VU VRIJE UNIVERSITEIT AMSTERDAM

# VU Research Portal

## Shaping film: A quantitative formal analysis of contemporary empathy-eliciting Hollywood cinema

Lankhuizen, Tess; Balint, Katalin; Konijn, Elly; Savardi, Mattia; Bartsch, Anne; Benini, Sergio

***published in***
Psychology of Aesthetics, Creativity and the Arts
2020

***DOI (link to publisher)***
10.1037/aca0000356

***document version***
Publisher's PDF, also known as Version of record

***document license***
Article 25fa Dutch Copyright Act

**Link to publication in VU Research Portal**

***citation for published version (APA)***
Lankhuizen, T., Balint, K., Konijn, E., Savardi, M., Bartsch, A., & Benini, S. (Accepted/In press). Shaping film: A quantitative formal analysis of contemporary empathy-eliciting Hollywood cinema. *Psychology of Aesthetics, Creativity and the Arts*. https://doi.org/10.1037/aca0000356

**General rights**
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

**Take down policy**
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**E-mail address:**
**vuresearchportal.ub@vu.nl**

Download date: 22. Jan. 2022

Psychology of Aesthetics, Creativity, and the Arts

# Shaping Film: A Quantitative Formal Analysis of Contemporary Empathy-Eliciting Hollywood Cinema

Tess Lankhuizen, Katalin E. Bálint, Mattia Savardi, Elly A. Konijn, Anne Bartsch, and Sergio Benini

Online First Publication, September 24, 2020. http://dx.doi.org/10.1037/aca0000356

CITATION
Lankhuizen, T., Bálint, K. E., Savardi, M., Konijn, E. A., Bartsch, A., & Benini, S. (2020, September 24). Shaping Film: A Quantitative Formal Analysis of Contemporary Empathy-Eliciting Hollywood Cinema. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. http://dx.doi.org/10.1037/aca0000356

AMERICAN PSYCHOLOGICAL ASSOCIATION
DIVISION 10
Psychology of Aesthetics, Creativity, and the Arts

© 2020 American Psychological Association
ISSN: 1931-3896
http://dx.doi.org/10.1037/aca0000356

# Shaping Film: A Quantitative Formal Analysis of Contemporary Empathy-Eliciting Hollywood Cinema

Tess Lankhuizen and Katalin E. Bálint
Vrije Universiteit Amsterdam

Mattia Savardi
University of Brescia

Elly A. Konijn
Vrije Universiteit Amsterdam

Anne Bartsch
University of Leipzig

Sergio Benini
University of Brescia

Previous research suggests that particular formal features of film, such as the use of close-ups, can affect the levels of empathy experienced by viewers. Because empathy is a key aspect of the audience's filmic experience, creative decisions in editing and cinematography may be motivated by the filmmaker's intention of eliciting empathy. The goal of this study was to investigate what film scenes intended to elicit empathy look like in terms of those visual formal features theoretically or empirically linked to viewer empathy and whether these features converge on something that might be dubbed an empathic style of cinema. Formal features included concern shot scale, face depiction, cut rate, camera perspective and angle, saturation, lighting, motion, and background clutter. Exploratory quantitative formal analyses of scenes sampled from contemporary popular empathy-eliciting Hollywood films ($N = 100$) revealed that such scenes are, at first glance, highly dissimilar in form. Further investigation through principal component analysis and correlational analysis, however, hinted not so much at a singular empathic style of cinema as it did at certain general principles, namely, the reduction of perceived distance through close-ups and face depiction, the balancing of arousing features with comprehensible levels of visual complexity, and the prioritization of coherence and reduced visual contrast to enable a smooth viewing experience.

*Keywords:* empathy, film, formal features, narrative, popular cinema

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

As almost anyone who frequents film theaters can testify, cinematic narratives can elicit powerful emotions for and heartfelt engagement with characters that have never lived a day off-screen. These affective responses to cinema are perhaps the most obvious example of the many types of gratification viewers may derive from film (Tan, 2008; see also Bartsch, 2012; Oliver & Raney, 2011; Zillmann, 1988). Evoking pleasurable or rewarding responses is often of prime importance to the success of mainstream fiction films, which aim to provide an entertaining narrative experience to mass audiences.

In the attempt to create such a fulfilling viewing experience, popular cinema often seeks to elicit empathic engagement from viewers, as empathy is believed to offer a way for the audience to emotionally latch on to the plot of the film (Grodal, 1999; Plantinga, 2009). This investment, in turn, is what allows viewers to be moved, to be thrilled, to laugh and cry. In the current neuropsychological literature, empathy is understood as a core affective-cognitive process through which individuals strive to understand the contents of the mind of another (Cuff, Brown, Taylor, & Howat, 2016; Decety & Jackson, 2006), with affective empathy relating to emotional contagion and shared affect, and cognitive empathy relating to the imagination and deliberate perspective-taking. As a fundamentally social species, empathy is an essential ability for humans, and consequently, it is a highly developed and sensitive process (Cosmides & Tooby, 1992; de Waal, 2008).

Tess Lankhuizen and Katalin E. Bálint, Department of Communication Science, Vrije Universiteit Amsterdam; Mattia Savardi, Department of Information Engineering, University of Brescia; Elly A. Konijn, Department of Communication Science, Vrije Universiteit Amsterdam; Anne Bartsch, Department of Communication and Media Studies, University of Leipzig; Sergio Benini, Department of Information Engineering, University of Brescia.

The authors have no known conflicts of interest to disclose. This study was funded through Project Grant 392211046 by the Deutsche Forschungsgemeinschaft, awarded to Katalin E. Bálint and Anne Bartsch. We thank the student assistants who contributed to the coding and data collection, Benedikt Aigner, Carolin Bruhn, Johanna Gill, and Deborah Kunze, as well as Freya Sukalla from the Department of Communication and Media Studies at the University of Leipzig for facilitating data collection. The supplemental materials for this article are available on the Open Science Framework.

Correspondence concerning this article should be addressed to Tess Lankhuizen, Department of Communication Science, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam. E-mail: t.lankhuizen@vu.nl

In addition to its central role in narrative enjoyment (Bartsch, 2012; Plantinga, 2009; Tan, 2008), empathy has also been established as an essential aspect of various narrative processes and responses (Busselle & Bilandzic, 2009; Coplan, 2008). By allowing the recognition and understanding of mental states in characters, it enables narrative comprehension (Kneepkens & Zwaan, 1995). It serves as a mediator between narrative exposure and media effects (Shen, 2010), predicts narrative persuasion (Green & Brock, 2000), and is an important aspect of narrative engagement (Busselle & Bilandzic, 2009).

The work of popular filmmakers, then, involves a series of creative decisions targeted at eliciting empathy in order to draw the audience into the narrative world they have created. In this view, filmmakers may be understood as "practical psychologists" (Taberham, 2018, p. 1) who direct not only their film but also the emotions of their audience. Taberham's conceptualization of the filmmaker echoes Tan's (1996) idea of film as an "emotion machine," carefully crafted to marshal viewer experiences. When considering the manipulation of audience emotions and empathic engagement, story and writing are often the elements receiving the most attention. However, as pointed out by various scholars of film, the audience experiences elicited by any given narrative arise from the interaction between story content and the medium-specific presentation of said content (Bordwell, 1985; Hasson et al., 2008; Zettl, 1990).

Indeed, a growing body of research shows that the *formal features* of film, that is, the stylistic and aesthetic properties of the presentational format, such as cinematography and editing (Detenber & Lang, 2011), play an important role in determining cognitive and emotional viewer responses. Examples include studies on the effect of shot scale on film mood (Benini, Savardi, Bálint, Kovács, & Signoroni, 2019) and on viewer empathy (Bálint, Klausch, & Pólya, 2018; Cao, 2013); the perceived affective character of a film based on lighting, color, and motion (Wang & Cheong, 2006); and the link between shot scale, shot duration, and narrative comprehension (Smith, Levin, & Cutting, 2012).

However, although research on these tools and their implementation is steadily expanding, only a few of these studies directly relate formal features to the empathy-eliciting nature of popular fiction films (notable exceptions include Bálint et al., 2018; Cao, 2013; Rooney & Bálint, 2018). Those that have tend to focus on only a small set of formal features in an equally small number of fiction films. This limits our understanding of the choices made by filmmakers and the meaning behind specific formal features.

The goal of this study was to address this lack of insight into what mainstream empathy-eliciting cinema looks like by describing a wide range of visual formal features that are theoretically or experimentally linked to empathy, with a focus on contemporary popular empathy-eliciting Hollywood (CPEEH) cinema. Quantitative style analysis of scenes from CPEEH films in the vein of Salt (1974) or Svanera, Savardi, Signoroni, Kovács, and Benini (2019) may reveal what visual formal features characterize empathy-eliciting cinema and indicate if Hollywood, in its attempt to direct the empathic experience of audiences, has converged on something that might be called an empathic style of cinema.

These questions were explored through manual and algorithmic analyses of the relevant visual formal features of 100 film scenes taken from successful Hollywood films with a dramatic element that were released between 2000 and 2018. The results of these detailed analyses may further our understanding of the stylistic choices made by directors, editors, and cinematographers in their effort to create a film that inspires empathic engagement. The results may also inform future studies on media effects and audience responses, as well as the workings of empathy and human cognition in relation to complex audiovisual stimuli. Additionally, the annotated film-scene database created over the course of this project may serve as a resource for other researchers.

## Film Form and Empathy

The formal features of any narrative directly relate to the affordances of the medium that the story is told in. In the case of film, an audiovisual medium, these formal features are associated with the aural and visual tracks through which story content is communicated. These formal features can be either *high-level features* relating to structure and semantics, such as narrative organization and mood, or *low-level features* relating to craft and technique, such as lighting and color grading (Rasheed, Sheikh, & Shah, 2003; Tarvainen, Westman, & Oittinen, 2015). This study focused on the low-level visual features relating directly to the camera and (digital) film that are part of the film's editing or cinematography. Previous research provides suggestions on what features may be theoretically or empirically linked to empathy in film viewers, as outlined in the following subsections.

### Face Depiction

The visual nature of film may affect empathy in and of itself because the mere observation of human faces and gazes can affect viewers both in real life and in film (Frischen, Bayliss, & Tipper, 2007). An evolutionary explanation for this observation is the prime importance of recognizing, processing, and responding to the faces of other people for a species that is as intensely social as humans are (Bruce & Young, 2013; Tomasello, 2009). Research into empathy has previously noted how simply watching other people go through an experience can resonate with observers on a neural level (Singer et al., 2004). Theories of embodied simulation recognize the observation of others as not just a visual process but also as a multimodal process involving mirror neurons and areas of the brain relating to emotional processing, motor processing, and somatosensory processing (Gallese & Guerra, 2019). Basic functional brain mechanisms normally used for direct interaction with the external world are also used for perception and imagination, making them relevant to the filmic experience as viewers understand both real-world others and characters on the screen through the reuse of the same neural circuits used for processing their own experiences (Gallese & Guerra, 2019). Thus, the depiction of actors' faces is likely to activate semiautomatic forms of affective empathy related to the observation of others.

### Shot Scale

With the observation of the faces and bodies of actors as a central aspect of the embodied simulation account of empathic engagement with film, it is only natural to consider the importance of shot scale to filmic experiences. The scale of a shot is determined by the distance between the subject of the shot and the camera, along with digital or lens zoom, and relates to the field of

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

view (Bordwell & Thompson, 2013). Although this distance can technically have an infinite amount of variation, shot scale is generally understood to have three main categories: (a) close-ups, in which only a small part of the mise-en-scène is shown in such detail that it almost fills the frame; (b) medium shots, in which the subject and setting occupy roughly equal space in the frame; and (c) long shots, which show all or most of the subject and much of the surroundings (Benini, Svanera, Adami, Leonardi, & Kovács, 2016). Shot scale has been established as an important cue for empathy, with research showing that close-ups increase viewer cognitive empathy (Bálint et al., 2018; Cao, 2013; Rooney & Bálint, 2018), mostly because they offer a detailed picture of the actor's face and simulate physical proximity.

Furthermore, Canini, Benini, and Leonardi (2011) found shot scale and shot scale patterns (or the likelihood of one shot scale type changing into another) to play an evident role in arousal elicitation. Less arousing scenes were found to be fairly consistent in terms of shot scale, whereas more arousing scenes were less predictable. Arousal is an important component of emotion in many emotion theories (as reviewed by Scarantino & de Sousa, 2018), and some researchers have hypothesized on the use of measures of arousal such as skin conductivity as indirect psychophysiological measures of empathy (Neumann & Westbury, 2011). In the case of film, increases in arousal through form could amplify the intensity of affective story responses, including empathy.

Finally, Cutting and Armstrong (2016) found both shot duration and the speed at which viewers made accurate emotion judgment to be related to shot scale, with close-ups being both the shortest in duration and in required time for the identification of the actor's emotion. These findings indicate that shot scale may affect viewers' empathic engagement not only through enhanced embodied simulation but also through the facilitation of emotion identification.

## Camera Angle

Previous studies in advertising have shown that the camera angle in the vertical plane may affect how favorably objects are evaluated (Meyers-Levy & Peracchio, 1992). Further research into the effect of camera angle on the evaluation of characters shows that characters shot from a low angle (forcing the viewer to look "up" at them) were thought to be taller, stronger, less afraid, bolder, and more aggressive than characters shot from a high angle (forcing the viewer to look "down" at them; Kraft, 1987). Camera angle may therefore affect empathic engagement by affecting audience attitudes toward and evaluations of characters. Additionally, the sense of parity that, according to Kraft (1987), follows from the use of neutral camera angles may increase empathy by enhancing perceived similarity to the subject of the camera.

## Camera Perspective

Previous research on perspective in video games found games that use a first-person perspective to be more engaging and immersive (Denisova & Cairns, 2015; Tamborini et al., 2001), and research on sports audiences found footage in the first-person perspective to elicit a greater sense of presence (Cummins, Keene, & Nutting, 2012). Work on empathic engagement and perspective-taking in novels has shown that readers adopt the spatiotemporal point of view of the protagonist while empathizing (Rinck & Bower, 1995). Internal shots (or first-person shots) in films may similarly facilitate empathic responses.

## Saturation

Saturation is the intensity or amount of gray of a color. Saturated colors, also described as "rich" or "intense," are known to be more arousing than unsaturated colors while also being evaluated more positively (Valdez & Mehrabian, 1994; Wilms & Oberfeld, 2018). As mentioned previously, such increases in arousal may enhance emotional processes such as empathy.

## Lighting

Valdez and Mehrabian (1994) also found that brightness, or the perception of light as per the luminance of the frame, correlates with both arousal and valence. Additionally, brightness and lighting are related to the emotional identity of films (Wang & Cheong, 2006). In the same study, shadow area, contrast, and use of high-key versus low-key lightning were found to influence the perceived affective character of the scene. The relative whiteness or blackness (or relative perceptions of light and darkness) and proportion of shadow area orient viewers in space and time and are thought to affect their feelings toward that which is shown (Zettl, 1990).

## Motion

As a visual medium, film is defined by its temporal dimension, allowing it to depict the movement of objects in front of the camera and the movement of the camera itself (Zettl, 1990). Lang's (2000) research has established motion as a potential orienting cue, enhancing mental resources available for interpretation. According to this research, motion is one of several formal features that automatically draw attention and thus bypass voluntary resource allocation. In another study, Visch and Tan (2009) found velocity to be related to perceived emotion: Participants interpreted the slow movement of abstract shapes to signify sadness. As noted by Wang and Cheong (2006), there is also a reliable correlation between perceived motion and arousal, likely a result of our instinctive association of movement with potential danger and excitement (see also Detenber & Lang, 2011; Ravaja, 2004; Simons, Detenber, Cuthbert, Schwartz, & Reiss, 2003). In the context of audiovisual media, Wang and Cheong propose that perceived motion can be measured not only as the (illusionary) distance traveled by objects on screen but also as visual excitement, or the changes in chromaticity and brightness of pixels that translate to the perception of exciting movement in viewers. Based on these observations, it is not unlikely for the amount of motion in a scene to influence viewer empathy through its effect on attention, arousal, and perceived emotion.

## Background Clutter

As previously established, recognizing and reading faces is an important factor in viewer empathy, making background clutter a relevant feature to consider. The limited capacity model (LCM) suggests that it is possible for a message to become so visually complex that it overtaxes the available mental resources for mes-

sage encoding, decoding, and processing, lowering the comprehension of and engagement with said message (Lang, 2000; Lang, Zhou, Schwartz, Bolls, & Potter, 2000). Background clutter, or the amount of visual distraction in the background of a frame, increases visual complexity and, past a certain threshold, may reduce interpretative resources. Furthermore, Cutting and Armstrong (2016) found that background clutter makes it harder for viewers to identify facial expressions in film frames.

### Shot Composition

Similar to background clutter, the complexity of the composition of the frame may affect viewers' ability to attend to and process all relevant information in the frame. Current algorithms can predict the visual focal point of the scene (i.e., the point where viewers are most likely to look; Kümmerer, 2017). Scenes with many such focal points may be more taxing on viewers' mental resources, as might scenes with focal points that strongly deviate from the center of the screen or that break the compositional rule of thirds. With this increased visual complexity, empathic engagement may decrease through a shortage of mental resources.

Because film is a temporal medium, the compositional complexity should not only be considered within frames but also between frames. The change of focal point between frames may also relate to visual complexity as the viewer is forced into many saccades to keep up with jumpy editing—a problem editors of action films frequently grapple with (Nedomansky, 2015). Similarly, cuts that introduce a large amount of new visual information, such as cuts to a different environment than previously established, may tax the viewer's resources for visual processing (Lang, 2000). Visual shot coherence, then, is another factor to consider in relation to the visual and compositional complexity of a film.

### Shot Length

The duration of a shot between two cuts or other transitions is an important factor determining the pacing of a film. It is closely related to the cut rate, the frequency with which the film changes from one shot to the next. Previous research has pointed to cuts as potential orienting cues that can enhance attention and increase available mental resources (Lang, 2000). Additionally, increases in the cut rate are commonly associated with higher levels of arousal in viewers (Lang, Chung, Lee, Schwartz, & Shin, 2005). There are indications that the effect of shot length on engagement depends on content; violent scenes were found to be more emotionally engaging with a briefer shot length (Bálint, Schoft, & Rooney, 2017), whereas scenes with prosocial content were found to be more emotionally engaging with longer shots (Sukalla, Bartsch, & Schnell, 2015). Finally, the patterning of shot duration across a scene was found to be a reliable indicator of authorship, indicating that directors make unique and recognizable decisions about the length of successive shots (Svanera et al., 2019). The same study also found that shots of any given length were more likely to be followed by shots of similar length and that cuts from very short to very long takes (and vice versa) were rare.

## Developments in Formal Feature Research

In sum, the previously presented overview of the literature indicates that the formal features of film can be expected to affect viewer empathy either directly or through other cognitive processes, such as comprehension, attention, and arousal, as predicted by the LCM and dimensional theories of emotion (Scarantino & de Sousa, 2018). The early 2000s saw a rise in research concerning the effect of formal features, following the formulation of the LCM (Lang, 2000), with an emerging focus on arousal and message-sensation value (e.g., Morgan, Palmgreen, Stephenson, Hoyle, & Lorch, 2003; Niederdeppe, Davis, Farrelly, & Yarsevich, 2007; Stephenson & Palmgreen, 2001). Two major complications, however, eventually stifled interest in such research. First, formal features could generally only be measured at the dichotomous level (i.e., the presence or absence of a particular feature), and second, the factorial design of most of these studies allowed only for a small number of features to be considered at once.

With the advent of sophisticated algorithmic analysis, communication and media research into the effect of formal features on audiences has experienced a revival (e.g., Benini et al., 2019; Cutting & Armstrong, 2016; Wang & Cheong, 2006). The goal of the present study was to continue this revived interest through an exploratory bottom-up approach, analyzing the formal features used in CPEEH films to understand what empathy-eliciting cinema looks like and to formulate suggestions on possible latent principles connecting form and empathy.

## Method

### Film Sampling

A sample of film scenes was compiled through two-stage cluster sampling to serve as the subject of a quantitative-style analysis. In the first stage, 100 films were randomly selected from the predefined sampling pool, followed by a second stage in which one suitable scene was randomly selected from each film.

Several criteria were observed in defining the sampling pool from which films were randomly selected to ensure that all films were representative of high-quality contemporary popular Hollywood cinema intended to elicit empathy (CPEEH films):

1. *Year of release.* Films had to be released between 2000 and 2018 to ensure they were produced in a similar technological production context, marked by the exponential increase of computer-generated imagery and digital recording following the turn of the century (Keating, 2014).

2. *Financial success.* Films had to rank among the Top 50 highest-grossing films at the U.S. box office in their year of release, out of the roughly 725 films that charted each year between 2000 and 2018 (Box Office Mojo, n.d.).

3. *Popularity.* Films had to have an Internet Movie Database (IMDb) score of at least 6.6 out of 10, the median score for this database (IMDb, 2019; Johnston, 2009). As the most widely used online film database, IMDb's audience rankings provide an indication of the film's popularity among general audiences over a wide period of time.

4. *Critical acclaim.* Films had to have a score of at least 61 out of 100 on the review aggregation platform Metacritic

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

(Metacritic, n.d.), corresponding to "mostly favorable reviews" according to the platform's score breakdown. This ensured that the sampling pool was reasonably homogenous in terms of the quality of the films included, reducing potential confounding factors in later analysis.

5. *Dramatic intent.* The film had to be labeled as "drama" on IMDb and/or Metacritic. Dramas generally have a strong emotional element to their story and can reasonably be presumed to be intended to evoke emotional and empathic engagement.

6. *Cultural background.* Films had to be an original English-language release primarily aimed at the Western market to ensure cultural consistency.

7. *Medium.* The film had to feature live-action material without animation, barring naturalistic computer-generated imagery, as animation and live-action film are fundamentally different in terms of stylistic affordances.

8. *Use of song.* The film could not be labeled as "musical" on either IMDb or Metacritic, as musicals regularly break the fourth wall through song and dance sequences, which may affect or interrupt empathic engagement.

With these criteria in place, the sampling pool ($N = 157$) was delimitated, from which 100 films were randomly selected in the first stage of sampling. This initial sample was then checked for its diversity. To ensure a maximum degree of stylistic variation, films were excluded from the initial sample if (a) they were part of a film series with more than one film featuring in the sample or (b) the director(s), cinematographer(s), or editor(s) of the film also worked on more than one other film in the sample. Through a process of iterative randomized resampling, a diverse sample of 100 films was secured, as summarized in the online supplemental materials (Lankhuizen et al., 2020).

## Temporal Scene Segmentation

To select suitable scenes from the 100 films, all films were viewed for temporal scene segmentation, that is, breaking down the film into its discrete narrative units. In this article, a scene is defined as a narrative unit depicting a single story beat, character goal, or plot development. It is a shot or collection of shots that comprise a single, complete, and unified dramatic element. Scene breaks may be recognized by a breach of unity of time, location, or action and often, but not always, involve a cut or other shot transition.

Segmentation was completed with the aid of four independent coders, trained to identify scenes and scene breaks using a coding manual, available in the online supplemental materials (Lankhuizen et al., 2020). During training, independent coders reached 73.7% agreement, which, given the complexity and the inherent ambiguity of identifying narrative units, was deemed sufficiently reliable for the coders to proceed individually. The 100 films were segmented into 5,225 film scenes, with one coder working on each film.

## Selection and Sampling of Empathy-Eliciting Film Scenes

In the next step, each scene was evaluated by the coders for its narrative coherence and potential to elicit empathy in viewers based on four criteria: First, the scene must feature a human or humanoid main character with recognizable facial expressions; second, the scene must feature conflict, meaning the main character must be facing danger, distress, or another challenge of sorts; third, the scene must be between 2 and 12 min long; and finally, the scene must not break the fourth wall through explicit references to the fictionality of the depicted narrative.

From the resulting pool of eligible scenes ($N = 1{,}614$), one empathy-eliciting scene was randomly selected per film to form the final sample of empathy-eliciting film scenes. A final check by the first author followed to determine if these scenes complied with the previously mentioned criteria:

1. The selected scene is a coherent and independent story unit.

2. Sufficient plot-relevant information pertaining to the narrative of the scene is contained within the scene itself.

3. The scene has at least one identifiable main character whose state of mind is communicated to the audience in some way.

If a sampled scene failed to pass these final checks, it was discarded from the sample, and a new scene from the same film was drawn. This process was repeated until a final sample of 100 empathy-eliciting film scenes from different films was established.

## Formal Feature Analysis

The formal feature analysis was largely completed using computer algorithms, with each feature measured second by second. Those features that could not be reliably coded using computer algorithms (indicated in Table 1) were coded manually using ELAN by four trained coders and Tess Lankhuizen using a coding manual, made available in the online supplemental materials (Lankhuizen et al., 2020). The visual formal features analyzed are summarized in Table 1.

## Statistical Analyses

Finally, the data on formal features were analyzed in three parts, reported in the next section. After descriptive analyses of individual features, a principal component analysis (PCA) was performed to identify clusters of formal features that might relate to latent constructs underlying filmmakers' stylistic decisions. The results of the PCA were elaborated on using correlational analysis to investigate potentially meaningful patterns in the variation of formal features.

# Results

## General Properties of Sampled Films

Descriptive statistics (see Table 2) revealed that the CPEEH films identified in the first step of the sampling procedure contained about

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 1
*Description of Formal Features*

| Name | Description |
|---|---|
| Shot scale | Perceived distance between the subject of the shot and the camera, measured as the duration of long shots, medium shots, and close-ups over the runtime of the scene. Possible range of .00 to 1.00. Also, the probability with which a shot from any given scale transitions to a shot of the same or of a different scale. After Savardi, Signoroni, Migliorati, and Benini (2018). |
| Face depiction | Visibility of actors' faces, measured as the average, minimum, and maximum number of faces in frame across the scene. Faces were only counted if visible for 50% or more. |
| Camera angle[a] | Camera angle in the vertical plane (high, neutral or eye-level, low), coded manually. Presented as ratio of the scene in milliseconds shot from that angle, ranging from .00 to 1.00. |
| Camera perspective[a] | Ratio of footage in milliseconds in which the camera approximates the physical or internal perspective of a diegetic character, ranging from .00 to 1.00. |
| Saturation | Color intensity, after Wang and Cheong (2006). Measured on a 0–255 scale, with 0 representing grayscale and 255 representing maximum saturation. |
| Saturation variance | Variation of color intensity within the frame, after Wang and Cheong (2006). Measured on a 0–255 scale, with 0 representing complete consistency in saturation within the frame and 255 representing maximum change in saturation within the frame. |
| Brightness | Pixel brightness, indicative of the relative perceived lightness or darkness of the frame, after Wang and Cheong (2006). Measured on a 0–255 scale, with 0 representing a completely black frame and 255 representing a completely white frame. |
| Shadow area | Average proportion of shadow area in the frame, ranging from .00 to 1.00. Shadow areas are areas darker than the shadow threshold, after Wang and Cheong (2006). |
| Contrast | Average dynamic range of each frame in the scene measured both as the standard deviation of pixel intensities (root mean square [RMS] contrast), normalized from 0 (no contrast; all pixels in the frame have the same value) to 1 (maximum contrast; values in the frame range from true black to true white) and as the information density of the frame (entropy contrast), measured in bits per pixel, with a range of 0 to 8. |
| Visual excitement | Change in pixel luminance and chromaticity between frames as a measure of perceived motion and visual change. Normalized from .00 (none of the pixels changed in 1 s) to 1.00 (all pixels changed in 1 s). After Wang and Cheong (2006). |
| Vector motion | Movement of superpixels (flattened macro areas of uniform color and brightness) across frames, expressed as the average distance in pixels covered per second, after Jeannin and Divakaran (2001). |
| Background clutter | Average amount of background clutter, measured as the average number of superpixels per frame, with reference to Cutting and Armstrong (2016). |
| Shot composition | Complexity of the shot composition, based on the average number of predicted focal points per frame (Kümmerer, 2017). Also includes compliance with the rule of thirds as the average distance between the main focal point (MFP) of each frame and the nearest of the four imaginary lines that divide the frame into equal thirds (horizontally or vertically). Normalized from .00 (minimum distance between MFP and a line of thirds) to 1.00 (maximum distance between MFP and a line of thirds). Similarly, the normalized distance between the MFP and the center of the screen was considered, along with the normalized distance between consecutive MFPs across frames. |
| Visual shot coherence | The proportion of cuts and transitions in which the frame before the cut is not visually consistent with the frame following the cut in terms of chromaticity and lighting, indicating the viewer is confronted with new or contrasting visual information across edits. Ranges from .00 to 1.00. |
| Shot length | Shot length and speed with which the video changes from one shot to the next, measured as the number of cuts over the scene duration, the number of total transitions (cut, fade, wipe, etc.) over the shot duration, and the average shot length. Also, the probability with which a shot of any given duration transitions to a shot of the same or a different duration. Based on Adami and Leonardi (1999). |

[a] Feature coded manually.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

52 scenes of 2 min and 40 s in length, with an average runtime of 125 min. Both film runtime and scene length were positively skewed, indicating that most films and scenes were sooner brief than they were long. Scenes that break the fourth wall or that lack a human main character were relatively rare, unlike scenes without conflict, which made up roughly 30% of all scenes in any given film.

## Descriptive Formal Feature Analysis of Scenes

Table 3 describes the sampled film scenes ($N = 100$) in terms of their formal features. All reported features are averaged for entire scenes.

**Shot scale.** On average, scenes in the sample mainly featured close-ups and only a very small amount of long shots (see Table 3). Using Spearman's rho as a nonparametric alternative to Pearson's $r$ to account for the nonnormal distribution of the shot scale ratios, a strong negative correlation was found between close-ups and medium shots ($\rho = -.88$, $p < .00$), indicating that an either/or principle might be at work in many of these scenes. Scenes appeared to be dominated either by medium shots or by close-ups; although the small standard deviation (*SD*) of the ratio of long shots indicated that the proportion of long shots did not vary much across scenes, there was a negative correlation between the ratio of long shots and close-ups ($\rho = -.46$, $p < .00$).

Table 4 shows the probability of one type of shot (close-up, medium shot, long shot) changing to any of the other types. Regardless of shot type, it was most likely for the following shot to match the

Table 2
*Descriptive Statistics of Sampled Films*

| Film features | Mean | Median | *SD* | Minimum | Maximum |
|---|---|---|---|---|---|
| Number of scenes | 52.25 | 52.00 | 16.67 | 11.00 | 104.00 |
| Average scene length (minutes) | 2.67 | 2.37 | 1.11 | 1.38 | 8.27 |
| Runtime (minutes) | 125.06 | 123.50 | 19.63 | 83.00 | 187.00 |
| Proportion of scenes without human main character | .02 | .00 | .04 | .00 | .22 |
| Proportion of scenes with fourth-wall breaks | .04 | .00 | .09 | .00 | .59 |
| Proportion of scenes without conflict | .30 | .30 | .15 | .00 | .62 |
| U.S. box office (millions of USD) | 123.50 | 98.80 | 85.44 | 50.93 | 700.06 |
| IMDb score | 7.47 | 7.50 | 0.45 | 6.60 | 8.60 |
| Metacritic score | 73.74 | 73.00 | 8.28 | 61.00 | 96.00 |

*Note.* IMDb = Internet Movie Database.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

preceding shot—that is, to cut from a close-up to another close-up, and so forth. Extreme changes in shot scale (e.g., from a close-up to a long shot) were improbable, indicating that the editors of CPEEH films prefer consistency and gradual progression in shot-scale changes. The high correlations between mirroring pairs of shot-scale changes confirm these results, indicating that if scale changed, it was likely to change back later in the scene (see Table 5).

**Depiction of faces.** As summarized in Table 3, every frame of a scene depicted an average of one or two human faces. Given that CPEEH films explicitly deal with narratives featuring human

Table 3
*Descriptive Statistics of Scene Formal Features*

| Formal features | Mean | Median | *SD* | Minimum | Maximum |
|---|---|---|---|---|---|
| Ratio of close-ups | .64 | .65 | .16 | .12 | .97 |
| Ratio of medium shots | .30 | .28 | .15 | .03 | .86 |
| Ratio of long shots | .06 | .02 | .07 | .00 | .25 |
| Average number of faces in frame across scene | 1.47 | 1.16 | 1.08 | 0.34 | 6.14 |
| Minimum number of faces in frame | 0.04 | 0.00 | 0.20 | 0.00 | 1.00 |
| Maximum number of faces in frame | 7.26 | 4.00 | 8.92 | 1.00 | 55.00 |
| Ratio of high angle footage | .23 | .18 | .19 | .00 | .83 |
| Ratio of low angle footage | .28 | .28 | .18 | .00 | .81 |
| Ratio of neutral angle footage | .48 | .49 | .24 | .04 | .94 |
| Ratio of footage from internal perspective | .01 | .00 | .02 | .00 | .11 |
| Average saturation of the scene | 74.86 | 72.72 | 25.27 | 36.29 | 205.11 |
| Average saturation variance of the scene | 63.27 | 61.36 | 14.14 | 34.15 | 103.13 |
| Average median pixel brightness of the scene | 34.03 | 30.66 | 20.86 | 1.75 | 109.31 |
| Average proportion of shadow area across all frames | .64 | .64 | .18 | .18 | .97 |
| Average RMS contrast of the scene | .17 | .17 | .05 | .05 | .27 |
| Average entropy contrast of the scene | 3.30 | 3.21 | .62 | 1.96 | 4.65 |
| Visual excitement at 1 frame per second | .09 | .00[a] | .13 | .00 | .39 |
| Vector movement at 1 frame per second | 4.91 | 4.86 | 2.13 | 1.34 | 11.26 |
| Average number of superpixels per frame | 31.92 | 30.74 | 8.92 | 13.58 | 59.22 |
| Standard deviation of number of superpixels per frame | 10.59 | 10.30 | 3.61 | 3.39 | 28.34 |
| Minimum number of superpixels per frame | 8.50 | 8.00 | 6.50 | 1.00 | 27.00 |
| Maximum number of superpixels per frame | 67.87 | 67.00 | 25.17 | 31.00 | 253.00 |
| Average number of focal points per frame | 1.41 | 1.38 | .23 | 1.04 | 2.48 |
| Normalized distance between MFP and center of frame | .08 | .08 | .03 | .02 | .17 |
| Normalized distance between MFP and line of thirds | .19 | .19 | .03 | .11 | .29 |
| Normalized distance between consecutive MFPs | .08 | .08 | .02 | .03 | .12 |
| Visually related shots across cuts | .89 | .90 | .09 | .57 | 1.00 |
| Cut rate | .22 | .21 | .10 | .03 | .44 |
| Shot rate | .24 | .23 | .11 | .03 | .50 |
| Ratio of very short takes (< 2 s) | .04 | .02 | .07 | .00 | .53 |
| Ratio of short takes (2–4.5 s) | .20 | .19 | .14 | .00 | .51 |
| Ratio of medium-short takes (4.5–7 s) | .21 | .21 | .09 | .00 | .50 |
| Ratio of medium takes (7–10 s) | .18 | .17 | .08 | .00 | .33 |
| Ratio of medium-long takes (10–22.5 s) | .25 | .25 | .12 | .04 | .57 |
| Ratio of long takes (22.5–40 s) | .08 | .05 | .09 | .00 | .42 |
| Ratio of very long takes (>40 s) | .05 | .02 | .07 | .00 | .50 |

*Note.* RMS = root mean square; MFP = main focal point.
[a] Value less than .00.

Table 4
*Probability of Cut Between Close-Ups (CS), Medium Shots (MS), and Long Shots (LS)*

| Shot scale | Mean (*SD*; minimum; maximum) | | |
|---|---|---|---|
| | CS | MS | LS |
| CS | .56 (.18; .08; .95) | .07 (.03; .00[a]; .16) | .01 (.02; .00; .09) |
| MS | .07 (.03; .01; .16) | .22 (.15; .01; .84) | .01 (.01; .00; .05) |
| LS | .01 (.02; .00; .08) | .01 (.01; .00; .05) | .03 (.04; .00; .17) |

[a] Value less than .00.

characters, this is unsurprising. The large variation of the maximum number of faces in any single frame of the scene, however, indicates that the sampled scenes were pluralistic in their settings (inhabited or deserted), characters (one or many), and framing of characters (shots showing only one or multiple characters).

**Camera angle and perspective.** The sampled scenes mainly featured footage shot from a neutral angle, though both high and low angles were common as well (see Table 3). While there were scenes entirely void of either high or low angles, all featured neutral angles, indicating that this angle is a vital and basic presentational tool. The ratio of neutral angle material was negatively correlated with both high angle material ($\rho = -.64$, $p < .01$) and low angle material ($\rho = -.55$, $p < .01$), although there was no significant correlation between the ratio of high and low angles ($\rho = -.08$, $p = .44$). Thus, high and low angles operated more or less independently from one another, possibly because they fulfill different narrative functions.

The vast majority of the scenes did not include shots that mimic the spatial point of view of a character. Those that did only featured an internal perspective for a small portion of the total video material—no more than 10.8%, even in the most extreme case.

**Lighting and color.** The data for lighting, contrast, and saturation (see Table 3) show that the majority of scenes were, on average, not particularly saturated but that this distribution was positively skewed; there was a small number of film scenes with a high average amount of saturation (up to 205.11 out of 255), whereas there were no extremely unsaturated or grayscale scenes.

The median pixel brightness similarly showed a positive skew, but here a greater number of scenes clustered toward the very low end of the scale, indicating CPEEH scenes could get very dark, while none, on average, got over 43% of the maximum possible brightness.

The scenes in the sample showed an average shadow area of 64% across all frames, indicating that most scenes were fairly shadowy. This distribution showed a slight negative skew; scenes clustered toward more rather than less shadow.

With an average root mean square (RMS) contrast of 0.17, most of these empathy-eliciting scenes were relatively low in contrast, indicating that low-key lighting (which creates a high-contrast chiaroscuro effect) was not a popular tool. However, the average scene was too dark to consider classic high-key lighting the standard. Instead, it seems most CPEEH scenes settle on a semi-natural lighting set-up. Entropy was similarly moderate to low, with a mean of 3.30 bit per pixel, indicating the average scenes consisted of frames that were neither high in contrast nor in information density.

**Motion.** Because visual excitement describes the proportionate amount of pixel change on a second-by-second basis, it is unlikely to see values approaching 1.00, as these would indicate that all pixels changed every second. The data showed that most scenes featured an average of 9% change in pixel values per second (see Table 3), meaning it would take approximately 11 s for all pixels in the frame to change significantly in luminosity and chromaticity. Vector movement averaged at 4.91 pixels per second, indicating superpixels in the frame traversed around 1/144th of the frame per second at a resolution of 720 pixels.

Both visual excitement and vector movement might appear to be rather low. However, perceived motion within a shot can have two sources: movement by the subjects in front of the camera or motion of the camera itself. In the former case, the background usually remains unchanged as the subject moves across the frame while in the latter case, the subject usually remains unchanged as the camera tracks a moving subject. In both cases, large parts of the frame remain unchanged even though viewers perceive motion, explaining how these scenes are not as slow as they might seem based on their averages.

**Background clutter.** The data on the average number of superpixels per frame are summarized in Table 3. Although the average amount of superpixels per frame ranged only from 13.58 to 59.21, the total variation of the number of superpixels per frame was much higher, ranging from 1 to 253 superpixels per frame. Thus, it would seem that while most frames clustered toward the lower end of the scale, there was a small set of frames with much higher average amounts of superpixels per frame. The average standard deviation of the number of superpixels per frame reflects this large degree of variation within scenes.

**Shot composition.** The average frame in the sample had but a single predicted focal point (see Table 3). High amounts of predicted focal points per frame were rare; only 2% of the sample had an average of more than two predicted focal points.

The average normalized distance between the main predicted focal point (MFP) and the center of the screen was .08, compared to an average distance of .19 between the MFP and the nearest line of thirds. This indicates that most frames had a central composition. Additionally, the average normalized distance between consecutive MFPs across frames was also .08, meaning the MFP would mostly remain constant or change gradually across frames.

**Visual shot coherence.** The percentage of visually coherent cuts was found to be very high at 89% (see Table 3). This indicates that the majority of scenes did not contain many cuts that introduced a large amount of new visual information, as indicated by the lack of remarkable changes in luminosity and chromaticity across cuts. As such significant changes in visual information

Table 5
*Partial Correlation Matrix for Probabilities of Cuts Between Close-Ups (CS), Medium Shots (MS), and Long Shots (LS)*

| Shot scale transitions | CS/MS | CS/LS | MS/LS |
|---|---|---|---|
| MS/CS | .97[a] | — | — |
| LS/CS | — | .90[a] | — |
| LS/MS | — | — | .86[a] |

[a] Spearman's rho significant at $p < .01$ (two-tailed).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

usually indicate a shift in time or place, it is likely they are more common to cuts between rather than within scenes.

**Shot rate.** On average, the sampled scenes had a cut or transition rate of .24, indicating one transition every 4.17 s (see Table 3). As indicated by the fairly large standard deviation, this rate tended to vary greatly across scenes. This variation, however, followed a normal distribution without outliers in the high range, indicating there might be a maximum shot rate at which CPEEH cinema can still meet audience expectations. The small difference between the shot and cut rate indicated that other transitions (wipes, dissolves, etc.) were an infrequently used tool.

**Shot duration.** Shot duration varied greatly between the different scenes, with the exception of medium takes (see Table 3). This may indicate that shorter or longer takes have more salient effects on story impact, making them noticeably more or less appropriate in specific contexts than the neutral medium-length take.

Table 6 summarizes the correlations between the probabilities of very short (VS), short (S), short-medium (SM), medium (M), medium-long (ML), long (L), and very long (VL) takes transitioning into any of the other types. Shorter takes tended to positively correlate with other types of short takes, while correlating negatively with longer takes, and vice versa. This shows that it was unlikely for the same scene to contain a mix of very short and very long takes. Instead, most scenes contained mainly brief or mainly long takes. The pacing of the scene, then, tended to be consistent throughout.

## Principal Component Analysis of Formal Features

An exploratory PCA was conducted on the 32 variables relating to the formal features described in Table 1 using oblique rotation (direct oblimin; see Field, 2009). The formal feature variables were standardized using the *z* scores to ensure all variables carried the same weight in the analysis, in spite of their different units of measurements (Jolliffe & Cadima, 2016). This initial set of variables was pared down during analysis. The average proportion of shadow area, the ratio of medium shots over scene length, and the cut rate were removed as they correlated very highly with the scene's median brightness ($\rho = -.98$, $p < .01$), ratio of close-ups ($\rho = -.88$, $p < .01$), and shot rate ($\rho = .99$, $p < .01$), respectively, causing singularity in the covariance matrix. A low Kaiser–Meyer–Olkin (KMO) value indicated instability in the model due to a lack of data, suggesting the sample size was not sufficient to establish reliable components (KMO = .46; Hutcheson & Sofroniou, 1999). Variables with an anti-image correlation of $r < .45$ were excluded in a stepwise manner, raising the KMO value to an acceptable level (KMO = .67).

Table 7 summarizes the factor loadings, explained variance, and eigenvalues of the final model. Model fit was acceptable, with 49% nonredundant residuals (Field, 2009). Kaiser's criterion of component eigenvalues over 1 would recommend four components (Kaiser, 1960). However, it should be noted that Kaiser's criterion may overestimate the number of relevant components in a data set of fewer than 30 variables with communalities smaller than .70 (Field, 2009; Nunnally & Bernstein, 1994), as was the case here (average communality = .69). Indeed, the fourth component recommended by Kaiser's criterion only contained one variable (shot rate) and, as such, does little to further understanding of possible latent constructs or data reduction. Based on this and on the scree plot, three components were extracted, offering meaningful combinations of variables that explained 58.01% of the variance in the data. The variables in each component may relate to visual complexity, visual contrast, and perceived closeness, respectively.

## Correlational Analysis of Formal Features

Correlational analysis was used to account for the relationships between features that were dropped from the PCA due to model instability resulting from a small sample size. Additionally, in using Spearman's rho, this follow-up also addressed the issue of nonnormally distributed data in the PCA. The full correlation matrix can be found in the online supplemental materials (Lankhuizen et al., 2020); relevant parts are reproduced or reported in the following subsections based on the strength and significance of the correlations in question.

**Correlations involving motion, clutter, and shot rate.** The first cluster of correlations involved many of the same features as the first component extracted in the PCA, relating to visual complexity. Background clutter, visual excitement, vector movement, and shot rate were found to be correlated (see Table 8). All these features were positively correlated—often strongly and significantly. Scenes high in these features also tended to be bright, low in shadow, and low in saturation variance. Saturation was also found to be positively correlated with visual excitement ($\rho = .23$, $p = .02$) and saturation variance ($\rho = .61$, $p = .00$), but it shared no significant correlation with the other variables and correlated

Table 6
*Correlational Matrix of Shot Type by Duration*

| | ρ (*p* value) | | | | | |
|---|---|---|---|---|---|---|
| Shot duration | S | SM | M | ML | L | VL |
| VS | .50 (.00)** | .04 (.72) | −.23 (.02)* | −.47 (.00)** | −.45 (.00)** | −.26 (.01)** |
| S | — | .44 (.00)** | −.23 (.02)** | −.74 (.00)** | −.69 (.00)** | −.58 (.00)** |
| SM | — | — | −.01 (.95) | −.47 (.00)** | −.46 (.00) | −.43 (.00)** |
| M | — | — | — | .17 (.09) | .01 (.96) | −.24 (.02)* |
| ML | — | — | — | — | .45 (.00)** | .27 (.01)** |
| L | — | — | — | — | — | .59 (.00)** |

*Note.* S = short; SM = short-medium; M = medium; ML = medium-long; L = long; VL = very long; VS = very short.
* $p < .05$. ** $p < .01$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 7
*Summary of Exploratory Principal Component Analysis (Direct Oblimin, N = 100; k = 10)*

| Formal features | Component | | |
|---|---|---|---|
| | Visual complexity | Visual contrast | Perceived closeness |
| Median pixel brightness | **.88** | .02 | .07 |
| Background clutter | **.85** | .07 | −.10 |
| Average vector movement | **.78** | .00 | −.20 |
| Visual excitement | **.61** | .00 | .25 |
| Saturation variance | **−.65** | .14 | −.01 |
| RMS contrast | −.07 | **−.84** | .11 |
| Visual shot coherence | −.01 | **.77** | .08 |
| Entropy contrast | .02 | **−.77** | −.01 |
| Ratio of close-ups | −.04 | .04 | **.82** |
| Ratio of long shots | −.07 | .07 | **−.77** |
| Eigenvalues | 3.11 | 1.94 | 1.33 |
| % of explained variance | 28.27 | 17.64 | 12.10 |

*Note*. RMS = root mean square. Highest component loadings in bold.

negatively with brightness ($\rho = -.22$, $p = .03$), indicating saturation operated independently from the cluster of variables described here.

**Correlations involving visual shot coherence.** The second component of the PCA, relating to visual contrast, was also confirmed in the correlational analysis. Scenes with high visual shot coherence were found to be less contrasting on both the entropy ($\rho = -.40$, $p = .00$) and RMS ($\rho = -.51$, $p = .00$) measures. Elaborating on this component, visual shot coherence negatively correlated with predicted focal-point change across cuts ($\rho = -.48$, $p = .00$), meaning that scenes with many noticeable changes in chromaticity and luminosity across cuts also featured many changes in predicted focal point across cuts.

**Correlations involving shot scale.** Complementary correlational analysis offered additional insight into the third component in the PCA, which consisted of two shot-scale variables, by showing how shot scale related to other formal features. It was previously reported that the distributions of long shots, medium shots, and close-ups tended to differ across scenes. The correlational analysis indicated that scenes with a high ratio of long shots were the most distinct of all shot-scale types; they featured fewer faces per frame ($\rho = -.43$, $p < .00$), used more high angle shots ($\rho = .20$, $p = .04$), were less bright ($\rho = -.22$, $p = .03$), and had a larger average proportion of shadow area across frames ($\rho = .23$, $p = .02$). They were also more likely to include shots with an internal perspective ($\rho = .20$, $p = .05$). Meanwhile, the ratio of close-ups or medium shots in a scene did not seem to correlate with other formal features, with the exception of shot rate. Close-ups shared a weak positive correlation with shot rate ($\rho = .22$, $p = .03$), whereas medium shots had a weak negative correlation with shot rate ($\rho = -.23$, $p = .02$), possibly indicating that the appropriateness of one or the other depended on the pacing of the scene. Additionally, scenes rich in close-ups also tended toward a higher average number of faces per frame ($\rho = .22$, $p < .03$).

As the most common shot scale type and the one most prominently highlighted in the theory as related to empathy, a comparison was made between the formal feature distributions of scenes with a high ratio of close-ups and a low ratio of close-ups based on a median split of the sample. The Mann–Whitney test confirmed previous results; scenes with a lot of close-ups had significantly fewer average predicted focal points per frame ($U = 757.00$, $z = -3.40$, $p < .00$, $r = -.34$), a higher average amount of faces per frame ($U = 1675.00$, $z = 2.81$, $p = .01$, $r = .28$), and a higher shot rate ($U = 1543.00$, $z = 2.02$, $p = .04$, $r = .20$) than scenes with few close-ups. Out of all shot-scale types, it appears close-ups were the most frequently associated with the depiction of human faces and the presentation of fast-paced material.

**Correlations involving camera angle and face depiction.** Scenes with a high ratio of neutral shots appeared to feature a higher average of faces per frame ($\rho = .20$, $p = .05$), while scenes with a high ratio of low angle shots tended to feature fewer faces per frame on average ($\rho = -.23$, $p = .02$), as did scenes with a high ratio of high angle shots, though this correlation was not significant ($\rho = -.16$, $p = .12$). It appears that neutral angles are most common for the depiction of faces, possibly to maintain a norm to which the more exceptional low and high angles can be a meaningful exception.

**Correlations involving shot composition.** It was previously established that central compositions were more representative of the sample than rule-of-thirds compositions; correlation analysis also found these two compositions to be negatively correlated ($\rho = -.64$, $p = .00$). The change in predicted focal points across cuts tended to be low ($\rho = -.28$, $p = .01$) in these common central compositions, indicating that most scenes maintained compositional stability.

**Scene duration and other formal features.** Mann–Whitney's test was used to establish whether longer scenes differed significantly in formal features compared to shorter scenes based on a median split of the sample on scene duration. It might be that

Table 8
*Correlational Matrix: Clutter, Motion, Brightness, Shot Rate, and Focal Points*

| Formal features | ρ (*p* value) | | | | | |
|---|---|---|---|---|---|---|
| | Visual excitement | Vector movement | Shot rate | Brightness | Shadow area | Saturation variance |
| Background clutter | .46 (.00)** | .74 (.00)** | .14 (.18) | .76 (.00)** | −.79 (.00)** | −.38 (.00)** |
| Visual excitement | — | .40 (.00)** | .27 (.01)** | .58 (.00)** | −.56 (.00)** | −.31 (.00)** |
| Vector movement | — | — | .40 (.00)** | .69 (.00)** | −.72 (.00)** | −.31 (.00)** |
| Shot rate | — | — | — | .24 (.02)* | −.24 (.02)* | −.03 (.80) |
| Brightness | — | — | — | — | −.98 (.00)** | −.51 (.00)** |
| Shadow area | — | — | — | — | — | .45 (.00)** |

* $p < .05$ (two-tailed). ** $p < .01$ (two-tailed).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

longer scenes serve a different narrative function or are better at eliciting empathy because they have more time to build engagement. It was found that longer scenes were significantly less bright than short scenes ($U = 891.00$, $z = -2.48$, $p = .01$, $r = -.25$), had a greater average proportion of shadow area across frames ($U = 1593.00$, $z = 2.37$, $p = .02$, $r = .24$), had a smaller average amount of faces per frame ($U = 918.00$, $z = -2.29$, $p = .02$, $r = -.23$), had a smaller ratio of close-ups ($U = 952.00$, $z = -2.05$, $p = .04$, $r = -.21$), and had a higher ratio of long shots ($U = 1633.00$, $z = 2.65$, $p = .01$, $r = .27$), but there was no significant difference in the ratio of medium shots ($U = 1460$, $z = 1.45$, $p = .15$, $r = .15$). Based on the negative correlations between the ratio of long shots, brightness, and average number of faces per frame established previously, the key difference between scenes of short and long duration appears to be the ratio of long shots, with shorter scenes mainly limiting themselves to close and medium shots, while longer scenes use long shots more frequently. Notably, shot rate did not significantly differ between scenes of short and long duration, indicating that the use of long shots does not relate to pacing.

## Discussion

### Summary of Main Results

The goal of this study was to analyze the formal features of contemporary popular Hollywood scenes that have the implicit intent to elicit empathy from viewers. One of the main assumptions of this research is that films that are intended to elicit empathy have a specific "look," or amalgamation of visual formal features, that may represent a cohesive empathic style of cinema intended to facilitate empathic engagement. Through quantitative stylistic analysis, the formal features defining empathy-eliciting film scenes were identified.

Characterizing the prototypical visual formal features of the average CPEEH scene proved to be a challenging feat, as the data did not converge into a single recurring pattern of variation. This implies that there is no such thing as a uniform empathic style for filmmakers to rely on, but rather that there is a multitude of ways in which formal features might be patterned to inspire the desired response. However, these patterns of formal features, though diverse, are not without their similarities and appear to converge on constructs related to empathy, such as perceived distance, visual complexity and arousal, and visual contrast and coherence. The next section considers the results for each formal feature and how these results relate to findings from previous research.

### Shaping Empathy

**Shot scale and face depiction.** Most CPEEH scenes consisted mainly of close-ups, followed by medium shots and only a small ratio of long shots. It was more likely for a shot to match the scale of the preceding shot than for the shot scale to change, and cuts between long shots and close-ups or back were very rare, indicating a certain consistency in or preference for a gradual change of shot scale in CPEEH scenes. These findings on the stability of shot scale are in line with studies by Canini et al. (2011).

Previous research suggests that close-ups and the depiction of human faces are directly related to the level of viewer empathy (Bálint et al., 2018; Cao, 2013). The fact that close-ups were most commonly used and that every frame, on average, depicted at least one face suggests that creators of CPEEH films apply this principle in their work.

The different shot types were found to have their own visual character, with scenes with a high ratio of long shots being the most distinct. These scenes featured a lower average number of faces per frame and used more high angle shots while being less bright, more shadowy, and more likely to use internal perspective shots. Though it is not necessarily long shots themselves that incorporate these correlating features (because other shots might use these features more when a scene contains many long shots), the distinctiveness of scenes rich in long shots might be due to the fact that long shots are often used to show more of the setting in which the action takes place (Bordwell & Thompson, 2013). The intent behind this might be to strategically distance the viewer from the depicted characters by increasing the illusion of physical distance and by making their faces harder to read, forcing viewers to speculate on the characters' emotions. Alternatively, the intent behind an increase in long shots may be expositional, offering viewers more contextual and environmental information on the action of the scene. Both of these stylistic strategies would explain why scenes rich on long shots would have a lower average amount of faces per frame. More targeted research would be necessary to understand the link between long shots, lower brightness, and the use of high angles, though it is possible these are features of environmental and establishing shots.

Scenes with a high ratio of close-ups and medium shots were less distinct, characterized only by shot rate. A great number of close-ups correlated with a faster pace, while the opposite held true for scenes with many medium shots. Findings by Cutting and Armstrong (2016) offer an explanation for the faster pacing of scenes rich in close-up; it takes viewers less time to visually process faces in close-up than faces in medium shot. This allows the editing of close-ups featuring faces to be tighter. Or, the other way around, it encourages the use of close-ups in fast-paced scenes.

The ratio of close-ups also increased with the average amount of faces per frame, and scenes rich in close-ups were more likely to be low in the average number of predicted focal points. This decrease in the number of focal points is likely because close-ups, by nature, focus on a single subject and direct all or most of the viewer's attention to that subject, whereas medium shots leave more space for the eye to roam. As faces are a common focal point, the fact that the average number of faces per frame was found to be higher in scenes with many close-ups likely indicates that such scenes had only very few shots that did not feature a face at all. All these factors suggest that close-ups are generally used to amplify empathy, intimately framing human characters without distractions and with only few shots not immediately related to the faces of the actors.

**Camera angle.** Most scenes featured mainly neutral camera angles, with a smaller ratio of low and high angles. Kraft (1987) suggested that neutral angles enhance parity between subject and viewer, possibly explaining their prevalence in empathy-eliciting scenes. Scenes with a high ratio of neutral shots also tended to have a higher average amount of faces per frame, a feature asso-

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

ciated with an increase in background clutter. It might be that neutral angles are popular for shots featuring many faces as they do not complicate the image with unusual viewpoints.

**Camera perspective.** Like close-ups, shots using an internal perspective were theorized to enhance empathic engagement by placing the viewer in the character's spatiotemporal point of view. However, this was not a popular tool among filmmakers of CPEEH films. The reason for this might be that internal shots directly challenge the self/other distinction that is maintained during empathic processes (Zaki & Ochsner, 2012). Alternatively, it might be that empathy requires frequent insight into the character's emotional state through depictions of the character's face, rather than (only) the merging of visual perspectives of character and viewer.

**Visual complexity.** One of the clusters identified in the PCA consisted of the features background clutter, visual excitement, vector movement, brightness, and saturation variance; further analysis similarly found strong correlations between these variables (a negative correlation, in the case of saturation variance), along with a positive correlation with shot rate and a negative correlation with the proportion of shadow area.

It may be the case that filmmakers use these features to manage the ease with which a scene is processed. Cuts, transitions, and motions can cue orienting responses and thereby increase attention and available mental resources for message processing and subsequent understanding (Lang, 2000). Both attention and comprehension are essential aspects of narrative engagement (Busselle & Bilandzic, 2009); the former may be considered a prerequisite for empathy, while the latter is tightly intertwined with it because understanding of the characters' emotions nurtures understanding of the story in its entirety (Kneepkens & Zwaan, 1995). The current results do not support the idea that creators "max out" on these features to create content that is highly attention-grabbing, as there was a large degree of scene-to-scene variation in shot rate, saturation, brightness, and to a lesser extent, visual excitement and vector motion. It is possible that filmmakers are intuitively aware of the risk of overwhelming viewers with visually complex information; the LCM predicts that if a message becomes too visually complex—as it well might with a high degree of motion, quick cuts, and a great amount of clutter and focal points—comprehension is compromised because the message overtaxes available mental resources (Lang, 2000). The fact that this cluster of visually complex variables includes brightness and excludes saturation variance may be indicative of attempts by filmmakers to keep such visually complex scenes comprehensible through clear lighting and consistent color intensity.

Several of the features associated with attention and visual complexity also relate to arousal. Previous research has established that a high shot rate (or high pacing) and fast motion are associated with arousal, as are brightness and saturation (Detenber & Lang, 2011; Valdez & Mehrabian, 1994). Arousal may increase with perceived emotional intensity, including the emotional intensity experienced through (affective) empathy, making it a seemingly attractive goal for filmmakers. However, as many of the formal features related to arousal are also related to visual complexity, they may need to be carefully balanced so that visuals do not become complicated rather than merely complex. Additionally, the variation in the use of these features may also be explained by the fact that not every narrative moment will benefit from a visual style focused on grabbing the viewer's attention and inducing arousal (Bálint et al., 2017; Sukalla et al., 2015).

**Shot composition and visual contrast.** The results indicated that CPEEH cinema can generally be typified as smooth and balanced in terms of shot composition within and across frames. Central shot compositions, in which the main predicted focal point is located at or near the center of the screen, were most common. Most shots only contained one predicted focal point, and this focal point remained fairly stable across shots. This means viewers generally only have to pay attention to one area of the screen without having to make frequent saccades.

Additionally, the visual shot coherence of most scenes was very high, indicating that extreme changes in luminosity and chromaticity between cuts were rare. This likely means that most cuts introduced new visual information piece by piece, rather than jumping to entirely novel visuals. Consideration for visual contrast appears to be important to filmmakers, as those rare scenes that were low in visual shot coherence were also high in focal-point changes and contrast (both RMS and entropy). It seems that while CPEEH cinema is mainly shot and composed with the intent to offer a smooth viewing experience and likely follows notions of continuity editing, there is a small subset of scenes that may be crafted to be visually contrasting and intentionally disruptive. Such disruptions are often intended to be narratively meaningful; for example, Zettl (1990) discusses how "complexity editing" calls attention to cuts in the film to foreground the juxtaposition of different shots, allowing viewers to meaningfully interpret these shots in light of narrative context. Thus, scenes high in contrast and low in visual shot coherence may invite viewers to engage in impactful processes of meaning-making. Their rarity in empathy-eliciting scenes, however, suggests that such disruptive techniques may interfere with empathic engagement if used to excess.

**Shot length.** While shot duration tended to vary greatly between scenes, pacing within scenes was found to be fairly consistent, as shorter takes correlated positively with other short takes and negatively with longer takes. Previous research by Svanera and colleagues (2019) showed similar results. Smooth editing may be relevant to empathic engagement because pacing is contingent on content, with certain story elements best presented at a specific pace (Bálint et al., 2017; Sukalla et al., 2015). Multiple changes in shot duration may also result in jarring viewer experiences, as viewers fail to identify the rhythm of the scene—an important aspect of viewers' filmic experience (Adams, Dorai, & Venkatesh, 2001; Wang & Cheong, 2006).

**Scene duration.** In comparing long and short scenes, it was found that longer scenes tended to be less bright, more shadowy, featured fewer faces per frame, and had a smaller ratio of close-ups with a higher ratio of long shots. A similar pattern was found for scenes with a high ratio of long shots, which tended to be less bright, more shadowy, featured fewer faces per frame, and used more high camera angles. Longer scenes with many long shots might focus more on the environment than on character action compared with shorter scenes with fewer long shots. This would allow for more visual exposition, but may not relate to empathy itself.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

## Strengths, Limitations, and Future Research

This study took an exploratory approach, describing the formal features of CPEEH films in an open-ended analysis. The inferences made based on these analyses provide an important first step in answering how the goal of eliciting empathy affects the aesthetic decisions made by filmmakers. Because the study used existing popular films as subject of analysis, rather than artificially manipulating an audiovisual stimulus in an experimental design, the results have high external validity and can serve as a stepping-stone to guide future research on the relation between formal features and empathy in film. Considering the results of the present study, such research should pay particular attention to the interplay between features relating to visual complexity, coherence and visual contrast, and perceived closeness or representation of the human face and body.

A limitation of this study was the focus on only the visual formal features of film. Previous research points to music and sound design as important aspects of the film experience (Cohen, 2001). Similarly, narrative context is expected to play an important role in filmmakers' stylistic decisions. Neither of these subjects fit the scope of this research, but both should be considered in future research to complete our understanding of empathy elicitation through form.

In addition, this study considered the stylistic properties of empathy-eliciting film by investigating patterns of formal feature variation in scenes assumed to have the *same* intended effect—namely, empathy elicitation. To ensure this shared intended effect, only film scenes from contemporary films with a dramatic genre element were considered. Future research on the link between formal features and intended viewer experience should assess whether scenes with *different* intended effects, from other genres or other time periods, display patterns of formal feature variation that are noticeably distinct from those found here. Such comparison between scenes with the implicit goal of empathy elicitation and scenes with other primary goals, such as thrilling viewers, making them laugh, or engaging them in cognitive play, would expand our knowledge on the different applications of form by creators.

Finally, this study considered the stylistic properties of empathy-eliciting film based on an assumption of intended effect. The next step in understanding the relationship between form and empathy would be to investigate if any of the features described do, in fact, affect the levels of empathy experienced by viewers. Based on the results of this study, it seems prudent for such a research project to take into consideration narrative responses related to empathy, particularly comprehension, attention, and arousal. Researchers should be prepared to rely on robust and nonparametric analyses in their studies on the narrative effect of the visual formal features of film, as the distributions of the majority of these formal features are by nature nonnormal.

## Closing Remarks

The results of this research indicate that most visual formal features associated with empathy show a great degree of variation when considered individually in the context of CPEEH film. However, correlational analysis and PCA showed that this variation is not random and suggest a complex network of interactions between these features, especially those relating to visual complexity, visual contrast and coherence, and perceived closeness. The weak to moderate nature of many of these relationships indicates that the patterns of influence between form and intended viewer response are likely to be subtle and sophisticated. As such, they may not always manifest in any given scene taken from contemporary popular empathy-eliciting Hollywood cinema.

Ultimately, it would appear that while there is no singular empathic style of cinema, there are patterns that suggest certain stylistic principles that may be applied by filmmakers intent on eliciting empathy. First, features directly related to the representation of actors' faces and bodies may be used to reduce the perceived distance between viewer and character, enhancing empathy without detours. Second, features related to arousal and visual complexity, such as cut rate, brightness, motion, and background clutter, may be balanced to engage viewers while safeguarding comprehension. Finally, through the manipulation of visual shot coherence, contrast, and focal points in frame, coherence may be improved to facilitate an accessible viewing experience. Thus, the formal features of empathy-eliciting film may serve as building blocks to a multitude of different roads to intended viewer effects, as empathy is shaped and crafted more than it is manufactured or produced.

## References

Adami, N., & Leonardi, R. (1999). *Identification of editing effect in image sequences by statistical modelling*. Presentation at the Picture Coding Symposium, Portland, OR.

Adams, B., Dorai, C., & Venkatesh, S. (2001). *Automated film rhythm extraction for scene analysis*. Presentation at the IEEE International Conference on Multimedia and Expo, 2001 (ICME 2001), Tokyo, Japan.

Bálint, K., Klausch, T., & Pólya, T. (2018). Watching closely: Shot scale influences theory of mind response in visual narratives. *Journal of Media Psychology: Theories, Methods, and Applications, 30,* 150–159. http://dx.doi.org/10.1027/1864-1105/a000189

Bálint, K., Schoft, C., & Rooney, B. (2017). *Depicting violence: The effect of shot scale, shot length and camera perspective on narrative engagement with violent films*. Presentation at the 67th International Communication Association Conference, San Diego, CA.

Bartsch, A. (2012). Emotional gratification in entertainment experience. Why viewers of movies and television series find it rewarding to experience emotions. *Media Psychology, 15,* 267–302. http://dx.doi.org/10.1080/15213269.2012.693811

Benini, S., Savardi, M., Bálint, K., Kovács, A. B., & Signoroni, A. (2019). On the influence of shot scale on film mood and narrative engagement in film viewers. *IEEE Transactions on Affective Computing*. Advance online publication. http://dx.doi.org/10.1109/TAFFC.2019.2939251

Benini, S., Svanera, M., Adami, N., Leonardi, R., & Kovács, A. B. (2016). Shot scale distribution in art films. *Multimedia Tools and Applications, 75,* 16499–16527. http://dx.doi.org/10.1007/s11042-016-3339-9

Bordwell, D. (1985). *Narration in the fiction film* (1st ed.). Madison, WI: University of Wisconsin Press.

Bordwell, D., & Thompson, K. (2013). *Film art: An introduction* (10th ed.). New York, NY: McGraw-Hill.

Box Office Mojo. (n.d.). *Home*. Retrieved from https://www.boxofficemojo.com/

Bruce, V., & Young, A. (2013). *Face perception*. London, England: Psychology Press. http://dx.doi.org/10.4324/9780203721254

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

Canini, L., Benini, S., & Leonardi, R. (2011). *Affective analysis on patterns of shot types in movies*. Presentation at the 7th International Symposium

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

on Image and Signal Processing and Analysis (ISPA), Trois-Rivières, Quebec, Canada.

Cao, X. (2013). The effects of facial close-ups and viewers' sex on empathy and intentions to help people in need. *Mass Communication & Society, 16,* 161–178. http://dx.doi.org/10.1080/15205436.2012.683928

Cohen, A. J. (2001). Music as a source of emotion in film. In P. N. Juslin & J. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 249–272). New York, NY: Oxford University Press.

Coplan, A. (2008). Empathy and character engagement. In P. Livingston & C. Plantinga (Eds.), *The Routledge companion to philosophy and film* (pp. 96–110). London, England: Routledge.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). New York, NY: Oxford University Press.

Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the concept. *Emotion Review, 8,* 144–153. http://dx.doi.org/10.1177/1754073914558466

Cummins, R. G., Keene, J. R., & Nutting, B. H. (2012). The impact of subjective camera in sports on arousal and enjoyment. *Mass Communication & Society, 15,* 74–97. http://dx.doi.org/10.1080/15205436.2011.558805

Cutting, J. E., & Armstrong, K. L. (2016). Facial expression, size, and clutter: Inferences from movie structure to emotion judgments and back. *Attention, Perception, & Psychophysics, 78,* 891–901. http://dx.doi.org/10.3758/s13414-015-1003-5

de Waal, F. B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology, 59,* 279–300. http://dx.doi.org/10.1146/annurev.psych.59.103006.093625

Decety, J., & Jackson, P. L. (2006). A social-neuroscience perspective on empathy. *Current Directions in Psychological Science, 15,* 54–58. http://dx.doi.org/10.1111/j.0963-7214.2006.00406.x

Denisova, A., & Cairns, P. (2015). First person vs. third person perspective in digital games: Do player preferences affect immersion? In B. Begole (Ed.), *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 145–148). New York, NY: Association for Computing Machinery. http://dx.doi.org/10.1145/2702123.2702256

Detenber, B. H., & Lang, A. (2011). The influence of form and presentation attributes of media on emotion. In K. Doveling, C. von Scheve, & E. Konijn (Eds.), *The Routledge handbook of emotions and mass media* (1st ed., pp. 275–293). London, England: Routledge.

Field, A. P. (2009). *Discovering statistics using SPSS: And sex, drugs and rock 'n' roll* (3rd ed.). Newbury Park, CA: Sage.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin, 133,* 694–724. http://dx.doi.org/10.1037/0033-2909.133.4.694

Gallese, V., & Guerra, M. (2019). *The empathic screen: Cinema and neuroscience*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/oso/9780198793533.001.0001

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79,* 701–721. http://dx.doi.org/10.1037/0022-3514.79.5.701

Grodal, T. (1999). *Moving pictures: A new theory of film genres, feelings, and cognition*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198159834.001.0001

Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., & Heeger, D. J. (2008). Neurocinematics: The neuroscience of film. *Projections, 2,* 1–26. http://dx.doi.org/10.3167/proj.2008.020102

Hutcheson, G. D., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. Newbury Park, CA: Sage.

Internet Movie Database. (2019). *IMDb: What is IMDb?* Retrieved from https://help.imdb.com/article/imdb/general-information/what-is-imdb/G836CY29Z4SGNMK5?ref_=helpsect_cons_1_1#

Jeannin, S., & Divakaran, A. (2001). MPEG-7 visual motion descriptors. *IEEE Circuits and Systems for Video Technology, 11,* 720–724. http://dx.doi.org/10.1109/76.927428

Johnston, N. (2009, October 9). *IMDb movie ratings over the years*. Retrieved from http://www.njohnston.ca/2009/10/imdb-movie-ratings-over-the-years/

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences, 374,* 20150202. http://dx.doi.org/10.1098/rsta.2015.0202

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141–151. http://dx.doi.org/10.1177/001316446002000116

Keating, P. (2014). *Cinematography* (1st ed.). New Brunswick, NJ: Rutgers University Press. http://dx.doi.org/10.36019/9780813563510

Kneepkens, E. W. E. M., & Zwaan, R. A. (1995). Emotions and literary text comprehension. *Poetics, 23,* 125–138. http://dx.doi.org/10.1016/0304-422X(94)00021-W

Kraft, R. N. (1987). The influence of camera angle on comprehension and retention of pictorial events. *Memory & Cognition, 15,* 291–307. http://dx.doi.org/10.3758/BF03197032

Kümmerer, M. (2017). *Index—DeepGaze II*. Retrieved from https://deepgaze.bethgelab.org/

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

Lang, A., Chung, Y., Lee, S., Schwartz, N., & Shin, M. (2005). It's an arousing, fast-paced kind of world: The effects of age and sensation seeking on the information processing of substance-abuse PSAs. *Media Psychology, 7,* 421–454. http://dx.doi.org/10.1207/S1532785XMEP0704_6

Lang, A., Zhou, S., Schwartz, N., Bolls, P. D., & Potter, R. F. (2000). The effects of edits on arousal, attention, and memory for television messages: When an edit is an edit can an edit be too much? *Journal of Broadcasting & Electronic Media, 44,* 94–109. http://dx.doi.org/10.1207/s15506878jobem4401_7

Lankhuizen, T., Bálint, K. E., Savardi, M., Konijn, E. A., Bartsch, A., & Benini, S. (2020, July 20). *Shaping film: A quantitative formal analysis of contemporary empathy-eliciting Hollywood cinema* [Supplementary materials]. Retrieved from https://osf.io/rptmg/?view_only=33df36ff88154174af85f14da888d42f

Metacritic. (n.d.). *How we create the Metascore magic—Metacritic*. Retrieved from https://www.metacritic.com/about-metascores

Meyers-Levy, J., & Peracchio, L. A. (1992). Getting an angle in advertising: The effect of camera angle on product evaluations. *Journal of Marketing Research, 29,* 454–461. http://dx.doi.org/10.1177/002224379202900406

Morgan, S. E., Palmgreen, P., Stephenson, M. T., Hoyle, R. H., & Lorch, E. P. (2003). Associations between message features and subjective evaluations of the sensation value of antidrug public service announcements. *Journal of Communication, 53,* 512–526. http://dx.doi.org/10.1111/j.1460-2466.2003.tb02605.x

Nedomansky, V. (2015, May 30). *The editing of Mad Max the movie [Blog post]*. Retrieved from https://vashivisuals.com/the-editing-of-mad-max-fury-road/

Neumann, D. L., & Westbury, H. R. (2011). The psychophysiological measurement of empathy. In D. J. Scapaletti (Ed.), *Psychology of empathy* (pp. 119–142). Hauppauge, NY: Nova Science.

Niederdeppe, J., Davis, K. C., Farrelly, M. C., & Yarsevich, J. (2007). Stylistic features, need for sensation, and confirmed recall of national smoking prevention advertisements. *Journal of Communication, 57,* 272–292. http://dx.doi.org/10.1111/j.1460-2466.2007.00343.x

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. New York, NY: McGraw-Hill.

Oliver, M. B., & Raney, A. A. (2011). Entertainment as pleasurable and meaningful: Identifying hedonic and eudaimonic motivations for entertainment consumption. *Journal of Communication, 61,* 984–1004. http://dx.doi.org/10.1111/j.1460-2466.2011.01585.x

Plantinga, C. (2009). *Moving viewers: American film and the spectator's experience* (1st ed.). Berkeley, CA: University of California Press. http://dx.doi.org/10.1525/9780520943919

Rasheed, Z., Sheikh, Y., & Shah, M. (2003). *Semantic film preview classification using low-level computable features*. Presentation at the 3rd International Workshop on Multimedia Data and Document Engineering (MDDE-2003), Berlin, Germany.

Ravaja, N. (2004). Effects of image motion on a small screen on emotion, attention, and memory: Moving-face versus static-face newscaster. *Journal of Broadcasting & Electronic Media, 48,* 108–133. http://dx.doi.org/10.1207/s15506878jobem4801_6

Rinck, M., & Bower, G. H. (1995). Anaphora resolution and the focus of attention in situation models. *Journal of Memory and Language, 34,* 110–131. http://dx.doi.org/10.1006/jmla.1995.1006

Rooney, B., & Bálint, K. E. (2018). Watching more closely: Shot scale affects film viewers' theory of mind tendency but not ability. *Frontiers in Psychology, 8,* 2349. http://dx.doi.org/10.3389/fpsyg.2017.02349

Salt, B. (1974). Statistical style analysis of motion pictures. *Film Quarterly, 28,* 13–22. http://dx.doi.org/10.2307/1211438

Savardi, M., Signoroni, A., Migliorati, P., & Benini, S. (2018). *Shot scale analysis in movies by convolutional neural networks*. Presentation at the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece.

Scarantino, A., & de Sousa, R. (2018). Emotion. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from https://plato.stanford.edu/archives/win2018/entries/emotion/

Shen, L. (2010). Mitigating psychological reactance: The role of message-induced empathy in persuasion. *Human Communication Research, 36,* 397–422. http://dx.doi.org/10.1111/j.1468-2958.2010.01381.x

Simons, R. F., Detenber, B. H., Cuthbert, B. N., Schwartz, D. D., & Reiss, J. E. (2003). Attention to television: Alpha power and its relationship to image motion and emotional content. *Media Psychology, 5,* 283–301. http://dx.doi.org/10.1207/S1532785XMEP0503_03

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science, 303,* 1157–1162. http://dx.doi.org/10.1126/science.1093535

Smith, T. J., Levin, D., & Cutting, J. E. (2012). A window on reality: Perceiving edited moving images. *Current Directions in Psychological Science, 21,* 107–113. http://dx.doi.org/10.1177/0963721412437407

Stephenson, M., & Palmgreen, P. (2001). Sensation seeking, perceived message sensation value, personal involvement, and processing of anti-marijuana PSAs. *Communication Monographs, 68,* 49–71. http://dx.doi.org/10.1080/03637750128051

Sukalla, F., Bartsch, A., & Schnell, C. (2015, May 21). *Time to indulge in feeling. the influence of narrative pace on empathy, perceived social distance, and intentions to help stigmatized groups*. Presentation at the International Communication Association 65th Annual Conference, San Juan, Puerto Rico.

Svanera, M., Savardi, M., Signoroni, A., Kovács, A. B., & Benini, S. (2019). Who is the film's director? Authorship recognition based on shot features. *IEEE MultiMedia, 26,* 43–54. http://dx.doi.org/10.1109/MMUL.2019.2940004

Taberham, P. (2018). *Lessons in perception: The avant-garde filmmaker* (1st ed.). Oxford, NY: Berghahn Books. http://dx.doi.org/10.2307/j.ctv3znzvc

Tamborini, R., Eastin, M., Lachlan, K., Skalski, P., Fediuk, T., & Brady, R. (2001). *Hostile thoughts, presence and violent virtual video games*. Presentation at the 51st Annual Convention of the International Communication Association, Washington, DC.

Tan, E. S. (1996). *Emotion and the structure of narrative film: Film as an emotion machine*. Mahwah, NJ: Lawrence Erlbaum Associates.

Tan, E. S. (2008). Entertainment is emotion: The functional architecture of the entertainment experience. *Media Psychology, 11,* 28–51. http://dx.doi.org/10.1080/15213260701853161

Tarvainen, J., Westman, S., & Oittinen, P. (2015). The way films feel: Aesthetic features and mood in film. *Psychology of Aesthetics, Creativity, and the Arts, 9,* 254–265. http://dx.doi.org/10.1037/a0039432

Tomasello, M. (2009). *Why we cooperate*. Cambridge, MA: MIT Press. http://dx.doi.org/10.7551/mitpress/8470.001.0001

Valdez, P., & Mehrabian, A. (1994). Effects of color on emotions. *Journal of Experimental Psychology: General, 123,* 394–409. http://dx.doi.org/10.1037/0096-3445.123.4.394

Visch, V. T., & Tan, E. S. (2009). Categorizing moving objects into film genres: The effect of animacy attribution, emotional response, and the deviation from non-fiction. *Cognition, 110,* 265–272. http://dx.doi.org/10.1016/j.cognition.2008.10.018

Wang, H. L., & Cheong, L.-F. (2006). Affective understanding in film. *IEEE Transactions on Circuits and Systems for Video Technology, 16,* 689–704. http://dx.doi.org/10.1109/TCSVT.2006.873781

Wilms, L., & Oberfeld, D. (2018). Color and emotion: Effects of hue, saturation, and brightness. *Psychological Research, 82,* 896–914. http://dx.doi.org/10.1007/s00426-017-0880-8

Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: Progress, pitfalls and promise. *Nature Neuroscience, 15,* 675–680. http://dx.doi.org/10.1038/nn.3085

Zettl, H. (1990). *Sight, sound, motion: Applied media aesthetics*. Belmont, CA: Wadsworth.

Zillmann, D. (1988). Mood management through communication choices. *American Behavioral Scientist, 31,* 327–340.

Received April 17, 2020
Revision received July 20, 2020
Accepted August 8, 2020 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.