Tracking emotional shifts during story reception: The relationship between narrative structure and affective responses

Marie-Luise C. R. Schmidt[1], Julia R. Winkler[2], Markus Appel[2], and Tobias Richter[1]

[1] *Department of Psychology IV, Julius-Maximilians-Universität Würzburg, Germany*

[2] *Psychology of Communication and New Media, Julius-Maximilians-Universität Würzburg, Germany*

Manuscript accepted for publication in the journal *Scientific Study of Literature*

Author Note

This research was supported by the German Research Foundation (grants RI 1100/11-1 and AP 207/4-1). Experimental stimuli are available from the first and the fourth author upon request. The complete data set and analysis scripts are available in the repository of the Open Science Framework (OSF) (https://osf.io/2wxze/?view_only=4c84f9a623584e0c9cae71a0bf90ef55). Parts of the data and findings reported in this manuscript were presented at the 22th congress of the German Psychological Society (DGPs) in Hildesheim, Germany (September 2022).

Correspondence concerning this article should be addressed to Marie-Luise Schmidt, University of Würzburg, Department of Psychology IV, Röntgenring 10, 97070 Würzburg, Germany. E-mail: marie-luise.schmidt@uni-wuerzburg.de

Marie-Luise C. R. Schmidt (M.Sc., University of Bamberg, 2016) is a doctoral student at the Department of Psychology IV at the University of Würzburg, Germany. https://orcid.org/0000-0001-5656-1841

Julia R. Winkler (M.A., University of Mannheim, Germany, 2017) is a doctoral student at the Department Psychology of Communication and New Media at the University of Würzburg, Germany. https://orcid.org/0000-0002-3185-5867

Markus Appel (PhD, University of Cologne, Germany, 2004) is a professor and chair of the Department Psychology of Communication and New Media at the University of Würzburg, Germany. https://orcid.org/0000-0003-4111-1308

Tobias Richter (PhD, University of Cologne, Germany, 2003) is a professor and chair of the Department of Psychology IV at the University of Würzburg, Germany.
https://orcid.org/0000-0002-0467-9044

Abstract

Varying emotional content guides audiences' attention during story reception. This study explored audiences' emotional responses to stories to examine how *emotional flow* is related to the narrative material in valence and structure. We investigated the match between textual and emotional valence trajectories and examined accompanying further immediate and retrospective affective responses. Responses were tracked to three auditory short stories of different genres (happy-ending story, tragedy, and thriller) over four plot units (introduction, complication, climax and end). Sympathetic and vagal activation were measured continuously. Valence, arousal, and emotional intensity were measured retrospectively. Textual valence trajectories acquired by sentiment analyses corresponded mostly to retrospective ratings of stories with textually clear content of happiness and sadness. Immediate and retrospective emotional responses corresponded evidently to structural narrative features such as the climax. We discuss our work as clarifying groundwork for future research on emotional flow that links textual analyses to emotional responses.

Keywords: arousal, emotion, emotional flow, narrative valence, psychophysiological measures

Tracking emotional shifts during story reception: The relationship between narrative structure and affective responses

Stories take us on mental journeys into the narrated world. The psychological reception of narrative media is based on the mental modeling of the story world and its events, a process called *narrative comprehension* (Brewer & Lichtenstein, 1982; Busselle & Bilandzic, 2008; Jacobs, 2015; Mar & Oatley, 2008; van Krieken, 2018). If it develops effortlessly, audiences can immerse in the story and lose themselves in experiencing the narrative events (Busselle & Bilandzic, 2008). Recent theoretical approaches are particularly concerned with the role of emotions for narrative processing (Dillard & Shen, 2018; Nabi & Green, 2015). In storytelling, stimulus level emotions are used to direct audiences' attention, e.g., to foreground facts or information, and thus contribute to the narrative effect of conveying complex content in a comprehensible way (Carroll, 2003). During the narrative experience, emotions serve as important carriers of information for updating the mental model (Bal & Veltkamp, 2013; Busselle & Bilandzic, 2009; Mar & Oatley, 2008; Mar et al., 2011; Miall & Kuiken, 2002; van Krieken, 2018).

Nabi and Green (2015) illustrate in their *emotional flow* theory that affective responses over the course of a narrative are dynamic. These responses, in turn, relate to the narrative stimulus material, which is also dynamic (the narrative arc). With our study, we aim to connect research on emotional flow to theory and research on narrative structure. To this end, we use stories of different genres (happy-ending, thriller, tragedy) as the audiences' affective responses likely differ characteristically between them. Secondly, there is a challenge to measure these responses in research focused on the scientific study of literature, especially given the different levels of emotions: those implied or explicated in the story versus those within the recipients' minds. We contribute to the methodological knowledge in this regard by tracking affective responses with psychophysiological measures and retrospective self-reports that we both connect to story arc progression and genre differences.

Narrative Emotions

We distinguish two levels of emotions in narrative reception: the text or stimulus material and the audience. Stimulus level emotions are emotions implied by the story (e.g., a twist of fate for the protagonist), as well as the emotions that characters themselves experience and display. In a broader definition, the narrator's emotions could be included as well. At the audience level, we distinguish emotions according to the source from which they arise: *narrative emotions* arise directly from experiencing the story or in direct response to its characters, *aesthetic emotions* arise from impressions about the aesthetics of the stimulus (e.g., judgement of quality; Menninghaus et al., 2019; Menninghaus et al., 2020), and *relived or remembered emotions* arise from triggers that make us re-experience past affects (Mar et. al, 2011). We understand emotional flow as a series of *emotional shifts* in response to the story (Nabi & Green, 2015), that is, as narrative emotions. Since these shifts between emotions of the same or different valence (or even shifts in emotional intensity) respond to the material, it is likely that they correspond to the emotions implied by the story (see also the notion of *event-congruent emotions*, Appel et al., 2019).

On the stimulus level, stories can be characterized by valence arcs, e.g., their course of valence. Words are assigned a so-called sentiment score and summed up over narrative time. Sentiment scores indicate the valence of a word as assessed by several independent raters. Valence is the affective quality of an object or event concerning its (un-)pleasantness and is typically understood as a one-dimensional scale from positive to negative or pleasant to unpleasant. In this view on core affect, positive and negative are opposite poles of the scale and neutral valence lies in between (Bradley & Lang, 1994; Lang, 1980; Russell, 1980).

Genre and Narrative Structure

Stereotypical narrative arcs imply a core affect and dominant emotions in specific succession (Carroll, 2003; Kim et al., 2017). Genres explicitly advertise the emotional responses they aim for, most prominent are the tragedy and the happy-ending story, often

found in romances (Eagleton, 2009; Mann, 1985). Their basic emotions are dominant in different plot stages and shift into one another. Overall, a connection between major shifts in the audiences' valence and structural elements of stories seems likely.

Narratives can be described by normative patterns of progression. Classical structures, often found in short stories, were described by Aristotle ([c. 350 BC] 2005) and Freytag (1863). They typically involve an exposition that introduces the audience to the story world and characters, a progression of the plot in which often a challenge arises and the action increases, a climax in which this progression culminates, and subsequently, the resolution or closure of the previous conflict or hardship. Boyd et al. (2020) confirmed structures that share a narrative "grammar" using text analyses. Their results supported coherent structural patterns of staging at the beginning of a story, followed by the progression of the plot towards a peak in cognitive tension during the middle-to-late story parts. As different stories proceed in different narrative pace, we propose a division in conceptual units (Bamberg, 2012) based on functionality for the plot. Bamberg proposes a division in orientation (or exposition), complication, action toward a resolution, and the resolution with subsequent closure (see also Cutting, 2016). We strive for a reasonable trade-off between unit specificity and generalizability. Therefore, we chose four conceptual units: the *introduction* relates to Bamberg's (2012) orientation or exposition. It sets the scene and involves the staging of the story. The *complication* is a distinct event or characterization that foreshadows the rising action. In the *climax* unit, the protagonist is faced with a hardship, decision, or task that involves an action or struggle up to a peak in cognitive tension due to uncertainty of a positive outcome. Finally, the end unit comprises the resolution and consequences of the attenuated action.

With regards to content, we compare a happy-ending story, a tragedy and a thriller story. Emotional dynamics differ with each genre (Kim et al., 2017). A happy-ending story typically follows a so-called "rags to riches" arc, from a negative initial state continuously

upward to a positive final state (Reagan et al., 2016). The tragedy, on the other hand, moves in the opposite direction, falling in valence instead of rising. Thriller stories often end with a resolution of the built-up tension or a change from fear to relief, following an overall trajectory of ups and downs that resolves in the "rags to riches" arc. Yet, the expected dominant emotions of a thriller differ significantly from the happy-ending genre, which usually bases its positive endings not on relief alone, but on happiness and joy. Different genres are characterized by different emotional dynamics. While happy-ending stories are predominantly about the typical hero's journey and overcoming challenges, tragedies tend to address the lost battle against adversity and the associated despair. Thriller stories, on the other hand, address our evolutionary drive of overcoming existential danger.

Levels of Audiences Emotional Responses

For our investigation, we draw on Barrett's (2016) theory of constructed emotion (Barrett, 2013; Barrett et al., 2015) that defines an emotion as a category or population of instances that is perceived in relation to a specific context. Barrett advocates mapping emotions across as many aspects as possible, by means of psychophysiological, behavioral, and self-report measures, and interpreting these responses in relation to their specific context situation. The theory states that emotions are mental constructions dependent on social reality. Individuals make sense of affective feelings that belong to a variety of perceptions (including physical sensations, thoughts, and motivational cues) by using concepts that organize their past experience to interpret the cause of the overall impression and its' implied consequences (e.g., action tendencies).

In this study, we focus on valence and arousal dimensions of emotional experience. We draw on psychophysiological measurements to track arousal over the course of the story and we measure subjective emotional experience in retrospection as not to disturb ongoing narrative processing (Alam & So, 2020; Coppin & Sander, 2016; Mauss & Robinson, 2009). As described by Barrett and Russell (2015), we assume that our physiological measures

capture the audiences' immediate, rapid affective responses, whereas the retrospective ratings represent much more interpretive assessments against the context of the overall story. Additionally, we explore distinct subjective emotion categories in relation to the valence and arousal dimensions. We use retrospectively rated dominant emotions to measure emotional intensity and to establish a more valid situational reference.

The autonomic nervous system displays immediate affective responses to stimuli (Kreibig, 2010; Potter & Bolls, 2012). Its activation is an interplay of sympathetic and vagal (or parasympathetic) activity (Fahr & Hofer, 2013; Svorc, 2018). They represent two core systems of the automatic nervous system that are typically involved in emotional states (Kreibig, 2010; Potter & Bolls, 2012). The sympathetic branch is considered activating or arousing in the literal sense. It increases our physical and psychological readiness to act, and thus, is associated with situations that we perceive as challenging or threatening (Fahr & Hofer, 2013). The vagal branch of the nervous system is considered regenerative, and thus, associated with situations of rest and relaxation (Fahr & Hofer, 2013). Psychologically, it is associated with executive functions as well as emotion regulation (Balzarotti et al., 2017; Kimhy et al., 2013; Pinna & Edwards, 2020; Stange et al., 2017), and social cognition (Zammuto et al., 2021).

Rationale of the Present Study

There is a profound body of research on audiences' emotional responses to narratives that focusses on single emotions (Bezdek et al., 2015; de Graaf & Hustinx, 2011; Nabi, 2002; So et al., 2016), overall emotions (Altmann et al., 2012; Carrera et al., 2010; Dunlop et al., 2010; Hoeken & Sinkeldam, 2014; Knobloch-Westerwick & Keplinger, 2006; Kreibig, 2010; Murphy et al., 2013; Prestin, 2013; Rodrigue et al., 2014), or an end emotion specifically (e.g., Hamby & Brinberg, 2016). However, Nabi and Green (2015) argue that the combination of all emotional shifts in response to a story (emotional flow) is a crucial factor in ongoing receptive processes such as narrative engagement. Emotional shifts can occur in different

dimensional aspects of emotion, such as valence or emotional intensity, or between distinct emotions (e.g., of the same valence). So far, research on emotional flow predominantly focuses on shifts in valence between broad story parts, without regard for possible structural dynamic differences induced by free combinations (e.g., Alam & So, 2020). The location of valence shifts in the narrative arc could play a role for the audiences' emotional responses, for example due to predictions and expectations towards the story arc. To our knowledge, there are no studies that track emotional responses to specific valence shifts comparing different underlying genres or emotions over comparable plot units. This study addresses this gap and tracks emotional experiences during stories of different genres over a generic story structure. Our aim is to empirically map the theoretical link between emotional shifts and story structure.

We still know little about how emotional shifts relate to basic properties of the narrative material. In our work, we focus on valence shifts independent of their direction. Using different genre-typical stories, we track such shifts over the course of the stories. We assume that the timing and combination of such major shifts within the narrative progression imply different affective responses. The arousal (or excitation) buildup associated with emotional shifts may alter aspects of the emotional experience depending on their occurrence in the narrative (Nabi & Green, 2015; Zillmann, 1996). For example, valence shifts at certain narrative points could be associated with an increase in affective intensity, such as the climax or end of a story. In our study, we divide stories with different emotional content into structurally similar units and track the occurrence of valence shifts by analyzing the text material in conjunction with affective responses. We measure audiences' immediate affective responses dynamically over different genres by coherent plot units and provide interpretative reference by retrospectively rated distinct emotions. We focus on two research questions: (1) To what extent do self-reported subjective emotional responses (valence, arousal, emotional intensity, distinct emotions) correspond to a story's narrative arc across three different genres

(happy-ending story, tragedy, thriller)? (2) To what extent do psychophysiological emotional responses (skin conductance level, sympathetic power, vagal power, respiratory sinus arrythmia) correspond to a story's narrative arc across three different genres (happy ending story, tragedy, thriller)?

Method

Participants

We used a linear mixed models analysis with random effects (random intercepts) of participant for analyzing the data (see section on Data Analysis Strategy below) and based the estimation of the required sample size on the method proposed by Westfall, Kenny and Judd (2014), as implemented in the accompanying web-based app (https://jakewestfall.shinyapps.io/crossedpower/). For the sample size estimation, we assumed that the variance components would distribute equally between the random effects of participants (0.5) and the residual term (0.5). All other variance components were assumed to be 0 since the random intercept of participants was the only random effect in the model. We further assumed the magnitude of the effect size (Cohen's *d*) to be 0.6 and specified the Type-I error probability α to be .05 (two-tailed) and the desired power 1-β to be .80 for pairwise comparisons between story parts. Given these specifications, the required sample size was estimated to be 17 participants per story. We chose to collect data by repeated measures, creating a sample that exceeded the required sample size per story by at least 15% because physiological measures often suffer from data loss.

We recruited 32 participants in the city of A, Germany (blinded), through the online participant platform of the Psychology Department at the University of A (blinded). One participant was excluded due to intoxication prior to the experiment, leaving a total sample of 31 participants (14 women) for the analyses (age: $M = 24.97$ years, $SD = 7.72$; *Min* = 18 years, *Max* = 51 years). Participants were predominantly students ($n = 24$) and native German speakers ($n = 30$). Experimental sessions had a mean duration of 1.5 hours and were

conducted individually for each participant in the A laboratory (blinded) (55%) or the B laboratory (blinded) at the University of A (blinded) (45%). They were each paid between €10 and €12.50 depending on the overall preparation time.

Text Materials

Three different short stories were selected from searches in German short-story forums on the Web that matched emotional arcs with three distinct emotional shifts: a happy-ending story, a tragedy, and a thriller. We partly changed and shortened the original stories to accommodate the basic genre structures. A comprehensive description per plot unit is provided in Table 1. The happy-ending story was a science-fiction story about the survival of bees in a post-apocalyptic world, narrated in third-person perspective (Frambach, 2018; 1691 words, 12:29 min). The tragedy was a short story about a bullying incident during a school trip, narrated in first person (Praetorius, n.d., 819 words, 5:54 min audio). The thriller story was a story about a sudden attack on the tram, narrated in third person (Eisenmann, 2008; 1641 words, 11:39 min audio). Using the R package syuzhet version 1.0.6 (Jockers, 2017) and the German sentiment lexicon *SentiWS* version 2.0 (Remus et al., 2010), we conducted sentiment analyses for each story and unit. Figure 1 shows sentiment scores over story units to represent valence trajectories of our stimulus stories.

In preparation of our experiments, we verified that the stories were similarly easy to read by computing Flesch-Reading-Ease scores (adapted scoring for German language according to Amstad, 1978; Flesch, 1948). They corresponded to middle to low reading difficulty levels. Every story was audio recorded to ensure timed mapping of online measures during presentations. One neutral text was used to compare with the effects of the emotion-laden stories (Kreibig et al., 2007; Piferi et al., 2000). We used an expository text about apples (Buth, 2014; 500 words, 3:28 min audio).

Design

Participants listened to two of three possible emotional stories in random order after listening to the neutral text, creating six possible story and order combinations that were balanced over the 31 participants. They split into five groups of five and one group of six over these combinations, resulting in a total sample size of 21 participants for the happy ending story and the tragedy and 20 participants for the thriller story.

Procedure

After obtaining written informed consent, participants were seated upright in a chair in front of a PC equipped with headphones. Their left hand, which was their non-dominant hand in 90% of cases, rested on an armrest at the table. Participants were to operate the mouse with their right hand during the questionnaire periods between story presentations. Physiological sensors were attached, and physiological signals were continuously sampled during the experimental session. The experimenter sat in the same room outside the participants' field of vision but was approachable in case of problems or questions.

We used Inquisit 5 (Millisecond Software, Inc.) as the experimental presentation software. After completing a short questionnaire, including demographics and control questions, participants were instructed to find a comfortable seating position and to avoid any unnecessary movements and speech during the tasks, relaxation periods, and stories. They were instructed to keep their eyes on the screen because their facial expressions were recorded by a webcam.

Before any stimulus presentation, participants completed a paced breathing task similar to the task used by Butler et al. (2006) in which they were asked to look at a sinusoidal wave passing through a vertical line onscreen. Participants were required to breathe in whenever the wave rose and breathe out whenever it fell. A short repeatable practice trial ensured their understanding of the instructions. The waveform pattern was designed to induce a respiratory frequency of nine cycles per min and was presented for 2:15 min. This

measurement period was later used for assessing individual differences in resting RSA (Grossman et al., 1990; Wilhelm et al., 2004).

Following a 3-min quiet sitting period that served as the resting baseline for autonomic nervous system measures, participants listened to the neutral text that served as a non-emotional, stimulated control period. Next, participants listened to two of three possible short stories in random order, each following a separate unstimulated baseline period of 3 min. Every story was preceded by instructions to listen carefully and to keep their eyes on the screen where a neutral visual animation of the narrator's voice was shown (for an exemplary screenshot, see online supplement S1). Immediately after each story, participants answered a set of self-report items. They were asked to rate their emotional experience for each story unit with the help of self-assessment manikins (SAM; Bradley & Lang, 1994; Lang, 1980) and a maximum of three out of a list of 15 emotions. A set of nine items targeting the general experience of emotional shifts was sampled for unrelated analyses. Additional SAM ratings of state valence and arousal were obtained after each resting baseline before participants listened to the next story. After the session, participants were debriefed and reimbursed.

Measures

Although every story could be separated into the four structural units of introduction, complication, climax and end, references of story units for the subjective ratings differed between genres. In the thriller story, we deemed it difficult for participants to mentally separate the complication and climax part after listening to the story. Therefore, we obtained ratings for only the introduction, one merged plot unit of complication and climax, and the end part of the story (compared to all four units in the other genres). The wording of each unit reference per story is provided in an online supplement (S2).

Emotion Experience

After a story was completed, participants were asked to think back to a specific story unit and to rate their experience. The *SAM ratings of state valence and arousal* were obtained

via seven pictures ranging from negative to positive valence and low to high arousal respectively (Bradley & Lang, 1994; Lang, 1980). For the rating of discrete emotions, participants were asked to choose and rate at most 3 of 15 emotions that they felt most intensely during the story unit on a 7-point Likert scale. The total list of emotions contained amusement, fear, anger, sadness, disgust, happiness/joy, surprise, relief, suspense, hope, fascination, compassion, boredom, annoyance, and contempt, with the option to add an emotion they felt most intensely but was not on the list. This approach aims to combine the advantages of free listing and the rating of selective distinct emotions. It complements ratings of valence and arousal by descriptive, contextual information. Due to inherent differences in rating incidences of single emotions, we exclusively analyzed overall *emotional intensity* as the mean of intensity ratings of all selected emotions per story unit.

Physiological Measures

Cardiac, electrodermal, and respiration activity were collected using Biopac acquisition systems and recording software version 4.4 and 5.0 at an acquisition rate of 2 kHz each. For the cardiovascular measures, disposable Ag–AgCl electrocardiogram electrodes were attached to the participants after careful skin preparation, using a Lead(II) configuration. To measure skin conductance, two disposable Ag–AgCl electrodermal electrodes were placed on the thenar and hypothenar site of the participant's palm on their left hand. The day before the experiment, participants were reminded to not wash their hands within an hour before the start of the experiment to ensure sufficient moisture of the dermal surfaces (Shaffer et al., 2016). Thoracic respiration was measured using a respiration belt that was placed on the chest.

Following data collection, the data were processed in AcqKnowledge software version 5.2 (Biopac Inc.; for detailed data processing procedures see online supplement S3). In addition to downsampling and filtering procedures, skin conductance waveforms were square root transformed to adjust for skew inherent to skin conductance data (Barraza et al., 2015;

Norris et al., 2007). RSA was normalized following the procedure of Grossman and Taylor (2007) to correct for the mean intervals between successive heartbeats.

We measure psychophysiological arousal as sympathetic nervous system activity, indicated by high values of *skin conductance level* (SCL) and the sympathetic component of *heart rate variability* (HRV), hereinafter referred to as *sympathetic power*. To distinguish activation patterns and further explore psychophysiological responses associated with valence shifts, we measure respiratory sinus arrhythmia (RSA) and the vagal component of HRV, hereinafter referred to as *vagal power*. SCL has the advantage of representing purely sympathetically innervated activation, and thus does not need to be corrected for vagal influences like HRV-based measures. However, it is more susceptible to disturbances during measurement. Vagal parameters included RSA, respiratory sinus arrhythmia. It constitutes heart rate changes in synchrony with respiration and is suspected to react to changes in valence and/or emotional regulatory efforts (Butler et al., 2006; Ravaja, 2004; Stange et al., 2017).

Data Scoring

Raw scores of SCL, heart rate, and respiration rate were extracted as the arithmetic mean of the respective story unit, HRV raw scores were obtained by spectral power analysis. Baseline periods were 3:00 min in length, the neutral text had a total duration of 3:30 min, and story units differed between 0:50 and 4:07 min in sequence length over the stories. To make these values comparable despite the underlying differences in recording time, each raw score was corrected for the respective lengths of each sequence, resulting in parameter values per second. Note that using period-length-corrected mean scores of the physiological data likely leads to conservative estimates of emotional responses.

One data set was excluded from the physiological data analyses because of severe disruptions during baseline recordings. Problems with data collection or non-responding left a total sample of 23 data sets for SCL analyses, 30 sets for heart rate and HRV analyses, and 26

sets for respiratory analyses, with at least one usable story or a minimum of 50% usable story units.

Data Visualization

For the visualization of psychophysiological data (based on different measurement units), we standardized response scores after Kreibig et al. (2007). We subtracted the prior baseline score from the raw score of the respective story unit. The paced breathing task served as baseline for RSA. Individual differences in each person's overall mean were eliminated by centering around the person mean separately for each parameter. RSA parameters were additionally divided by respiration rate. Finally, all scores were subjected to a *z*-normalization transformation separately for each story (and the preceding neutral text) and mapped onto a *T* scale with a mean of 50 and a standard deviation of 10.

Data Analysis

Raw scores were analyzed with linear mixed effects models by using the lmer function of the R package lme4 version 1.12 (Bates et al. 2015). Linear mixed models are the method of choice for analyzing our data due to the cross-dependencies that originate from one person listening to two of three possible stories with multiple measures per story and random missing values (stories nested within participants; Judd et al., 2017; Richter, 2006). For each parameter and story, we used a linear mixed model with random effects (random intercepts) of subjects and fixed effects of baseline level (grand-mean centered) and story part (dummy-coded), comparing every subsequent part to the introduction. The models for RSA differed by including RSA levels of the paced breathing period as baseline and respiration rate as an additional predictor, both grand-mean centered. We further used least-square means (lsmeans package version 2.3; Lenth, 2016) for multiple comparisons with the Kenward-Roger approximation of degrees of freedom and Bonferroni-Holm method for *p* value adjustment to test differences between the remaining story parts.

Availability of Data, Code, and Materials

Data and R-code for the analyses reported in the present paper are available at the repository of the Open Science Framework (https://osf.io/2wxze/?view_only=4c84f9a623584e0c9cae71a0bf90ef55). The experimental materials are available from the authors upon request.

Results

We report results from three data sources: psychophysiological measures, subjective arousal and valence ratings, and subjective intensity ratings of discrete emotions. Figures 2 to 4 illustrate our findings per genre. Please see online supplement S4 and S5 for the detailed trajectories of valence and arousal ratings for each genre, online supplement S6 to S9 for the detailed trajectories of psychophysiological measures, online supplement S10 for the detailed trajectories of emotion intensity ratings and online supplement S11 for detailed ratings of distinct emotions.

Research Question 1: Narrative Arcs and Subjective Emotional Responses

Subjective emotional responses were acquired after the stories. Participants were given a short description of the particular story unit to guide their recall and asked to rate their emotional experience during listening. They had to choose the three emotions they felt most intensely during the respective story unit and rate their intensities. Intensity scores were computed across all rated emotions per unit. The respective emotions selected serve as descriptive context.

Valence Trajectories

The sentiment analysis of the happy-ending story showed a progression of the "man in a hole" type. Stories of this progression start happy, until the protagonist “falls in a hole”, i.e., something unfortunate happens that the protagonist ultimately overcomes. In the case of our story, we found a progression from neutral to negative to positive valence (Figure 1a). In comparison, the subjectively experienced valence was of the "rags to riches" type, i.e., from a negative initial state continuously upward to a positive final state (introduction to

complication unit, $t(60) = 2.55$, $p = .013$, $d = 0.68$, complication to climax unit, $t(60) = 2.55$, p_a[1] $= .040$, $d = 0.68$, climax to end unit, $t(60) = 1.64$, $p_a = .107$, $d = 0.43$; Figure 2a and online supplement S4). Thus, the introduction was experienced more negatively than the textual basis indicates, and the following trajectory of subjective valence appears to be shifted relative to the textual valence (a premature positive experience). Since hope was rated second most intense in the complication unit, we assume that activated expectations regarding the course of the story triggered anticipated positive feelings much earlier than the sentiment analysis indicated (see online supplement S11a). Additionally, due to the positive experience of the expected happy ending, there might have been a contrast effect in the following rating of the introduction unit.

In the tragedy, a very high correspondence of the text analysis with subjective valence ratings was noticeable. Overall, the story followed a "riches to rags" progression (introduction to complication unit, $t(60) = -4.07$, $p < .001$, $d = -1.09$, complication to climax unit, $t(60) = -4.07$, $p_a < .001$, $d = -1.09$, climax to end unit, $t(60) = 0.87$, $p_a = .386$, $d = 0.24$; Figures 1b and 2b and online supplement S4). The transition from predominantly positive to predominantly negative valence was also identifiable in both plots between the introduction and the complication. The distinct emotion ratings were consistent with this progression, especially within the first three units (see online supplement S11b).

In the thriller story, we found the least correspondence between sentiment scores and subjective valence. The sentiment scores fluctuated slightly below the neutral value of zero in a tendential "Oedipus" type (fall - rise – fall; Figure 1c). Subjective ratings shifted from positive initial valence to a negative middle part ($t(57) = -4.70$, $p < .001$, $d = -1.30$) back into positive valence ($t(57) = 8.81$, $p_a < .001$, $d = 2.44$), that is, a "man in a hole" progression (Figure 2c). Thus, in contrast to the textual analysis, there were two major shifts in overall

[1] Bonferroni-Holm-adjusted p values are indicated as p_a.

valence that participants experienced within the story. This was also reflected in the selected emotion, which showed the typical fear - relief progression from the middle to the end of the story. In the initial unit, we found instead great ambivalence (boredom, amusement, disgust, see online supplement S11c).

Overall, in the genres that relied on shifts of happiness and sadness, namely the happy-ending story and tragedy, we found a notably higher correspondence of the sentiment scores with the subjective valence ratings.

Arousal and Emotional Intensity

Based on the excitation transfer theory (Zillmann, 1988, 1996), we can assume arousal as an important determinant of the intensity of emotional experience. Within genres, different dominant emotions account for the observed shifts in overall valence. Thus, the emergence of activating emotions such as happiness, fear, and anger has the potential to result in intensity increases in emotional experience. This is equally true for emotional shifts associated with suspense (Nabi & Green, 2015). In this section, we compare subjective arousal trajectories with trajectories of emotional intensity in the context of the different genres.

In the happy-ending story, subjective arousal did not vary significantly over the course of the story (introduction to complication unit, $t(60) = -0.84$, $p = .405$, $d = -0.19$, complication to climax unit, $t(60) = 0.33$, $p_a = .999$, $d = 0.08$, climax to end unit, $t(60) = -0.84$, $p_a = .999$, $d = -0.19$; Figure 2a and online supplement S5). Neither did emotional intensity (introduction to complication unit, $t(56) = -1.38$, $p = .173$, $d = -0.16$, complication to climax unit, $t(56) = 1.11$, $p_a = .999$, $d = 0.22$, climax to end unit, $t(56) = 0.36$, $p_a = .999$, $d = 0.07$; Figure 3a and online supplement S10). Suspense was selected most frequently and rated most intense in the introduction, $M = 4.80$ ($SE = 0.36$) by 50% of participants, and the climax unit, $M = 4.78$ ($SE = 0.32$) by 45% of participants (Table 2).

In the tragedy, we found a significant rise in subjective arousal and emotional intensity towards the climax unit (arousal, $t(60) = 4.22$, $p_a < .001$, $d = 1.00$, intensity, $t(58) = 3.11$, $p_a =$

.014, $d = 0.68$; Figure 2b and Figure 3b and online supplements S5 and S10). Again, suspense was selected most often and rated most intense in the introduction, $M = 5.00$ ($SE = 0.31$) by one third of participants, and the climax unit, $M = 4.80$ ($SE = 0.58$) by 24% of participants (Table 2).

In the thriller story, arousal was rated highest in the middle parts of the story, $t(57) = 7.31$, $p < .001$, $d = 1.63$ (Figure 2c and online supplement S5). Emotional intensity did not vary significantly over the course of the story, but peaked at the end, $t(38) = 2.37$, $p < .05$, $d = 0.29$ (introduction to end unit; see Figure 3c and online supplement S10). The middle units were also rated as highly suspenseful by 70% of participants, $M = 5.79$ ($SE = 0.24$; Table 2).

Especially in the stories with intense negative emotions in the middle part (anger, fear), we found increased arousal in the climax unit that corresponded to suspense ratings. Only the thriller story showed an emotional intensity peak at the end, following a rise in suspense and arousal. Research Question 2 answers the question if and how psychophysiological activity corresponds to these findings.

Research Question 2: Narrative Arcs and Psychophysiological Emotional Responses

The subjective experience of arousal is a time-delayed and interpretive component of emotional experience. Psychophysiological activation, on the other hand, represents an immediate stimulus-related response of our body. The activation of the sympathetic and the parasympathetic branch of the autonomic nervous system was measured, each by means of two parameters. The skin conductance level (SCL) and the sympathetic part of the heart rate variability (in the following sympathetic power) are measures of sympathetic activation, i.e., arousal in the activating sense. The two measures of parasympathetic activation were both obtained from heart rate variability: parasympathetic power and the respiratory sinus arrhythmia (RSA) parameter, both representing attenuating, regulatory activation. The overview across the stories (Figure 4) showed a high degree of consistency between the parameters within a story.

Based on the previous findings, an increase in arousal was to be expected for the happy-ending story well before the end unit, since the ratings of hope and increasing valence already set in earlier than the sentiment analysis suggested. Indeed, we found the most significant activation increase over all physiological measures from the complication to the climax unit (Figure 4a and Table 3). Despite the highest suspense score in the introduction unit, $M = 4.80$ ($SE = 0.36$) of 50% of the participants, we found a relative deactivation compared to the neutral story. In the context of subjective ratings, we could attribute this activation pattern to sadness (Kreibig, 2010), but a link to compassion would also be conceivable. Thus, it seemed plausible that sympathetic activation associated with suspense was overshadowed by dampening negative emotions. Accordingly, this would not indicate a negativity bias in the subjective ratings of the introduction unit (if we compared them to the sentiment scores), but rather that the sentiment analysis did not capture the aggregated sentiment of the introduction in a way that was representative of the audience's feelings. In comparison, we found a high correspondence of the suspense ratings to the physiological activation in the climax unit (Table 3). Here, we notice a clear double activation of both nervous system branches that is frequently found in states of stress or when facing a challenge (Berntson et al., 1991; El-Sheikh et al., 2009).

In the tragedy, a high activation level in the introduction unit was initially noticeable in comparison to the neutral story (Figure 4b). In this story, suspense was also rated highest at the beginning, but here in connection with happiness and amusement, which are physiologically associated with high activation instead of curbed activation (Kreibig, 2010). The following shift to dominantly negative emotions such as sadness and compassion resembled the introduction unit of the happy-ending story (Figure 4 and Table 3). Another parallel was the significant double activation towards the climax, which in this story also fits the increased ratings of arousal and emotional intensity. Towards the end, vagal power particularly increased. Although suspense subjectively decreased here, we found a rise in

physiological activation in all parameters. The outstanding increase of vagal power could indicate frustration and a heightened control of negative emotions, which seems plausible in terms of the tragic ending.

In the thriller story, subjective ratings suggested an arousal increase in one of the middle story parts, associated with the emergence of fear (Kreibig, 2010). Consistent with the other two stories, we also found a strong significant increase in physiological activation in all parameters toward the climax unit (Figure 4c and Table 3). This corresponded highly with subjective ratings of arousal, emotional intensity, and suspense. The collective drop in physiological activation at the end of the story also clearly corresponded to the subjective ratings of relief and drops below the baseline level of the neutral story.

In sum, we found that physiological activation corresponded more with our structural unit classification and with activation patterns associated with dominant subjective emotions and less with suspense ratings. Across all genres, there was a consistent activation toward the climax unit. The activating dominant emotions of the stories (e.g., happiness, fear) were indeed most often associated with higher physiological activation (mostly of both nervous system branches). The exception is the end of tragedy, where we suspect increased regulatory activation. In terms of excitation transfer, the high correspondence of increased physiological arousal with emotional intensity ratings may suggest that it happened within our narrative units rather than between them.

General Discussion

Our work aimed to combine research on narrative structure with research on emotional flow, thereby connecting and extending theoretical and empirical approaches. We examined the dynamics of stories of different genres over comparable narrative units via multiple lenses: textual vs. experiential, and within the emotional experience, immediate vs. retrospective responses. Second, we wanted to demonstrate a way to dynamically measure the

different levels of emotional experience separately and to relate them to each other. This allows for tangible comparisons between different stories and their experience by audiences.

We began by arguing for an inherent link between narrative structure and emotional responses. There are several theories that have already addressed the intersection of connecting emotional trajectories on a textual level with affective responses, such as Brewer and Lichtenstein's (1982) Structural Affect Theory or Lang's (2000) Limited Capacity Model of Mediated Message Processing. Nevertheless, for research dealing with dynamic emotional responses, it is important to keep basic assumptions in mind: (1) The dynamic qualities of the narrative stimulus shape the elicited emotional responses. (2) Audiences' emotional responses to a narrative can arise from different processes (e.g., aesthetic emotions, remembered emotions, narrative emotions). (3) An emotional response is multi-facetted - dynamically, we can distinguish an immediate response and a retrospective, interpretive one. (4) These are interrelated, although they need not coincide.

We have demonstrated a method to compare stories of different contents and dynamics via theoretically derived units with respect to both immediate emotional responses and retrospective ones. We especially succeeded in comparing stories of different lengths based on the pre-defined units that share a narrative "grammar". In another preliminary step, we applied quantitative analyses of the stimulus material to obtain a fundamental mapping of the stories' emotional dynamics. Previous research on emotional flow does not yet incorporate sentiment analyses, although research on narrative language processing (NLP) is already making progress with many different text types, lexicons, and analysis methods (e.g., Christ et al., 2022; Nandwani & Verma, 2021; Piper et al., 2021). In our study, sentiment analysis proved to be suitable for examining valence trajectories of different stories.

The Research Questions

We posed two research questions: Research Question 1 addressed the extent of correspondence between each story's narrative arc (as mapped by the sentiment analysis) and

audiences' self-reported emotional responses (valence, arousal, emotional intensity, distinct emotions). Research Question 2 addressed the extent of correspondence between each story's narrative arc (as mapped by the previous analyses) and audiences' psychophysiological emotional responses (in sympathetic and vagal autonomous nervous system activity).

Apart from the limitations of sentiment analysis, which we discuss in detail below, a more differentiated subjective experience of valence was evident in the ratings. Whenever a story section contained within itself a dynamic variation of valence, as in the introduction of the happy-ending story, or the literal content of sections was ambiguously explainable, as in the thriller story, prominent deviations of the rated valence from the sentiment analyses were noticeable. We additionally noticed higher salience of negatively arousing emotions in the subjective ratings than positively arousing emotions. Retrospective emotional responses are constructed by memory and interpretation and therefore commonly affected by a negativity bias, which has been widely demonstrated in previous emotion studies (e.g., Baumeister et al., 2001; Bohanek et al., 2004; Egidi & Gerrig, 2009). This effect is particularly evident in our findings when comparing the emotion ratings with the arousal ratings. Intense and negative arousing emotions were more consistent with the separate arousal ratings than intense and positive arousing emotions.

Generally, we found a higher correspondence of subjective valence trajectories and sentiment arcs in the genres that implied shifts between happiness and sadness, the happy-ending story and the tragedy compared to the thriller story. Subjective valence ratings corresponded much more consistently with structural elements such as the story climax. This is in line with preceding research on suspense trajectories over structural story patterns (Boyd et al., 2020). The immediate, physiological responses also confirm this: Consistently across all stories, we found sympathico-vagal double activation in the climax unit.

Within each story, we found very homogenous trajectories of the psychophysiological parameters, which speaks for the reliability of our method. In relation to the preceding

analyses, we found that physiological activation does not correspond strongly with ratings of suspense at any given narrative unit. Although the data was highly consistent in the climax units between genres, we found high variance at other points that were rated as particularly suspenseful, namely the introductions of the happy-ending story and the tragedy. There, it seems, the sympathetic activation that is associated with an intense feeling of suspense is overlayed by activation patterns of other dominant narrative emotions, such as happiness and sadness. In terms of excitation transfer, our findings support the idea that cumulation of arousal happens in successions of arousing states, but not from affective states of low arousal. Thus, future studies on excitation transfer in narrative reception should pay attention to emotional successions and the excitation potentials of dominant emotions apart from suspense. Furthermore, we sought to connect occurrences of high physiological activation with higher ratings of emotional intensity, as we would expect based on the Excitation Transfer Theory. We found rather little support for this between story units. No particularly high emotional intensity was rated in the end units after evidently arousing climax units. However, this does not necessarily mean that excitation transfer had not occurred within the respective climax unit. Our method is based on measures of aggregated physiological activation within story units. Here, we found high correspondence of increased physiological activation with increased emotional intensity. Therefore, we cannot draw a firm conclusion with the unit division that we used.

Limitations

Despite our substantial findings, at least four limitations of the present study must be noted. First, we were able to use only one story for each of the three narrative genres. Thus, whereas our assessment regarding emotional responses was rather broad and multi-faceted, our stimulus set was limited. Future research is encouraged to include multiple stories per genre and further controlled (or experimentally manipulated) properties such as narrator perspective.

Second, although the sentiment analysis we used offers an easily implementable approach to comparing textual and individual valence trajectories, it is also limited by its' word-for-word evaluation. In sections of possibly mixed or rapidly changing valence, the reduction of complexity becomes particularly apparent in aggregations over several sentences. However, it is also striking that the valence ratings of the audience can lead to markedly different conclusions regarding the overall valence rating of a story unit than the mere summation of the sentiments would lead one to expect. In our opinion, this shows that valence experience is more than the sum of its parts, and that research on emotional flow could profit from more elaborate text analytic methods. We encourage future research to test different text analytic approaches.

Third, we examined self-reported emotional responses on an aggregated level of prototypical plot units and forewent the investigation of moment-to-moment responses. Examining emotional responses as they occur is challenging, alternative approaches such as real-time-response measurement (Winkler et al., 2022) or self-probed emotional retrospections (Winkler et al., 2023) have their own limitations. That said, an investigation of smaller story segments or even moment-to-moment analyses would further clarify our findings. It should be noted, however, that reducing the size of the story segments will be especially difficult in terms of comparability of physiological and subjective data. Participants need mental landmarks for their retrospective judgements. With reduced segments sizes comes the obstacle of reduced possibilities for guidance into comparable mental units. In our study, for example, the subjective measurements for the thriller story did not differentiate between complication and climax unit for this reason. This, in turn, made it more difficult to determine where exactly the overall subjective valence shifted in this story. In this case, our determination was based on substitutional information of arousal profiles of the dominant underlying emotions.

Fourth, our methodological approach did not allow us to clearly differentiate between narrative and aesthetic emotions. The emotion word list included concepts that are more or less likely to be fiction- rather than aesthetics-related, and per instruction it explicitly referred to the story content. We therefore assume the retrospective measures to be mostly fiction-related but cannot rule out an unwanted influence of artefact-related emotions as a by-product of aesthetic appraisals of the story (Menninghaus et al., 2019; Menninghaus et al., 2020; Tan, 2013; Wassiliwizky et al., 2017). Future research could profit from a more detailed assessment of comprehension-related affect, such as suspense, curiosity, surprise (as suggested by Brewer & Liechtenstein, 1982), or confusion (as suggested, e.g., by D'Mello & Graesser, 2014). Clarifying the role of narrative comprehension for emotional flow could broaden our understanding of the relationship between narrative structure and emotional responses considerably.

Conclusion

Prior theory and research based on emotional flow (Nabi & Green, 2015) conceptualized the narrative reception process as dynamic, yet involved little explicit theorizing of narrative structure and its implications for emotional responses. This paper fills this gap by connecting these two branches of research. Based on Barrett's (2016) theory of constructed emotion (Barrett, 2013; Barrett et al., 2015) we investigated different response levels in narrative reception. We worked with stories of several genres and combined text-based analyses with analyses of emotional responses, measured immediately and retrospective. Our work shows that predictions of emotional trajectories based on the narrative material are reasonably possible and should be integrated into research on emotional flow. In line with previous work, we were able to demonstrate that emotional responses strongly relate to structural narrative elements, such as the rising action in the climax. The story plays a major role in classifying physiological and subjective data on emotional responses. Not least, we were able to clarify the trade-off between measurement accuracy and ease of

implementation based on our methods for future research and to point out a reasonable middle ground, which is especially important for follow-up research on emotional flow and excitation transfer.

References

Alam, N., & So, J. (2020). Contributions of emotional flow in narrative persuasion: An empirical test of the emotional flow framework. *Communication Quarterly*, *68*(2), 161-182. https://doi.org/10.1080/01463373.2020.1725079

Altmann, U., Bohrn, I. C., Lubrich, O., Menninghaus, W., & Jacobs, A. M. (2012). The power of emotional valence—from cognitive to affective processes in reading. *Frontiers in Human Neuroscience*, *6*, 192. https://doi.org/10.3389/fnhum.2012.00192

Amstad, T. (1978). *Wie verständlich sind unsere Zeitungen?* [How understandable are our newspapers?] (Unpublished doctoral dissertation). University of Zürich.

Appel, M., Schreiner, C., Haffmans, M. B., & Richter, T. (2019). The mediating role of event-congruent emotions in narrative persuasion. *Poetics*, *77*, 101385. https://doi.org/10.1016/j.poetic.2019.101385

Aristotle ([c. 350 BC] 2005). *Poetics*. Barnes & Noble.

Bal, P. M., & Veltkamp, M. (2013). How does fiction reading influence empathy? An experimental investigation on the role of emotional transportation. *PloS One*, *8*(1), e55341. https://doi.org/10.1371/journal.pone.0055341

Balzarotti, S., Biassoni, F., Colombo, B., & Ciceri, M. R. (2017). Cardiac vagal control as a marker of emotion regulation in healthy adults: A review. *Biological Psychology*, *130*, 54-66. https://doi.org/10.1016/j.biopsycho.2017.10.008

Bamberg, M. (2012). Narrative analysis. In H. Cooper (Ed.), *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 85-102). American Psychological Association. https://doi.org/10.1037/13620-006

Barraza, J. A., Alexander, V., Beavin, L. E., Terris, E. T., & Zak, P. J. (2015). The heart of the story: Peripheral physiology during narrative exposure predicts charitable giving. *Biological Psychology*, *105*, 138-143. https://doi.org/10.1016/j.biopsycho.2015.01.008

Barrett, L. F. (2013). Psychological construction: The Darwinian approach to the science of emotion. *Emotion Review*, *5*(4), 379-389. https://doi.org/10.1177/1754073913489753

Barrett, L. F. (2016). Navigating the science of emotion. In H. Meiselman (Ed.), *Emotion measurement* (pp. 31-63). Woodhead Publishing. https://doi.org/10.1016/B978-0-08-100508-8.00002-3

Barrett, L. F., & Russell, J. A. (2015). *The psychological construction of emotion*. Guilford Publications.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323-370. https://doi.org/10.1037/1089-2680.5.4.323

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1991). Autonomic determinism: the modes of autonomic control, the doctrine of autonomic space, and the laws of autonomic constraint. *Psychological Review, 98*(4), 459. https://doi.org/10.1037/0033-295x.98.4.459

Bezdek, M. A., Gerrig, R. J., Wenzel, W. G., Shin, J., Revill, K. P., & Schumacher, E. H. (2015). Neural evidence that suspense narrows attentional focus. *Neuroscience*, *303*, 338-345. https://doi.org/10.1016/j.neuroscience.2015.06.055

Bohanek, J. G., Fivush, R., & Walker, E. (2005). Memories of positive and negative emotional events. *Applied Cognitive Psychology*, *19*(1), 51-66. https://doi.org/10.1002/acp.1064

Boyd, R. L., Blackburn, K. G., & Pennebaker, J. W. (2020). The narrative arc: Revealing core narrative structures through text analysis. *Science Advances*, *6*(32), eaba2196. https://doi.org/10.1126/sciadv.aba2196

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49-59. https://doi.org/10.1016/0005-7916(94)90063-9

Brewer, W. F., & Lichtenstein, E. H. (1982). Stories are to entertain: A structural-affect theory of stories. *Journal of Pragmatics, 6*(5-6), 473-486. https://doi.org/10.1016/0378-2166(82)90021-2

Busselle, R., & Bilandzic, H. (2008). Fictionality and perceived realism in experiencing stories: A model of narrative comprehension and engagement. *Communication Theory*, *18*(2), 255-280. https://doi.org/10.1111/j.1468-2885.2008.00322.x

Busselle, R., & Bilandzic, H. (2009). Measuring narrative engagement. *Media Psychology*, *12*(4), 321-347. https://doi.org/10.1080/15213260903287259

Buth, C. (2014, April 22). *Äpfel* [Apples]. Planet Wissen (WDR). https://www.planet-wissen.de/gesellschaft/lebensmittel/aepfel_vom_paradies_in_jede_obstschale/pwwbaepfelvomparadiesinjedeobstschale100.html

Butler, E. A., Wilhelm, F. H., & Gross, J. J. (2006). Respiratory sinus arrhythmia, emotion, and emotion regulation during social interaction. *Psychophysiology*, *43*(6), 612-622. https://doi.org/10.1111/j.1469-8986.2006.00467.x

Carrera, P., Muñoz, D., & Caballero, A. (2010). Mixed emotional appeals in emotional and danger control processes. *Health Communication*, *25*(8), 726-736. https://doi.org/10.1080/10410236.2010.521914

Carroll, N. (2003). *Engaging the moving image*. Yale University Press.

Christ, L., Amiriparian, S., Baird, A., Tzirakis, P., Kathan, A., Müller, N., Stappen, L., Meßner, E.-M., König, A., Cowen, A., Cambria, E., & Schuller, B. W. (2022). The MuSe 2022 Multimodal Sentiment Analysis Challenge: Humor, emotional reactions, and stress. In *MuSe' 22: Proceedings of the 3rd International on Multimodal*

Sentiment Analysis Workshop and Challenge (pp. 5-14). Association for Computing Machinery. https://doi.org/10.1145/3551876.3554817

Coppin, G., & Sander, D. (2016). Theoretical approaches to emotion and its measurement. H. Meiselman (Ed.), *Emotion measurement* (pp. 3-30). Woodhead Publishing. https://doi.org/10.1016/B978-0-08-100508-8.00001-1

Cutting, J. E. (2016). Narrative theory and the dynamics of popular movies. *Psychonomic Bulletin & Review*, *23*(6), 1713-1743. https://doi.org/10.3758/s13423-016-1051-4

de Graaf, A., & Hustinx, L. (2011). The effect of story structure on emotion, transportation, and persuasion. *Information Design Journal*, *19*(2), 142-154. https://doi.org/10.1075/idj.19.2.05gra

Dillard, J. P., & Shen, L. (2018). Threat appeals as multi-emotion messages: An argument structure model of fear and disgust. *Human Communication Research, 44*(2), 103-126. https://doi.org/10.1093/hcr/hqx002

D'Mello, S., & Graesser, A. (2014). Confusion. In R. Pekrun & L. Linnenbrink-Gracia (Eds.), *International handbook of emotions in education* (pp. 299-320). Routledge. https://doi.org/10.4324/9780203148211

Dunlop, S. M., Wakefield, M., & Kashima, Y. (2010). Pathways to persuasion: Cognitive and experiential responses to health-promoting mass media messages. *Communication Research*, *37*(1), 133-164. https://doi.org/10.1177/0093650209351912

Eagleton, T. (2009). *Sweet violence: The idea of the tragic*. John Wiley & Sons.

Egidi, G., & Gerrig, R. J. (2009). How valence affects language processing: Negativity bias and mood congruence in narrative comprehension. *Memory & Cognition*, *37*(5), 547-555. https://doi.org/10.3758/MC.37.5.547

Eisenmann. (2008, July 11). *Die Typen, bei denen man wegsieht* [The guys you look away from]. [Online forum post]. Wortkrieger.

https://www.wortkrieger.de/index.php?threads/die-typen-bei-denen-man-wegsieht.57260/

El-Sheikh, M., Kouros, C. D., Erath, S., Cummings, E. M., Keller, P., & Staton, L. (2009). Marital conflict and children's externalizing behavior: Pathways involving interactions between parasympathetic and sympathetic nervous system activity. *Monographs of the Society for Research in Child Development, 74*(1), vii. https://doi.org/10.1111/j.1540-5834.2009.00501.x

Fahr, A., & Hofer, M. (2013). Psychophysiologische Messmethoden. In W. Möhring & D. Schlütz (Eds.), *Handbuch standardisierte Erhebungsverfahren in der Kommunikationswissenschaft* [Handbook of standardized survey procedures in communication science] (pp. 347-365). Springer VS. https://doi.org/10.1007/978-3-531-18776-1_19

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*(3), 221. https://doi.org/10.1037/h0057532

Frambach, S. (2018, February 09). *Fiction Friday – Noahs Bienen* [Fiction Friday – Noah's bees]. Tor Online (Fischer Verlag). https://www.tor-online.de/fiction/2018/02/noahs-bienen-sabine-frambach/

Freytag, G. (1863). *Die Technik des Dramas*. Verlag von S. Hirzel.

Grossman, P., Stemmler, G., & Meinhardt, E. (1990). Paced respiratory sinus arrhythmia as an index of cardiac parasympathetic tone during varying behavioral tasks. *Psychophysiology*, *27*(4), 404-416. https://doi.org/10.1111/j.1469-8986.1990.tb02335.x

Grossman, P., & Taylor, E. W. (2007). Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions. *Biological Psychology*, *74*(2), 263-285. https://doi.org/10.1016/j.biopsycho.2005.11.014

Hamby, A., & Brinberg, D. (2016). Happily ever after: How ending valence influences narrative persuasion in cautionary stories. *Journal of Advertising*, *45*(4), 498-508. https://doi.org/10.1080/00913367.2016.1262302

Hoeken, H., & Sinkeldam, J. (2014). The role of identification and perception of just outcome in evoking emotions in narrative persuasion. *Journal of Communication*, *64*(5), 935-955. https://doi.org/10.1111/jcom.12114

Jacobs, A. M. (2015). Neurocognitive poetics: Methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience*, *9*, 186. https://doi.org/10.3389/fnhum.2015.00186

Jockers, M. (2017). *Package „syuzhet“* (Version 1.0.6) [R software package]. https://cran. r-project. org/web/packages/syuzhet

Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology*, *68*, 601-625. https://doi.org/10.1146/annurev-psych-122414-033702

Kim, E., Padó, S., & Klinger, R. (2017). Investigating the relationship between literary genres and emotional plot development. In B. Alex, S. Degaetano-Ortlieb, A. Feldman, A. Kazantseva, N. Reiter, & S. Szpakowicz (Eds.), *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature* (pp. 17–26). Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-2203

Kimhy, D., Crowley, O. V., McKinley, P. S., Burg, M. M., Lachman, M. E., Tun, P. A., Ryff, C. D., Seeman, T. E., & Sloan, R. P. (2013). The association of cardiac vagal control and executive functioning – Findings from the MIDUS study. *Journal of Psychiatric Research*, *47*(5), 628-635. https://doi.org/10.1016/j.jpsychires.2013.01.018

Knobloch-Westerwick, S., & Keplinger, C. (2006). Mystery appeal: Effects of uncertainty and resolution on the enjoyment of mystery. *Media Psychology*, *8*(3), 193-212. https://doi.org/10.1207/s1532785xmep0803_1

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology*, *84*(3), 394-421. https://doi.org/10.1016/j.biopsycho.2010.03.010

Kreibig, S. D., Wilhelm, F. H., Roth, W. T., & Gross, J. J. (2007). Cardiovascular, electrodermal, and respiratory response patterns to fear-and sadness-inducing films. *Psychophysiology*, *44*(5), 787-806. https://doi.org/10.1111/j.1469-8986.2007.00550.x

Lang, P. J. (1980). *Self-assessment manikin*. University of Florida.

Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, *50*(1), 46-70. https://doi.org/10.1111/j.1460-2466.2000.tb02833.x

Lenth, R.V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software, 69*(1), 1-33. https://doi.org/10.18637/jss.v069.i01

Mann, P. H. (1985). Romantic fiction and its readership. *Poetics*, *14*(1-2), 95-105. https://doi.org/10.1016/0304-422X(85)90006-3

Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, *3*(3), 173-192.

Mar, R. A., Oatley, K., Djikic, M., & Mullin, J. (2011). Emotion and narrative fiction: Interactive influences before, during, and after reading. *Cognition & Emotion*, *25*(5), 818-833. https://doi.org/10.1111/j.1745-6924.2008.00073.x

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & Emotion*, *23*(2), 209-237. https://doi.org/10.1080/02699930802204677

Menninghaus, W., Schindler, I., Wagner, V., Wassiliwizky, E., Hanich, J., Jacobsen, T., & Koelsch, S. (2020). Aesthetic emotions are a key factor in aesthetic evaluation: Reply to Skov and Nadal (2020). *Psychological Review*, *127*(4), 650-654. https://doi.org/10.1037/rev0000213

Menninghaus, W., Wagner, V., Wassiliwizky, E., Schindler, I., Hanich, J., Jacobsen, T., & Koelsch, S. (2019). What are aesthetic emotions? *Psychological Review*, *126*(2), 171-195. https://doi.org/10.1037/rev0000135

Miall, D. S., & Kuiken, D. (2002). A feeling for fiction: Becoming what we behold. *Poetics*, *30*(4), 221-241. https://doi.org/10.1016/S0304-422X(02)00011-6

Murphy, S. T., Frank, L. B., Chatterjee, J. S., & Baezconde-Garbanati, L. (2013). Narrative versus nonnarrative: The role of identification, transportation, and emotion in reducing health disparities. *Journal of Communication*, *63*(1), 116-137. https://doi.org/10.1111/jcom.12007

Nabi, R. (2002). Anger, fear, uncertainty, and attitudes: A test of the cognitive-functional model. *Communication Monographs*, *69*(3), 204-216. https://doi.org/10.1080/03637750216541

Nabi, R. L., & Green, M. C. (2015). The role of a narrative's emotional flow in promoting persuasive outcomes. *Media Psychology*, *18*(2), 137-162. https://doi.org/10.1080/15213269.2014.912585

Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, *11*(1), 81. https://doi.org/10.1007/s13278-021-00776-6

Norris, C. J., Larsen, J. T., & Cacioppo, J. T. (2007). Neuroticism is associated with larger and more prolonged electrodermal responses to emotionally evocative pictures. *Psychophysiology*, *44*(5), 823–826. https://doi.org/10.1111/j.1469-8986.2007.00551.x

Piferi, R. L., Kline, K. A., Younger, J., & Lawler, K. A. (2000). An alternative approach for achieving cardiovascular baseline: Viewing an aquatic video. *International Journal of Psychophysiology*, *37*(2), 207-217. https://doi.org/10.1016/S0167-8760(00)00102-1

Pinna, T., & Edwards, D. J. (2020). A systematic review of associations between interoception, vagal tone, and emotional regulation: Potential applications for mental

health, wellbeing, psychological flexibility, and chronic conditions. *Frontiers in Psychology*, *11*, 1792-1792. https://doi.org/10.3389/fpsyg.2020.01792

Piper, A., So, R. J., & Bamman, D. (2021). Narrative theory for computational narrative understanding. In M.-F. Moens, X. Huang, L. Specia, S. Wen-tau Yih (Eds.), *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (pp. 298-311). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.emnlp-main.26

Potter, R. F., & Bolls, P. (2012). *Psychophysiological measurement and meaning: Cognitive and emotional processing of media*. Routledge.

Praetorius, C. (n.d.). *Die Steinigung des Klassentrottels* [The stoning of the class idiot]. PraetoriusCC. https://praetoriuscc.de/ccunst/geschichten/die-steinigung-des-klassentrottels/

Prestin, A. (2013). The pursuit of hopefulness: Operationalizing hope in entertainment media narratives. *Media Psychology*, *16*(3), 318-346. https://doi.org/10.1080/15213269.2013.773494

Ravaja, N. (2004a). Contributions of psychophysiology to media research: Review and recommendations. *Media Psychology*, *6*(2), 193-235. https://doi.org/10.1207/s1532785xmep0602_4

Ravaja, N. (2004b). Effects of a small talking facial image on autonomic activity: The moderating influence of dispositional BIS and BAS sensitivities and emotions. *Biological Psychology*, *65*(2), 163-183. https://doi.org/10.1016/S0301-0511(03)00078-4

Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, *5*, Article 31. https://doi.org/10.1140/epjds/s13688-016-0093-1

Remus, R., Quasthoff, U., & Heyer, G. (2010). SentiWS – A publicly available German-language resource for sentiment analysis. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, & D. Tapias (Eds.), *Proceedings of the 7th International Language Resources and Evaluation (LREC'10)* (pp. 1168-1171). European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2010/pdf/490_Paper.pdf

Richter, T. (2006). What is wrong with ANOVA and Multiple Regression? Analyzing sentence reading times with Hierarchical Linear Models. *Discourse Processes, 41*(3), 221-250. https://doi.org/10.1207/s15326950dp4103_1.

Rodrigue, J. R., Fleishman, A., Vishnevsky, T., Fitzpatrick, S., & Boger, M. (2014). Organ donation video messaging: Differential appeal, emotional valence, and behavioral intention. *Clinical Transplantation*, *28*(10), 1184-1192. https://doi.org/10.1111/ctr.12449

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161. https://doi.org/10.1037/h0077714

Shaffer, F., Combatalade, D., Peper, E. & Meehan, Z. M. (2016). A guide to cleaner electrodermal activity measurements. *Biofeedback*, 44(2), 90–100. https://doi.org/10.5298/1081-5937-44.2.01

So, J., Kuang, K., & Cho, H. (2016). Reexamining fear appeal models from cognitive appraisal theory and functional emotion theory perspectives. *Communication Monographs*, *83*(1), 120-144. https://doi.org/10.1080/03637751.2015.1044257

Stange, J. P., Hamilton, J. L., Fresco, D. M., & Alloy, L. B. (2017). Flexible parasympathetic responses to sadness facilitate spontaneous affect regulation. *Psychophysiology*, *54*(7), 1054-1069. https://doi.org/10.1111/psyp.12856

Svorc, P. (2018). Autonomic nervous system: What we know about it. In P. Svorc (Ed.), *Autonomic Nervous System* (pp. 3-12). IntechOpen. https://doi.org/10.5772/intechopen.81026

Tan, E.S. (2013). *Emotion and the structure of narrative film: Film as an emotion machine*. Routledge.

van Krieken, K. (2018). How reading narratives can improve our fitness to survive: A mental simulation model. *Narrative Inquiry*, *28*(1), 139-160. https://doi.org/10.1075/ni.17049.kri

Wassiliwizky, E., Koelsch, S., Wagner, V., Jacobsen, T., & Menninghaus, W. (2017). The emotional power of poetry: Neural circuitry, psychophysiology and compositional principles. *Social Cognitive and Affective Neuroscience*, *12*(8), 1229-1240. https://doi.org/10.1093/scan/nsx069

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General, 143*(5), 2020-2045. https://doi.org/10.1037/xge0000014

Wilhelm, F. H., Grossman, P., & Coyle, M. A. (2004). Improving estimation of cardiac vagal tone during spontaneous breathing using a paced breathing calibration. *Biomedical Sciences Instrumentation*, *40*, 317-324. https://pubmed.ncbi.nlm.nih.gov/15133978/

Winkler, J. R., Appel, M., Schmidt, M.-L. C. R., & Richter, T. (2022). The experience of emotional shifts in narrative persuasion. *Media Psychology*. Ahead-of-print. https://doi.org/10.1080/15213269.2022.2103711

Winkler, J. R., Mengelkamp, C., & Appel, M. (2022). Real-time responses to stories: Linking valence shifts to post-exposure emotional flow and transportation. *Communication Research Reports*, *39*(5), 237-247. https://doi.org/10.1080/08824096.2022.2119380

Zammuto, M., Ottaviani, C., Laghi, F., & Lonigro, A. (2021). The heart in the mind: A systematic review and meta-analysis of the association between Theory of Mind and cardiac vagal tone. *Frontiers in Physiology*, *12*. https://doi.org/10.3389/fphys.2021.611609

Zillmann, D. (1988). Mood management: Using entertainment to full advantage. In L. Donohew, H. E. Sypher & E. T. Higgins (Eds.), *Communication, social cognition, and affect* (pp. 147–171). Erlbaum.

Zillmann, D. (1996). The psychology of suspense in dramatic exposition. In P. Vorderer, H.J. Wulff & M. Friedrichsen (Eds.), *Suspense: Conceptualizations, theoretical analyses, and empirical explorations* (pp. 199-231). Routledge.

Tables

Table 1

Description of Plot Units per Story

Plot Unit	**Happy-ending Story** ***"Noah's Bees"***	**Tragedy** ***"The class idiot"***	**Thriller Story** ***"Guys that make you look the other way"***
Introduction - Setting of the scene (place, time, backstory) - Introduction of protagonist	The year is 2046. Lissa is a scientist who works at the plantation 'Green Delta' that was built to cultivate plants for diverse climatic environments. In 2025 a huge storm had decimated the unknown country's population and resources. After that, the seasons of the northern hemisphere started to shift. Lissa's job is to mathematically predict these changes. Her colleague, Noah, and his bee population that pollinates the plants depend on her calculations.	A person narrates about a school trip in tenth grade to a youth hostel in a small village in South Tyrol (with games, barbeques, and secretly smoking cigarettes and having drinks). The class hikes into a canyon. They arrive at a lovely place in nature and take a break to eat. Their teacher reads and the 'class idiot' goes to collect rocks.	The protagonist awakes in a tramway car in the middle of the night. He had celebrated his promotion at a work party, fallen asleep and had missed his stop. The terminal stop is ahead and he is alone in the tramway car. The protagonist monologizes about having proven his mother wrong, who thinks of him as a no-good. When the tram stops, he goes to the door and pushes the 'open' button.
Complication - An event foreshadows the rising action	Noah calls a staff meeting to discuss the problem that this year's spring will already begin while his bees are still hibernating. When they wake later, they will likely starve and die, and the crop and people will suffer the consequences of the late	The protagonist describes the 'class idiot' with autistic features. He recounts how the 'idiot' ate all the leftovers at breakfast. The other kids were disgusted and made fun of him, but the 'idiot' did not understand them. The protagonist once tried to	The door does not open. A brute looking man enters the tramway car at the other end. The protagonist tries to talk to him, but the man only insults him and pulls out a hammer. The protagonist is in disbelief and hopes for it to be a joke. The man

- Protagonist is faced with a hardship, decision or task	pollination. After some brainstorming, a colleague suggests to awaken the bees artificially.	connect with him, but the 'idiot' did not care. The class watches him pick up rocks from a river.	approaches him. The protagonist threatens to call the police, but the man knocks his phone out of his hand. The protagonist tries to run away.
Climax - Action rises to a critical point for the protagonist	Lissa has to calculate an accurate date for the beginning of spring with only a few days tolerance. Her calculations do not match those of the calculation software. When asked about this, she admits that she corrects the programs values by intuition and chance (using a dice and deciding to subtract or add this number according to her instinct).	One kid starts to bully the 'class idiot', picks up a rock and throws it into his direction. Some other kids laugh and join in. The 'idiot' laughs. More kids throw more rocks at him. The protagonist is paralyzed by his feeling of danger, but does nothing as the situation exacerbates. A rock hits the 'idiot' in the head and he falls over.	The man laughs and follows him. He hits the protagonists arm with his hammer. He screams in pain, trips over and falls to the floor. As he tries to stand up, the man hits his ankle, and comes close. He explains that there are no witnesses and the police will never know it was him. The man grabs his hair and pulls his head up to hit him.
End - Resolves the critical point - Final consequences	Noah has started the artificial awakening process and Lissa waits. After a while, the bees start to move and fly. A leap in time reveals the success of the mission. Noah leaves for another research station and entrusts his local bee population in Lissa's care.	The teacher lies her book aside and helps the 'idiot' up. She looks at his head, decides that it is not a big deal, and gives the class a reprimanding look. They pack up and leave the canyon. No one mentions the happenings further. In the evening, the 'idiot' eats the leftovers and laughs.	A bearded man hits the man's head with a bottle and he goes down. As the man tries to get up and takes his hammer, the bearded man stabs him with the broken bottleneck. The protagonists awakes under a reeking rag. The bearded man is a homeless person that witnessed the attack, send his friend to call the police and helped the protagonist. He thanks him for not looking the other way.

Table 2

Means and Standard Errors of Suspense Ratings per Story Unit and Genre

		Genre		
		Happy-Ending Story	Tragedy	Thriller Story
		M (SE), n_k	*M (SE), n_k*	*M (SE), n_k*
Story Unit	Introduction	4.80 (0.36), *n* = 10	5.00 (0.31), *n* = 7	2.00 (NA), *n* = 1
	Complication	4.43 (0.48), *n* = 7	5.00 (0.00), *n* = 4	5.79 (0.24), *n* = 14
	Climax	4.78 (0.32), *n* = 9	4.80 (0.58), *n* = 5	
	End	4.00 (2.00), *n* = 2	4.50 (0.50), *n* = 2	4.75 (0.48), *n* = 4

Note. Happy-ending story, *N* = 20; tragedy, *N* = 21; thriller story, *N* = 20. n_k represent total numbers of emotion ratings per story unit.

Table 3

Changes of Psychophysiological Measures Between Each Story Unit and Genre

Happy-Ending Story

		Introduction - Complication			Complication - Climax			Climax - End		
		t value	*d*		*t value*	*d*		*t value*	*d*	
Sympathetic Activation	SCL	$t(34) = 1.89$		*n.s.*	$t(33.7) = 7.04$	2.76	***	$t(33.1) = -0.72$		*n.s.*
	Sympathetic Power	$t(53.4) = 3.40$	0.86	**	$t(53.3) = 7.73$	1.95	***	$t(53) = -0.64$		*n.s.*
Vagal Activation	Vagal Power	$t(53.4) = 0.10$		*n.s.*	$t(53.3) = 6.83$	1.73	***	$t(53) = -1.33$		*n.s.*
	RSA	$t(51.6) = 2.38$	2.10	*	$t(64) = 4.64$	2.10	***	$t(49.1) = 1.49$		*n.s.*

Tragedy

		Introduction - Complication			Complication - Climax			Climax - End		
		t value	*d*		*t value*	*d*		*t value*	*d*	
Sympathetic Activation	SCL	$t(43.6) = -14.07$	-4.80	***	$t(45.7) = 4.93$	1.69	***	$t(44.3) = 12.55$	4.29	***
	Sympathetic Power	$t(60) = -8.99$	-2.32	***	$t(60) = 3.56$	0.92	**	$t(60) = 1.86$		*n.s.*

Vagal Activation	Vagal Power	$t(60) = -7.53$	-1.95	***	$t(60) = 4.09$	1.06	***	$t(60) = 16.72$	4.32	***
	RSA	$t(64.5) = -4.29$	-2.45	***	$t(60.9) = 3.31$	1.10	**	$t(64.6) = 4.68$	2.91	***

Thriller Story

		Introduction - Complication			Complication - Climax			Climax - End		
		t value	*d*		*t value*	*d*		*t value*	*d*	
Sympathetic Activation	SCL	$t(35) = -2.14$	-0.96	*	$t(27) = 7.28$	3.25	***	$t(27) = -11.00$	-4.92	***
	Sympathetic Power	$t(45) = -3.24$	-0.74	**	$t(45) = 10.28$	2.35	***	$t(45) = -13.97$	-3.19	***
Vagal Activation	Vagal Power	$t(45) = 0.28$		*n.s.*	$t(45) = 4.37$	1.00	***	$t(45) = -8.53$	-1.95	***
	RSA	$t(32.6) = -2.02$		*n.s.*	$t(32.1) = 6.33$	1.81	***	$t(32.4) = -9.74$	-2.78	***

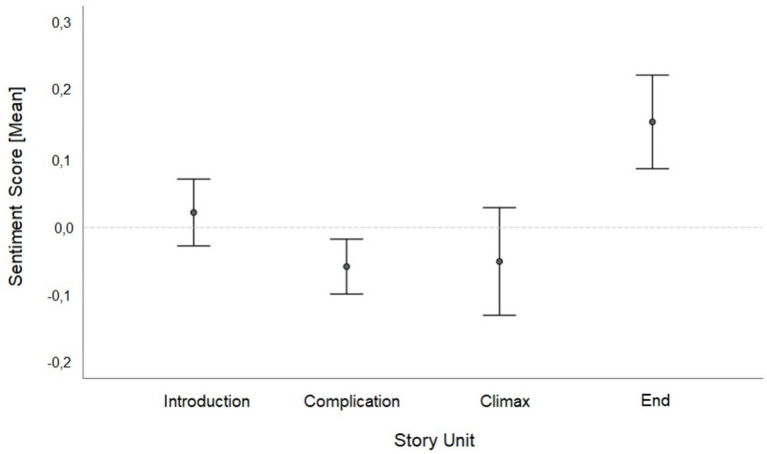
Note. Valence shifts are marked in grey, positive values indicate increases.

Figures

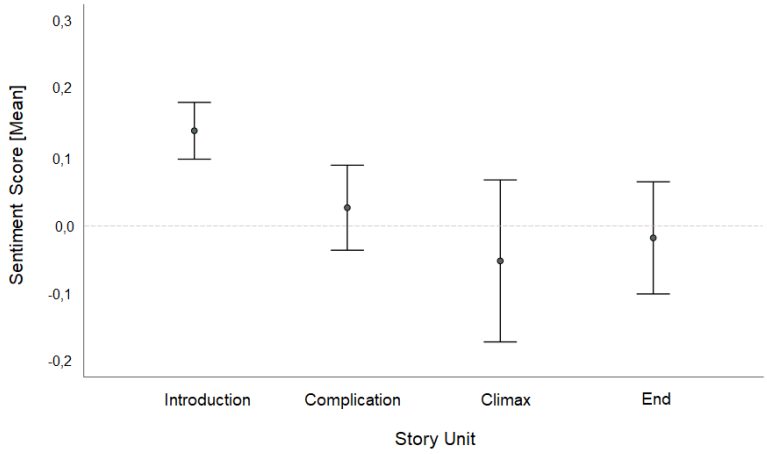
Figure 1

Sentiment scores of each genre story

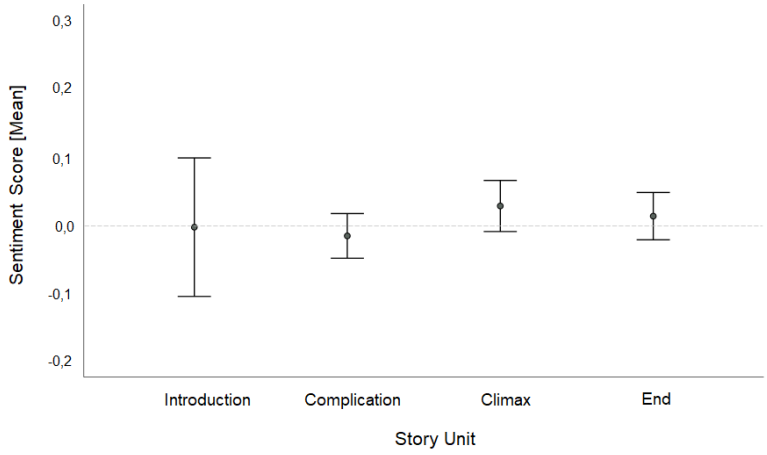
a) Happy-ending story



b) Tragedy



c) Thriller story

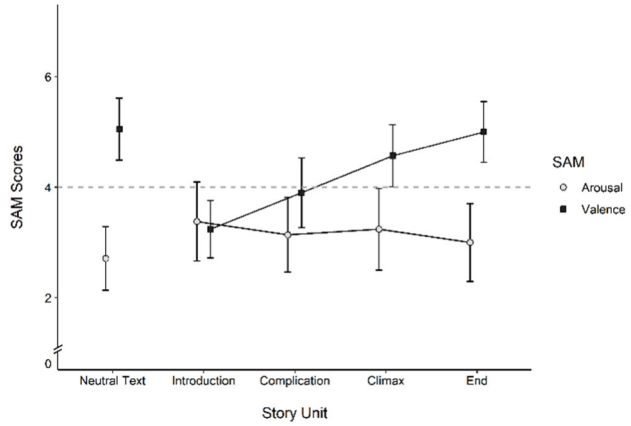


Note. Error bars represent the standard error of the mean; dotted line marks neutral valence; happy-ending story, $N = 111$; tragedy, $N = 50$; thriller story, $N = 106$.

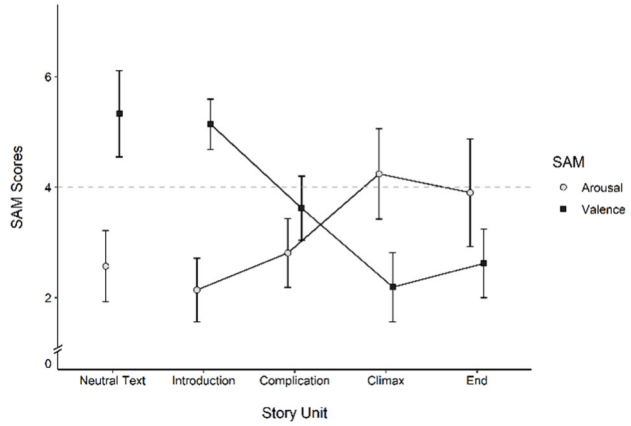
Figure 2

Subjective valence and arousal scores of each genre

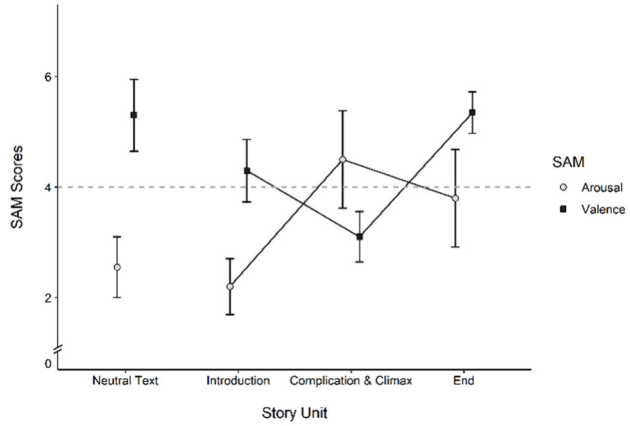
d) Happy-ending story



e) Tragedy



f) Thriller story

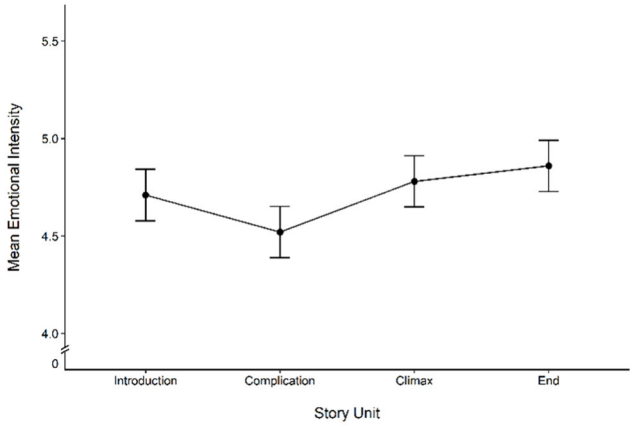


Note. Error bars represent the standard error of the mean; dotted line marks neutral valence; happy-ending story, $N = 21$; tragedy, $N = 20$; thriller story, $N = 20$.

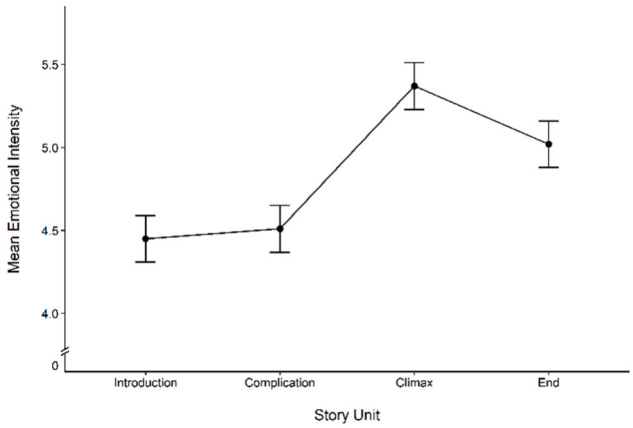
Figure 3

Emotional intensity scores of each genre

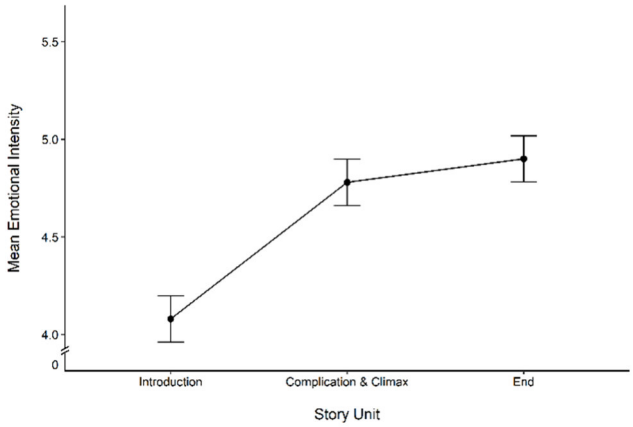
a) Happy-ending story



b) Tragedy



c) Thriller story

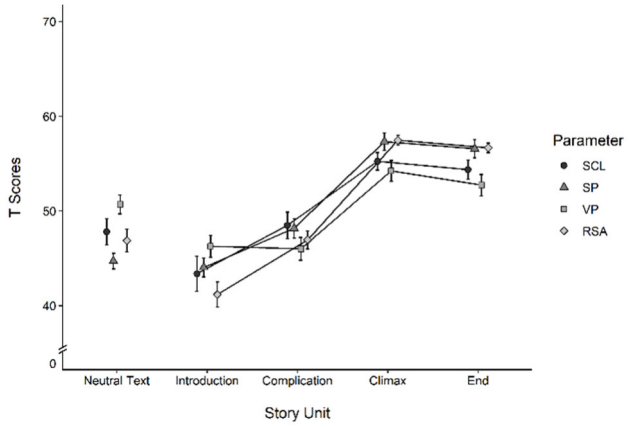


Note. Error bars represent the standard error of the mean; happy-ending story, $N = 20$; tragedy, $N = 21$; thriller story, $N = 20$.

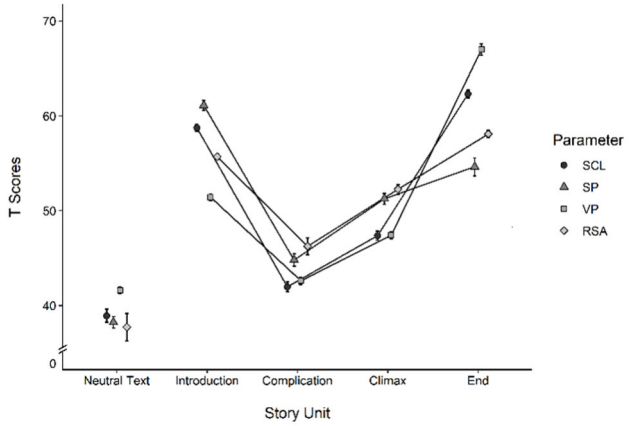
Figure 4

T scores of psychophysiological measures of each genre

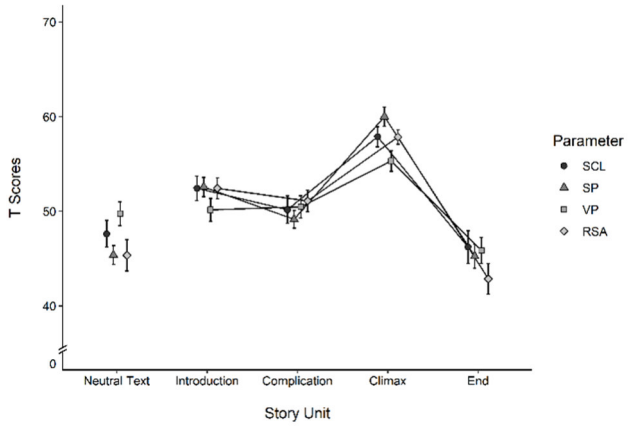
a) Happy-ending story



b) Tragedy



c) Thriller story



Note. Mean T scores of parameter change per second, error bars represent the standard error of the mean; SCL: skin conductance level, SP: sympathetic power, VP: vagal power, RSA: respiratory sinus arrythmia; happy-ending story, SCL: $N = 13$; SPR & VPR: $N = 19$; RSA_N: $N = 17$; tragedy, SCL: $N = 17$; SPR & VPR: $N = 21$; RSA_N: $N = 18$; thriller story, SCL: $N = 10$; SPR & VPR: $N = 16$; RSA_N: $N = 12$.