




ARTICLE



<https://doi.org/10.1057/s41599-022-01461-5>

OPEN

Building blocks of suspense: subjective and physiological effects of narrative content and film music

Gary Bente¹, Kevin Kryston², Nolan T. Jahn¹  [✉] & Ralf Schmäzle¹

The current study explores the role of narrative content and non-diegetic music, as well as their interplay in the creation of film suspense. In a between-subjects design, three audience groups watched either the full version (audio-video) of a suspenseful short movie (*Love Field*) or a video-only version or listened to the audio track only. Audience responses were assessed with a combination of continuous response measurement (CRM) and psychophysiological measures, comprising heart rate (i.e., inter-beat interval, IBI), pulse volume amplitude (PVA), and skin conductance level (SCL). Frame-by-frame content coding was performed to identify distinct plot segments and musical moods (tense vs. relaxing) and mark the critical visual and auditory events that directed the audience's inferences about the nature of the plot and elicited specific outcome expectations. Results showed that continuous self-report data and objective physiological measures were largely dissociated, suggesting different processing modalities. Tense music alone induced feelings of suspense (CRM) even in the absence of any clues about the story content. Overall, the audio-only version led to the highest arousal levels, as indicated by SCL and PVA, while the video-only version led to the lowest arousal levels. IBI data revealed short-term heart rate deceleration responses to salient narrative clues that could be interpreted in terms of heightened attention and cognitive resource allocation. Results are discussed in light of a multidimensional framework of tension and suspense and a terminological differentiation of both concepts.

¹Michigan State University, East Lansing, MI, USA. ²Sam Houston State University, Huntsville, TX, USA. ✉email: jahnola@msu.edu

Introduction

Suspense is a crucial factor in audience involvement with dramatic narratives, but how suspense is created, and the underlying psychological processes remain insufficiently understood (Bálint et al., 2017). In the context of media psychology, suspense is characterized by uncertainty and psychological tension that results from worries about the fate of a valued protagonist and the dwindling hope for a desirable story ending (Carroll, 1996; Vorderer and Knobloch, 2000; Zillmann, 1980, 1996). However, movie directors and media scholars agree that beyond factual plot-level information, formal presentation features, such as sound, cinematography, and specifically so-called *nondiegetic music*¹ can play an important role in the establishment and maintenance of film suspense (Bullerjahn and Güldenring, 1994). To include these variables into a broader concept of suspense, Lehne and Koelsch (2015) recently defined “... tension and suspense as affective states that (a) are associated with conflict, dissonance, instability, or uncertainty, (b) create a yearning for resolution, (c) concern events of potential emotional significance, and (d) build on future-directed processes of expectation, anticipation, and prediction” (p. 2). Importantly, these processes can be brought about by narrative content as well as formal features, such as music.

Indeed, there is ample evidence that music alone—at least Western tonal music—can induce tension and elicit feelings of suspense comparable to those experienced during a captivating narrative (Lehne and Koelsch, 2015). In brief, research shows that musical properties such as contour, harmony, tempo, pitch, volume, etc.—can raise expectations about how a tonal sequence will continue, which can lead to tension and suspense if these expectations are uncertain or if their fulfillment is suspended (Bannister and Eerola, 2018; Farbood, 2012; Pearce and Wiggins, 2012; Pressnitzer et al., 2000; Steinbeis et al., 2006). These findings suggest that musical tension and narrative suspense may share some degree of generative mechanisms. Specifically, both plot-level information and non-diegetic film music involve incomplete event sequences with uncertain continuation or outcomes that invoke anticipatory processes, which in turn prompt psychological tension and arousal. However, empirical research on these issues is scarce because limited work has examined suspense in the context of film and emotional phenomena more broadly (Tan, 2018). Therefore, it remains unclear how both components contribute to suspense, and how they interact when combined in film. The current study addresses these questions in a direct comparison of a full version of a suspenseful short movie with an audio-only and a video-only version of the same film.

The study is based on two major assertions. The first assertion holds that the effects of film music are context specific. It is well known that film directors and composers artfully tailor film music to create, amplify or sustain suspense, and the musical features used, and effects induced likely differ depending on the respective narrative context. For instance, tension can result from different levels of uncertainty and varying audience expectations during the unfolding film narrative. Lehne (2014) notes: “In its most diffuse form, tension can just arise from the expectancy that “something” significant will happen. Depending on what then actually happens, the tension resolves into positive or negative emotions” (p. 133). In this sense, music can be the sole driver of suspense when images are neutral or even relaxing (e.g., a nature walk), frame ambiguous narrative cues and events, or amplify narrative suspense when aligned with emotionally challenging content (Cohen et al., 2015). To identify critical narrative events and plot phases that represent different levels of context-based uncertainty and tension the current approach consequently includes a frame-by-frame content coding procedure.

The second assertion holds that suspense is a dynamic process that includes bodily responses experienced such as chills, tension, and arousal. We further assert that suspense cannot be fully understood without assessing these underlying physiological processes. Research has shown that audiences easily agree upon what is suspenseful and what is not, whether they judge narratives in literature and film or music (Lehne and Koelsch, 2015) or combinations of both (Bezdek et al., 2017). It is important to note that there might be differences in the judgment of tension or suspense as stimulus features and the judgment of the feelings one experiences while exposed to visual or auditory stimuli (see, for instance, Gabrielsson, 2002; Lehne and Koelsch, 2015). Yet, independent of whether audiences are instructed to rate features or feelings, judgments are always at least partially dependent on prior media exposure and acquired genre knowledge (Steinbeis et al., 2006), and, as a result, do not always accurately reflect their spontaneous feelings or bodily sensations. As such, we assessed subjective self-reports and physiological processes independently from each other to identify convergent and divergent response patterns in both domains. Moreover, we asked which physiological parameters are most suitable to capture differential bodily response patterns associated with the relevant stimulus features (narrative content vs. non-diegetic music) and how these relate to deliberate audience judgments.

Physiological correlates of suspense. Physiological audience responses have been documented for suspenseful narratives in film and literature (Chun et al., 2020; Lehne and Koelsch, 2015; Zillmann et al., 1975) as well as for the reception of tensional music (Bannister and Eerola, 2018; Krumhansl, 1997; Lehne et al., 2013). In terms of parameters, various physiological measures have been deployed in studies of suspense, including heart rate (inter-beat interval), electrodermal activity (skin conductance level or galvanic skin response), and, less frequently, skin temperature and pulse volume amplitude, which indicate vasoconstriction and vasodilation. For instance, Nomikos et al. (1968) found that ambiguous narratives elicited increased skin conductance levels (SCL) over time. Likewise, prolonged, and discrete increases in SCL have been found to occur when (a) narratives create uncertainty and ambiguity (Sparks, 1989) and (b) when that uncertainty is suddenly resolved (Roberts and Hoetzel, 2007). Corroborating the role of uncertainty in physiological audience responses, Chun et al. (2020) showed that SCL systematically decreased during repeated exposure to the same suspenseful film clips, indicating lower arousal when the ending was known.

Zillmann et al. (1975) found that heart rate (HR) commonly rises during exposure to suspenseful media stimuli and drops when narrative suspense is resolved. The authors also found that the resolution of suspense led to an increase in skin temperature (vasodilation), also suggesting lower arousal. However, Hubert and deJong-Meyer (1991) found that film suspense was associated with increased skin conductance, but also heart rate deceleration, a pattern often associated with attention and cognitive effort (Fisher et al., 2018; Lang, 2000, 2006; Potter et al., 2018). However, in the study by Chun et al. (2020), there was no difference in heart rate during first- and second-time exposure despite the expectation of dwindling attention during re-viewing. The authors attribute this result to the fact that heart rate is a parameter with low specificity, as it is “...associated with multiple cognitive and emotional processes, such as attention, information encoding, emotional arousal, and positive/negative emotional responses” (p. 10). Recently, Baldwin and Bente (2021) applied pulse volume amplitude (PVA), which is used less often in film studies, to analyze audience responses to suspenseful scenes in

two different films in a drama series (*Rocky, Rocky II*). They found PVA to be more sensitive than SCL or heart rate, respectively, for different types of suspense.

In the domain of music perception, Krumhansl (1997) showed that the experienced degree of tension (amongst other emotions) when listening to different musical excerpts, was associated with lower finger temperature and lower pulse volume amplitude (PVA) reflecting vasoconstriction, an index of sympathetic arousal. Counterintuitively, though, the author reported a negative correlation between tension and SCL, indicating that higher perceived tension corresponded with lower SCL. Heart rate, on other hand, was uncorrelated with musical tension in this study. Similarly, several studies found that “chill-inducing” music was associated with decreases in PVA (Benedek and Kaernbach, 2011; Salimpoor et al., 2009; Salimpoor et al., 2011). In these studies, lower PVA was associated with lower SCL (lower arousal), a pattern that resembles the findings reported by Krumhansl (1997), but also with faster heart rate. Bannister and Eerola (2018), on the other hand, found a significant drop in the frequency of phasic skin conductance responses (lower arousal) when tensional sequences were removed from musical pieces. In sum, just like for narrative suspense, findings on musical tension reveal a somewhat mixed pattern of results across different physiological parameters.

Few studies have directly analyzed the interplay of film content and music or compared their relative impact on subjective experiences. Perhaps the most often-cited work in this area is that of Thayer and Levenson (1983), who investigated the effects of music on physiological responses to a stressful film (workplace accidents). They found that electrodermal activity was highest when the film was shown with suspenseful music compared to no music or relaxing music. Using a continuous self-report measure (CRM), Bezdek et al. (2017) examined the influence of congruent or incongruent tensional music on the perception of various emotional film clips. Results showed that feelings of suspense were significantly elevated when the film clips were presented with congruent, tensional music compared to presentations with incongruent or no music. Accompanying brain imaging data also suggested that “...perceptual, attentional, and memory processes respond to the suspense on a moment-by-moment basis” (p. 73). As there was no audio-only condition in this study, specific contributions of music to suspense remained unclear. In another study, Aly et al. (2017) found that suspenseful audio features alone led to increases in heart rate, whereas video features alone did not. Again, these distinct response patterns might be related to the possibility that HR acceleration can be associated with arousal, while HR deceleration can also be associated with attention and cognitive effort (Mullen et al., 2012). Furthermore, deceleration may be more pronounced when processing visual content compared to music.

In conclusion, at this stage of research, it is difficult to clearly predict how narrative content and non-diegetic music would affect physiological and subjective responses to suspenseful films. Physiological responses appear to vary across stimulus features as well as parameters. PVA consistently responds to suspenseful narratives or tensional music, whereas SCL and HR measures show contradictory results. Regardless, these different physiological measures are likely uniquely sensitive to different sensory modalities and stimulus components and corresponding cognitive and affective audience responses.

The current study

To identify the relative impact of narrative content and non-diegetic film music on felt suspense, the current study compares the subjective and physiological audience responses to a

suspenseful short movie with respective responses to an audio-only and a video-only version of the same film in a between-subjects design. The study builds on the notion that suspense is not a linearly progressing phenomenon, but rather co-varies with specific narrative events, different plot segments, or stages of the narration. In line with Alwitt (2002), who stated, “suspense resides in both the global plot of the narrative as well as in the events that comprise the narrative” (p. 36, Zillmann, 1996), the current study aimed to identify physiological patterns associated with (1) local cognitive and affective responses to specific narrative events and (2) cascading or enduring changes in arousal levels related to the overall narrative arc or extended plot phases. We also aimed (3) to study the specific influence of non-diegetic music on these responses. A frame-by-frame annotation of the stimulus material served as a basis to define the respective units of analysis. A combination of continuous self-reports (CRM), heart rate (HR), skin conductance level (SCL), and pulse volume amplitude (PVA) were applied to study interdependencies between subjective impressions of suspense and their bodily correlates. This combination has not been used in the study of suspense so far but has recently proved particularly informative to disentangle cognitive and affective processes during entertainment media consumption (Baldwin and Bente, 2021). Importantly, inexpensive sensors can record responses to media stimuli in naturalistic settings, and their widespread availability and decreasing cost offer new opportunities for psychophysiological research on entertainment effects, especially suspense.

Method

Participants. Students from a large US university ($N=105$; 73.3% female; mean age = 20.33) participated in the study for extra course credit. All procedures were approved by the university’s institutional review board and all participants provided informed consent before participating.

Stimulus material. The experimental stimulus was the suspenseful short movie *Love Field* (Raithe, 2008). *Love Field* is a 5:30 min long film that contains the key plot elements *diffuse tension, story framing, anticipatory suspense, and resolution/relief*. To disentangle the specific effects of content and music on film suspense, we compare the continuous audience responses to the full version of a suspenseful stimulus film to the responses to the separately presented audio and video track of the same film.

The plot of *Love Field* is roughly as follows: The film starts with a tracking shot of a wheat field. The camera then follows a beeping sound and shows a mobile phone lying on the ground. An open wallet and spread bills and a bloody handkerchief appear. A whimpering female voice then further promotes the suspicion that something malicious is happening here. Then, we see the convulsing foot of the woman and finally a male hand sticking a knife into the ground. As these cues are introduced, the audience forms more specific expectations about a bad ending. A farmer appears and runs to his car, searching for a blanket as if to cover up his deed. As he returns to the woman, however, we learn that the farmer was helping a woman to give birth. Throughout the film, the image or sound of a crow serves as a recurrent theme between scenes. The two major building blocks of the film—the creation of suspense and its resolution—are accompanied by different types of music, that were characterized by an expert (graduate student with music education) as tensional music (musical gestures, dissonant intervals, abrupt cut-offs) and cheerful music (melodic, soft open chords), respectively. A detailed description can be found in the Supplement Materials.

Three stimulus versions were produced as MP4 clips: a full audio-visual version, an audio-only (black screen), and a video-



Fig. 1 Experimental setup. The physiological sensors worn by participants and stimulus screen.

only version (no sound). All versions were appended to the same 2'4" relaxation film, showing different nature scenes accompanied by relaxing music, that served for baseline measurement (Baldwin and Bente, 2021).

Stimulus annotation and plot phase identification. Four graduate student annotators received two hours of training to identify the onset of the occurrence of objects, people, or sounds in the film with video frame accuracy. Then, each independently performed the time-series annotation of the narrative event structure by manually recording the time stamp for the appearance and change of a new object, person, and sound, respectively. After completing their annotation independently, all four and the study coordinator met to referee and settle disagreements to create the final annotation. Only minimal discrepancies in the number of the marked events and the respective timestamps occurred and were settled in a joint review process. Based on the event annotation, the researchers further delineated four major plot phases and included their start and end times in the annotation protocol. These phases were named: (1) *diffuse tension* (see Lehne and Koelsch, 2015), in which any visual clues to the storyline are absent (tracking shot across a wheat field) and the audience experience is exclusively driven by sound and music, (2) *story framing* (see Tan et al., 2008), in which tensional music accompanies specific events (mobile phone, money, blood, knife) that give direction to audiences' interpretation of the plot (violent crime), (3) *anticipatory suspense* (see Cohen et al., 2015), in which the audience is interpreting the events in this narrative framework and music might amplify the affective responses, and finally (4) *resolution/relief* (Lehne, 2014; Zillmann et al., 1975), in which narrative uncertainty is efficiently resolved (baby born, happy faces, laughter) and audience's relief is supported by cheerful music. Event and phase annotations were used as a basis for the time-related analysis procedures.

Audience response measures

Continuous response measure (CRM). Participants' self-reported felt suspense was continuously assessed using a 9-point slider

rating displayed on the stimulus screen to the right side of the video window (Fig. 1). The instructions were "Please continuously indicate the level of suspense you feel while watching the short movie", respectively "Please continuously indicate the level of suspense you feel while listening to the musical sequence". The rating scale ranged from -4 (low suspense) to $+4$ (high suspense) and could be adjusted with the keyboard's up-down keys. The nine dots of the scale indicating the suspense level were colored, with a yellow dot signifying the neutral scale middle, up to four red dots (highest suspense), and down to four blue dots (lowest suspense).

Psychophysiological measures. Physiological data were acquired using a *Lightstone iom1* device² consisting of three finger sensors, a preprocessing unit, and a USB2-transmission module (Fig. 1). The device features a photoplethysmographic (PPG) sensor (applied to the middle finger of the non-dominant hand) to assess finger pulse and two electrodes (ring and pointer finger of the same hand) to assess skin conductance level (SCL). Pulse volume amplitude (PVA) and inter-beat interval (IBI; the inverse of heart rate) were extracted from the PPG signal. The *iom1* device was developed as a commercial system for relaxation exercises.

Data recording. Physiological data were recorded at a frequency of 30 Hz using a custom C++ program based on the *iom1* SDK. A python program that displayed the stimulus and captured CRM data ran in parallel with the physiological recording program via the Windows *PostMessageA* function (*winuser.h*), continuously sending CRM data and event codes to mark the start and the end of the stimulus. Physio data, CRM data, and event markers were stored in CSV format by the C++ program with a shared time code for all channels.

Procedure. Upon arrival, participants were introduced to the purpose of the study and its general procedures. Participants were randomly assigned to one of the three conditions: audio-visual (AV), video-only, or audio-only. They sat in a cushioned chair, placed inside a booth, about 60–70 cm in front of a 21" computer screen. Physiological sensors were attached to the non-dominant hand and participants were instructed to keep this hand still during the experiment. Then, participants were familiarized with the CRM rating procedure (see Fig. 1). After instruction the experimenter left the room, and the participants independently completed the experimental procedures. The film stimulus was presented via a custom python program that also handled the data collection, providing precise temporal synchronization stimulus display and response recordings (see below).

To acquire physiological baseline data, all participants viewed a two-minute and four-second relaxation video before the experimental stimulus started. This baseline video contained nature scenery and was accompanied by relaxing music (Baldwin and Bente, 2021). All participants experienced the same baseline stimulus. After the relaxation phase, participants in the *audio-visual* condition watched the original version of the film. Participants in the *video-only* condition watched a silent film version without background music or sound. Finally, participants in *audio-only* condition only listened to the audio track, comprising background music and sound effects, with a black screen replacing the video. Overall, the session lasted 7.5 min. Upon completion, participants were debriefed and received course credit confirmation.

Data preprocessing. The program *HeartPy* (<https://github.com/paulvangentcom/heartpy>) was used for peak detection in the raw PPG signal. Using a custom graphical

Table 1 By-condition Pearson correlations between the dependent variables across the whole movie.

		PVA	SCL	IBI
AV	SCL	−0.50***		
	IBI	0.05	−0.28***	
	CRM	−0.10	−0.16***	−0.24***
Video	SCL	−0.72***		
	IBI	−0.20***	0.13*	
	CRM	0.71***	−0.71***	−0.27***
Audio	SCL	−0.57***		
	IBI	0.45***	−0.321***	
	CRM	0.06	0.05	−0.20***

$p(df = 82) = * < 0.05, *** < 0.001$.

inspection and editing program (*Visual Basic 6.0*), the detected peaks were checked against the raw data curve to identify and manually correct misses or false detections. After this data-cleaning procedure, the inter-beat interval (IBI) was then calculated as the distance between the PPG signal peaks (in milliseconds). Pulse volume amplitude (PVA) was calculated as the amplitude difference between maximum at peak time and minimum between consecutive peaks. *Skin conductance level* (SCL) was preprocessed in the *iom1* device by applying a low pass filter, with unpublished constants

To account for individual differences in the physiological parameters we applied baseline correction to the individual datasets calculating the difference of all data points to the mean of the second minute of the relaxation phase. The first minute of relaxation was dropped to leave participants time to come to a rest. Further, *z*-transformation was applied to the individual data sets to adjust scales. Last, we applied a low pass filter with a constant of 5 Hz to eliminate high-frequency jitter, using the *sciPy* library. CRM data was used in raw format.

Results

Correlations between dependent measures. To identify inter-dependencies between the dependent variables within each different experimental condition (audio-visual, video-only, audio-only), we conducted two-tailed Pearson correlations for the averaged response time series. To reduce *p*-value inflation, time-series data were resampled from 25 to 1 Hz. Further, PVA and SCL time series were de-trended using linear regression to exclude the effects of overall drifts due to environmental conditions (e.g., room temperature). The alpha level for all tests in our manuscript was $p = 0.05$.

As expected, SCL and PVA showed consistent negative correlations across conditions, meaning vasoconstriction, in general, was associated with higher SCL, and vasodilation was associated with lower SCL (see Table 1). Thus, both measures can be consistently interpreted as arousal indexes. Correlations between IBI and PVA as well as IBI and SCL were unsystematic. While there was no significant correlation between IBI and PVA in the AV condition, there was a negative correlation in the video condition and a positive correlation in the audio condition. Also, SCL and IBI correlated negatively in the AV and audio-only conditions and positively in the video-only condition.

IBI was the only parameter that consistently correlated negatively with CRM across conditions, showing that the subjective experience of suspense was associated with lower IBI (faster pulse). In the video-only condition, all three physio parameters correlated significantly with CRM. Correlations between PVA, SCL, and CRM were strong, but interestingly not in the expected direction. PVA correlated positively with CRM,

suggesting that higher PVA (vasodilation) was associated with higher CRM values (more suspense), and vice versa. SCL correlated negatively with CRM, meaning the higher skin conductance level was associated with lower CRM ratings.

Differences between conditions and plot segments. Two-tailed repeated measure ANOVAs were conducted for all dependent variables with *stimulus condition* as the between-subject factor and the four *plot phases* as the within-subject factor. Analyses were based on averages of the individual data sets across the plot phases (Table 2). Furthermore, we conducted frame-by-frame (the MP4 movie had 25 FPS consistent with the data rate of 25 Hz) two-sided pairwise *t*-tests comparisons to identify the specific differences between all condition pairings and to inspect minute temporal dynamics of the audience response over time (Fig. 2). Results from ANOVA, *t*-tests, and graphical inspection are reported in conjunction below.

Comparisons of the subjective self-report data (CRM) revealed significant effects of plot segments and conditions as well as a significant interaction effect (Table 2). Frame-by-frame comparisons detailed in Fig. 2 show that the audience was immediately captivated already during the intro when the tensional music starts. This effect is also visible in the audio-only condition where audience responses appear to be exclusively driven by the music when visual clues are absent. This finding was corroborated by the data suggesting that CRM ratings for the video-only version stay at the lowest level for the entirety of *diffuse tension*. Self-reported suspense rose in the second phase (*story framing*) in tandem with six consecutive narrative clues that progressively nurture the suspicion of an ongoing violent crime. During this phase, subjective suspense (CRM) also built up in the audience of the video-only condition before finally reaching the response level of the audio-only and AV conditions at the end of the *story framing* phase. From there on, during the following phases of *anticipatory suspense* and *resolution/relief*, CRM levels of the three conditions are mostly identical. CRM ratings for all three conditions then show a marked drop in the phase of *resolution/relief* resulting in a significant difference from the phase of *anticipatory suspense*.

ANOVA results for SCL also reveal significant differences between conditions and plot phases as well as a significant interaction effect. As Fig. 2 indicates, the most pronounced and enduring differences occurred between the audio and the video version. While all conditions show a strong initial SCL response contingent on the stimulus onset, SCL in the video-only and AV condition dropped in the phase of *diffuse tension*. This drop was more pronounced for the video-only version. Conversely, SCL showed no significant changes in the audio-only condition. Consequently, only the video-only and audio-only conditions showed significant differences in sequential *t*-tests. In the AV condition, SCL started to rise in the *story framing* phase, converging with the audio condition, then peaking at the end of this phase in response to the last dramatic clue (bloody knife). After the stab, both the AV and the audio versions showed significantly higher SCL levels than the video-only version. SCL remained roughly unchanged across all conditions during the phase of *resolution/relief* (childbirth, happy faces) besides a slight increase in the AV condition.

ANOVA for PVA reveals a significant main effect for the condition only. Throughout the whole film, PVA in the video-only condition differed significantly from the other conditions, which both contain audio (Fig. 2). While PVA in the video-only condition remained slightly above the baseline level (i.e., relative calm; vasodilation) for the whole film, PVA indicated arousal (vasoconstriction) in the AV and audio-only conditions.

Table 2 ANOVAs for the segment means of the four measures: within-subject factor = film segment; between-subjects factor = condition (AV, video, audio).					
Measure	Factors	df	F	p	η^2
CRM	Segment	1.98, 196.25	178.99	<0.001	0.64
	Condition	2, 99	6.15	0.003	0.11
	Segment*Condition	3.97, 196.25	8.24	<0.001	0.14
SCL	Segment	1.77, 175.42	7.92	0.001	0.07
	Condition	2, 99	8.00	0.001	0.14
	Segment*Condition	3.54, 175.42	4.06	0.005	0.08
PVA	Segment	1.89, 187.23	1.23	0.30	0.01
	Condition	2, 99	6.41	0.002	0.12
	Segment*Condition	3.78, 187.23	1.21	0.30	0.02
IBI	Segment	2.51, 247.97	10.82	<0.001	0.10
	Condition	2, 99	1.01	0.37	0.02
	Segment*Condition	5.01, 247.97	0.49	0.82	0.01

Main effects of Segment and Segment-by-Condition interactions are Greenhouse-Geisser corrected for sphericity.

Similar to SCL, PVA also showed a marked local response in the AV condition and a moderate response in the audio-only condition to the critical event around minute 4.5 (bloody knife). In contrast to the permanent level change in SCL, after this event, PVA showed a short-lived decrease (vasoconstriction indicating tension) that started seconds before the event and a swift increase (vasodilation, indicating relief) immediately after the event. PVA rose in the audio-only condition slightly before the peak in the AV condition. In fact, the visual event was preceded by a dramatic peak in the tensional music followed by a moment of silence that was broken by the appearance of the ‘knife’ accompanied by a shrill sound.

ANOVA for IBI reveals only a significant main effect for the factor ‘segment’. This is mainly due to a lower IBI level (faster heart rate) in the phase of *anticipatory suspense* (Fig. 2). Overall, all three conditions show similar variation patterns, with only short moments of significant differences that coordinate with critical narrative events. To further explore these local responses, we conducted event-related analyses of the IBI responses.

Event-related analyses of heart-rate responses. A two-step procedure was applied in event-related analyses of the IBI data. First, we pursued a *stimulus-based approach* that focused on significant differences between conditions associated with the five critical events coded for the *story framing* segment of the film which gave direction to the audience’s expectations of a violent crime. Second, based on these events we analyzed the local IBI responses in the three groups, we pursued a *response-based approach* that further examined significant local IBI response differences between the groups and explored which contingent narrative events sparked said differences. Fig. 3 visualizes the event structure underlying these analyses.

Stimulus-based analysis. With regard to the *five-story framing* events (see Fig. 3), we submitted individual pre-post differences of the IBI means of the five-second intervals before and after the onset of the events to a repeated measurement ANOVA, with the between-subject factor *condition*, and the within-subject factor *event number*. Results (Greenhouse–Geisser corrected) revealed a significant, yet weak, effect of *event number* ($F(3.57, 353.38) = 4.46, p = 0.002; \eta^2 = 0.04$) as well as a medium-size significant effect for *condition* ($F(2,99) = 4.55, p = 0.01; \eta^2 = 0.08$; Cohen, 1988). No interaction effect was observed ($F(8,396) = 0.9, p = 0.52; \eta^2 = 0.02$). While the first two events did not show any significant effects, the pre-post difference drastically increased in the AV condition for the third event (*bloody fabric*) and then

stayed on level for the fourth event (*twitching foot*; Fig. 4). Both difference values were significantly different from zero as revealed by one-sample *t*-test (event #3: $t(33) = 3.62; p < 0.001$; event #4: $t(33) = 2.10, p = 0.02$). In the audio-only condition, a significant deviation from zero was only observable for event #4, i.e., when the *woman’s whimpering* turns into a *loud screaming* ($t(34) = 2.54, p = 0.02$). Overall, the events causing significant pre-post-differences were associated with heart rate acceleration.

Response-based analysis. Figure 4 also marks the significant IBI differences between conditions (circled areas). All these differences, except one, were contingent on visual events that introduced relevant narrative information. For every event, IBI responses in both visual conditions (AV and video-only) showed heart-rate deceleration. However, there was one exception: at the beginning of the story framing, an acoustic event relevant to the story framing (screaming of a woman) prompted significantly and pronounced heart-rate deceleration in both versions with audio (AV and audio-only). The first significant response difference occurred in the phase of *diffuse tension* when the camera moves from a tracking shot to a zoom-in motion to a spot close to the viewer’s virtual position. The second event that caused significantly different responses was the woman’s screams and the third was the knife stab, both in the *story framing phase*. During the phase of *anticipatory suspense*, all events are associated with the appearance of the crow and the farmer’s face. In the *resolution/relief* phase, the four marked response differences were associated with the appearance of human faces (woman, baby, farmer), as well as with the appearance of the police.

Discussion

The current study investigated the impact of narrative content (visual and auditory) and non-diegetic music on physiological audience responses and felt suspense during a short movie drama. Specifically, we compared the original version of a suspenseful short movie with an audio-only and a video-only version of the same film. We applied an integrated measurement approach combining a set of physiological measures comprising inter-beat-interval (IBI, equivalent to heart rate), pulse-volume-amplitude (PVA), and skin conductance level (SCL) with continuous self-report measures (CRM). Frame-by-frame content coding was performed to identify distinct plot segments and musical moods (tense vs. relaxing) and to mark the critical visual and auditory events that directed the audience’s inferences about the nature of the plot and elicited specific outcome expectations. Based on this coding four major film segments were identified: (1) *diffuse*

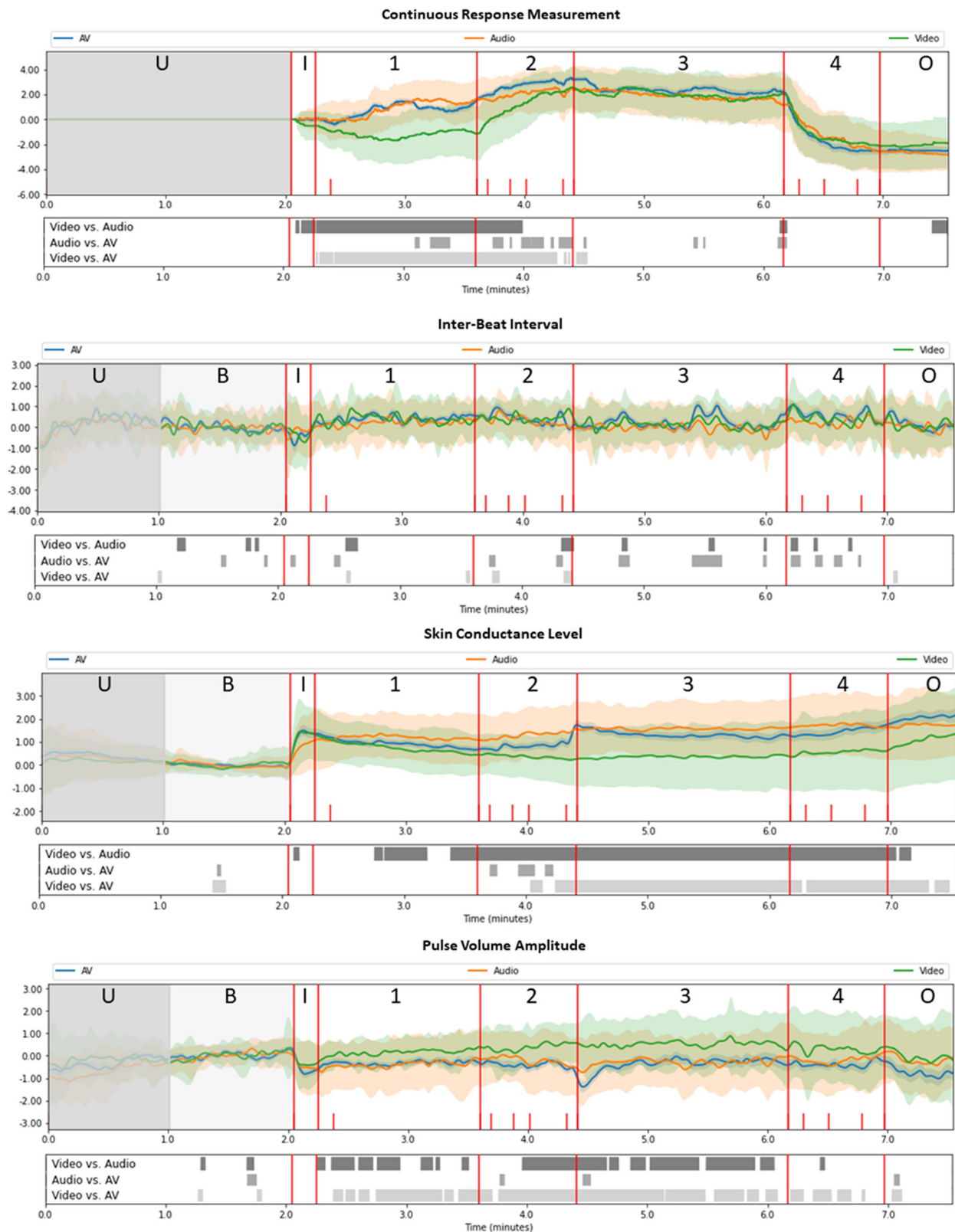


Fig. 2 Average curves and t-test comparisons of subjective and physiological audience responses under three stimulus conditions (AV: audio-visual; Audio: audio only; Video: video-only). For each dependent variable, the upper graphs show the filtered time series, averaged across participants with the standard error of the means as shaded areas around the mean. The vertical red lines mark the beginning and end of the major stimulus segments: U = unused; B = baseline; I = Intro; 1 = diffuse tension; 2 = story framing; 3 = anticipatory suspense; 4 = resolution/relief; O = Outro. The lower parts of each sub-graph show the results of the frame-by-frame pairwise t-test comparisons between the experimental groups ($p < 0.05$, uncorrected) are marked as gray blocks).

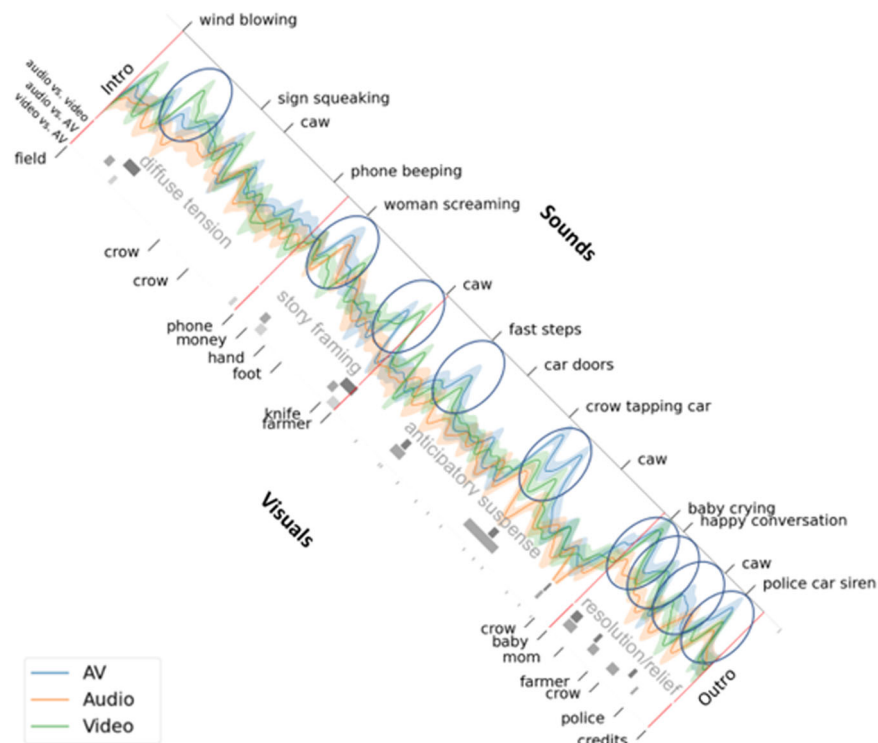


Fig. 3 Visualization of event structure, defined by the coding protocol (y-axis ticks for the events and gray labels for the plot phases) and the significant IBI differences between conditions (blue circles). The small letters on the minor primary x-axis indicate the appearance of the two recurrent themes: the crow (C) and the farmer (F).

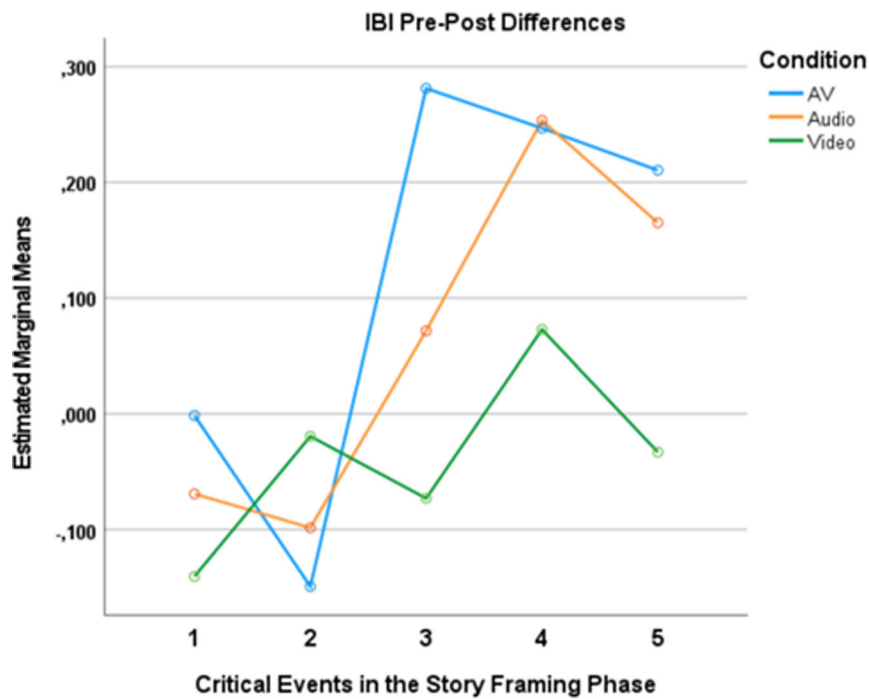


Fig. 4 Estimated marginal means of pre-post IBI differences (mean of 5 s before the critical event—mean 5 s after the event). Positive values indicate that IBI was lower after than before the event. Circles mark a significant difference from zero in the one-sample-t-test. Positive differences indicate that IBI was lower after the onset of the event than before, i.e., heart rate increased relative to the pre-value. See Supplementary Materials for a version of the figure that contains error bars.

tension (ambiguous visuals + tensional music), (2) *story framing* (clues to a violent crime + tensional music), (3) *anticipatory suspense* (ambiguous activities of the suspect + tensional music), and (4) *resolution/relief* (narrative twist from a crime to childbirth + relaxing music). The results suggest that non-diegetic music has a strong effect on physiological audience responses as well as deliberate self-reports and that the different measures exhibit a degree of divergence that suggests different processing pathways. The findings are discussed in depth below.

Self-reported suspense (CRM). Audiences' deliberate judgments of suspense—measured through CRM—revealed significant main effects and an interaction effect for stimulus condition and plot. During the phase of *diffuse tension*, we found nearly identical response patterns in the audio-visual and the audio-only condition, both containing auditory information. CRM ratings started to rise in these two conditions in conjunction with the onset of music, whereas audience ratings in the video-only condition dropped below the scale mean. In the *story framing* phase (i.e., when visual narrative clues were introduced), CRM ratings in both conditions featuring sound rose to near-maximum levels. Importantly, however, CRM ratings in the video-only condition converged with the two other conditions only after visual clues appeared. After the story framing stage, all three response curves evolved in a parallel fashion, including the final phase of *resolution/relief* when all conditions showed the expected drop in suspense ratings. Thus, the temporal evolution of the CRM data was perfectly aligned with the plot structure identified via content-coding, adding convergent validity to our independent coding procedure.

The musical tone, as well as visual narrative clues, evoked similar levels of perceived suspense dependent on the time point at which this information became available. Because CRM peak levels for the full AV version were not higher than the peak levels in both other conditions and there was no substantial variation within the plot phases nor systematic correlations between CRM and physiological data, it is unlikely that CRM data reflects only automatic emotional responses. Rather, the data suggest that both types of information (music and narrative clues) feed into the same categorial judgment process, which draws on the audience's genre knowledge. In this sense, CRM ratings might at least be partly driven by the audiences' perceptions of whether content items and formal stimulus features match their preconceptions of "suspense as a genre" instead of an actual 'introspective suspense thermometer' (Steinbeis et al., 2006). From that perspective, we suggest that scholars consider the way that CRM measures act as a "typicality" rating, i.e., the judgment of the stimulus features "graded category membership" (Rosch and Mervis, 1975; Folstein and Dieciuc, 2019), rather than assuming automatic readouts of bodily responses, such as tension and arousal. This line of reasoning is compatible with various theoretical accounts of the cognitive processes during question answering, introspective self-reporting, and especially reporting on affective experiences that are difficult to verbalize (e.g., Schwarz, 1999; Nisbett and Wilson, 1977; Graesser and Black, 2017; Larson and Fredrickson, 1999), and these challenges to the CRM approach have been discussed repeatedly over the methods' storied history (e.g., Levy, 1982).

Physiological correlates of suspense. The physiological data point to the necessity of measuring physiological response when studying suspense and tension. Specifically, physiological responses revealed distinct roles of narrative content and non-diegetic music. Consistent with earlier findings, we found a correlation between heart rate and suspense ratings in the expected direction such that higher pulse rates coincided with higher

suspense ratings (Zillmann et al., 1975). However, interpreting this finding as coherence between subjective and objective arousal measures is questionable since no corresponding correlations could be found for either PVA or SCL. Also, data inspection revealed that the IBI effect was driven by a lower heart rate level (higher IBI) during the phase of *resolution/relief*. This significant difference, however, resulted from a series of short-term HR-deceleration patterns in both visual conditions (AV and video-only). These patterns were associated with emotionally significant events (expressive faces, arrival of the police) as identified in the content-coding, which did not affect suspense ratings (CRM). Taken together, the data suggest that heart rate deceleration more likely reflected an attentional effect instead of lowered arousal, which is consistent with previous literature (Lang, 1990, 2000, 2006; Potter et al., 2018). This interpretation is also supported by the fact that no such effects were observed for SCL and PVA, which both remained stable during the *resolution/relief* phase. Under the premise that both parameters are straightforward indicators of arousal, this result is surprising, as one could expect arousal to decrease when the story comes to a happy ending accompanied by relaxing music. One possible explanation could be derived from *excitation transfer theory* (Zillmann, 1971, 1983), which claims that bodily arousal that builds up during emotionally challenging phases of a media offering can persist during phases of relief, such as a happy ending, thus intensifying the audiences' positive feelings.

Overall, a marked difference between IBI and both other physiological parameters were observed across the whole movie. While PVA and SCL reflect more ambient differences between the conditions, with little variations on the micro-level, IBI marked short-term responses to significant narrative events occurring in either of the channels (visual or auditory). PVA and SCL were consistently negatively correlated, such that vasoconstriction and vasodilation were associated with higher and lower skin conductance levels, respectively, thus justifying the interpretation of both parameters in terms of arousal.

Arousal differences between conditions. Both parameters (PVA and SCL) reveal the strong influence of sound and music on the audiences' arousal levels. Though slightly different for the four plot segments, the predominant pattern found for both parameters was a significant difference between the video-only version and both other versions that included audio. Interestingly, the lowest arousal levels were found for the video-only condition. Together with the fact that no systematic trends and, with one exception (see below), no short-term variations were found in PVA and SCL in any condition, this data suggests that we might have observed a sensory modality effect rather than story-related arousal differences. However, research on modality-specific arousal effects when processing emotional stimuli shows equivocal, partly contradicting results, making it difficult to interpret exactly why we observed less arousal in response to a video-only stimulus. For instance, Bradley and Lang (2000) reported that acoustic stimuli (IADS) and visual stimuli (IAPS) elicited similar physiological arousal responses, but overall responses to auditory stimuli were comparatively weaker. Brouwer et al. (2013) found no modality differences in physiological arousal measures when presenting emotional pictures and sounds (IADS, IAPS) as well as their combinations. Using pieces of classical music with different emotional tonality and emotional pictures, however, Baumgartner et al. (2006) found similar to our result that physiological arousal—measured through skin conductance, heart rate, and respiration—was significantly higher in the combined and sound conditions compared to the picture condition. Interestingly, the authors also reported that the accuracy of the emotion recognition was highest

in the combined conditions, followed by the picture conditions, and lowest in the sound conditions.

With these findings in mind, one might conclude that non-diegetic music more globally revs up the audiences' tonic arousal level and that the phasic effects of content features build upon this foundation. However, such an interpretation must be formulated with great caution and needs further exploration. Our stimulus material was quite different from the one used by Baumgartner et al. (2006). In particular, we used one coherent stimulus in which the tensional music in contrast to the narrative clues was persistent over several minutes. Yet, the fact that arousal in the audio condition exceeded arousal in the video condition after only a few seconds of stimulus presentation again might speak for a modality effect. Future work could consider the extent to which the modality differences we identified in this study are specific to music or might reflect a general modality effect, as found for auditory versus visual information processing in other contexts (e.g., Keene and Lang, 2016). Relatedly, some evidence suggests that auditory information can also elicit varying degrees of mental imagery (e.g., Bolls and Lang, 2003), which in turn may influence physiological response patterns. While this may seem almost trivial for verbal auditory descriptions of specific scenes, it remains an open question of how far specific musical patterns might be able to evoke different inner pictures.

Differences in the arousal measures (PVA and SCL). As mentioned above, there was only one significant local difference in the two arousal measures: the knife stab. The knife stab completes the *story framing* and leads into the phase of *anticipatory suspense*, and evoked pronounced variation in the full AV condition. Responses to this event are slightly different for PVA and SCL data. While vasoconstriction response (tension/arousal) visible in PVA was followed by immediate vasodilation (resolution/relief), SCL accordingly increased with the occurrence of this event but then stayed on a higher level for the whole phase of *anticipatory suspense*. We wonder whether the observed difference between the two measures might be indicative of two different suspense mechanisms: one (PVA) related to local uncertainty and its resolution (independent of its emotional valence) and the other (SCL) related to an emotionally negatively toned solution and resulting—potentially fearful—expectations associated with arousal (Lehne and Koelsch, 2015).

Event-based analyses (IBI). Interestingly, the knife stab event was also associated with significant IBI responses. Heart rate deceleration was observed immediately before the event when the dramatic music was paused for a moment thus announcing the upcoming dramatic event. Heart rate acceleration was observed immediately after the event. This pattern could be interpreted in terms of cognitive resource allocation in attendance of a narrative clue followed by a momentary rise in arousal when the emotionally salient clue was provided (Lang, 1990, 2000, 2006; Potter et al., 2018). In fact, the event-related analysis further corroborated this interpretation; similar response patterns were found in response to other salient narrative clues, including close-ups of the protagonists' faces. We applied a *stimulus-* and a *response-driven* analysis procedure to better understand these response dynamics in terms of cognitive and emotional responses (Figs. 3 and 4).

The *stimulus-based* approach was built on the critical story-framing events identified in the content coding procedure. Comparing five seconds before the critical events with 5 s post-event, we found significant IBI differences in the AV conditions for two of the five events (bloody fabric and twitching foot), revealing a heart rate deceleration in response to these events. A

similar result was found for the audio-only condition for the screaming event that co-occurred with the twitching foot. Presumably, these events were perceived as crucial to the understanding of what was going on and the anticipation of further events and therefore may have led to increased attention.

This interpretation of heart rate deceleration is supported by the finding from the *response-based analysis*. Here, we observed short-lived, significant IBI differences between conditions. The differences pointed to heart rate deceleration patterns in response to local events. With one exception (the woman screaming), all significant differences occurred between the audio-only and conditions that contained visual information (AV and video-only). Importantly, during the phase of *anticipatory suspense*—when nothing violent happens but the audience is expecting a conclusion or resolution—heart rate deceleration responses were associated with the farmer's face, which served as the sole emotional information source. In the *relief/resolution* phase, deceleration patterns were again contingent on the presence of human faces/characters (baby, mother, farmer) and significant events (police arriving on the scene). In line with earlier research, we, therefore, interpret the observed heart rate deceleration in terms of attention and cognitive resource allocation prompted by salient narrative clues (Lang, 1990, 2000, 2006; Potter et al., 2018).

Conclusions and limitations

Overall, the results indicate that suspense is a complex phenomenon—implying different stimulus features, time scales, and response patterns. Evidently, tense music alone can trigger audiences' subjective feelings of suspense and lead to a persistent rise in physiological arousal. Images alone, on the other hand, can boost deliberate suspense ratings when they contribute to the audiences' understanding of the storyline and the anticipation of emotionally salient outcomes. While salient visual clues evoke systematic short-term effects in the heart rate (IBI), i.e., heart rate deceleration putatively associated with cognitive resource allocation, they do not elicit longer-term level changes in arousal as measured through PVA and GSR. Consequently, response patterns in the full AV condition were similar to the audio-condition condition with regard to longer-term arousal effects and more similar to the video-only condition in short-term attentional responses (indicated by IBI).

The partial dissociation of physiological responses and subjective reports justifies the assumption of two processing pathways. On the one hand, certain automatic responses result from either formal stimulus characteristics (tense music) or the anticipation of emotionally challenging events. On the other hand, people seem to make categorical judgments based on feature typicality (tense music, screaming, blood). The latter process is clearly more deliberative in nature and may depend on the instruction to continuously evaluate perceived suspense via CRM, although one could argue that this mindset also comes online when we expect a film to 'live up to' its genre description (e.g., when selecting them based on the tag 'suspense' in the Netflix library). Future studies could analyze to which degree these automatic responses inform self-report data and may even override the categorical judgments. In particular, such studies should also help to differentiate between bottom-up processes (for instance automatic response to overlearned threat indicators) and top-down processes (for instance looking for confirmative evidence in line with the elicited anticipations).

Unlike a recent study (Baldwin and Bente, 2021), we did not find robust evidence for the capacity of PVA to indicate local tension and relief patterns. Such a pattern was only identified in one case: the knife stab. However, it is important to note a key difference between the narrative structure of the stimulus film in

the current study versus that of Baldwin and Bente (2021). In contrast to Baldwin and Bente's (2021) study, which featured long narrative sequences consistent with most Hollywood movies that contain multiple micro-situations wherein anticipatory processes are triggered towards local outcomes (e.g., the end of a fight scene, the rescue of a puppy from a fire), our stimulus featured a linearly advancing story with a final solution. Yet, this is not the only way to create suspense. For instance, Doust and Piwek (2017) distinguish two types of suspense: *revelatory suspense* (e.g., where audiences witness ambiguous events and race to demystify the uncertainty) and *completion-based suspense* (e.g., where audiences dread an anticipated outcome). While the first type corresponds to the linear narrative structure of *Love Field*, the second type could have multiple recurring tensional situations, such as pulling the trigger of a one-round loaded gun in Alfred Hitchcock's *Bang Bang. You're Dead* that each spark the same anticipation (Schmälzle and Grall, 2020). The latter type might provide more promising data regarding short-term anticipatory tension and relief patterns. Without a doubt, many other forms of narratives exist, and we caution generalizing our findings to explain suspense and tension responses to all narrative structures.

Finally, considering the different building blocks of suspense as well as the multiple response dimensions, it is an open question of how and when the term 'suspense' should be used. For instance, some may ask whether short-term patterns of tension and relief might be subsumed under suspense. Others may question suspense's necessary and boundary conditions: do anticipated outcomes have to possess some kind of emotional valence? Should emotional valence associated with the likely outcome have to be negative and result from an affective disposition towards a protagonist (Vorderer et al., 2001; Zillmann, 1996)?

At the time being, we prefer two potential uses for the term. The first is a more minimalist definition of suspense that describes the affective component of an anticipatory cognitive process striving for the completion of an incomplete form or process. Such a definition would be in line with principles formulated in Gestalt psychology (see Lewin, 1935), propagating the rise of *psychological tension* when confronted with an open Gestalt. This tension implies negative affect and arousal and motivates the organism to find a resolution (closure). Unlike many definitions of suspense, the effect would be bound to the tension and not to the emotional valence of likely outcomes (Zillmann, 1996). Prescribing such a definition would make suspense a fundamental mechanism and suspenseful narratives would be a special way to create this experience (Lehne and Koelsch, 2015) that applies to micro levels of narrative suspense and the overall narrative arc. Furthermore, this definition would not require fears and hopes to result from an affective disposition towards a protagonist as a necessary condition (Zillmann, 1980, 1996).

Alternatively, one might suggest a definition of suspense that differs from tension. In their model of tension and suspense, Lehne and Koelsch (2015) refer to the pioneer of modern experimental psychology Wilhelm Wundt, who defined the dualism between "Spannung" and "Lösung" as a basic dimension of emotional experience (Bente and Feist, 2000). In fact, in German the term "Spannung" stands equally for both "tension" and "suspense" and "Lösung" stands for "relief/relaxation (Entspannung)" as well as for "resolution." We consider these dichotomies in the English dictionary as helpful to resolve terminological confusion, as they could provide a consistent conceptual framework that distinguishes between putatively distinct processes components of suspense. We suggest reserving the tuple "suspense/resolution" for higher-level cognitive processes strictly bound to media narratives that can create uncertainty, induce affect-laden expectations (fears and hopes), and elicit anticipatory

processes striving for resolution. We further suggest using the tuple "tension/relief" for low-level, automatic, effectively toned responses that are measurable as physiological arousal and that strive for relief/relaxation following a homeostatic principle. Tension, in contrast to suspense, can be induced by formal features of the media offerings, including auditory (sounds, music) as well as visual cues (e.g., darkness, colors, motion). Importantly, however, tension can also be a consequence of suspense as induced by a narrative. In this sense, tension can contribute to and/or result from suspense. Reading a book, for instance, might elicit tension and physiological arousal, exclusively driven by its narrative content. On the other hand, visual or sound effects and music in a suspenseful movie can induce tension and physiological arousal in a separate process and work in parallel with narrative cues to support the feeling of suspense. Moreover, tension-inducing presentation features could potentially prime cognitive processing and thus determine the interpretation of the narrative structure. The two approaches each hold advantages and disadvantages and exploring the utility of each definition of suspense would benefit from further discussion and scientific inquiry.

Data availability

Data can be found through this OSF link (<https://osf.io/ny9cw/>) and scripts can be furnished upon request.

Received: 15 June 2022; Accepted: 23 November 2022;

Published online: 13 December 2022

Notes

- 1 Non-diegetic sounds and music are added to a film during post-production to support the effects of the narrative—in contrast to diegetic sounds and music that are part of the plot (Wakefield et al., 2017). The music accompanying the shark attack in Spielberg's *Jaws* or the shower scene in Hitchcock's *Psycho* are examples of non-diegetic music. Sam playing 'As time goes by' in Ross's *Casablanca* and the car-radio playing contemporary pop songs in Steven King's *Christine* are examples of diegetic music.
- 2 *iom1* is no longer produced or sold by its parent company, but many similar low-cost devices are available, including *iom2* (Unyte Health Inc., Toronto, Canada, Unyte.com), which uses a PPG sensor integrated into an ear clip, but does not include SCL.

References

- Alwitt LF (2002) Suspense and advertising responses. *J Consum Psychol* 12(1):35–49. https://doi.org/10.1207/S15327663JCP1201_04
- Aly B, Eversten F, Omoba O, Schaller R, Terlap E (2017) Effects of auditory stimuli on blood pressure, respiration rate, and heart rate changes while watching a suspenseful video. *J Adv Student Sci* 17:1893–1897
- Baldwin J, Bente G (2021) The winner doesn't take it all: analyzing audience responses to an inspirational sports movie. *Media Commun* 9:202–214. <https://doi.org/10.17645/mac.v9i2.3840>
- Bálint K, Kuijpers M, Doicaru M (2017) The effect of suspense structure on felt suspense and narrative absorption in literature and film. In: Hakemulder F, Kuijpers M, Tan ES, Bálint K, Doicaru M (eds) *Narrative absorption*. John Benjamins Publishing Company, Amsterdam, pp. 177–197
- Bannister S, Eerola T (2018) Suppressing the chills: effects of musical manipulation on the chills response. *Front Psychol* 9:2046. <https://doi.org/10.3389/fpsyg.2018.02046>
- Baumgartner T, Esslen M, Jäncke L (2006) From emotion perception to emotion experience: emotions evoked by pictures and classical music. *Int J Psychophysiol* 60(1):34–43. <https://doi.org/10.1016/j.ijpsycho.2005.04.007>
- Benedek M, Kaernbach C (2011) Physiological correlates and emotional specificity of human piloerection. *Biol Psychol* 86:320–329. <https://doi.org/10.1016/j.biopsycho.2010.12.012>
- Bente G, Feist A (2000) Affect-talk and its kin. In: Zillmann D, Vorderer P (ed) *Media entertainment: the psychology of its appeal*. Lawrence Erlbaum Associates, Mahwah, pp. 113–134

- Bezdek MA, Wenzel WG, Schumacher EH (2017) The effect of visual and musical suspense on brain activation and memory during naturalistic viewing. *Biol Psychol* 129:73–12981. <https://doi.org/10.1016/j.biopsycho.2017.07.020>
- Bolls PD, Lang A (2003) I saw it on the radio: the allocation of attention to high-imagery radio advertisements. *Media Psychol* 5:33–55. https://doi.org/10.1207/S1532785XMEP0501_2
- Brouwer AM, van Wouwe N, Mühl C, van Erp J, Toet A (2013) Perceiving blocks of emotional pictures and sounds: effects on physiological variables. *Front Hum Neurosci* 7:295. <https://doi.org/10.3389/fnhum.2013.00295>
- Bullerjahn C, Gildenring M (1994) An empirical investigation of effects of film music using qualitative content analysis. *Psychomusicology* 13:99–118. <https://doi.org/10.1037/h0094100>
- Carroll N (1996) The paradox of suspense. In: Vorderer P, Wulff HJ, Friedrichsen M (ed) *Suspense: conceptualizations, theoretical analyses, and empirical explorations*. Lawrence Erlbaum Associates, Inc, Mahwah, pp. 71–91
- Chun C, Park B, Shi C (2020) Re-living suspense: emotional and cognitive responses during repeated exposure to suspenseful film. *Front Psychol* 11:558234. <https://doi.org/10.3389/fpsyg.2020.558234>
- Cohen AL, Shavalian E, Rube M (2015) The power of the picture: how narrative film captures attention and disrupts goal pursuit. *PLoS ONE* 10(12):e0144493. <https://doi.org/10.1371/journal.pone.0144493>
- Cohen J (1988) Statistical power analysis for the behavioral sciences. Lawrence Erlbaum Associates, New York
- Doust R, Piwek P (2017) A model of suspense for narrative generation. In: Alonso JM, Bugarin A, Reiter E (eds) 10th International conference on natural language generation. Association for Computational Linguistics, pp. 178–187
- Farbood MM (2012) A parametric, temporal model of musical tension. *Music Percept* 29(4):387–428. <https://doi.org/10.1525/mp.2012.29.4.387>
- Fisher JT, Keene JR, Huskey R, Weber R (2018) The limited capacity model of motivated mediated message processing: taking stock of the past. *Ann Int Commun Assoc* 42(4):270–290. <https://doi.org/10.1080/23808985.2018.1534552>
- Folstein JR, Dieciuc MA (2019) The cognitive neuroscience of stable and flexible semantic typicality. *Front Psychol* 10. <https://doi.org/10.3389/fpsyg.2019.01265>
- Gabrielsson A (2002) Emotion perceived and emotion felt: same or different? *Music Sci* 5:123–147. <https://doi.org/10.1177/102986490200505105>
- Graesser AC, Black JB (2017) The psychology of questions. Routledge
- Hubert W, de Jong-Meyer R (1991) Autonomic, neuroendocrine, and subjective responses to emotion-inducing film stimuli. *Int J Psychophysiol* 11:131–140
- Keene JR, Lang A (2016) Dynamic motivated processing of emotional trajectories in public service announcements. *Commun Monogr* 83(4):468–485. <https://doi.org/10.1080/03637751.2016.1198040>
- Krumhansl CL (1997) An exploratory study of musical emotions and psychophysiology. *Can J Exp Psychol* 51(4):336–353. <https://doi.org/10.1037/1196-1961.51.4.336>
- Lang A (1990) Involuntary attention and physiological arousal evoked by structural features and emotional content in TV commercials. *Commun Res* 17(3):275–299. <https://doi.org/10.1177/009365090017003001>
- Lang A (2000) The limited capacity model of mediated message processing. *J Commun* 50(1):46–70. <https://doi.org/10.1111/j.1460-2466.2000.tb02833.x>
- Lang A (2006) Motivated cognition (LC4MP): The influence of appetitive and aversive activation on the processing of video games. In: Messaris P, Humphries L (ed) *Digital media: transformation in human communication*. Peter Lang, New York, pp. 237–256
- Larson RJ, Fredrickson BL (1999) Measurement issues in emotion research. In: Kahneman D, Diener E, Schwarz N (eds) *Well-being: the foundations of hedonic psychology*. Russell Sage Foundation, pp. 40–60
- Lehne M, Rohrmeier M, Koelsch S (2013) Tension-related activity in the orbito-frontal cortex and amygdala: an fMRI study with music. *Soc Cogn Affect Neurosci* 9:1515–1523. <https://doi.org/10.1093/scan/nst141>
- Lehne M (2014) Emotional experiences of tension and suspense: psychological mechanisms and neural correlates. Freie Universität Berlin
- Lehne M, Koelsch S (2015) Toward a general psychological model of tension and suspense. *Front Psychol* 6. <https://doi.org/10.3389/fpsyg.2015.00079>
- Levy MR (1982) The Lazarusfeld–Stanton program analyzer: an historical note. *J Commun* 32(4):30–38. <https://doi.org/10.1111/j.1460-2466.1982.tb02516.x>
- Lewin K (1935) *A dynamic theory of personality*. McGraw-Hill, New York and London
- Mullen R, Faull A, Jones ES, Kingston K (2012) Attentional focus and performance anxiety: effects on simulated race-driving performance and heart rate variability. *Front Psychol* 3. <https://doi.org/10.3389/fpsyg.2012.00426>
- Nisbett RE, Wilson TD (1977) Telling more than we can know: Verbal reports on mental processes. *Psychol Rev* 84(3):231–259. <https://doi.org/10.1037/0033-295X.84.3.231>
- Nomikos MS, Opton E, Jr, Averill JR, Lazarus RS (1968) Surprise versus suspense in the production of stress reaction. *J Pers Soc Psychol* 8(2): 204–208
- Pearce MT, Wiggins GA (2012) Auditory expectation: the information dynamics of music perception and cognition. *Top Cogn Sci* 4(4):625–652. <https://doi.org/10.1111/j.1756-8765.2012.01214.x>
- Potter RF, Sites J, Jamison-Koenig E, Zheng X (2018) The impact of cognitive load on the cardiac orienting response to auditory structural features during natural radio listening situations. *J Cogn* 1(1):39. <https://doi.org/10.5334/joc.43>
- Pressnitzer D, McAdams S, Winsberg S, Fineberg J (2000) Perception of musical tension for nontonal orchestral timbres and its relation to psychoacoustic roughness. *Percept Psychophys* 62:66–80
- Raithe M (2008) Lovefield [Video file]. Retrieved from <https://www.youtube.com/watch?v=4meeZifCVro>
- Roberts CM, Hoetzel KM (2007) Psychophysiological responses to visual fear stimuli. *Psi Chi J Undergrad Res* 12(1):31–34
- Rosch E, Mervis CB (1975) Family resemblances: studies in the internal structure of categories. *Cogn Psychol* 7:573–605. [https://doi.org/10.1016/0010-0285\(75\)90024-9](https://doi.org/10.1016/0010-0285(75)90024-9)
- Salimpoor VN, Benovoy M, Larcher K, Dagher A, Zatorre RJ (2011) Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nat Neurosci* 14(2):257–262. <https://doi.org/10.1038/nn.2726>
- Salimpoor VN, Benovoy M, Longo G, Cooperstock JR, Zatorre RJ (2009) The rewarding aspects of music listening are related to degree of emotional arousal. *PLoS One* 4(10):e7487. <https://doi.org/10.1371/journal.pone.0007487>
- Schmälzle R, Grall C (2020) The coupled brains of captivated audiences: an investigation of the collective brain dynamics of an audience watching a suspenseful film. *J Media Psychol* 32:1–13. <https://doi.org/10.1027/1864-1105/a000271>
- Schwarz N (1999) Self-reports: how the questions shape the answers. *Am Psychol* 54(2):93–105. <https://doi.org/10.1037/0003-066X.54.2.93>
- Sparks GG (1989) Understanding emotional reactions to a suspenseful movie: The interaction between forewarning and preferred coping style. *Commun. Monogr* 56(4):325–340. <https://doi.org/10.1080/03637758909390268>
- Steinbeis N, Koelsch S, Sloboda JA (2006) The role of harmonic expectancy violations in musical emotions. *J Cogn Neurosci* 18:1380–1393. <https://doi.org/10.1162/jocn.2006.18.8.1380>
- Tan ET (2018) A psychology of the film. *Palgrave Commun* 4(82) <https://doi.org/10.1057/s41599-018-0111-y>
- Tan SL, Spackman MP, Wakefield EM (2008) Effects of diegetic and non-diegetic presentation of film music on viewers' interpretation of film narrative. In: Conference Proceedings for the 2008 international conference of music perception and cognition. Hokkaido University, Japan. Causal Productions, Australia
- Thayer JF, Levenson RW (1983) Effects of music on psychophysiological responses to a stressful film. *Psychomusicology* 3(1):44–52
- Vorderer P, Knobloch S, Schramm H (2001) Does entertainment suffer from interactivity? The impact of watching an interactive TV movie on viewers' experience of entertainment. *Media Psychol* 3:343–363
- Vorderer P, Knobloch S (2000) Conflict and suspense in drama. In Zillmann D, Vorderer P (ed), *Media entertainment: the psychology of its appeal*. Lawrence Erlbaum Associates Publishers, Mahwah, pp. 59–72
- Wakefield EM, Tan SL, Spackman MP (2017) The effects of diegetic and non-diegetic music on viewers' interpretations of a film scene. *Music Percept* 34:605–623. <https://doi.org/10.1525/mp.2017.34.5.605>
- Zillmann D (1971) Excitation transfer in communication-mediated aggressive behavior. *J Exp Soc Psycho* 7:419–434
- Zillmann D (1980) Anatomy of suspense. In: Tannenbaum PH (ed) *The entertainment functions of television*. Psychological Press, London, pp. 133–163
- Zillmann D (1983) Transfer of excitation in emotional behavior. In: Cacioppo JT, Petty RE (ed) *Social psychophysiology: a sourcebook*. Guilford Press, New York, pp. 215–240
- Zillmann D (1996) The psychology of suspense in dramatic exposition. In Vorderer P, Wulff HJ, Friedrichsen M (ed) *Suspense: conceptualizations, theoretical analyses, and empirical explorations*. Mahwah, Lawrence Erlbaum, pp. 199–231
- Zillmann D, Hay TA, Bryant J (1975) The effect of suspense and its resolution on the appreciation of dramatic presentations. *J Res Pers* 9:307–323. [https://doi.org/10.1016/0092-6566\(75\)90005-7](https://doi.org/10.1016/0092-6566(75)90005-7)

Acknowledgements

We thank Lorraine Kuch for providing us with a trained musician's description of the film music.

Competing interests

The authors declare no competing interests.

Ethical approval

Approval was obtained from the ethics committee of the Internal Review Board at Michigan State University. The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Informed consent

Consent was obtained from all participants to participate in the study and have their anonymous data shared. All biological data was de-identified.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-022-01461-5>.

Correspondence and requests for materials should be addressed to Nolan T. Jahn.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022