How Act Structure Sculpts Shot Lengths and Shot Transitions in Hollywood Film

James E. Cutting, Kaitlin L. Brunick, and Jordan E. DeLong

Abstract: Cinematic tradition suggests that Hollywood films, like plays, are divided into acts. Thompson (1999) streamlined the conception of this large-scale film structure by suggesting that most films are composed of four acts of generally equal length—the setup, the complicating action, the development, and the climax (often including an epilog). These acts are based on the structure of the narrative, and would not necessarily have a physical manifestation in shots and transitions. Nonetheless, exploring a sample of 150 Hollywood style films from 1935 to 2005, this article demonstrates that acts shape shot lengths and transitions. Dividing films into quarters, we found that shots are longer at quarter boundaries and generally shorter near the middle of each quarter. Moreover, aside from the beginnings and ends of films, the article shows that fades, dissolves, and other non-cut transitions are more common in the third and less common in the fourth quarters of films.

Keywords: cinemetrics, film acts, Hollywood film, shots, transitions

Film analysis can be done in many, sometimes overlapping, ways—through historical study, through cultural study, through the lens of a large-scale theory, through assessment of psychological responses, and through statistical analysis. The statistical analysis of film—the approach we adopt here and sometimes called cinemetrics—begins with units. One cannot measure and compare films without having separable units that are appropriate to the cinemetric task. It is generally agreed that available units in film form a nested hierarchy: acts, sequences, scenes, shots, and frames (e.g., Field 1979; Harrington 1973).

Frames are easy to measure. The film industry long ago settled on the projection rate of twenty-four frames/second, enhanced by an episcotister that presents each frame three times. The flicker rate of 72 Hz is sufficiently above critical flicker fusion (~60 Hz; e.g., Anderson and Anderson 2005), the point above which a train of discrete flashes is seen as continuous. Beyond this

doi: 10.3167/proj.2011.050102 ISSN 1934-9688 (Print), ISSN 1934-9696 (Online)

flicker value most motion in a film is psychologically and physiologically indistinguishable from real motion. One can easily count frames, although their number is often somewhat different for analog film and its digital form on DVD.[1] Regardless, we have counted frames in digital media to document the increasing amounts of motion and movement in films from 1935 to 2005 (Cutting, DeLong, and Brunick 2010).

Shots are also easy to determine. They are continuous runs of film frames taken from a stationary or moving camera interrupted by transitions of various kinds, where the camera adopts a new position or lens configuration. Cuts are instantaneous; one frame belongs to one shot, the next frame belongs to another. Most films often have a number of longer-duration transitions—dissolves, fades, wipes, and the like, which in older films can take up to several seconds to complete. However, these rarely amount to more than 10 percent of all transitions in the films we have sampled. Straight cuts are by far the most common and they make up almost 99 percent of all transitions in contemporary film. Interestingly, when given the task to detect cuts, viewers often miss a large number of them (Smith and Henderson 2008) and computer techniques are scarcely any better (Over, Ianeva, Kraiij, and Smeaton 2006). This can be a problem as average shot length (ASL) has become by far the most common physical metric for film (Bordwell, Staiger, and Thompson 1985; Salt 1992, 2006; see also http://www.cinemetrics.lv). We (Cutting, DeLong, and Nothelfer 2010) have used a hybrid, relatively intensive technique—a user-driven computer algorithm—to achieve better than 99 percent success in finding transitions benchmarked against manual, frame-by-frame analysis. We found that there are fluctuations in shot length that, since about 1960, have begun to mimic the natural fluctuations of human attention.

Conversely, scenes and sequences are more difficult to determine. Messaris (1994) defined a scene as taking place at a particular time and location, with transitions across time, location, or both defining scene changes; sequences are often like scenes but with continuous change in location, as in chase sequences. These definitions fit well with the notion of continuity editing, and Messaris's framework is fine for most television programs, where he carried out his analysis. Unfortunately, in our experience, it does not work uniformly well for Hollywood film.[2] Because scenes and sequences are composed of events, it may be that a focus on event structure (Zacks and Magliano 2010) rather than scene structure, may reveal interesting physical patterns in film. Nevertheless, because scenes and sequences are ancillary to our analyses here, we do not deal with them further.

Acts are larger units, and are most evident in guides to writing screenplays. In *Screenplay* (1979) Syd Field systematized long-term Hollywood practice and suggested that there are three acts: act 1—the setup; act 2—the confrontation; and act 3—the resolution. Field's act 2 is twice as long as the other acts, and is subdivided around a turning point in the middle. Interestingly, this framework is copied in a recent manual for writing novels (Ingermanson and Economy 2010). Kristin Thompson (1999) suggested a more straightforward framework for most films with four generally equal-length acts: act 1—the setup; act 2—the complicating action; act 3—the development; and act 4—the climax, which often includes a short epilog. Although it has not always met with complete approval (see Deutelbaum 2000), Thompson's is the scheme we will adopt in this article.

Thompson (1999) documented the four acts in many films. The setup, much as the term suggests, establishes the world in which the narrative takes place and introduces the characters and their initial goals. The complicating action typically derails those goals or refocuses the characters and their actions on a new set of problems. The development often broadens the action (or creates more delays) or introduces subplots while the protagonist struggles to resolve the complications that arose before. Finally, the climax forces the protagonist into action toward the possible achievement of the new goals. The climax is often followed by an epilog that may tie up subplots and generally restore a stable worldview, but it is considered part of the climax and is too short to be a separate act. This article uses a sample of films and information about shot lengths and transitions to explore the possible physical signature of film acts in shot lengths and on transitions.

Films and Analyses

The analyses here are part of a larger project on the investigation of the detailed physical structure of Hollywood film (Cutting, DeLong, and Brunick 2011; Cutting, DeLong, and Nothelfer 2010). For this project we selected 150 English-language films, ten in each of fifteen years, every five years from 1935 to 2005. For films of 1980 and later, we selected from among those with the highest US gross proceeds of their release year. Because those data were not systematically kept for films from before 1977, we selected from among films rated by the most viewers on the Internet Movie Database (IMDb, www.imdb.com). We also sampled films from different genres generally defined by the first listed genre for each film on the IMDb.[3] The supplementary materials accompanying Cutting, Delong, and Brunick (2011) and Cutting, DeLong, and Nothelfer (2010) list these films and a number of their cinemetric characteristics. Among the genres were thirty-two action films, twenty adventure films, forty-seven dramas, forty-one comedies, and ten animations. The numbers of these vary across years due to changes in the industry and viewers' tastes. Initially, all candidate transitions between shots were flagged by a computer algorithm and then checked by hand. In a second, manual pass through all films we coded each transition as a cut, fade out, fade in, dissolve, or other (wipe, iris in or out, etc.). Again, we estimate that we found at least 99% of all transitions.

We are focally interested in pacing within films. We have previously reported on two kinds. First, Cutting, DeLong, and Nothelfer (2010) found that the sequence of shot lengths in films has a waxing and waning structure (shot lengths getting shorter then longer) that occurs simultaneously at different scales—that is, fluctuations on the orders of tens of seconds, minutes, tens of minutes, and longer. Interestingly, this structure has been evolving in Hollywood film, particularly since 1960, so that it has come nearly to match the fluctuating pattern of human attention as measured in laboratory experiments. Second, Cutting, DeLong, and Brunick (2011) also posited a similar need for the pacing of the motions by actors and objects and the movements of cameras, sometimes escalating to high intensity but almost always returning to calmer shots and scenes. The dynamic patterns we investigate here are more global and concern shot lengths and the use of various transitions over acts.

Average Shot Lengths across Acts

Average Shot Length has been an important and pervasive metric in the study of films (Bordwell, Staiger, and Thompson 1985; Salt 1992, 2006). A general assumption that follows from its use, however, has been that ASLs are more or less uniform throughout a film. Given our interests in film pacing, we thought this not particularly likely. Thus, our first question was: do acts differ in their ASLs?

Thompson (1999) detailed the analysis of acts in ten films, and also reported act lengths in ninety others, ten per decade from the 1910s to the 1990s (Appendix A). As suggested earlier, the acts showed a remarkable near-uniformity in their relative lengths. From these one hundred films, we selected the eighty that fit her four-act scheme. She divided two longer films into five acts (they had two development sections), and eighteen others into only three acts (they had no development section). Of the latter, nine were silent films. From the residual eighty films, we aggregated her data for statistical analysis. We found that the median setup takes up 24.7 percent of the run time of the film (ignoring initial credits without scenic backdrop), the complicating action takes up 26 percent, the development 24.9 percent, and the climax and epilog 24.1 percent (ignoring trailing credits). The average standard error of each mean, the typical statistical index of variability used in these contexts, for the four act percentages was 0.4 percent—impressively small.

We find Thompson's argument persuasive and her data remarkably clean. Nonetheless, the division into acts is built on a detailed analysis of the narrative, and we work from the physical properties of film. Many observers might agree on act divisions, but these divisions would not necessarily be reflected in any physical measure of a film's shots and transitions. Thus, without prejudice as to what we might find, we sought data in shot lengths and in shot transitions that might corroborate Thompson's analysis.

Cutting, Delong, and Nothelfer (2010) measured all shots and ASLs in each of the 150 films. As a first effort from this database, we divided the array of the shots in these films into four equal-numbered shot segments. Because our subdivisions are not based on the film narratives, we will continue to call them quarters, not acts. Consider a hypothetical example: if a film has 1,600 shots we marked the quarters as ending with shots 400, 800, 1,200, and 1,600. We then computed the four ASLs, one for each shot quarter. We immediately worried about averaging across large differences in whole-film ASLs and how these might affect our results. After all, some of our sample films had whole-film ASLs of more than 20 seconds (e.g., *Harvey* 1950, 20.9 seconds), and others less than 3 seconds (e.g., *Mission: Impossible II* 2000, 2.6 seconds). To combat this problem, we normalized the values by dividing the ASL for each shot quarter by the ASL for the entire film. In this manner, values near 1.0 for shot quarters would have ASLs near that for the whole film, regardless of what the latter value might be. We first looked at potential quarter variations across time in films from 1935 to 2005, but found no differences, a result consistent with the notion that there is considerable stability in certain aspects of Hollywood film.

We next looked for differences across genre. Figure 1 shows the results. There were no significant differences across shot quarters for four genres—adventure, animation, comedy, and drama. That is, mean ASLs were essentially uniform across quarters, and likely the four acts. In contrast, however, there was a substantial decline in ASL for action films and this decline interacted with that of the four other genres,[4] denoting that action films systematically differ from those of other genres. It does not seem particularly surprising that ASL should decline as an action film reaches climax, but the modal drop is quite dramatic. In general, the ASLs of the last quarter of shots in action films are as much as 50 percent shorter than those of the first quarter. Moreover, no such systematic drop in other genres suggests that filmmakers may not systematically manipulate ASLs in those films as a function of the act structure.

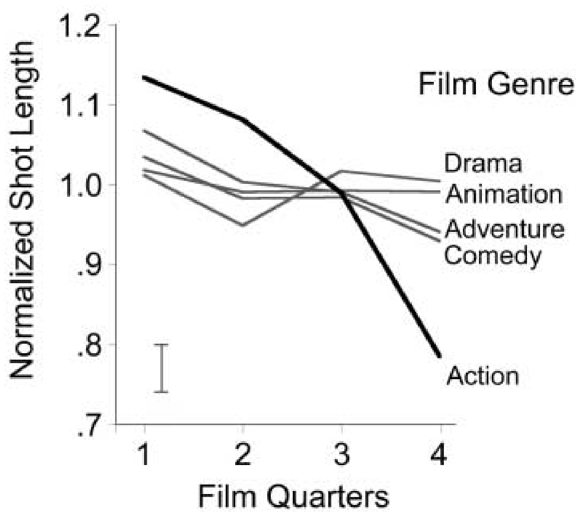

Figure 1: Normalized average shot lengths (ASLs) in five film genres across four equal quarters (in terms of number of shots) of 150 films from 1935 to 2005. Values near 1.0 indicate that the ASL for a given quarter is about the same as for the film as a whole; values above 1.0 indicate they are longer; and values below indicate they are shorter. The bracket at the lower left indicates average standard error of the mean for all data points.

Average Shot Lengths within Acts

We find the data in Figure 1 interesting, and the difference between action films and other genres compelling. However, these data also raise a measurement problem. We realized that to do a more detailed analysis of shot lengths within quarters we could not simply divide the number of shots by four because numerical quarters (of shots) and temporal quarters (of film) are not the same. To use this method would mean the last quarter of the shots of an action film might be only two-thirds the duration of the first quarter. We also wanted to compare individual shot lengths across films more directly, and this presented a second problem because films vary widely in their number of shots. For example, the films in our sample had between a bit over 200 (*The Seven Year Itch* 1955) and just under 3,100 (*King Kong* 2005) shots, with a mean of just over 1,100.

We solved these problems in the following manner: imagine a hypothetical 1,600-shot action film from the 1980s that is 90 minutes long. Following the pattern in Figure 1, the first 22.5-minute temporal quarter might have 350 shots (ASL = 3.86 seconds), the second quarter 375 shots (3.6 seconds), the third 400 shots (3.37 seconds), and the last 475 shots (2.8 seconds). We then interpolated the waveform of shots so that those of each temporal quarter of the film would fit into 250 bins of what we call *adjusted shots*, yielding a total of 1,000 adjusted shots per film. That is, we created a common timeline for each film and interpolated the length of a given adjusted shot from the lengths of the actual shots.

Consider an adventure film from our sample: *Cast Away* (2000) has 878 shots and an overall ASL of 9.21 seconds. This is a long ASL for a film made in 2000, but sampling full-length films with less than 500 or more than a 1,200, 2,000, or 3,000 shots is done in the same general way. The four temporal quarters of *Cast Away* show considerable variation; their ASLs are 8.72, 7.23, 8.42, and 16.06 seconds and they have 232, 280, 240, and 126 shots, respectively. Again, our method converts each one of these into 250 adjusted shots. How might this be done?

Imagine a serial plot of all of *Cast Away*'s fourth-quarter shots, 753 through 878, spread along a horizontal axis (the abscissa) and registering the length of each shot along the vertical axis (the ordinate) as a dot. Then connect the dots, creating a continuous if quite bumpy function. Next, rather than moving along the shot sequence (the abscissa again) from 753 upward in units of 1.0, move upward in steps of only .504 units (.504 = 126/250, or the number of fourth-quarter shots divided by the adjusted shots the data will be put into). Once one has the value on the abscissa, one then moves up vertically until intersecting the line drawn between the values of the actual shot lengths. This method is called linear interpolation and it is a standard statistical procedure.

The lower panel of Figure 2 gives a more concrete idea of how this is done. The durations of the last thirteen shots (866 through 878) of *Cast Away* (2000) are shown as circles with gray centers. Interpolated values are shown as black diamonds; sometimes an interpolated value lies close to one of the

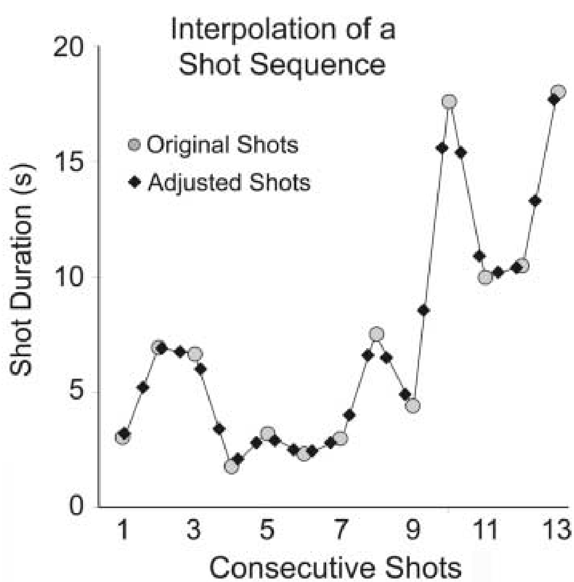

Figure 2: The top panel shows a still from the last shot of Cast Away *(2000) where Chuck Nolan (Tom Hanks) decides in which direction his future lies. The bottom panel shows a representation of the shot lengths of the last thirteen shots in the film, and the method for interpolating values for the adjusted shot lengths. The still in the top panel is part of shot 13 in the bottom panel.*

actual values, sometimes it lies between two. By this method, the waveform of the thirteen actual shot lengths of *Cast Away* has been resampled to become twenty-four new adjusted shot lengths. In this manner we created a pattern of 1,000 adjusted shots for every film that represent the pattern of its shot structure.[5] Having done this, we could then compare all films because each would have been resampled to have exactly the same number of shots.

From our 150 films we removed seven that were longer than 2.5 hours. Thompson (1999) suggested that these would likely be divided into five or more acts. Thus, we used the remaining 143 films for further analysis. Again, because ASLs vary so much across films, we normalized every adjusted shot within a film by a two-step process. We first subtracted the whole-film adjusted ASL from every adjusted shot value, then divided that value by the standard deviation of all adjusted shot lengths. This creates what is called a unit normal distribution (with a mean of zero and a standard deviation of 1.0). This too is a standard statistical procedure, one used to compare patterns in data sets with different means and variances. We then calculated the median value for each of the 1,000 adjusted shot lengths across the 143 films. Figure 3 shows a scatter plot of the resulting values.

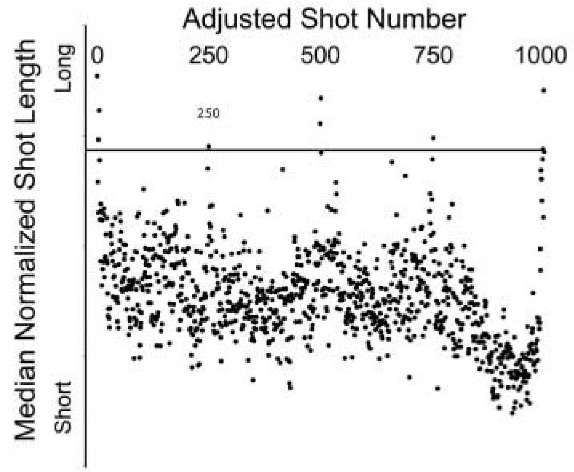

Figure 3: A scatter plot of the 1,000 median adjusted shot lengths (interpolated as in Figure 2) for 143 films. Shot 1 has a value that is off the top of the scale shown.

We recognize that these are highly massaged data—shots arranged by temporal quarter in each film, new shot values interpolated from them and placed into 1,000 bins, and then medians taken across all films. Nonetheless, we think the results are interpretable, interesting, and important. Again, there were no obvious differences in these patterns for films of different eras; thus we take our analyses to represent the whole of seventy years—1935 to 2005—of popular film.

Notice first the upper dots in Figure 3. The first three adjusted shots (shots 1, 2, and 3, as well as shot 7) to the far left are quite long, and above the arbitrary horizontal line in this normalized plot. This group makes sense. It is often the case that the first shots of a film are lengthy establishing shots that begin the setup of the film. Perhaps more important, at least since about 1960, they also carry opening credits. Thus, viewers may need more time to absorb both the credit information and the scenic content. The next median-adjusted shot above the arbitrary cutoff is near the boundary of the first and second temporal quarters (shot 249). In addition, two of the shots near the boundary of the second and third quarters (shots 499 and 501) are above the cutoff. One shot at the boundary of the third and fourth quarters (shot 751) is above the cutoff, as are two at the very end (shots 999 and 1,000). It is important to note that no other shots anywhere among the 1000 group adjusted shots are in this range.[6]

Longer than average shots are likely to occur at the boundaries of a film divided into temporal quarters.

We interpret these results in the following way: although acknowledging that the alignment of quarters to acts across films is imperfect, we think we can generalize from these results and assume that an act often begins or ends with longer shots. Without a narrative account of these 143 films, however, we cannot be sure. Nevertheless, the pattern is clear: Longer than average shots are likely to occur at the boundaries of a film divided into temporal quarters.

One must then ask: How common is this pattern in our films? As it turns out it is relatively rare. Consider the two adjusted shots nearest the quarter boundaries (shots 1 & 2, 249 & 251, 499 & 501, 749 & 751, and 999 & 1,000). Where one would expect one-quarter of the films (35.75) to have each separate shot pair longer than their ASL, 81, 66, 57, 45, and 50 did. The first three of these values are reliable. Moreover, one would expect no films to have all ten of these shots above their ASL ($143/2^{10}$), but in fact seven films did.[7]

Of the seven films that fit our pattern, *Inherit the Wind* (1960) offers a particularly good example of this structure. Among its fifteen longest shots are five that straddle the quarter boundaries of the film. Stills from these shots are shown in Figure 4. The film's duration is 126 minutes, it has 513 shots, and its overall ASL is 15.2 seconds. Shot 1 is 79 seconds long. Covered with credits, it begins outside the Hillsboro Courthouse where the Cates trial (the nominal stand-in for the Scopes trial) takes place. The camera then dollies back, cranes upward while looking down at the courthouse steps, and then pans left across the town square setting up the larger space of the film's action. Shot 115, 31.5 minutes into the film, is 90 seconds long and initiates the trial and its courtroom deliberation. Shot 272 is 63 minutes into the film, is 45 seconds long, and has Henry Drummond (the Clarence Darrow character played by Spencer Tracy) cross-examining a student who had been taught

Figure 4: Five stills from five different shots in Inherit the Wind *(1960). Still* a *is part of the setup and shows the steps of the courthouse where the Cates trial (a stand-in for the Scopes trial) will be held. Still* b *shows the opening scene of the court case from the judge's point of view, with Henry Drummond (the Clarence Darrow character played by Spencer Tracy) on the left and Matthew Harrison Brady (the William Jennings Bryan character played by Frederic March) on the right. Still* c *shows Drummond examining Howard (Jimmy Boyd), a student in Cates's class in which evolution was discussed, with Judge Mel Coffey (Harry Morgan) in the background. Still* d *shows Drummond examining Brady on the veridicality of the Bible, and still* e *shows Drummond with E. K. Hornbeck (the H. L. Mencken character played by Gene Kelly) after the case. These stills are from the shots that straddle the four quarters of the film. Consistent with the pattern shown in Figure 3, they are among the longest in the film.*

evolution by Bertram Cates (Dick York), a science teacher in the local school. It winds up with Drummond defending the right to think, refocusing the courtroom debate away from questions of scripture. Shot 415 is 31.5 minutes later, takes 150 seconds, and has Drummond cross-examining Matthew Harrison Brady (the William Jennings Bryan character played by Frederic March), quizzing him on the veridicality of the Bible. Finally, shot 513 is 74 seconds long and shows Drummond and E. K. Horbeck (the H. L. Menken character played by Gene Kelly) conferring after Drumond has lost the case (but will appeal) and after Brady has died.

Nonetheless, we do not think the outliers in Figure 3 are the most important aspect of these data. Instead, notice first the prominent scallop in the fourth-quarter data. That is, the shot lengths roughly in the middle (actually a bit later in this quarter of the film) are shorter than those nearer the beginning or end. One can test for a scallop-like pattern by fitting the data with a quadratic equation, and the curvilinear fit to these data reveals a highly reliable trend.[8] One might think that action films cause (or at least dominate) this pattern, with shot lengths getting shorter across the climax. However, this cannot be. Our method of taking medians across films guarantees that no genre can dominate this pattern. Moreover, the rising tail at the end of the fourth quarter seems likely to reflect longer shot lengths in epilogs, which can occur in all genres. Thompson's data show that some epilogs are longer than others and that a few films have none at all. That variation could easily contribute to the upward curl in the fourth quarter data shown in Figure 3.

The large-scale rhythmic change in shot lengths, and hence film intensity, could signal to viewers, however subtly, where they are in the narrative and help guide their expectations.

Setting the fourth quarter aside and ignoring those data points above the baseline we have already discussed, we found reliable scallop-like trends in the first, second, and third quarters as well,[9] although only the second and third are salient in scatter of data in Figure 3. These scallops, we suggest, represent a tendency toward intensification of films (Bordwell 2002, 2006) near the middle of acts, where shot lengths become slightly shorter.

We can estimate the depth of these scallops by examining the quadratic fits to each quarter, averaging the end points, and comparing that value with the minimum. The differences between these are 1.34, .81, 1.07, and 1.20 seconds. This means that the mean smoothed minimum of each scallop corresponds to mid-act shots being as much as 1.1 seconds shorter than those at the beginnings and ends. The large-scale rhythmic change in shot lengths, and hence film intensity, could signal to viewers, however subtly, where they are in the narrative and help guide their expectations.

Temporal Locations of Dissolves, Fades, and Their Like

From our data presented thus far it is clear that acts influence shot lengths across Hollywood films. Might acts affect the nature of the transitions between shots as well? For our final analysis we categorized the transitions in all 150 films by type (e.g., cut, fade in, fade out, dissolve, wipe, iris in, iris out), yielding more than 170,000 transitions in all. In this analysis, we were particularly interested in the almost 5,400 non-cuts and their locations across the length of each film. Their importance is that each typically marks a discontinuity in the narrative, an ellipsis when going from one thread of the story to another, a jump ahead or back in time, or a change to a different place (Bordwell and Thompson 2004).

We had two predictions. First, we thought that Thompson's third act, the development section where subplots broaden the story line, would likely have more non-cut transitions—particularly fades and dissolves—than the other acts. The reason is that there would necessarily be more narrative movement back and forth among subplots and these shifts would, in classical Hollywood film, be crafted with fades, dissolves, and wipes. Second, we thought that the climax act would have the fewest non-cuts. This is because the narrative hurtles toward a conclusion in a generally linear manner and without sidetracks to independent activity by characters outside the main thrust of the story. However, we anticipated an exception to this pattern in the epilog. In separating itself from the climax proper, and in tying together any loose ends of the narrative, we would expect a higher than normal level of non-cut transitions there.

We marked the presence of each non-cut transition by frame number from the beginning of each film. Fade outs were marked when scene luminance appeared to fall halfway to black, fade ins where the luminance had risen halfway to that of the subsequent shot; and dissolves, wipes, and iris ins and outs by the frame in which they appeared half complete (equal luminance for the two overlapping scenes of the dissolves, and equal areas for the two frames covered by the wipes and irises). The number of this frame was then recorded as a proportion of the number of frames in the film after having removed final credits and introductory credits without scenic content. In this manner, frame 30,500 in a film with 150,000 frames would be assigned a proportion of 0.2033 (30,500/150,000) of the film's length. These proportions were then placed into twenty equal sized bins for each film (.00 to .05, .05 to .10, .15 to .20 . . . .90 to .95, .95 to 1.0, where 1.0 represents the length of the film). We chose twenty because, in our experience, it would be a large enough number to register small differences within each of the four acts but not large enough to increase substantially the variability in the data.

Next, each non-cut transition was given a weight according to their number in a film. Consider two examples. *King Solomon's Mines* (1950) has sixty-three dissolves and no other type of non-cut transition. Thus each would count 0.016 (1/63). *The Usual Suspects* (1995) has three dissolves, four fade ins, and five fade outs, for a total of twelve non-cut transitions. Here, each would count 0.0833 (1/12). In this manner, across the twenty bins, the sum of the weights for all non-cut transitions would add to 1.0 for each film. This procedure guarantees that the distribution of these transitions would be of equal importance for each film, regardless of the number or type of non-cut transitions or total number of transitions in it. Finally, the seven films longer than 150 minutes were again excluded, as were four films that had no non-cut transitions (for example, *M*A*S*H* 1970). This left 139 films contributing to the mean distribution shown in Figure 5.

Four trends can be seen. First, notice that the initial bin has nearly 16 percent of all non-cut transitions. This is hardly a surprise, although we neglected to predict it, as the setup of a movie often needs to present several different but related places that are relevant to the narrative. More concretely, many films begin with a fade in and often have a string of dissolves before settling down to the story line. Moreover, for films in our sample from 1985 to 2005, 22 percent of all non-cut transitions in films occurred within the first twentieth of the film (the first bin shown in Figure 5). In contrast and largely because earlier films have proportionately more dissolves, fades, and wipes, our films from 1935 to 1955 had only 8 percent of their non-cut transitions occur within this bin. But other than the first bin, there were few relative differences in films from different eras.

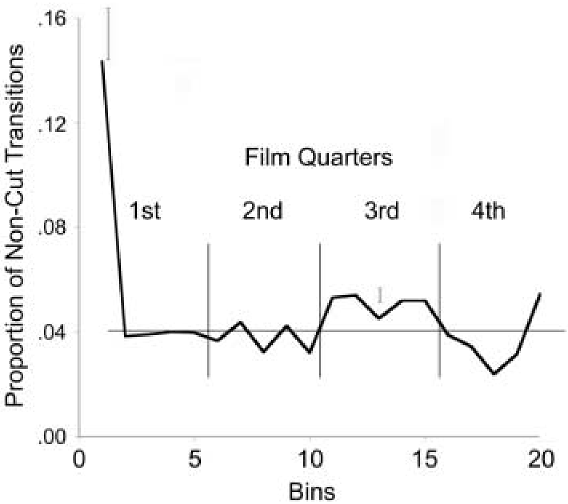

Figure 5: Fades, dissolves, wipes, and other non-cut transitions as they are distributed across the length of films. Each of 139 films was divided into twenty equal-length bins; non-cuts placed within them according to their proportional temporal distance from the beginning of the film; each then weighted and summed within bins; averaged across films within release years (1935 to 2005); and averaged across all release years. There are five bins to a film quarter. The horizontal line represents the average of bins 2 through 20. The bracket at the first bin indicates one standard error of the mean of the films by release year for the first bin only; the bracket at bin 13 indicates the average standard error of the mean for bins 2 through 20 by release year.

Second, notice that the last or twentieth bin also has a larger proportion of non-cut transitions, although its increase above the mean is more modest than for the first bin.[10] It seems likely that this is due to dissolves and fades being used in epilogs, where narrative details are wrapped up or a general world picture is restored.

Third, notice that bins 11 through 15 have an elevated proportion of non-cut transitions.[11] As suggested above, if the third act (the development) broadens the narrative and carries out the story line in subplots, this is exactly what we might expect. Again, dissolves and fades generally indicate changes in time and place (Bordwell and Thompson 2004; Messaris 1994), and subplots are carried out across these dimensions.

Finally, notice that the first part of the fourth quarter—bins 16, 17, 18, and 19—has fewer non-cut transitions, and that bins 18 and 19 have the fewest non-cut transitions across all films. Again, with the film coming to climax, it seems unlikely that non-cut transitions will occur because it is unlikely that the film will move to another story line.

Conclusions

By analyzing shots and transitions in 150 Hollywood films released from 1935 to 2005, we found several striking patterns that appear to be driven by the signature of a film's act structure. First, we found that shot lengths at the beginnings and ends of four equal length quarters of a film—and perhaps near the boundaries of the four acts as outlined by Thompson (1999)—tend to be longer than average. Second, within each quarter and perhaps each act of a film there is a subtle pattern of general shortening and then lengthening of shots. We suggest this reflects a fluctuating intensification of continuity within each act that builds across the four acts of a film. Third, non-cut transitions (dissolves, fades, and the like) occur in relative abundance in three locations—at the very beginning of the setup of the film, before the very end and likely in epilogs, and more important for our argument throughout the third quarter (the third act or development section) of a film. Epilogs aside, these non-cut transitions are also fewest in a film's fourth quarter.

We find these results compelling. Our data suggest that the narrative structure of film at the level of acts has physical instantiation—in the lengths of shots within each act, at act boundaries, and in the transitions at several locales. We also take these results as providing strong physical evidence in support of the thematic division of most Hollywood films into four acts. Part of the reason for these general patterns must be due to cultural transmission among filmmakers as they created and followed established norms. It seems possible that these patterns, which are larger than the scale of narrative events (Zacks and Magliano 2010) also serve to guide film viewers about what to expect in films and when. These patterns are part of a scaffold on which viewers can affix their large-scale expectations about a film and its narrative progress.

Our data suggest that the narrative structure of film at the level of acts has physical instantiation—in the lengths of shots within each act, at act boundaries, and in the transitions at several locales.

As a caveat, however, one must remember that our data do not mean that every film has a shot-length profile like that shown in Figure 3 with long shots at act boundaries and shorter ones at mid-act, or a transition-type profile like that shown in Figure 4 with more dissolves in the development section (act 3) and fewer in the climax (act 4). However, because we sampled widely across Hollywood films from 1935 to 2005 and found statistically robust generic patterns, we feel confident that these data are representative of the physical structure of the average popular film.

James E. Cutting is professor in the Department of Psychology at Cornell University. He is the author of *Perception with an Eye for Motion* (1986) and *Impressionism and Its Canon* (2006), as well as more than 100 scientific articles on form, motion, depth, and related areas of perception.

Kaitlin L. Brunick is a graduate student in the Department of Psychology at Cornell University. She received her BA in Psychology and Linguistics from the College of William and Mary and is interested in film, TV, and cognitive development.

Jordan E. DeLong is a graduate student in the Department of Psychology at Cornell University. He received his BS in Psychology and in Cognitive Science from Indiana University and is interested in film, faces, fingerprints, and the visual system.

Notes

[1] Digitization is often not done at exactly twenty-four frames/second, but instead fractionally more rapidly. Without appropriate synchronization this can create digital frames that are blends of analog frames.

[2] We thank Michael Masucci, an undergraduate summer intern from George Washington University, for his hours of labor successfully convincing us of the difficulty of segmenting films into scenes.

[3] Deviations from IMDb listings were few. For example, we designated *Ace Ventura: When Nature Calls* (1995) as a comedy, not an action or adventure film. The IMDb considers it "action, adventure, comedy, mystery." In addition *The Asphalt Jungle* (1950) is listed as "crime, film-noir, drama" but we considered it a drama.

[4] The decline for action films was reliable ($F(3,112) = 17.2$, $p < .0001$), as was its interaction with that of the four other genres ($F(4,447) = 4.58$, $p < .003$).

[5] In our computations, this method also makes shots 250, 500, and 750 unusable because their sampling is compromised.

[6] The a priori likelihood of having one adjusted shot from regions within ten shots of a quarter boundary (shots 1–10, 240–260, 490–510, 740–760, and 990–1,000) being among the longest seven shots is $1\text{-}(1\text{-}.08)^7$, or .44. But the likelihood of all seven of these being among the longest seven is $(.08)^7$, or $p < 2.09{*}10^{-8}$.

[7] For the five boundary pairs, χ^2s(1) = 29.63, $p < .0001$; 13.96, $p < .001$; 7.205, $p < .007$, 1.47, $p < .22$, and 3.38, $p < .066$, respectively; and $\chi^2(1) = 6.76$, $p < .009$, for the ten boundary pairs in seven of the films.

[8] The quadratic fit to the fourth quarter data is reliable, $R^2 = .233$, $t(245) = 7.36$, $p < .0001$.

[9] The tests for the quadratic fits to the first, second, and third quarters are reliable, $R^2 = .194$, $t(245) = 7.7$; $R^2 = .082$, $t(245) = 4.71$; and $R^2 = .131$, $t(245) = 6.09$, respectively, all ps < .0001.

[10] Setting aside the first bin and the bins of the third quarter, the mean of the last bin is 2.59 standard deviations above that for the remaining bins, $p < .005$.

[11] Let us set bins 1 and 20 aside, and ordinally rank the remaining eighteen bins, choosing the five with the highest probability. This is a combinations and permutations problem yielding 18!/(5!*13!) = 8,568 possible combinations of five bins. The a priori likelihood of five adjacent bins occurring among these is 13 (2–6, 3–7 . . . 15–19) = 13. Thus, ignoring the locations of those bins (which all happen to be in the third act) the probability of all adjacent bins being chosen is 13/8568, or $p < .0015$. Also, consider the number of release years for which the values in bins 11 through 15 are above the grand mean for bins 2 through 20 (0.45). Of the seventy-five data points (15 release years x 5 bins), forty-six are above that mean, $z = 2.43$, $p < .007$ by binomial expansion.

References

Anderson, Joseph, and Barbara F. Anderson. 2005. *Moving Image Theory: Ecological Considerations*. Carbondale: Southern Illinois University Press.

Bordwell, David. 2002. "Intensified Continuity." *Film Quarterly* 55 (3): 16–28.

———. 2006. *The Way Hollywood Tells It*. Berkeley: University of California Press.

Bordwell, David, and Kristin Thompson. 2004. *Film Art: An Introduction*. 7th ed. Boston: McGraw-Hill.

Bordwell, David, Janet Staiger, and Kristin Thompson. 1985. *The Classical Hollywood Cinema: Film Style and Mode of Production to 1960*. New York: Columbia University Press.

Cutting, James E., Jordan E. DeLong, and Christine E. Nothelfer. 2010. "Attention and the Evolution of Hollywood Film." *Psychological Science* 21 (3): 440–447.

Cutting, James E., Jordan E. DeLong, and Kaitlin L. Brunick. 2011. "Visual Activity and Hollywood Film: 1935 to 2005 and Beyond." *Psychology of Aesthetics, Creativity, and the Arts,* in press.

Deutelbaum, Marshall. 2000. "Review of *Storytelling in the New Hollywood*." *Film Criticism* 25 (1): 87–90.

Field, Syd. 1979. *Screenplay: The Foundations of Screenwriting*. New York: Dell.

Harrington, John. 1973. *The Rhetoric of Film*. New York: Holt, Rinehart, and Winston.

Ingermanson, Randy, and Peter Economy. 2010. *Writing Fiction for Dummies*. Hoboken, NJ: Wiley.

Messaris, Paul. 1994. *Visual Literacy: Image, Mind, and Reality.* Boulder, CO: Westview Press.

Over, Paul, Tzveta Ianeva, Wessel Kraaij, and Alan F. Smeaton. 2006. "TRECVID 2005—An overview." Proceedings of TRECVID 2005. http://www-nlpir.nist.gov/projects/tvpubs/tv5.papers/tv5overview.pdf (accessed 18 May 2010).

Salt, Barry. 1992. *Film Style and Technology: History and Analysis*. 2nd ed. London: Starword.

———. 2006. *Moving into Pictures*. London: Starword.

Smith, Tim J. 2006. "An Attentional Theory of Continuity Editing." PhD diss., University of Edinburgh. http://homepages.inf.ed.ac.uk/s9732397/publications/smith_ATOCE_2006.pdf (accessed 18 May 2010).

Smith, Tim J., and John M. Henderson. 2008. "Edit Blindness: The Relationship between Attention and Global Change Blindness in Dynamic Scenes." *Journal of Eye Movement Research* 2 (2): 6, 1–17.

Thompson, Kristin. 1999. *Storytelling in the New Hollywood*. Cambridge, MA: Harvard University Press.

Zacks, Jeffrey M., and Joseph P. Magliano. 2010. "Film, Narrative, and Cognitive Neuroscience." Forthcoming in *Art and the Senses,* eds. Francesca Bacci and David Melcher. New York: Oxford University Press.

Filmography

Altman, Robert. 1970. *M*A*S*H*. USA.

Bennett, Compton, and Andrew Marton. 1950. *King Solomon's Mines*. USA.

Huston, John. 1950. *The Asphalt Jungle*. USA.

Jackson, Peter. 2005. *King Kong*. New Zealand, USA, and Germany.

Koster, Henry. 1950. *Harvey.* USA.

Kramer, Stanley. 1960. *Inherit the Wind*. USA.

Oedekerk, Steve. 1995. *Ace Ventura: When Nature Calls*. USA.

Singer, Bryan. 1995. *The Usual Suspects*. USA and Germany.

Wilder, Billy. 1955. *The Seven Year Itch*. USA.

Woo, John. 2000. *Mission: Impossible II*. USA and Germany.

Zemeckis, Robert. 2000. *Cast Away.* USA.