EMPIRICAL STUDIES OF THE ARTS, Vol. 30(2) 143-166, 2012

MATURE CINEMATIC CONTENT FOR IMMATURE MINDS: "PUSHING THE ENVELOPE" VERSUS "TONING IT DOWN" IN FAMILY FILMS

DEAN KEITH SIMONTON
University of California, Davis

LAUREN ELIZABETH SKIDMORE

JAMES C. KAUFMAN

California State University San Bernardino

ABSTRACT

How does a film's content influence its reception by moviegoers and critics? What movie qualities result in better reviews, a higher box office, and more awards? This study investigates these questions in the specific genre of family films. One strategy is to "push the envelope" by intensifying adult themes and hints of sex and violence. An alternative strategy is to "tone it down," and keep any adult content to a minimum. A sample of 220 family films was assessed on (a) 15 measures of mature content; (b) multiple measures of film evaluations (3), box office performance (4), and movie honors (3, including children and teenager awards); and (c) 5 control variables. Broadly, this study supports the "pushing the envelope" strategy, especially regarding violence, topics to talk about, jump scenes, blood/gore, and inappropriate music. The optimal mature content for a family film differs markedly from that needed for films in general.

Films have been created for over 100 years, and have even become recognized as a major form of art, and yet psychologists have been somewhat slow to apply their methods to the scientific study of cinematic creativity and aesthetics.

© 2012, Baywood Publishing Co., Inc.
doi: http://dx.org/10.2190/EM.30.2.c
http://baywood.com

To be sure, Münsterberg (1916) and Arnheim (1957) pioneered the psychological study of film. Even so, these classic investigations, however insightful, remain far removed from the measurement and statistical techniques available today. For the most part, psychologists have been content to study film more indirectly, such as conducting laboratory experiments on how soundtracks influence the affective or cognitive responses to brief clips (e.g., Vitouch, 2001). This is not to say that cinema has not been subjected to direct scientific analyses. It is only that such research tends to concentrate on other cinematic variables besides those concerning creativity and aesthetics. For example, a huge literature has emerged concerning the economics of the film industry, with much of this work coming from the hands of researchers in cultural economics (for review, see Hadida, 2008). The focus is most often on whether a film makes money rather than whether the film counts as a creative work of art (e.g., De Vany, 2004). These two aspects of film are by no means equivalent and are often antithetical (Simonton, 2005a, 2005b, 2009b; see also Holbrook, 1999).

Nonetheless, psychologists have taken increased interest in cinematic creativity and aesthetics in the last decade or so (e.g., Beckwith, 2009; Plucker, Holden, & Neustadter, 2008; Plucker, Kaufman, Temple, & Qian, 2009; Simonton, 2002, 2009a; Zickar & Slaughter, 1999). This research has already addressed a diversity of methodological and substantive issues (for review, see Simonton, 2011). But here we would like to concentrate on one particular question: How is a film's cinematic impact influenced by its content? By "cinematic impact," we mean three sets of broad criteria, namely, the film's evaluation by critics and consumers, the film's short- and long-term box office performance, and the major awards or nominations that the film receives. These are the main criteria by which films are normally judged (Simonton, 2009b). By "content," however, we mean something more restricted, namely, the content that can be considered most suitable for "mature" viewers—especially content that some parents might not want their children or teenagers to encounter. Moreover, we are not interested in mature content in films intended for adult audiences, but rather our interest is in mature content as it appears in family films (cf. Cerridwen & Simonton, 2009). Although such films will clearly exhibit less "objectionable" content than other mainstream films, they can still contain material that might not be considered suitable for viewing by the whole family.

To illustrate, consider one of the most successful family films ever produced: the 1982 *E.T.: The Extra-Terrestrial*. This picture received a rating of PG ("Parental Guidance Suggested: Some Material May Not Be Suitable For Children") from the Motion Picture Association of America (MPAA; www.mpaa.org) because it included some content that denied it an MPAA rating of G ("General Audiences: All Ages Admitted"). In particular, *E.T.* included some "gore" (a close-up of a bleeding cut finger), profanity (both four-letter words and sexual references), alcohol consumption (*E.T.* gets drunk), suggested teenage smoking (in the card game scene), and some frightening or intense scenes

(involving guns at the film's end). When *E.T.* was re-released in 2002, it was cleaned up to remove the more objectionable material (e.g., the expression "penis breath" and the guns), but it still received a PG rating for what remained. Some family films will even contain content that earns an MPAA rating of PG-13 ("Parents Strongly Cautioned: Some Material May Be Inappropriate For Children Under 13"). An example is the 2005 *Harry Potter and the Goblet of Fire*, which contains some amount of sex and nudity, violence and gore, profanity, alcohol consumption, and frightening or intense scenes. As these two instances imply, mature content is a quantitative rather than qualitative feature of film (Cerridwen & Simonton, 2009; Thompson & Yokota, 2004). The sex and violence in *Harry Potter* is stronger than that in *E.T.*, yet the latter contains noticeably stronger content than the 2008 *WALL·E*, which earned a G rating for rather mild "sex" (viz. *WALL·E* putting on as eyeglasses a bra he had found) and minimal violence (e.g., EVE's ray gun that only hits inanimate objects so that nobody is killed or injured). Stated differently, even family films vary immensely concerning the amount or intensity of maturity expected of the movie-going audience.

Now this variation in mature content raises a theoretically interesting and pragmatically important question: In making a family film, how far should the filmmaker go? What content yields the greatest overall impact across all audiences? For instance, when watching the film at home, G-rated material might maintain a young child's interest, but the innocent content could send the teens off yawning to their bedrooms or put the parents fast asleep on the couch. More PG-rated content might keep the whole family in the living room together, but put parents in the uncomfortable position of explaining to their puzzled toddlers what some problematic expression or action signifies. Presumably, the most successful family films will find just the right mix of mature content that keeps everybody happy at least most of the time. For some content, the filmmaker will try to "push the envelope" by maximizing the intensity without going beyond a certain threshold (e.g., *E.T.* showing teenagers with cigarettes in ashtrays but without actually showing them smoking). For other content, the filmmaker will strive to "tone it down" as much as possible without reducing the film to the soporific maturity level of the cartoons on Saturday morning television. Thus, *WALL·E's* bra episode probably caused parents to laugh just as much as their children, but the reaction would have been very different if the bra were used in a more sexually explicit fashion (e.g., placed over his chest rather than his eyes). In the push-the-envelope case, a content measure might have a positive correlation with at least some measures of cinematic impact, whereas in the tone-it-down case, the correlation might be negative. The size and direction of these correlations are expected to be contingent not only on the specific content attribute but also on the particular impact criterion. The content predictors of critical acclaim may differ from the predictors of consumer praise, for example.

It is the purpose of this investigation to examine the relative importance of "pushing the envelope" versus "toning it down" for all the principal forms of mature content (sex and violence, suggestive or scary music, alcohol, drug, or tobacco use, profanity, etc.) and for all major types of cinematic impact (viz. evaluations, earnings, and awards). Nevertheless, before doing so we must acknowledge that a considerable body of research has addressed this question for films in general (e.g., Cerridwen & Simonton, 2009; Holbrook, 1999; Ravid & Basuroy, 2004; Simonton, 2005b; Thompson & Yokota, 2004). This research has most often used either the MPAA ratings as measures of mature content or the assessments of specific content categories provided by various websites (e.g., www.kids-in-mind.com; www.screenit.com). Even so, from the standpoint of our current interests, these earlier studies have two drawbacks.

First, because these investigations examined films in general, the samples were dominated by films that received the R rating ("Restricted: Children Under 17 Require Accompanying Parent or Adult Guardian"), an MPAA assignment whose frequency is rivaled only by PG-13. Yet R-rated films are decidedly not family films.

Second, many of these prior inquiries deliberately excluded animations because of the need to look at cinematic attributes that are only relevant to live-action films (e.g., awards and nominations for direction and acting). Although not all animations are considered family films (e.g., the R-rated 1999 *South Park: Bigger Longer & Uncut*), a very large proportion of high-impact family films are indeed animated. Accordingly, the deletion of animations seriously undermines how much these results can be generalized to the genre of interest here.

All told, the present study will provide a perspective on family films that cannot be inferred from previous studies. That said, later we will compare our main findings with earlier results to learn what generalizes and what does not.

METHOD

Sample

The sampling began by identifying 239 films that satisfied one or both of the following criteria: (a) the film belonged to the "family" genre as designated by the Internet Movie Database (IMDb.com) and (b) the film was a Disney production (a studio devoted to family films). This initial sample was further reduced to 220 films by imposing two additional criteria, namely, that (a) the film had to receive an MPAA rating of G, PG, or PG-13 and (b) the film had to be released between 1996 and 2009, inclusively. The latter restriction ensured that some key impact variables would be available for most or all of the sampled films. The resulting sample was distributed fairly evenly across the consecutive years, a fact that suggests there exists a certain fixed market for about 15-16 family films per year.

Measures

To the extent possible, all 220 films were assessed on variables regarding mature content, cinematic impact, and statistical control.

Mature Content

Screen It! (www.screenit.com) is a website that scores feature films according to mature content. The ratings are done at the time they are released and thus before there can be feedback regarding the film's impact. Because the site does not have any political or religious affiliation, it offers a suitable source for ratings of mature content. Furthermore, it rates films on more distinct content attributes than any alternative site (cf. www.kids-in-mind.com). To be specific, Screen It! evaluates films on 15 content attributes using a six-level scale: none, minor, mild, moderate, heavy or extreme. Rank categories were here converted to a numerical score from 1 = none to 6 = extreme. The resulting measures and their basic statistics were as follows (quotations from www.screenit.com/about_us.asp):

1. Alcohol/drugs: "any consumption, presence of and/or references to or comments about alcohol or drugs and their effects" (M = 2.53, SD = 1.19, range = 1-6);
2. Blood/gore: blood and/or gore plus "the sight and/or sound of bodily functions and instances of crude humor" (M = 2.38, SD = 1.00, range = 1-6);
3. Disrespectful/bad attitude: "dishonorable, contemptible, ill-tempered and other such behavior" that children might mimic (M = 4.73, SD = 1.07, range = 2-6);
4. Frightening/tense scenes: including PG- or G-rated instances (M = 3.40, SD = 1.30, range = 1-6);
5. Guns/weapons: "the use and effects of any guns or other weapons (knives, bombs, etc...)" (M = 2.98, SD = 1.72, range = 1-6);
6. Imitative behavior: "irritating or dangerous behavior and dialogue that kids may imitate . . . ranging from playing with electrical wiring to performing karate moves to using phrases such as 'shut up,' 'idiot,' and 'butt head'" (M = 3.34, SD = 0.93, range = 1-5);
7. Inappropriate music: "movies aimed toward teens usually contain popular music . . . [that may contain] objectionable material" (M = 1.28, SD = 0.66, range = 1-4);
8. Jump scenes: "any scene from any genre (such as a person suddenly being grabbed) that will make you and/or your child jump" (M = 1.46, SD = 0.78, range = 1-5);
9. Scary/tense music: "suspenseful movie scenes are usually more suspenseful due to the use of tense music" (M = 3.57, SD = 1.68, range = 1-6);

10. Profanity: "a detailed listing of expletives and religious phrases used as exclamations" (M = 2.29, SD = 1.23, range = 1-5);
11. Sex/nudity: "any scenes that show nudity (including scantly clad people, cleavage, etc.) and/or sexual behavior (partial, full or implied), as well as related dialogue and any other sexual/sensual material" (M = 2.84, SD = 1.27, range = 1-5);
12. Smoking: "Any scenes where characters smoke (cigarettes, cigars, pipes)" (M = 1.58, SD = 0.84, range = 1-6);
13. Tense family scenes: "scenes that show or imply family problems . . . [including] any such scenes or discussion of family bickering, fights or upheavals (such as divorce and death)" (M = 2.74, SD = 1.21, range = 1-6);
14. Topics to talk about: "scenes or subject material that you may want to discuss with your children" (M = 3.47, SD = 0.76, range = 2-5);
15. Violence: "any scenes of violence—hitting, kicking, stabbing, shooting, exploding, and every other possible act of harming people, animals, or property is listed, as is physical slapstick style humor" (M = 3.88, SD = 1.22, range = 1-6).

Interestingly, two content variables did not descend to the lowest possible score of 1, whereas a third of the content variables attained a maximum score of 5. Scores of 6 in these content categories would probably push the film toward an MPAA rating of R. Remarkably, the highest levels of violence, blood/gore, and guns/weapons can still appear in family films.

The Screen It! ratings have been shown to coincide closely with the alternative evaluations provided by Kids-in-Mind (www.kids-in-mind.com), where the latter provides evaluations on three somewhat more inclusive dimensions: sex/nudity, violence/gore, and profanity (Thompson & Yokota, 2004). The same scores also correspond with the MPAA ratings films receive (Cerridwen & Simonton, 2009).

Cinematic Impact

Cinematic impact can be assessed by three sets of variables: film evaluations, box office performance, and movie awards and nominations.

Film Evaluations

These evaluations take two forms (cf. Holbrook, 1999; Plucker et al, 2008).

First are the evaluations published by professional film critics. Two websites have consolidated these diverse critical assessments into composite scores that were adopted unchanged here: Metacritic (www.metacritic.com; M = 56.08, SD = 17.47, range = 9-96; n = 199) and Rotten Tomatoes (www.rottentomatoes.com; M = 55.92, SD = 25.86, range = 0-100; n = 218). Judging from the ranges, the cinematic merit of the sampled films ranges from

"five-star" masterpieces to "turkeys" or "bombs," with the average film being slightly above average. These two composites also correlate .93 ($p < .001$, $n = 199$), a high correlation that reflects considerable overlap in the film criticisms used to compute the scores (albeit the two sites apparently use different computational formulas). Although these critical evaluations are published early during the film's theatrical release, such critical evaluations correlate highly with movie-guide ratings published after the films come out in video-DVD format (Simonton, 2009a).

The second form of film evaluations are those provided by consumers themselves, especially the users of film websites (Plucker et al., 2008). Here we will use the ratings posted as of January 2010, on the Internet Movie Database (IMDb; www.imdb.com; $M = 6.20$, $SD = 1.28$, range = 1.40-8.90; $n = 220$). Despite the common belief that critics and consumers have very different judgments of cinematic merit, the two alternative evaluations tend to exhibit positive correlations (Chang & Ki, 2005; Plucker et al., 2009). Although novices and experts generally rate creative or aesthetic stimuli differently (Kaufman, Baer, & Cole, 2009; Kaufman, Baer, Cole, & Sexton, 2008), Plucker et al. suggest that IMDb raters are not actual novices but can instead be considered "quasi experts" or "gifted novices." These terms refer to a particular level of expertise. Quasi experts typically have had some professional experience in an area (i.e., taken advanced college classes or participated at a lower level) but are not at the level of traditional experts. Gifted novices do not have expertise yet show strong potential for eventual expertise (such as a high school student with an aptitude for a particular creative domain). Both quasi experts and gifted novices do generally show rater agreement with experts in other creative contexts (Baer, Kaufman, & Riggs, 2009; Kaufman, Gentile, & Baer, 2005).

In the present case, the IMDb score correlates .81 ($p < .001$, $n = 199$) with Metacritic and .81 ($p < .001$, $n = 218$) with Rotten Tomatoes. Nonetheless, these correlations are not so high as to prevent some divergences in their respective correlations with the mature content indicators. Critics and consumers view movies much the same way, but not identically (Chang & Ki, 2005; Holbrook, 1999).

Box Office Performance

Financial data were obtained from the Internet Movie Database (www.imdb.com). The raw data consisted of total domestic gross, the gross earnings on the opening weekend, and the production budget. However, because these three variables are highly skewed, they were all subjected to a logarithmic transformation, yielding new values for gross ($M = 18.37$, $SD = 0.87$, range = 15.07-20.10, $n = 214$), opening weekend ($M = 16.92$, $SD = 1.23$, range = 10.61-18.88, $n = 210$), and budget ($M = 17.80$, $SD = 0.86$, range = 14.77-19.52). Although the log-transformed variables will be used in all statistical analyses

reported later, it is worth noting that the average family film in the sample cost about $78.4 million to produce, made about $35.5 million on opening weekend, and grossed about $128.8 million. In any event, budget correlated .41 ($p < .001$, $n = 190$) with opening weekend and .60 ($p < .001$, $n = 192$) with gross, while the latter two variables correlated .74 ($p < .001$, $n = 209$).

Finally, a fourth variable was computed, namely, estimated profit, which was defined as the total gross expressed in millions of U.S. dollars minus two times the budget also expressed in millions (cf. Cerridwen & Simonton, 2009), that is, profit = (gross – 2 × budget)/1,000,000 ($M = –19.80$, $SD = 87.93$, range = –308.38-189.44). In other words, as a rough estimate, a film does not start becoming truly profitable until it earns twice as much as it cost to make (De Vany & Walls, 1999). The reason is that films also have marketing and promotion costs that are proportional to the production costs (Prag & Casavant, 1994). According to this formula, the average family film loses millions at the box office. Although such films often have higher gross earnings, they also cost much more to make, market, and promote. For example, although the 2008 *The Chronicles of Narnia: Prince Caspian* grossed $141.6 million, its budget was $225 million! In any case, because the distribution of estimated profit was approximately normal, it was not necessary to impose the logarithmic transformation.

Movie Awards and Nominations

Many empirical studies have used movie awards and nominations to assess a film's cinematic impact (e.g., Deuchert, Adjamah, & Pauly, 2005; Faulkner & Anderson, 1987; Simonton, 2004b; Sochay, 1994). Although such honors are very numerous and varied, research has indicated that the awards and nominations bestowed by the Academy of Motion Picture Arts and Sciences (the "Oscars"; http://www.oscars.org) have the highest general reliability and validity (Ginsburgh, 2003; Simonton, 2004a; cf. Collins & Hand, 2006; Pardoe & Simonton, 2008). Nonetheless, these honors reflect the judgments of adult professionals, namely, the approximately 6,000 members of the Academy. Consequently, we also sought forms of recognition that represented the judgments of less mature minds, that is, children and teenagers. For this purpose, we used the Kids' Choice (www.nick.com) and the Teen Choice (teenchoiceawards.com). For each set, a composite measure was created that consisted of a count of nominations and awards, where a nomination earned one point and an award two points (Basuroy, Chatterjee, & Ravid 2003; Cattani & Ferriani, 2008; Simonton, 2004b). The resulting composites represent cinematic merit as judged by children ($M = 1.20$, $SD = 1.24$, range = 0-8, $n = 220$), teenagers ($M = 1.80$, $SD = 3.37$, range = 0-21, $n = 190$), and adults ($M = 0.77$, $SD = 1.88$, range = 0-17, $n = 220$). Significantly, although the children's and teenagers' honors are slightly correlated ($r = .19$, $p < .01$, $n = 190$), neither of these measures was correlated with

the adults' honors (r = .04 and r = .09, respectively, both ps > .1). So, the first two youngest age categories have more common cinematic tastes than either share with the oldest age category.

Statistical Controls

Three variables were introduced to control for potential statistical artifacts (cf. Cerridwen & Simonton, 2009; Simonton, 2009a). The first was the *date* that the film was released (M = 2002.66, SD = 4.02, range = 1996-2009). This variable allows correction for linear trends over the 14-year interval. The second control was the film's MPAA rating, which was recorded using three 0-1 dummy variables for the three ratings in this sample: *G* (M = 0.21, SD = 0.41), *PG* (M = 0.46, SD = 0.50), and *PG-13* (M = 0.33, SD = 0.47). The means for the three dummy variables inform us that 46% of the films were rated PG, followed by 33% for PG-13 and 21% for G. Hence, PG is by far the most popular rating for family films, a major departure from films in general (De Vany & Walls, 2002). Finally, another 0-1 dummy variable recorded whether the film was an animation rather than relying primarily on live-action characters (M = 0.40, SD = 0.49). Thus, 40% of the films in the sample were animations, a highly popular category for family showing. The utility of all of these control variables will become immediately obvious in the first part of the Results section.

RESULTS

Before discussing the results for the various cinematic impact criteria, it is advisable to make two sets of observations. The first set of observations concerns the correlations between the 15 indicators of mature and the 5 control variables. These correlations are shown in Table 1. Besides indicating what will be partialled out when these control variables are introduced into the multiple regression equations to be reported later, these coefficients have intrinsic substantive interest. In particular, please note the following three points:

1. Judging from the correlations with date of release, some mature content has been increasing in intensity even during the relatively short period represented by this sample (viz. 1996-2009). This increase is seen for blood/gore, disrespectful/bad attitude, frightening/tense scenes, guns/weapons, imitative behavior, jump scenes, scary/tense music, sex/nudity, tense family scenes, topics to talk about, and violence. Apparently, the films that emerge each year are increasingly pushing the envelope. Indeed, the only mature content to decline over the years is profanity and smoking, which have been toned down in family films. Alcohol/drugs and inappropriate music, in contrast, have been stable over this 14-year period.

2. The MPAA ratings of G, PG, and PG-13 exhibit rather contrasting relationships with mature content. The contrasts are especially conspicuous for G and PG-13: The statistically significant correlations that these ratings have with

Table 1. Mature Content and Control Variables: Pearson Product-Moment Correlations

Content assessment	Date	G	PG	PG-13	Animated
Alcohol/drugs	.09	–.42***	–.19**	.56***	–.40***
Blood/gore	.32***	–.35***	–.09	.40***	–.17*
Disrespectful/bad attitude	.44***	–.26***	.04	.19**	–.03
Frightening/tense scenes	.26***	.20**	.10	–.27***	.41***
Guns/weapons	.26***	–.20**	–.03	.20**	.01
Imitative behavior	.46***	–.28***	.05	.18**	.16*
Inappropriate music	.07	–.17*	–.07	.22**	–.05
Jump scenes	.25***	.03	–.06	.04	.03
Profanity	–.18**	–.49***	–.29***	.73***	–.61
Scary/tense music	.31***	.12	.11	–.22**	.35***
Sex/nudity	.22**	–.48***	–.17*	.60***	–.42***
Smoking	–.22**	–.21**	–.05	.24***	–.12
Tense family scenes	.33***	–.16*	.14*	–.01	–.13
Topics to talk about	.72***	–.05	.12	–.09	.12
Violence	.22**	–.22**	–.08	.28***	–.02

Note: $N = 220$ for all correlations.
*$p < .05$, **$p < .01$, ***$p < .001$.

mature content will have opposite signs. This polarity is most striking for alcohol/drugs, blood/gore, disrespectful/bad attitude, frightening/tense scenes, guns/weapons, imitative behavior, inappropriate music, profanity, sex/nudity, smoking, and violence. For most of these content measures, G has less and PG-13 has more, but there is one exception: frightening/tense scenes, where G has more and PG-13 has less. In the main, PG-rated films fall between these two family-film extremes, and thus the correlations are often small and nonsignificant. In the few cases where the PG correlations are substantial, they usually line up with the G correlations, with the lone exception of tense family scenes—a trait attribute that has no relevance for films rated PG-13.

3. Animated family films have a distinctive mature content. On the one hand, they are less prone to feature alcohol/drugs, blood/gore, and sex/nudity. On the other hand, animations are more likely to contain frightening/tense scenes, imitative behavior, and scary/tense music. Plainly, animated films maintain audience interest by means of more innocent forms of content stimulation.

The second set of observations has to do with the correlations between the three sets of impact criteria. To what extent do film evaluations, box office performance, and movie honors agree with respect to cinematic merit? The answers are revealed in the following correlations: (a) the Metacritic, Rotten Tomato, and IMDb scores correlate .35 ($p < .001$, $n = 197$), .29 ($p < .001$, $n = 214$), and .40 ($p < .001$, $n = 214$), respectively, with domestic gross, whereas only the IMDb scores correlate with opening weekend ($r = .18$, $p < .01$, $n = 210$); (b) the Metacritic, Rotten Tomato, and IMDb scores correlate .47 ($p < .001$, $n = 199$), .41 ($p < .001$, $n = 218$), and .46 ($p < .001$, $n = 220$) with the adults' movie honors, but do not correlate at all with the children's and teenagers' honors; (c) the children's and teenagers' movie honors correlate .35 ($p < .001$, $n = 210$) and .30 ($p < .001$, $n = 186$) with opening weekend and .34 ($p < .001$, $n = 214$) and .31 ($p < .001$, $n = 187$) with gross, whereas the adults' honors correlate only with gross ($r = .34$, $p < .001$, $n = 214$); and (d) although films that win children's and adults' honors are profitable ($r = .29$, $p < .001$, $n = 192$, and $r = .14$, $p < .05$, $n = 192$), the films that win the teenagers' honors do not satisfy our profitability criterion. The latter finding may seem peculiar, but it is in fact very difficult to identify correlates of profit (Simonton, 2011). There is also no correlation between profit and film evaluations. Finally, even if budget is not a direct criterion of cinematic impact, it correlates strongly with Metacritic ($r = .33$, $p < .001$, $n = 181$), Rotten Tomatoes ($r = .27$, $p < .001$, $n = 193$), and IMDb ($r = .40$, $p < .001$, $n = 193$) ratings, as well as the adult movie honors ($r = .28$, $p < .001$, $n = 193$). Given the correlations given earlier between budget and box office, it should be evident that money can buy some degree of cinematic success (cf. De Vany & Walls 1999; Sedgwick & Pokorny, 1999; Simonton, 2005a).

Film Evaluations

Table 2 gives the correlations between the measures of mature content and the three evaluation indicators, the first two by critics and the third by consumers. From these statistics, we can draw the following major conclusions. First, all three evaluations are positively correlated with frightening/tense scenes, jump scenes, scary/tense music, and topics to talk about. Second, the two critic evaluations (Metacritic and Rotten Tomatoes), but not the consumer evaluation (Internet Movie Database), are negatively correlated with imitative behavior, profanity, and sex/nudity. Third, the consumer evaluation is distinguished by being positively associated with blood/gore, disrespectful/bad attitude, and violence. Lastly, it is notable that smoking and tense family scenes have no relevance no matter which impact criterion is chosen.

The statistics in Table 2 are misleading insofar as the correlations are not corrected for the control variables listed in Table 1 (release date, the three MPAA ratings, and animation status). In addition, some of the mature content indicators correlate highly enough that it is difficult to tease out their relative importance

Table 2. Mature Content and Film Evaluations: Pearson Product-Moment Correlations

Content assessment	MC	RT	IMDb
Alcohol/drugs	–.13	–.14*	.02
Blood/gore	.06	.06	.20**
Disrespectful/bad attitude	.00	.03	.15*
Frightening/tense scenes	.35***	.33***	.33***
Guns/weapons	.18*	.13	.27***
Imitative behavior	–.15*	–.15*	.01
Inappropriate music	–.15*	–.11	–.07
Jump scenes	.23**	.20**	.28***
Profanity	–.23**	–.23**	–.08
Scary/tense music	.31***	.29***	.30***
Sex/nudity	–.24**	–.21**	–.03
Smoking	.08	.06	.06
Tense family scenes	.04	.02	.09
Topics to talk about	.21*	.19**	.29***
Violence	.11	.12	.25***

Note: MC = Metacritic score (n = 199), RT = Rotten Tomatoes score (n = 218), and IMDb = Internet Movie Dtabse user ratings (n = 220).
*p < .05, **p < .01, ***p < .001.

(e.g., blood/gore and violence tend to occur together; see also Cerridwen & Simonton, 2009). To remedy this problem, each film evaluation was regressed on both the control variables and the mature content assessments. Content assessments that were not statistically significant (by the .05 criterion) were then progressively deleted by backward elimination. The findings for the three cinematic impact criteria are as follows (where the βs are standardized partial regression coefficients):

1. Metacritic ratings were a positive function of smoking (β = .18, p = .011) and topics to talk about (β = .24, p = .006), but a negative function of imitative behavior (β = –.16, p = .037) and sex/nudity (β = –.27, p = .003). The only statistically significant control variable was animation, which was a positive predictor of these evaluations (β = .22, p = .008). The equation as a whole explained 23% of the variance in the impact criterion (R^2 = .23, p < .001, n = 199). It should be observed that smoking was not a correlate but still emerged as a

predictor, an example of statistical suppression (Maassen & Bakker, 2001). The zero-order correlation between smoking and these ratings was obscured by the content and/or control variables. Other examples of suppression effects will appear below.

2. Rotten Tomatoes ratings were a positive function of smoking ($\beta = .17$, $p = .017$), topics to talk about ($\beta = .22$, $p = .016$), and frightening/tense scenes ($\beta = .17$, $p = .024$), but a negative function of sex/nudity ($\beta = –.20$, $p = .039$). Once again, the only significant control variable was animation ($\beta = .20$, $p = .018$). Here 20% of the variance was accounted for ($R^2 = .20$, $p < .001$, $n = 218$). Smoking again came in to the equation owing to statistical suppression. Although both Metacritic and Rotten Tomatoes were correlated with frightening/tense scenes, this content only predicted the latter criterion in the regression equation.

3. The IMDb ratings were also a positive function of topics to talk about ($\beta = .20$, $p = .025$), but there was also a positive function of guns/weapons ($\beta = .17$, $p = .009$) and jump scenes ($\beta = .17$, $p = .008$). Animated films also rated more highly ($\beta = .31$, $p < .001$). Now 24% of the variance was explained by the regression equation ($R^2 = .24$, $p < .001$, $n = 220$). For this criterion, there were no suppression effects.

It is apparent that one content variable cuts across all three evaluations, namely, topics to talk about. For this attribute, it is best to push the envelope. Although the critics gave better reviews to movies containing smoking, sex/nudity had a negative impact. It is also of interest that all three evaluations rate animated family films more highly than life-action films.

Box Office Performance

Table 3 shows how the 15 content variables relate to the financial success indicators. Let us begin with the correlates that are common to opening weekend and gross, namely, blood/gore, disrespectful/bad attitude, guns/weapons, imitative behavior, scary/tense music, sex/nudity, and violence—all of which are positive. Yet these positive associations must be offset by production costs, which has positive relationships with most of the same variables (with the notable exception of sex/nudity). For this reason, we should look at profit. Now only profanity has a positive relation (because swear words are cheap), and negative relations are obtained for disrespectful/bad attitude, frightening/tense scenes, guns/weapons, and scary/tense music.

To clarify the connections between mature content and box office, multiple regression analyses were again run with the control variables and with backward elimination of the nonsignificant control variables. The results are as follows:

1. Performance on opening weekend is a positive function of violence ($\beta = .31$, $p < .001$) and a negative function of smoking ($\beta = –.27$, $p < .001$) and tense family scenes ($\beta = –.15$, $p = .023$). With respect to the control variables, PG-13 films did better by this criterion ($\beta = .27$, $p = .013$) and animated films did worse ($\beta = –.17$,

Table 3. Mature Content and Box Office Performance: Pearson Product-Moment Correlations

Content assessment	Opening weekend	Gross	Budget	Profit
Alcohol/drugs	.13	.11	.04	.10
Blood/gore	.31***	.35***	.42***	–.08
Disrespectful/bad attitude	.25***	.20**	.22**	–.19**
Frightening/tense scenes	.07	.17*	.33***	–.16*
Guns/weapons	.20**	.22**	.39***	–.23**
Imitative behavior	.18*	.16*	.09	–.00
Inappropriate music	.07	.02	.00	.08
Jump scenes	.13	.26***	.24**	.04
Profanity	.18**	.13	–.04	.19**
Scary/tense music	.17*	.23**	.41***	–.25***
Sex/nudity	.25***	.22**	.10	.02
Smoking	–.15*	–.09	.01	.00
Tense family scenes	–.03	.09	.13	–.02
Topics to talk about	.06	.16*	.19**	–.13
Violence	.35***	.36***	.39***	–.12

Note: Opening weekend earnings (n = 210), gross (n = 214), and budget (n = 193) are all subjected to a logarithmic transformation to render their distributions more normal. That transformation was not necessary for profit (gross minus twice the budget; n = 192).

*p < .05, **p < .01, ***p < .001.

p = .037). Altogether, 29% of the variance was explained (R^2 = .29, p < .001, n = 210). Because tense family scenes did not display any association with this criterion, its presence in the prediction equation constitutes a suppressor effect.

2. Domestic gross was also a positive function of violence (β = .26, p < .001), but with jump scenes as an added positive predictor (β = .135, p = .040). Smoking continued to have a negative effect (β = –.14, p = .039). But given the nonsignificant correlation in Table 3, the presence of smoking as a predictor might be considered the result of statistical suppression. Among the control variables, only release date proved important: As films become more recent, they earn more (β = .16, p = .021), an increase that may reflect the inflation of ticket prices. This time, 25% of the variance was accounted for (R^2 = .25, p < .001, n = 214).

3. Profit was more difficult to predict, with only 13% of the variance explained ($R^2 = .13$, $p < .001$, $n = 192$). No control variable was statistically significant, and only one content variable emerged as a predictor: Guns/weapons had a negative impact ($\beta = -.25$, $p = .001$).

To put these findings in context, the production budget was also subjected to the same analysis, yielding an equation that explained 29% of the variance ($R^2 = .29$, $p < .001$, $n = 193$). The films that cost the most to produce have lots of blood/gore ($\beta = .24$, $p = .004$) and scary/tense music ($\beta = .32$, $p < .001$). Evidently, horror films at the family level are expensive to produce. Lastly, films are also becoming increasingly more costly to make ($\beta = .15$, $p = .025$).

Movie Awards and Nominations

Table 4 provides the parallel correlations for the three movie honors. At once, it becomes obvious that not a single content variable exhibits a consistent correlation across all three impact measures, albeit the children and teenagers concur on the positive effect of alcohol/drugs, profanity, and sex/nudity, and the negative effect of frightening/tense scenes. The adult honors actually disagree dramatically for frightening/tense scenes, profanity, and sex/nudity—instances where the relationships have opposite signs. In one case only, violence, do the teenagers and adults share the same cinematic tastes.

The multiple regression analyses implemented for film evaluations and box office performance were applied here, too, producing the following findings:

1. The children's awards and nominations are a positive function of sex/nudity ($\beta = .22$, $p = .025$), blood/gore ($\beta = .17$, $p = .047$), and inappropriate music ($\beta = .14$, $p = .040$), but a negative function of frightening/tense scenes ($\beta = -.17$, $p = .045$). The only statistically significant control variable was release date, indicating that more recent family films are less likely to earn these honors ($\beta = -.16$, $p = .033$). Just 18% of the variance was explained ($R^2 = .18$, $p < .001$, $n = 220$). Given the correlation shown in Table 4, the impact of blood/gore was apparently boosted by statistical suppression.

2. It is easier to predict the corresponding teenagers' honors, with 35% of the variance accounted for ($R^2 = .35$, $p < .001$, $n = 190$). These honors are more likely for films that feature violence ($\beta = .28$, $p = .004$) and sex/nudity ($\beta = .18$, $p = .046$), but less likely for those that feature guns/weapons ($\beta = -.21$, $p = .034$), the latter emerging owing to statistical suppression. In addition, the probability of an award or nod has been increasing over the years ($\beta = .19$, $p = .004$). However, animated films are less likely to be so honored ($\beta = -.27$, $p = .001$).

3. Prediction of the adults' honors was comparable to that of the children's: just 20% ($R^2 = .20$, $p < .001$, $n = 220$). The films that enjoy higher odds of receiving these awards and nominations contain alcohol/drugs ($\beta = .19$, $p = .031$), guns/weapons ($\beta = .20$, $p = .004$), jump scenes ($\beta = .21$, $p = .002$), and topics to talk about ($\beta = .19$, $p = .039$), but are less likely to contain sex/nudity ($\beta = -.35$,

Table 4. Mature Content and Movie Awards-Nominations: Pearson Product-Moment Correlations

Content assessment	Children	Teenagers	Adults
Alcohol/drugs	.19**	.37***	.04
Blood/gore	.13	.29***	.10
Disrespectful/bad attitude	–.00	.21**	.05
Frightening/tense scenes	–.27***	–.17*	.29***
Guns/weapons	–.03	.11	.23**
Imitative behavior	.11	.23**	–.08
Inappropriate music	.23**	.12	–.07
Jump scenes	–.14*	.06	.30***
Profanity	.30***	–.36***	–.14*
Scary/tense music	–.23**	–.06	.26***
Sex/nudity	.33***	.44***	–.17*
Smoking	.05	.02	.10
Tense family scenes	–.05	.08	.10
Topics to talk about	–.15*	.14	.17*
Violence	–.02	.23**	.19**

Note: Children = Kids' Choice Awards (n = 220), Adolescents = Teen Choice Awards (n = 190), and Adults = Academy Awards ("Oscars"; n = 220). In each case the award measures are summations of awards and nominations, where awards = 2, nominations = 1, and neither = 0.

*$p < .05$, **$p < .01$, ***$p < .001$.

$p < .001$). The alcohol/drugs relation only appears in the regression equation and thus is the consequence of statistical suppression. None of the control variables had statistically significant effects.

The regression results reinforce the conclusion from the correlation findings: Adults award very different mature content than do children and teenagers. Particularly striking is the divergent reaction to sex/nudity. Within a given MPAA rating, for children and teenagers there cannot be too much (pushing the envelope), whereas for adults there cannot be too little (toning it down).

DISCUSSION

The discussion consists of three parts: consolidations, comparisons, and implications.

Consolidations

To get a better conception of the empirical results, we decided to reorganize the main findings into a more structured recapitulation. This reorganization will concentrate on the multiple regression analyses. Because these analyses include the control variables, they avoid possible spurious relationships. Owing to occasional statistical suppression, these analyses were also able to detect new predictors not apparent in the correlation coefficients. We will start with the overall prediction equations and then turn to the specific predictors.

Prediction Equations: How Much Can We Explain?

The first point to note is that the several criteria of cinematic impact vary greatly in their predictability given the mature content and statistical control variables. The criterion least susceptible to prediction—and this should be no surprise to producers—is net profit. Only 13% of the variance was accounted for, and only one content variable (guns/weapons) emerged as a significant predictor! The best advice that we can give to a moviemaker who wants to create a moneymaking family film is "leave out the guns and other weapons!" Nevertheless, this poor predictability may be partly ascribed to the fact that profit is a difference score, and hence will have lower reliability, which will attenuate the relationship (Lord, 1958). This lowered reliability is probably accentuated by the fact that profit is computed using production budget, which in most cases constitutes only a very rough estimate.

Fortunately, because the remaining impact criteria are not difference scores, they do not suffer from this methodological problem. Still, only 18% of the variance in the children's evaluations, represented by the Kids' Choice Awards, can be predicted, despite the inclusion of many more predictors. This lower predictability may partially reflect the nature of the voting, which is conducted like a popularity contest similar to *American Idol*.

Predictability moves up a notch when we turn to film assessments by older moviegoers. Specifically, we can predict 20% of the variance in the Rotten Tomatoes ratings and in the Academy Awards (Oscars), 23% of the variance in the Metacritic evaluations, and 24% in the IMDb ratings. Because the last criterion is based on consumer ratings that span a wide spectrum of age groups, it may be most representative of moviegoers as a whole. In line with this conjecture, we can account for 25% of the variance in domestic gross earnings and 29% of the variance in the earnings on the opening weekend. Yet with respect to predictability, the most remarkable finding probably concerns the teenagers' assessments, as indicated by the Teen Choice Awards: 35% of the variance is explained! That the cinematic opinions of teens are so predictable may help adults appreciate why it so often seems that family films (and, indeed, most films) cater so much to this age group.

Criterion Predictors: What Content Variables Work?

The foregoing discussion just looked at percentage of variance explained without regard to the specific content predictions. Now we will consolidate the earlier findings by grouping the indicators of mature content into four groups. Again, the focus is on the multiple regression findings. The resulting predictors fall into four groups:

1. *Null effects*—Some content measures had no predictive utility for any criterion of cinematic impact. Three variables fall into this category: disrespectful/bad attitude, profanity, and scary/tense music. This is not to say that these three gauges of mature content exhibit no correlations with cinematic impact. For correlations that involve these three measures do appear in Tables 2 thru 4. It is only the case that the three do not have any residual predictive power after adjusting for other content variables and the statistical controls. For example, all three of these variables are associated with other content variables that had stronger predictive relations. Even so, with respect to the filmmaker's decision to push the envelope or tone it down, these three aspects of mature content can be largely ignored.

2. *Single effects*—In the next category are those indicators of mature content that only appeared as predictors with a single impact criterion. These predictive associations may be either positive or negative. The former includes the relationships between alcohol/drugs and the Academy Awards, and two predictors of the Kids' Choice Awards, namely, blood/gore and inappropriate music. Here the filmmaker can push the envelope, albeit what is specifically pushed depends on whether the movie is aimed at the adults or the children in the family. The negative predictive relations are imitative behavior and the Metacritic ratings and tense family scenes and opening weekend. So this content might be toned down.

3. *Multiple but inconsistent effects*—In the third group are content variables that make the filmmaker's life more difficult if he or she wishes to make a cinematic product that is perfect for the holidays. These particular content variables have inconsistent predictive relations with two or more criteria of cinematic impact. The same content can have a positive impact on one criterion and a negative impact on another criterion. There are four content measures in this category: (a) frightening/tense scenes is a positive predictor of Rotten Tomatoes ratings but a negative predictor of the Kids' Choice Awards; (b) guns/weapons is a positive predictor of IMDb consumer ratings and the Academy Awards but a negative predictor of Teen Choice Awards and box office profit; (c) smoking is a positive predictor Metacritic and Rotten Tomatoes ratings but a negative predictor of opening weekend and domestic gross; and, most strikingly; (d) sex/nudity is a positive predictor of the Kids' Choice and Teen Choice Awards but a negative predictor of the Academy Awards and the Metacritic and Rotten Tomatoes ratings. The ambivalent impact of sex/nudity must really put a filmmaker in a

quandary if the aim is truly to produce a film "for the whole family." Soften this content, and the children and teenagers may respond negatively and the adults positively, but crank it up, and reverse reaction unfolds. No wonder, then, that both websites and v-chips are devoted to helping parents keep sex and nudity out of their living rooms!

4. *Multiple and consistent effects*—The last category is the most helpful: mature content that has the same positive impact on all impact criteria with which significant predictions were identified. From the perspective of the push-the-envelope versus tone-it-down issue, the former more-is-better option wins. Never is there an adverse cost of pushing-the-envelope, so the filmmaker is advised to put as much strong content possible within the confines of the MPAA rating (G, PG, or PG-13) and the genre (animation versus life action). This filmmaking advice applies to three variables: (a) jump scenes for the IMDb ratings, domestic gross, and the Academy Awards; (b) topics to talk about for the Metacritic, Rotten Tomatoes, and IMDb ratings as well as the Academy Awards; and (c) violence for opening weekend and domestic gross earnings, as well as the Teen Choice Awards. Increase the presence and intensity of jump scenes, topics to talk about, and violence, and the result is family film that pleases almost everyone except the children, but without alienating the latter. Moreover, the latter can be appeased by introducing (sub-R level) blood/gore and inappropriate music, two content variables in the second category. That combination of pushing-the-envelope factors—jump scenes, topics to talk about, violence, blood/gore, and inappropriate music—will produce a family film with the broadest appeal with respect to critics, consumers, awards, and box office.

Comparisons

The above results can be directly compared to those reported in Cerridwen and Simonton (2009) because the latter investigators also used the Screen It! ratings of mature content to predict box office, critical evaluations, and movie awards, albeit in the latter case the Oscars and the Golden Globes. Here we concentrate on the standardized regression coefficients presented in their Table 2 (Cerridwen & Simonton, 2009, p. 205). In addition, we focus on those impact criteria that are shared with those used in the current study, namely, domestic box office, Metacritic ratings, and the Academy Awards. Thus, both independent and dependent variables are identical, the only difference being that the earlier study used a far more inclusive sample of films (N = 914 films released between 2001 and 2005) only a small subset of which included the family films studied here (N = 220 films released between 1996 and 2009). That in mind, we note the following:

First, for family films, the Metacritic ratings were a positive function of smoking and topics to talk about, but a negative function of imitative behavior and

sex/nudity. For the more inclusive films in Cerridwen and Simonton (2009), the same criterion was a positive function of frightening/tense scenes, smoking, and topics to talk about, but a negative function of imitative behavior, inappropriate music, jump scenes, profanity, and violence. Hence, the two studies agree on the positive impact of topics to talk about and the negative impact of imitative behavior, but totally disagree on the impact of smoking, which has a positive influence for family films but a negative influence for films in general. Nonetheless, because the mean score for smoking in family films is only 1.58 on a 6-point scale, we are not talking about the same intensity of mature content.

Second, for the family film, domestic gross was a positive function of violence and jump scenes but a negative function of smoking. In contrast, for the broader sample in Cerridwen and Simonton (2009), the same financial criterion was positively influenced by blood/gore, imitative behavior, inappropriate music, and scary/tense music, but negatively influenced by smoking. Clearly, smoking again provides a central content predictor, but in this case its consequence is uniformly negative. Cigarettes, cigars, and pipes do not encourage ticket sales regardless of the type of film being shown at the local theater.

Third, family films that have a higher likelihood of receiving Oscar awards and nominations are prone to contain alcohol/drugs, guns/weapons, jump scenes, and topics to talk about, but are less prone to contain sex/nudity. However, when the Academy voters are judging all eligible films, they are positively affected by frightening/tense scenes and smoking but negatively affected by imitative behavior, jump scenes, scary/tense music, and sex/nudity (Cerridwen & Simonton, 2009). Now it is the last content variable that emerges as the consistent predictor—negative in both samples. In comparison, the disagreement about smoking is striking.

All told, the contrasting prediction results indicate that family films can be easily differentiated from films in general. Even though certain content predictions replicate across the two samples (e.g., the negative consequence of sex/nudity on Oscar recognition), most of the predictions go their separate ways. The contrasts with respect to smoking are perhaps the most intriguing because this content attribute so often appears in the prediction equations.

Implications

This investigation was only preliminary. Not only have few studies examined the relation between content and impact, but also this investigation was the very first to concentrate on family films. Furthermore, the amount of variance explained never exceeds 35%, and can dip as low as 13%—for box office profit, a very critical variable! No doubt many crucial variables have been left out of the equations (for inventory, see Simonton, 2011). Besides being necessary to boost the predictability of any given criterion, these missing variables may also serve

as moderating variables between content and impact. Just to provide the most obvious example, except for family film and animation, we did not specifically examine genre. Yet, certainly, the effects of some content factors on a given cinematic criterion would depend on, say, whether the film was drama or comedy, science fiction or fantasy adventure.

Beyond the implications for future research on film, and especially family film, we might also ask whether this investigation has any relevance for any of the other arts. Here the potential implications are threefold. First, whenever a particular art form has more than one standard of impact, then content predictors of artistic success can (probably will) diverge across criteria. Thus, novels also may be judged as business (best-seller lists) or as art (critical acclaim and book awards; cf. Harvey, 1953). Second, where appropriate, it would be useful to scrutinize how the specific content indicators used here predict impact in comparable genres. For instance, perhaps mature content has some predictive value with respect to the relative popularity of the plays attributed to William Shakespeare (cf. Simonton, 2009c). Third and last, one must ask what kinds of contrasts appear in art forms that, like film, feature both family and mature genres. For example, presumably mature content measures would display contrasting consequences for Nancy Drew versus Agatha Christie detective fiction. The artistic tension between toning it down versus pushing the envelope might be even more conspicuous for storybooks designed for parents to read to their children. In short, the results reported here can serve to inspire research in other art forms, especially in literary genre that can share many of the same mature content.

ACKNOWLEDGMENT

The authors would like to thank Even Braudaway for assistance with data retrieval.

REFERENCES

Academy of Motion Picture Arts and Sciences. The Official Academy Awards® Database. Retrieved April 21, 2010, from http://www.oscars.org/awardsdatabase/index.html

Arnheim, R. (1957). *Film as art*. Berkeley: University of California Press.

Baer, J., Kaufman, J. C., & Riggs, M. (2009). Rater-domain interactions in the Consensual Assessment Technique. *International Journal of Creativity and Problem Solving, 19,* 87-92.

Basuroy, S., Chatterjee, S., & Ravid, S. A. (2003). How critical are critical reviews? The box office effects of film critics, star power, and budgets. *Journal of Marketing, 67*, 103-117.

Beckwith, D. C. (2009). Values of protagonists in best pictures and blockbusters: Implications for marketing. *Psychology and Marketing, 26*, 445-469.

Cattani, G., & Ferriani, S. (2008). A core-periphery perspective on individual creative performance: Social networks and cinematic achievements in the Hollywood film industry. *Organization Science, 19*, 824-844.

Cerridwen, A., & Simonton, D. K. (2009). Sex doesn't sell—nor impress: Content, box office, critics, and awards in mainstream cinema. *Psychology of Aesthetics, Creativity, and the Arts, 3*, 200-210.

Chang, B.-H., & Ki, E.-J. (2005). Devising a practical model for predicting theatrical movie success: Focusing on the experience good property. *Journal of Media Economics, 18,* 247-269.

Collins, A., & Hand, C. (2006). Vote clustering in tournaments: What can Oscar tell us? *Creativity Research Journal, 18,* 427-434.

Deuchert, E., Adjamah, K., & Pauly, F. (2005). For Oscar glory or Oscar money? Academy Awards and movie success. *Journal of Cultural Economics, 29,* 159-176.

De Vany, A. (2004). *Hollywood economics*. London and New York: Routledge.

De Vany, A., & Walls, W. D. (1999). Uncertainty in the movie industry: Does star power reduce the terror of the box office? *Journal of Cultural Economics, 23*, 285-318.

De Vany, A., & Walls, W. D. (2002). Does Hollywood make too many R-rated movies? Risk, stochastic dominance, and the illusion of expectation. *Journal of Business, 75*, 425-451.

Faulkner, R. R., & Anderson, A. B. (1987). Short-term projects and emergent careers: Evidence from Hollywood. *American Journal of Sociology, 92*, 879-909.

Ginsburgh, V. (2003). Awards, success and aesthetic quality in the arts. *Journal of Economic Perspectives, 17*, 99-111.

Hadida, A. L. (2008). Motion picture performance: A review and research agenda. *International Journal of Management Reviews, 11*, 297-335.

Harvey, J. (1953). The content characteristics of best-selling novels. *Public Opinion Quarterly, 17*, 91-114.

Holbrook, M. B. (1999). Popular appeal versus expert judgments of motion pictures. *Journal of Consumer Research, 26*, 144-155.

Internet Movie Database. Retrieved March 17, 2010 from http://IMDb.com/

Kaufman, J. C., Baer, J., & Cole, J. C. (2009). Expertise, domains, and the Consensual Assessment Technique. *Journal of Creative Behavior, 43,* 223-233.

Kaufman, J. C., Baer, J., Cole, J. C., & Sexton, J. D. (2008). A comparison of expert and nonexpert raters using the Consensual Assessment Technique. *Creativity Research Journal, 20*, 171-178.

Kaufman, J. C., Gentile, C. A., & Baer, J. (2005). Do gifted student writers and creative writing experts rate creativity the same way? *Gifted Child Quarterly, 49,* 260-265.

Kids in Mind. Available online at www.kids-in-mind.com

Lord, F. M. (1958). The utilization of unreliable difference scores. *Journal of Educational Psychology, 49*, 150-152.

Maassen, G. H., & Bakker, A. B. (2001). Suppressor variables in path models. *Sociological Methods & Research, 30*, 241-270.

Metacritic. Retrieved April 16, 2010 from http://www.metacritic.com

Münsterberg, H. (1916). *The photoplay: A psychological study*. New York: Appleton.

Pardoe, I., & Simonton, D. K. (2008). Applying discrete choice models to predict Academy Award winners. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 171*, 375-394.

Plucker, J. A., Holden, J., & Neustadter, D. (2008). The criterion problem and creativity in film: Psychometric characteristics of various measures. *Psychology of Aesthetics, Creativity, and the Arts, 2*, 190-196.

Plucker, J. A., Kaufman, J. C., Temple, J. S., & Qian, M. (2009). Do experts and novices evaluate movies the same way? *Psychology and Marketing, 26,* 470-478.

Prag, J., & Casavant, J. (1994). An empirical study of the determinants of revenues and marketing expenditures in the motion picture industry. *Journal of Cultural Economics, 18*, 217-235.

Ravid, S. A., & Basuroy, S. (2004). Managerial objectives, the R-rating puzzle, and the production of violent films. *Journal of Business, 77*, S155-S192.

Screen It! Retrieved May 18, 2010 from http://www.screenit.com

Sedgwick, J., & Pokorny, M. (1999). Movie stars and the distribution of financially successful films in the motion picture industry. *Journal of Cultural Economics, 23*, 319-323.

Simonton, D. K. (2002). Collaborative aesthetics in the feature film: Cinematic components predicting the differential impact of 2,323 Oscar-nominated movies. *Empirical Studies of the Arts, 20*, 115-125.

Simonton, D. K. (2004a). Film awards as indicators of cinematic creativity and achievement: A quantitative comparison of the Oscars and six alternatives. *Creativity Research Journal, 16*, 163-172.

Simonton, D. K. (2004b). Group artistic creativity: Creative clusters and cinematic success in 1,327 feature films. *Journal of Applied Social Psychology, 34*, 1494-1520.

Simonton, D. K. (2005a). Cinematic creativity and production budgets: Does money make the movie? *Journal of Creative Behavior, 39*, 1-15.

Simonton, D. K. (2005b). Film as art versus film as business: Differential correlates of screenplay characteristics. *Empirical Studies of the Arts, 23*, 93-117.

Simonton, D. K. (2009a). Cinematic success, aesthetics, and economics: An exploratory recursive model. *Psychology of Aesthetics, Creativity, and the Arts, 3*, 128-138.

Simonton, D. K. (2009b). Cinematic success criteria and their predictors: The art and business of the film industry. *Psychology and Marketing, 26*, 400-420.

Simonton, D. K. (2009c). The literary genius of William Shakespeare: Empirical lessons drawn from his dramatic and poetic creativity. In S. B. Kaufman & J. C. Kaufman (Eds.), *The psychology of creative writing* (pp. 131-145). New York: Cambridge University Press.

Simonton, D. K. (2011). *Great flicks: Scientific studies of cinematic creativity and aesthetics*. New York: Oxford University Press.

Sochay, S. (1994). Predicting the performance of motion pictures. *Journal of Media Economics, 7*, 1-20.

Thompson, K. M., & Yokota, F. (2004). Violence, sex, and profanity in films: Correlation of movie ratings with content. *Medscape General Medicine, 6*, 3-54. Retrieved February 8, 2008, from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1435631

Vitouch, O. (2001). When your ear sets the stage: Musical context effects in film perception. *Psychology of Music, 29*, 70-83.

Zickar, M. J., & Slaughter, J. E. (1999). Examining creative performance over time using hierarchical linear modeling: An illustration using film directors. *Human Performance, 12*, 211-230.

Direct reprint requests to:

Dean Keith Simonton
Department of Psychology
One Shields Avenue
University of California
Davis, CA 95616-8686
e-mail: dksimonton@ucdavis.edu