unaddressed	Develop fatigue-aware guidelines for short-form audiovisual platforms

In this regard, acoustic homogenization poses a bigger question than sound. It is about the type of sensory worlds that digital platforms are creating and what type of human attention they are presuming to have as limitlessly available. Provided that the background sound continues to be at a minimum standardized in the short video platforms around direct effectiveness, it can be concluded that auditory fatigue will probably continue to be an unacknowledged price of everyday media consumption. In contrast, when platform design starts to consider listening comfort as a constituent of user well-being, then sound could not only be an object of interaction, but also a place of more human media practice [91–94].

6. Conclusion

This review has analyzed the conceptual, theoretical, and practical aspects of acoustic homogenization of background sound effects as an influencer of user auditory fatigue in short video platforms. With the growing influence of short video sites in daily media consumption, their auditory spaces should be taken more seriously as a scholarly object. Although the focus of previous studies has been on visual overstimulation, algorithmic dependence, and behavioral immersion, the current analysis provides that repetitive and standardized background sound effects are another significant but under-researched aspect of platform experience. The main argument made in this article is that acoustic homogenization is not a simple aesthetic fad or a by-product of content imitation, but a structurally generated aspect of platform ecosystems and can be one of the causes of cumulative auditory exhaustion.

It has been demonstrated in the review that acoustic homogenization is a product of the interplay between creator practices, platform-native editing technologies, viral audio culture, and recommendation algorithms that continue to prioritize acoustically similar content. In this case, the users are not merely exposed to discrete repeated sounds, but to a larger loss of auditory diversity in the long term through scrolling. This decrease is significant since the accelerated time turnover, the compressing structures of the loudness, the repetitiveness of the emotional signals, and the ongoing audiovisual feedback all condition short video listening. The sound environment thus created is very efficient in the capture of attention, but may be inefficient in maintaining sound comfort. In this respect, the sonic rationale of short video platforms indicates even further conflict between the optimization of engagement and sensorial health.

The major impact of this article has been to conceptualize auditory fatigue as a multidimensional experience, which is not limited to the traditional physiological concept of hearing strain. Auditory fatigue in the context of short video platforms encompasses auditory perceptual saturation, auditory listening effort, cognitive, affective, and decreased tolerance to continued sonic stimulation. Such a wider framing is required since the sound pressure related to the use of short videos can be under harmful volume levels and thus not even evident in the traditional acoustic safety paradigms. This review has suggested that repetition, spectral crowding, dynamic compression, feed level similarity, and cross-modal overload have a cumulative interaction, resulting in fatigue in ways that are subtle, distributed, and even obscured by the visual novelty of platform content.

At the empirical level, the article has pointed out both potential and weaknesses in the existing evidence base. Available literature in the fields of psychoacoustics, media psychology, soundscape studies, and digital media research overwhelmingly indicates that repetitive and compressed sound environments may enhance annoyance, listening effort, and sensory burden. Nonetheless, there is still scanty evidence that is particular to the short video platforms where recommendations are used. Techniques of acoustic homogenization measurement, auditory fatigue measurement, and realistic user exposure are yet to be optimally combined. Future research that integrates feed-level audio analysis, experiments that have ecological validity of exposure, multi-modal fatigue measurement, and platform-sensitive behavioral evidence is therefore indicated in this review. It is only through such integration that the connection between platform sound design and user fatigue will be felt suggestive and not completely proven.

It is also highlighted in the review that the influence of acoustic homogenization depends on the characteristics of the users, the content genre, listening conditions, and platform designs. Auditory fatigue does not occur equally, or even solely, in individual sensitivity. Some users might be more susceptible due to sensory profiles, usage intensity, or situational cognitive load, whereas some platforms might enhance the occurrence of fatigue due to recommendation loops, default audio templates, and auto-play formats that reduce auditory recovery. Important practical implications can be made from these findings. To creators, they propose that more sonic diversity and tampered sound-system background design could be useful in facilitating more sustainable audience interaction. Regarding platforms, they signal the necessity to go beyond exclusive engagement-based measures and view acoustic diversity, recovery, and listening comfort as valid dimensions of responsible design. To regulators and researchers of digital well-being, they provide the significance of the broader discussion on screen time and content exposure to the quality of mediated sensory environments.

Several promising directions for future research emerge from this review. First, longitudinal studies are needed to track how auditory fatigue develops over months or years of regular short video platform use, moving beyond single-session measurements. Second, the development of real-time acoustic diversity metrics integrated into recommendation algorithms could enable empirical testing of whether increased feed-level sonic variation reduces fatigue accumulation. Third, controlled experimental designs that systematically manipulate acoustic homogenization while holding visual content constant are essential to isolate sound-specific fatigue effects. Fourth, research should investigate differential vulnerability to acoustic homogenization across user populations, including neurodivergent individuals, older adults, and users with pre-existing auditory sensitivities. Fifth, cross-platform comparative studies examining how different platform architectures (e.g., algorithmic curation, and creator tool design) influence homogenization rates would provide valuable design guidance. Finally, intervention studies testing fatigue-aware audio design modifications, such as dynamic range preservation, suggested pause periods, or acoustic diversity alerts, could establish evidence-based guidelines for healthier platform sound ecologies.

Funding: This work received no external funding.

Institutional review board statement: Not applicable.

Informed consent statement: Not applicable.

Data availability statement: All data is included within the article.

Conflict of interest: The author declares no conflict of interest.

AI use statement: The author declares that no artificial intelligence (AI) tools were used in the preparation of this manuscript.

References

1. Zeng C. Unveiling the impact of short videos: Consumption behavior and decision-making in the digital age. Highlights in Business, Economics and Management. 2023; 21: 469–474.
2. Yang Q, Wang Y, Song M, et al. Sonic strategies: Unveiling the impact of sound features in short video ads on enterprise market entry performance. Journal of Business-to-Business Marketing. 2025; 32(1): 95–116.
3. Yang Q, Wang Y, Wang Q, et al. Harmonizing sight and sound: The impact of auditory emotional arousal, visual variation, and their congruence on consumer engagement in short video marketing. Journal of Theoretical and Applied Electronic Commerce Research. 2025; 20(2): 69.
4. Wang Y, He Z. Standardization of music loudness in the new media era. In: Proceedings of the 2023 IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology (WI–IAT); 26–29 October 2023; Venice, Italy.
5. Zhang D, Chen X, Liu X, et al. Evaluation of the sound environment in a metro platform space based on audio-visual interaction experiments. Tunnelling and Underground Space Technology. 2025; 159: 106476.
6. Warp R, Zhu M, Kiprijanovska I, et al. Moved By Sound: How head-tracked spatial audio affects autonomic emotional state and immersion-driven auditory orienting response in VR Environments. In: Proceedings of the AES Europe Spring 2022, 152nd Audio Engineering Society (AES) Convention; 7–8 May 2022; The Hague, The Netherlands; 16–19 May 2022, Online.
7. Saroka V. The role of sound in the immersive experience. Avant. 2024; 14(3).
8. Kääriä P. Designing Immersive and Adaptive Wind Soundscapes for Interactive Media: Workflow Strategies for Recording, Stylization, and Spatialization in Wwise [Bachelor's thesis]. JAMK University of Applied Sciences; 2025.
9. Filimowicz M. Designing Interactions for Music and Sound. CRC Press; 2022.
10. Lakhani N. Musical Acoustics Fundamentals. Educohack Press; 2025.
11. Geng Y. Exploring the Sonic Identity of Popular Music: Detailed Characteristics and Expressions. Journal of Humanities, Arts and Social Science. 2025; 9(7).
12. Cubides Gutierrez JM. Developing Compositional Methods by the Integration of Popular Music Creation Techniques into the Electroacoustic Music Practice [PhD thesis]. University of West London; 2024.
13. Hutchinson JC, Karageorghis CI, Jones L. See hear: Psychological effects of music and music-video during treadmill running. Annals of Behavioral Medicine. 2015; 49(2): 199–211.
14. Mutlu Aİ, Yüksel M. Listening effort, fatigue, and streamed voice quality during online university courses. Logopedics Phoniatrics Vocology. 2025; 50(1): 11–18.
15. Dobrucki A, Kin MJ, Kruk B. Various aspects of auditory fatigue caused by listening to loud music. Advances in Clinical Audiology. 2017; 1: 167–185.
16. Mas L, Bolls P, Rodero E, et al. The impact of the sonic logo's acoustic features on orienting responses, emotions and brand personality transmission. Journal of Product & Brand Management. 2021; 30(5): 740–753.
17. Vrhovnik MM, Prislan R. Recent Developments in Investigating and Understanding Impact Sound Annoyance—A Literature Review. Acoustics; 2025; 7(2): 21.
18. Erfanian M, Mitchell AJ, Kang J, et al. The psychophysiological implications of soundscape: A systematic review of empirical literature and a research agenda. International Journal of Environmental Research and Public Health. 2019; 16(19): 3533.
19. Zhang B. Effect of Soundscape on Emotional Response in an Urban Acoustical Environment [PhD thesis]. University of Sheffield; 2021.
20. Weisser A, Buchholz JM, Keidser G. Complex acoustic environments: Review, framework, and subjective model. Trends in Hearing. 2019; 23.
21. Yang Q. Effects of Sound Masking Noise on Workers' Perception and Performance [PhD thesis]. Concordia University; 2024.

22. Rajas-Fernández M, Gértrudix-Barrio M, Baños-González M. Knowledge in images and sounds: Informative, narrative and aesthetic analysis of the video for MOOC. Publications. 2021; 9(3): 32.
23. Cavegn M, Douglas R, Akkermans G, et al. Study of an ultrasound-based process analytical tool for homogenization of nanoparticulate pharmaceutical vehicles. Journal of Pharmaceutical Sciences. 2011; 100(8): 3374–3385.
24. Mei CC, Auriault JL, Ng CO. Some applications of the homogenization theory. Advances in Applied Mechanics. 1996; 32: 277–348.
25. Reed L. Empowering Film Sound Practice: Countering Visual Hegemony and Industrial Ideology with Reference to the Short-Film 'Fade' [PhD thesis]. University of Huddersfield; 2015.
26. Cubides-Gutierrez J. Electroacoustic Pop: Developing Compositional Methods by the Integration of Popular Music Creation Techniques into the Electroacoustic Music Practice [PhD thesis]. University of West London; 2023.
27. Skjerseth AE. Music's Visual Waves: Popular Music Technology and Audiovisual Aesthetics. The University of Chicago; 2022.
28. Vinokurtseva A, Akhmedova A. The impact of short-form videos on the future of musical clips. Journal of Sciences.. 2025; 2(12 (93)): 2024–2034.
29. Chattopadhyay B. Auditory Setting: Environmental Sounds in Film and Media Arts. Edinburgh University Press; 2021.
30. Salselas I, Penha R, Bernardes G. Sound design inducing attention in the context of audiovisual immersive environments. Personal and Ubiquitous Computing. 2021; 25(4): 737–748.
31. Heikkilä J. An Ode to Listening: Sound Diegesis and the Perception of Space in Linear and Non-Linear Narrative [Master's thesis]. Aalto University; 2022.
32. Antons JN, Schleicher R, Arndt S, et al. Too tired for calling? A physiological measure of fatigue caused by bandwidth limitations. In: 2012 Fourth International Workshop on Quality of Multimedia Experience; 5–7 July 2012; Melbourne, Australia.
33. Mahardhika J. Cognitive Load in Auditory-Only Narrative Processing: A Theoretical Framework for Duration Limits in Audio Cinema. Available online: https://philarchive.org/archive/MAHCLI
34. Rankin CH, Abrams T, Barry RJ, et al. Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. Neurobiology of Learning and Memory. 2009; 92(2): 135–138.
35. Dissegna A, Turatto M, Chiandetti C. Context-specific habituation: A review. Animals. 2021; 11(6): 1767.
36. Schmid S, Wilson DA, Rankin CH. Habituation mechanisms and their importance for cognitive function. Frontiers in Integrative Neuroscience; 2015; 8: 128974.
37. Miller M. Exploring the Relaxation Potential of Nature Soundscapes and Photography through Field Recordings and Meditation Nature Walks [PhD thesis]. University of Wolverhampton; 2025.
38. Sharma A. A Comparative Analysis of Analog and Digital Audio: Technologies, Characteristics, and Applications. In: Kashyap SV, Bajaj S, Bagchi S, et al. (editors). The Media Mosaic: Exploring Diverse Artistic Forms. Innovative Scientific Publication; 2023. pp. 61–65.
39. Qiao X. Soundscape Ecology and Posthumanist Music Practice: An Approach towards Ecological Healing. Journal of Dharma. 2025; 50(2).
40. Long Z, Yap JH, Koning SI. Harmonising Tradition and Technology: A Review of Multimedia Integration in Guizhou's Vocal Music. Asian Pendidikan. 2024; 4(1): 14–22.
41. Delikara A. Implementation approaches and strategies in the use of equalizer in music production: A systematic review. RAST Muzikoloji Dergisi. 2025; 13(3): 329–353.
42. Andringa TC. Appraising the Sonic Environment: A Conceptual Framework for Perceptual, Computational, and Cognitive Requirements. Behavioral Sciences. 2025; 15(6): 797.
43. Chin S. Towards Perceptual Augmentation [PhD thesis]. Massachusetts Institute of Technology; 2024.
44. Wu T, Lin H, Wang H, et al. Music as a tourism soundscape: Music scape in sensory marketing and destination experience design. Current Issues in Tourism. 2025; 1–23.
45. Rosenweg PN. Evaluation of a Self-Regulating Skill for Rapid Attention Recovery (rART) [PhD thesis]. Swinburne University of Technology; 2018.
46. Li J, Wang Y, Lin Y, et al. The power of sound: The impact of auditory features in hotel short influencer-generated videos on viewer engagement. International Journal of Hospitality Management. 2025; 131: 104341.
47. Merchant ND, Firstrup KM, Johnson MP, et al. Measuring acoustic habitats. Methods in Ecology and Evolution. 2015; 6(3): 257–265.
48. Abramowitz JC, Goupell MJ, DeRoy Milvae K. Cochlear-implant simulated signal degradation exacerbates listening

effort in older listeners. Ear & Hearing. 2024; 45(2): 441–450.
49. Smith ML, Winn MB. The influence of momentary listening effort on accumulated fatigue. Journal of the Acoustical Society of America. 2025; 157(4_Supplement): A119.
50. Levey K. The Role of Platform Design in Excessive Short Form Video Use and the Resulting Impacts on Cognitive Processes [Bachelor's thesis]. University of South Carolina; 2025.
51. Akbarzadeh S, Lobarinas E, Kehtarnavaz N. Online personalization of compression in hearing aids via maximum likelihood inverse reinforcement learning. IEEE Access. 2022; 10: 58537–58546.
52. Mahmoudi-Nejad A, Guzdial M, Boulanger P. Arachnophobia exposure therapy using experience-driven procedural content generation via reinforcement learning (EDPCGRL). In: Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment; 11–15 October 2021; Online. pp. 164–171.
53. Maystre L, Russo D, Zhao Y. Optimizing audio recommendations for the long-term: A reinforcement learning perspective. arXiv preprint; 2023. doi: 10.48550/arXiv.2302.03561
54. Jones RH. The text is reading you: Teaching language in the age of the algorithm. Linguistics and Education. 2021; 62: 100750.
55. Li S, Liu G, Hu P, et al. Acoustic sensing for contactless health monitoring: Technologies, algorithms, and emerging applications. IEEE Access. 2025; 13, 163907–163935.
56. Mobin S. Interpreting Nonverbal Cues Through Sound: Exploring Accessibility for BLV Individuals in Online Communication Using Cinematic Media and Sound Design [PhD thesis]. OCAD University; 2025.
57. Natalevich D. The Sensuality of Sound. Sound Design Concepts, Strategies and Techniques Developed and Used in the Short Film "Closure" [Master thesis]. Tartu University Viljandi Culture Academy; 2016.
58. La Rosa M. Contemporary Post-Production: Create, Cut, Collaborate, Color, Deliver. Routledge; 2025.
59. Snyder J. The Visual Made Verbal: A Comprehensive Training Manual and Guide to the History and Applications of Audio Description. Æ Academic Publishing; 2020.
60. Chen CJ, Dai YT, Sun YM, et al. Evaluation of auditory fatigue in combined noise, heat and workload exposure. Industrial Health. 2007; 45(4): 527–534.
61. Mehrotra A, Shukla SP, Shukla AK, et al. A comprehensive review of auditory and non-auditory effects of noise on human health. Noise and Health. 2024; 26(121): 59–69.
62. Xu C. The Determinants of Creator Performance on Creative Content Platforms: Evidence from Xiaohongshu and Bilibili [PhD thesis]. Singapore Management University; 2023.
63. Brown N. Rationalist Empiricism: A Theory of Speculative Critique. Fordham University Press; 2021.
64. Peña M. Robust clustering methodology for multi-frequency acoustic data: A review of standardization, initialization and cluster geometry. Fisheries Research. 2018; 200: 49–60.
65. Trillò T. "PoV: You are reading an academic article." The memetic performance of affiliation in TikTok's platform vernacular. New Media & Society. 2025; 28(1): 380–400.
66. Gao K, Van Dommelen J, Geers M. Investigation of the effects of the microstructure on the sound absorption performance of polymer foams using a computational homogenization approach. European Journal of Mechanics-A/Solids. 2017; 61: 330–344.
67. Proppe DS, Sturdy CB, St. Clair CC. Anthropogenic noise decreases urban songbird diversity and may contribute to homogenization. Global Change Biology. 2013; 19(4): 1075–1084.
68. Gustafson SJ, Key AP, Hornsby BW, et al. Fatigue related to speech processing in children with hearing loss: Behavioral, subjective, and electrophysiological measures. Journal of Speech, Language, and Hearing Research. 2018; 61(4): 1000–1011.
69. Hornsby BW, Naylor G, Bess FH. A taxonomy of fatigue concepts and their relation to hearing loss. Ear and Hearing. 2016; 37: 136S–144S.
70. Laufs C, Herweg A, Antink CH. Methods and evaluation of physiological measurements with acoustic stimuli—A systematic review. Physiological Measurement. 2023; 44(11): 11TR01.
71. Alhanbali S, Dawes P, Lloyd S, et al. Self-reported listening-related effort and fatigue in hearing-impaired adults. Ear and Hearing. 2017; 38(1): e39–e48.
72. Wang Y, Naylor G, Kramer SE, et al. Relations between self-reported daily-life fatigue, hearing status, and pupil dilation during a speech perception in noise task. Ear and Hearing. 2018; 39(3): 573–582.
73. Hornsby BWY, Gustafson SJ, Lancaster H, et al. Subjective fatigue in children with hearing loss assessed using self-and parent-proxy report. American Journal of Audiology. 2017; 26(3S): 393–407.
74. Nazarchuk Z, Skalskyi V, Serhiyenko O. Acoustic Emission: Methodology and Application. Springer; 2017.

75. Holford KM, Eaton MJ, Hensman JJ, et al. A new methodology for automating acoustic emission detection of metallic fatigue fractures in highly demanding aerospace environments: An overview. Progress in Aerospace Sciences. 2017; 90: 1–11.
76. Islam MA, Haque EA. Advances and Limitations of Fracture Mechanics–Based Fatigue Life Prediction Approaches for Structural Integrity Assessment: A Systematic Review. American Journal of Interdisciplinary Studies. 2022; 3(3): 68–98.
77. de la Prida D, Larrosa-Navarro M, Azpicueta-Ruiz LA, et al. On the need for homogeneous and standardized methodologies for perceptual testing in acoustics. In: Proceedings of the 11th Convention of the European Acoustics Association Forum Acusticum/EuroNoise 2025; 23–26 June 2025; Málaga, Spain. pp. 3505–3511
78. Venet T, Thomas A, Merlen L, et al. Parameters influencing auditory fatigue among professionals working in the amplified music sector: Noise exposure and individual factors. International Journal of Audiology. 2024; 63(9): 686–694.
79. Moessinger M, Stürmer R, Mühlensiep M. Auditive beta stimulation as a countermeasure against driver fatigue. PLoS ONE. 2021; 16(1): e0245251.
80. Venkova J, Nkpoikanna E. Optimizing User Experience across Platforms: A Study on Enhancing Music-Streaming Applications' Cross-Device Unification [Bachelor's thesis]. Jönköping University; 2024.
81. Riionheimo J. Hearing as Intended: How Differences in Listening Conditions Affect Sound Translation [PhD thesis]. Aalto University; 2025.
82. Mitchell A, Erfanian M, Soelistyo C, et al. Effects of soundscape complexity on urban noise annoyance ratings: A large-scale online listening experiment. International Journal of Environmental Research and Public Health. 2022; 19(22): 14872.
83. Kamehkhosh I, Bonnin G, Jannach D. Effects of recommendations on the playlist creation behavior of users. User Modeling and User-Adapted Interaction. 2020; 30(2): 285–322.
84. Sreedevi N, Saranya V, Ramudu K, et al. A Privacy-Conscious Federated Reinforcement Learning Framework for Affect-Aware English Listening. International Journal of Advanced Computer Science & Applications. 2026; 17(1).
85. Altomonte S, Schiavon S, Kent MG, et al. Indoor environmental quality and occupant satisfaction in green-certified buildings. Building Research & Information. 2019; 47(3): 255–274.
86. Kalchev E. Beyond the sound waves: A comprehensive exploration of the burn-in phenomenon in audio equipment across physiological, psychological, and societal domains. Cureus. 2024; 16(1).
87. Gorichanaz T, Schiavon S, Kent M.G, et al. The impacts of online experience on health and well-being: The overlooked aesthetic dimension. Virtual Worlds; 2023; 47(3): 255–274.
88. Hagemann R, Huddleston Skees J, Thierer A. Soft law for hard problems: The governance of emerging technologies in an uncertain future. Colorado Technology Law Journal. 2018; 17: 37.
89. Elden B. Human-Artificial Intelligence Collaboration Fatigue: Causes, Consequences, and Mitigation Strategies. Socrates Journal of Interdisciplinary Social Researches; 11(59): 150–160.
90. Trivedi B. Avoiding personalization fatigue: A Heuristics-Based Framework for Retail Digital commerce. Journal of Information Systems Engineering and Management. 2025; 10: 263–273.
91. Wang Y. Fatigue-Aware Adaptive Interfaces for Wearable Devices Using Deep Learning. arXiv preprint; 2025. doi: 10.48550/arXiv.2506.13203
92. Karim R, Guo X, Wu H, et al. Advancing Research on Workers' Fatigue. Physical Ergonomics and Human Factors. 2025; 185: 77.
93. Chiaraviglio L, Cuomo F, Listanti M, et al. Fatigue-aware management of cellular networks infrastructure with sleep modes. IEEE Transactions on Mobile Computing. 2017; 16(11): 3028–3041.
94. Mike AA, Paul A, Ntwari R. Human Factors Shaping Cybersecurity Behavior in Work from Home Environment in Ugandan Universities. International Journal of Research and Innovation in Applied Science; 2026; 11(1): 1487–1500.