12

Entertainment Communication Decisions, Episode 2: "Earned" Channels

"Earned" media is actually a hodgepodge of quite different kinds of communication about an entertainment product. It encompasses the word of mouth that consumers articulate via various channels, quality signals from consumer "herds" (as reflected by the chart position of a new product—because what sells must be good, right?), and the buzz a product receives on the Internet and elsewhere. Consumers' judgments and behaviors are also the input that automated recommendation engines transform into personalized predictions (i.e., algorithm-generated "earned" media). Beyond consumers, there are other stakeholders of an entertainment product whose evaluations feed "earned" communication channels, such as those critics who judge a new product's quality in their reviews and those who hand out awards for entertainment products that they consider to be of outstanding quality.

The common element of all these kinds of communication is that they cannot be controlled by the producer, or at least to a much lesser degree than paid and owned media. Whether consumers share their opinions about a new product, whether buzz develops, whether critics decide to write reviews (and what they write), and what the members of award committees see in the product, all of these are outside the realm of the producer. Of course, producers *try* to stimulate and steer such communication through targeted advertising, by offering press screenings, or by sending out samples (or by not doing so); we will discuss the effectiveness of such steps in this chapter (along with some more dubious ones). But despite all such efforts, it is the others outside the firm who eventually decide whether an

entertainment product "earns" such communication, just as these others decide the valence of such communication. That is the essence of most kinds of "earned" media: if the stakeholders love your product, you will *earn* their praise, but if they do not like it, you might *earn* harsh criticism instead.

In what follows, we will discuss what *Entertainment Science* scholars can tell us about the mechanisms of each kind of earned media and what we know about their respective roles for entertainment product success. We begin with the different ways consumers communicate about entertainment. One important way is definitely word of mouth. We refer to the word of mouth that is exchanged among consumers about entertainment products as "informed cascades": those consumers who talk or write about a product make an explicit "informed" assessment of a entertainment product's quality, based on their own experience with the product. (Please note that this implies that "chatter" about a product by someone who has not yet seen, read, played, or listened to it does *not* fall into our definition of word of mouth. We get back to this in a moment.)

When consumers use the success of a product as their choice criterion, we refer to such herding behavior as "uninformed cascades" because high sales only tell us that a lot of people have bought a product, but not whether they actually *liked* it. The same logic applies to the consumer buzz that exists for an entertainment product prior to its release: none of those who sent the 700,000 tweets about Jurassic World in the week before its release had actually seen the film, nor had any of the 6.6 million who had become a "fan" of the film on Facebook by then. (This includes the "chatter" we have spoken of above, by the way.) We then discuss automated recommendation systems developed from consumer data and, finally, other stakeholders' communication about entertainment: what is it worth to receive a high "Tomatometer" rating, or to win an Oscar?

Informed Cascades: The Power of Word of Mouth

> "[He] only goes to the movies when at least five people in whom he has complete confidence have recommended the film to him as worth seeing."
> —*Nobel prize winner* Heinrich Böll *(1963, p. 226) describing his fictitious character* Leo Schnier *in the novel The Clown [Translated from the German original by* Leila Vennewitz. *Courtesy of Melville House Publishing]*

Marketers have long considered word of mouth (which we will refer to as WOM on the next pages) to be a powerful source of influence on the behavior

of others. And scholars have agreed, after Johan Arndt's (1967) seminal scholarly article added the phenomenon to their theories. However, until recently WOM has been widely considered the "mysterious force" that Arndt (1967, p. 291) named it, rather than something that can be systematically researched and understood. The main problem was that WOM was mostly invisible to the outside observer, as it is shared between consumers in personal conversation, and surveys shed relatively little light on its nature and effects. But with the rise of digitally mediated communication, things have changed quite fundamentally. On the Internet and in social media, WOM is not only visible, but can also be tracked in detail by managers and scholars alike, both on an aggregate and an individual level. As a consequence, WOM has almost overnight become one of the most intensively studied topics in marketing—and *Entertainment Science* in particular.[310]

So, what exactly is WOM after all? Because the concept is so intuitive, an unintended result is a lack of precision when it comes to defining it. Because WOM is all about supposedly objective quality judgments, it is important to limit the concept to personal communication by consumers about a new product *which they have already consumed*. We insist that WOM should not be confused with the behaviors that constitute (pre-release) buzz, as there is a huge difference between WOM's experience-based communication versus chatting about a product which one has not yet consumed. All pre-release communication is necessarily anticipatory and speculative and, thus, lacks the experience component which is so crucial for WOM.[311] It is also this experiential nature which makes the cascades that WOM triggers conceptually different from other cascades that we discuss in this book (De Vany and Lee 2001): only WOM cascades are based on consumers' "true" quality perceptions (which makes them "informed" or "quality-based" cascades), not inferred from the *actions* of other consumers (as is the case for "uninformed" or "action-based" cascades).

---

[310]In their meta-analyses of WOM effects, You et al. (2015) included 51 empirical articles that link (Internet) WOM with product success, and Rosario et al. (2016) compile close to 100 (!) studies from between 2004 and 2014.

[311]Maybe the words of those who have gained access to a new product prior to its release via illegal sources mark an exception, but we ignore them here—we have dedicated a whole section of our book to their actions in our chapter on entertainment distribution.

And what it the logic that spurs the idea that WOM is influential and powerful? WOM's main "power source" is that it can initiate cascades. Remember the crucial role of personal recommendations in the classic diffusion model by Frank Bass. The recommendations by consumers who have already experienced the product ("Innovators") influence the adoption decisions of other consumers ("Imitators") who can then, after experiencing the product on their own, affect still others. All these recommendations are a central element of the WOM concept.

Let us add that scholars sometimes study the mere amount ("volume") of WOM for a product, in large part because of the ready availability of this type of WOM data from online product reviews posted by consumers on websites, such as Amazon. Such volume information, if studied in isolation, is conceptually different from the *valenced* opinion that triggers informed WOM-based cascades. Volume information does not reveal whether consumers *like* a certain product, only how much *awareness* the product has received among them—it thus falls into the "uninformed" category, very similar to success-related information such as charts. Consequently, our discussion in this section focuses on the valence element of WOM.[312]

In what follows, we will overview what scholars have learned regarding the important, but complex, role of such WOM valence for consumer decisions and entertainment product success. But before we do so, let us take a quick look at the drivers of WOM: what makes a consumer engage in WOM and share his or her thoughts and feelings about entertainment product with friends and others?

What Makes Us Articulate Word of Mouth?

As consumers, we have dozens of consumption experiences every single day, but we only spread WOM about a few of them. Why for some products and not for others? Whereas "inner forces" motivate us to publicly express our experiences with products and services, we have to hold product factors responsible for our selection process for articulating WOM.

Scholars have argued that WOM is "goal-driven." Based in part on an empirical investigation of more than 2,000 online community members we conducted (Hennig-Thurau et al. 2004), the following six basic motivations,

---

[312]We will get back to volume-related insights in our discussion of "herding" behavior in a few pages.

or psycho-social functions, can be considered as crucial for articulations of WOM regarding entertainment consumption:[313]

- *Impression management.* Consumers engage in WOM to demonstrate their expertise about a product or product category, aiming to enhance their self-worth. Whereas this usually takes the form of positive WOM through which one demonstrates the ability to spend time and money wisely ("This way I can express my joy about a good buy;" Hennig-Thurau et al. 2004), explaining why a popular movie or song does *not* deserve its reputation also falls in this category.
- *Emotion regulation.* Negative entertainment experiences cause negative emotions to accrue, and articulating negative WOM about the responsible product can help a consumer by providing a way to vent those emotions ("I like to get anger off my chest").
- *Concern for other consumers.* Some consumers have a genuine interest in helping others, and WOM is a means to live out such altruistic tendencies. Positive WOM can help others experience joy, while negative WOM can help others avoid hurt or anger ("I want to save others from having the same negative experiences as me").
- *Social bonding.* Whereas the previous motives focus on psychological aspects of a consumer's personality, WOM can also be driven by social motives, serving as a means to address the all-too-human need for social interaction. As consumers, we enjoy talking about a new album, and such conversation can be an integral part of social relationships ("It is fun to communicate this way with other people in the community"). In our study of different Internet platforms, we find social bonding to be the strongest internal driver of writing comments about products, with an effect that is twice as strong as for any other motive.
- *Information acquisition.* WOM can also address consumers' interests in "knowing more." What is the meaning of the TV show's ending last night? How can I master the next level in the new video game? Engaging in WOM can help provide answers to such pressing questions.
- *Persuasion.* And finally, Berger (2014) stresses WOM's role of as a means to convince others to choose one entertainment product (or activity) over alternatives. This is particularly relevant in group consumption settings. When you are going to the movies with friends and want to see the new

---

[313]See also Berger's (2014) summary of WOM research.

Kevin Costner film instead of the next Avengers episode, praising Mr. Costner's previous efforts might be the way to go.

In addition to such inner drivers, scholars have compiled evidence that WOM activities are also influenced by the characteristics of the object of communication, i.e., the product, or, more specifically, the consumer's perception and evaluation of it. Berger and his colleagues have shed light on such external WOM drivers in a number of studies (Berger and Milkman 2012; Berger and Schwartz 2011). Their findings, which are not specific for entertainment, highlight the following characteristics:

- *Interestingness.* The more consumers find a product "interesting," the higher the probability that they will engage in WOM about it. Naturally, what is "interesting" lies in the eye of the beholder; the term encompasses attributes such as novel, exciting, and unusual. It is closely related to the concept of "involvement," which refers to the importance (or personal relevance) of a product as perceived by a consumer (Jain and Srinivasan 1991). Berger and Schwartz (2011) stress that "interestingness" triggers immediate WOM; when we have seen a movie that has fascinated us, we want to talk about it immediately. Entertainment, per se, scores highly in interestingness, compared to other product categories. That is why we talk (and write) so much about it.
- *Surprisingness.* This concept relates to interestingness, or may be a facet of it: when consuming a product, do consumers perceive it as surprising? Surprise sparks WOM even when controlling for interest (Berger and Milkman 2012)—and it does so both immediately after consuming a product *and* later. So we can expect entertainment experiences that surprise us to get extra conversation, probably one of the reasons why The Sixth Sense was such a big hit movie.
- *Positivity.* Although negative events stick deeply in our memory, we prefer to talk about positive experiences with others. Ratings on Internet forums tend to be largely positive, and the same is true for WOM shared on social media. According to Berger and Schwartz (2011), good products receive more WOM recommendations than bad products receive warnings.
- *Arousal/emotionality.* What affects us emotionally, both in positive ways (e.g., excites us) and negative ways (e.g., scares us or makes us cry) spurs a higher level of WOM than what does not do so. This means that strong emotions associated with an entertainment product not only affect its

success directly (via creating anticipatory/anticipated emotions that lead us to consume a product, or by stimulating us to watch a movie more than once), but also indirectly by attracting *other* consumers via WOM cascades.[314]

The film MY BIG FAT GREEK WEDDING addressed several of these aspects and consumers' inner motives. In particular, the emotional reactions of many who had seen the film were quite extreme, and the film thus became a stunning hit mostly through WOM. Its lead actress Nia Vardalos tied WOM to pure chance, quite in line with the "Nobody-Knows-Anything" mantra: "We got lucky. You can't manufacture word of mouth. You can't pay people to tell their 10 cousins" (quoted in Strause 2016).

We kindly object, with all due respect: not only did the film itself provide the content that fueled WOM, but its producers also provided the conditions in which WOM could blossom. They carefully orchestrated cast appearances, packed their screenings, and zoned in on the female audience and the Greek community (e.g., the star traveled extensively making appearances at Greek organizations and bridal shows). All of this, of course, could not *guarantee* that WOM would blossom (it's a probabilistic world after all!), but in entertainment, as in other parts of life, luck favors those who work hard (and know the right things).

Does Word of Mouth Influence Entertainment Product Success? Yes. But It's Complicated

"If the picture is bad, you might as well shoot everybody coming out of the theater—they will quickly enough kill any film."
  —John Friedkin, *former vice president for advertising and promotion at 20th Century Fox (quoted in* Austin 1989, p. 3)

As noted earlier, empirical studies that address the role of WOM for consumer decision making usually distinguish between the valence of the WOM for a product (such as the average rating at Amazon that a game received from consumers) and the volume of such WOM (i.e., the *number*

---

[314]See also this book's section on consumer emotions and their role in the sensations-familiarity framework in the entertainment consumption chapter.

of consumer reviews written about the product).[315] We focus on valence here as it is the valence element that captures how consumers judge the quality of a product—the source of informed cascades. We will take a look at average effects across products first, before diving deeper into contingencies that determine how strongly the success of a particular product is affected by WOM.

Average Effects: Word of Mouth (Valence) Matters!

In their meta-analysis of WOM effects on product sales across products and industries, You et al. (2015) find substantial average effects for WOM valence; they also report that this effect tends to be higher for entertainment products than for "others." *Entertainment Science* scholars, taking a more fine-grained look, have provided evidence that WOM valence is influential for all forms of entertainment that we feature in this book, except for music (something we return to when discussing context effects). Here's a summary of what we know about the role of WOM for the success of books, movies, TV shows, and games.

Books. The seminal study on WOM effects is by Chevalier and Mayzlin (2006), who analyze how consumer reviews on the websites of retailers Amazon.com and Barnes and Noble affect a book's *relative* sales rank (the difference in ranks between the two sites).[316] Using a data set of about 1,100 books with at least one consumer review (which combines a random

---

[315]Let us note that the separation of "valence" and "volume" carries a lot of analytical problems. It mixes the WOM about a product with the product's popularity (which is the source of *un*informed cascades) and also its success (as experiences are required for WOM, more successful products get reviewed more often). Also, valence and volume are systematically inter-related, as the quality of a product is a source of its popularity (the more consumers like a product, the more WOM they will share about it). Scholars have found that WOM valence is a main driver of the amount of WOM for entertainment products; when both "facets" of WOM are included in the same study, WOM volume thus tends to absorb the impact of valence on success (Duan et al. 2008; Karniouchina 2011). Empirical results on WOM valence effects depend strongly on whether a scholar accounts for these problems (e.g., by using instrumental variables for WOM volume and changes in WOM valence ratings over time) and also controls for other drivers of success (such as advertising); they are thus far from consistent across studies (see also for example Forman et al. 2008 and Chintagunta et al. 2010—the latter authors also provide empirical evidence of the consequences of (not) accounting for these aspects). In our coverage of WOM here, we focus on studies that address such challenges in a powerful way.

[316]Note how smart this approach by Judith Chevalier and Dina Mayzlin is: because the books sold at the two sites are identical except for the WOM and some other factors like price (for which the authors control in their analysis), looking at the differences in sales ranks (rather than at absolute sales) eliminates the effect of all product/book characteristics on sales and allows the analytical spotlight to be put on the WOM on the sites.

selection from 1998 to 2002 and bestsellers from 1991 to 2002)[317] and ana-
lyzing rank changes at three points in time in 2003 and 2004 with regres-
sion analysis, the scholars find that a one-star increase in rating at Amazon
corresponds with a 52% increase in rank-difference. Also, if all reviews for
a book give it five stars (versus none do so), the book's relative sales rank
improves by more than 100%. A similar rise in 1-star ratings has an even
stronger sales effect (albeit negative, of course).

Since then, several other scholars have provided additional evidence for
the role of WOM valence for books. Sun (2012), applying the same statis-
tical approach to a more recent sample (892 randomly selected books pub-
lished 2002–2006), also finds that WOM valence is influential. She, though,
reports a smaller improvement in relative ranks—a one-star increase at
Amazon corresponds with "only" a 21% higher relative sales rank. This result
might indicate that the WOM effect weakens in this context, or could simply
be the result of the data set (or of changes in Amazon's assortment). For her
large set of some 30,000 books, Meiseberg (2016) finds both five- and one-
star consumer reviews to influence *absolute* book sales ranks at Amazon. Her
effects are again substantial, but also somewhat weaker than the ones reported
by Chevalier and Mayzlin: a change from zero to 100% 5-star WOM corre-
sponds with an average improvement in sales ranks of about 52%.

Jabr and Zheng (2014), with a GMM approach, estimate that a one-
star improvement in book ratings at Amazon leads to a 26–34% improve-
ment in the book's sales rank. Their data consist of consumer reviews at
Amazon for 1,740 randomly selected (non-fictional) books from 2007 to
2009; each book had at least 25 reviews.[318] This is quite similar to Li and
Hitt's (2008) finding of a 27% increase in (estimated) sales at Amazon in
response to a one-star rise in WOM valence; they estimate this effect for
about 2,600 books published in 2000–2004 using a fixed effects regression.
And when Schmidt-Stölting et al. (2011) link consumer reviews at Amazon.
de to *nationwide* German book sales for their a large data set, they still find
a WOM effect—they estimate that, on average, an increase of one star on
Amazon.de is associated with 4–7% higher nationwide sales.

When judging the size of these effects, keep in mind that, because
WOM is predominantly positive and usually shows relatively little vari-
ation between consumers, a one-star change in reviews is quite enormous.

---

[317]The scholars also conduct analyses with a larger data set of almost 2,400 books, but we focus here on
the (more robust) results for those books which had at least one consumer review at the beginning of
the investigation period.

[318]The 34% improvement is found when the authors use earlier reviews written by a "WOM giver" as a
statistical instrument for that person's WOM valence about a book.

In the Li and Hit data set for example, the average WOM valence is 4, whereas the average deviation from this value is only 0.60—thus, a one-star change would cover about 75% of *all* books in the data set.

Movies. Studies examining movie success that focus on post-release revenues (the time frame in which WOM can actually matter), and thoughtfully account for the role of alternative information sources that are available at this time, also find that experience-based WOM affects product success.

Specifically, Chintagunta et al. (2010) use a data set of 148 movies released in the U.S. around 2004 and analyze box-office results for different geographic areas, not only domestic. Their logic is that WOM spreads over the Internet, so consumers in one location can make use of such information on a movie's release day when the movie is already being shown elsewhere. Using consumer ratings from Yahoo Movies and a large set of controls, such as advertising and distribution, as well as an instrument for the volume of WOM (the two major reasons for biased results in other studies), the scholars estimate a movie's first-day box office in a geographic market, using a GMM approach. They find that a one-unit increase in WOM valence on Yahoo's 1-to-13 rating scale corresponds with a 10% increase in opening-day sales in a specific market.

In a follow-up investigation, the same team of scholars uses a subset of this data for a regression approach in which they study the box office generated during a four-week post-release window (Gopinath et al. 2013). They again account for endogeneity and include several controls. This time the scholars find a WOM valence elasticity for Yahoo consumer ratings of 0.22, which means that a 10% increase in the WOM valence for a film is associated with roughly 2% higher revenues in a movie's first four weeks.

A study by Niraj and Singh (2015) suggests that the role of WOM valence (measured here as a "positivity ratio" of consumer reviews on several websites, portals, and forums) for movie success also exists at the Indian box office. Their investigation, which applies a panel regression to a small data set of 48 Bollywood movies (released in 2010–2011), also provides tentative evidence that the link between WOM valence and the success of movies might not be linear.[319] When they include a squared term of WOM valence, they find that it is negative (and significant). In other words, the value of positive WOM decreases with the number of people who give it—if *everyone* loves a film, that might not necessarily help it. We get back to this idea when discussing the role of the variance in WOM valence in the next section.

---

[319]Take note that there are certain limitations though—their work does not consider a number of key controls (such as advertising and distribution) and also does not account for the endogeneity of WOM.

TV shows. In their survey panel of TV audiences, Lovett and Staelin (2016) find that remembering WOM is closely linked with TV watching: having been exposed to WOM for a show increases the likelihood that a consumer watches the next episode by 6%. This might not sound like much, but it is 25% higher than the same effect the scholars detected for paid advertising, and almost three times as high as the effect of visits to the broadcaster's website (which was the scholars' measure of "owned" communication).

Games. In our own investigation of 100 Xbox 360 games, we find that WOM valence also makes a difference for games (Marchand et al. 2016). In our 3SLS estimation, we control for several other factors, such as advertising, and also address the endogenous nature of WOM volume. The results teach us that the valence of WOM for a game posted on Amazon.com in the weeks after a game's release influences the game's sales quite strongly; the average valence elasticity in our study is 0.47, suggesting that an improvement of 10% in WOM valence translates into 4.5% higher game sales in the following three weeks.

These results provide clear evidence that WOM matters for entertainment success. But all effects we have reported so far are averages, aggregations over heterogeneous sets of products. Let us now see if contingencies exist that alter the role that WOM plays, as we have shown contingencies to do so for advertising, sequels, stars, and other things in this book. We will shed light on a number of such context factors, ranging from product types, to WOM types, to different groups of consumers. Contingencies will also help us to understand why scholars such as Dewan and Ramaprasad (2012) were unable to find support for an "average" effect for the role of WOM in *music* success. Music has a non-verbal, non-visual character and triggers highly subjective consumer judgments, which will limit the effect of verbal recommendations (and warnings) by other consumers for music, in general. But there are indications in Dewan and Ramaprasad's study that the effect exists at least for some types of music.[320]

The Product Type Matters

We have already shown that certain product characteristics, such as the "interestingness" of a product, influence how much WOM is articulated

---

[320]In addition and consistent with this logic, we will also see that "uninformed" action-based cascades, which are intuitive to grasp and aggregate the "judgments" of many people, matter a lot for consumers' music choices.

for a product. Scholars also point out that the *impact* that WOM has on product success varies between types of entertainment products. In particular, they argue that WOM valence plays a more prominent role for the success of smaller "independent" products, whereas it tends to be less influential for high-budgeted "commercial" products.

For games, Zhu and Zhang (2010) find, in an analysis of 141 console games (released on both PS2 and Xbox from 2003 to 2005), that an interaction of the valence of WOM posted on the GameStop retail website with a game's "popularity" is strongly negative: WOM valence plays a stronger role for the demand of less-popular games among U.S. consumers.[321] And when Dewan and Ramaprasad (2012) analyze a data set of 1,762 songs that were posted in MP3 format on music blogs in 2006, their 2SLS estimations show no impact of WOM valence (consumer ratings of the songs at Amazon) on song sales for the whole data set. But when they split their data set, the regression coefficient for WOM valence is then clearly higher for niche songs that were ranked below 5,000 than for more highly ranked songs: for niche songs, a one-unit increase in ratings corresponds to a 20% increase in song sales. Neither parameter reaches statistical significance though.

Finally, the book-related results by Meiseberg (2016) are also in line with the prominent role of WOM for niche titles. She finds that the effect for positive WOM (namely 5-star reviews of books on Amazon in her analysis) is highest for the lowest selling quantile (i.e., the least popular products in her data set). The effect is not linear for the other quantiles, so that her results point to an "awareness effect" of positive WOM that makes books with (good) reviews stand out from the vast number of similar titles (that are not written by star authors). WOM tells us that at least *someone* likes them.[322]

Overall, these empirical insights on differing WOM effects for niche and commercial products correspond with industry wisdom among entertainment managers who often consider smaller movies to be much more "WOM sensitive" than bigger productions. It also provides the basis for

---

[321]Zhu and Zhang measure popularity as the above/below average sales of a game in a respective month, compared to all games in the data set in the same period.

[322]We speculate that this effect is further enhanced by Amazon's search engine, which might put products with positive recommendations in a more prominent place when presenting search results to consumers.

entertainment's two main strategic concepts which we will discuss later in this book: the blockbuster concept (for which WOM plays only a marginal role) and the niche concept (for which WOM is essential).[323]

Not All Word of Mouth is Created Equal

"Brüno's box office decline from Friday to Saturday indicates that…[it] could be the first movie defeated by the Twitter effect."
    —Corliss *(2009)*

With the rise of the Internet, WOM has not only become observable, but also increasingly complex and heterogeneous. The empirical studies we listed above all use consumer reviews on Amazon and similar sites, but this particular kind of WOM functions differently than other kinds of personal communication, and thus also affects product success differently. Today, three main kinds of WOM co-exist, each with unique characteristics, and are of particular interest for entertainment managers (Hennig-Thurau et al. 2015). Figure 12.1 overviews what they have in common and what sets them apart.

Traditional (or offline) word of mouth (a.k.a. TWOM) is the face-to-face communication between consumers about a product, based on personal experience with the product. TWOM's main characteristics are that it is shared with an individual consumer or a small group, is transmitted in real time, assumes a personal connection between those who exchange it, and that it enables feedback, combining a "push" element (information whose transmission is initiated by the sender) and a "pull" element (information that is requested by the receiving consumer).

Electronic (or online) word of mouth (or EWOM) is experience-based communication that a consumer makes available to a potentially very large group of anonymous others over the Internet, on forums such as Yahoo Movies or sites such as Amazon. Traditionally the main source for scholars when analyzing WOM effects, EWOM is "pull-only"—we have to actively search for information on a blog or review, the blog does not provide us with information on its own. It is also asynchronous and allows no (or very limited) feedback—but it is often stored for a long time. EWOM also usually offers "summary statistics," such as average ratings and the number of those who have rated a product.

---

[323]We discuss the two strategic concepts and the role that WOM plays for each of them in detail in our chapter on integrated entertainment marketing.
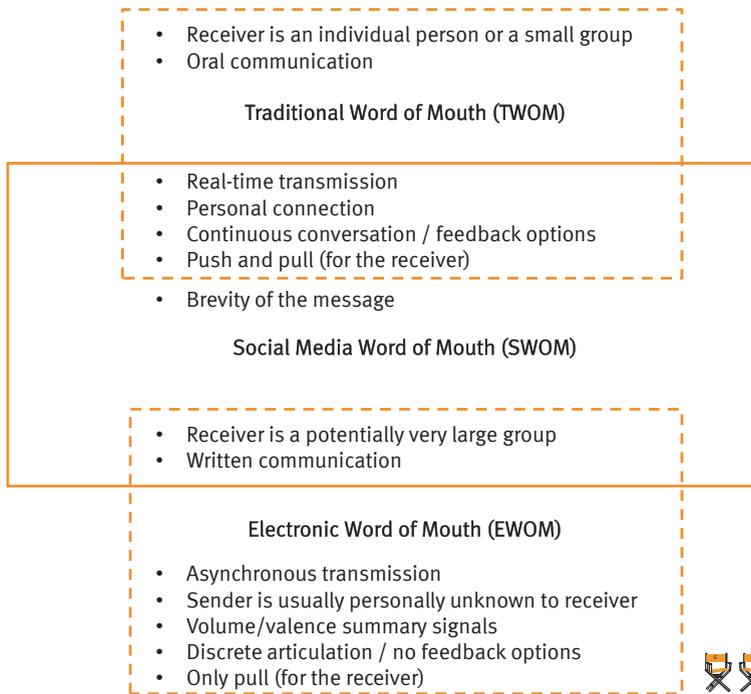
Fig. 12.1 Three major kinds of word of mouth in the digital age
Note: Reprinted with minor modifications with permission by the Journal of the Academy of Marketing Science/Springer from Hennig-Thurau et al. (2015, p. 389).

Finally, *social media (or microblogging) word of mouth* (SWOM) includes statements about a product experience that are broadcast to the sender's social network (select members or all) through a specific web-based service such as Twitter or Facebook. SWOM combines elements of both TWOM and EWOM—like TWOM, it involves a real-time exchange that combines "push" and "pull," a personal connection between sender and receiver, and the chance to provide feedback as part of ongoing exchanges), and just as EWOM, it can reach a very large group of potential receivers. But SWOM is not just a mixture of the two, but adds its own characteristics, namely a medium-specific brevity because of technical or usage-based restraints (nobody reads epic-length rants on Facebook). The unique brew of these elements allows consumers to push their evaluations almost instantly to very large numbers of related others—the basis of the so-called "Twitter effect" that we discuss below.

The distinction of these kinds of WOM is not purely abstract and conceptual, but has manifest consequences for which *Entertainment Science*

scholars have compiled initial evidence. In their analysis of movie success, Shyam Gopinath and his colleagues (2013) included a measure of SWOM next to their EWOM data from Yahoo.[324] They find an SWOM-valence elasticity of 0.35, which suggests that a 10% improvement in SWOM valence for a film is linked to about 3% higher revenues in the month after the movie's release. This effect is higher than the effect for EWOM, but more importantly, SWOM and EWOM are both significant influencers of movie success when analyzed *jointly*. The WOM types *carry unique information* and/or *reach unique customer segments*.

In our own study of 100 console games (Marchand et al. 2016), we also included SWOM (expressed in the tweets about each game) in addition to EWOM (via Amazon). Again, the two kinds of WOM work differently: in contrast to the substantial effects of EWOM valence in the weeks after a game's release, the valence of tweets does not significantly link to game sales. Our conclusion: whereas social media communication is best suited to transmit "social" information (such as excitement), consumer reviews are most effective for providing information about a product's performance.[325]

But the crucial time for SWOM is not weeks or even months after a product's release, but way earlier: its real-time character, in conjunction with its ability to push information on the smartphones of large consumer groups, can affect the diffusion of a product *very early* when the impact of both TWOM and EWOM is still systematically limited. In a separate study (Hennig-Thurau et al. 2015), we investigated how SWOM via Twitter affects a film's destiny in the *hours and days* after it has been launched—something journalists and entertainment managers have labeled the "Twitter effect" (see for example the introductory quote by Corliss 2009).

To do so, we collected all four million tweets posted by consumers during the North American opening weekends of 105 wide-release movies between October 2009 and October 2010 and linked them to the movies' daily box office during their initial weekend. Combining the manual coding of 51,000 tweets and machine learning, we separated *evaluative* SWOM from *anticipatory* chatter; Fig. 12.2 shows their respective distribution over the movies' first three days in release. We then ran an OLS regression in which we explained the box office distribution over the first three days (i.e., the drop

---

[324]Actually, Gopinath et al. do not measure SWOM in its raw form, but data from Google Blogs, which share several features with social media. They determine the valence of blog comments with human coders.

[325]We also probed for interaction effects between the two kinds of WOM, but found none.
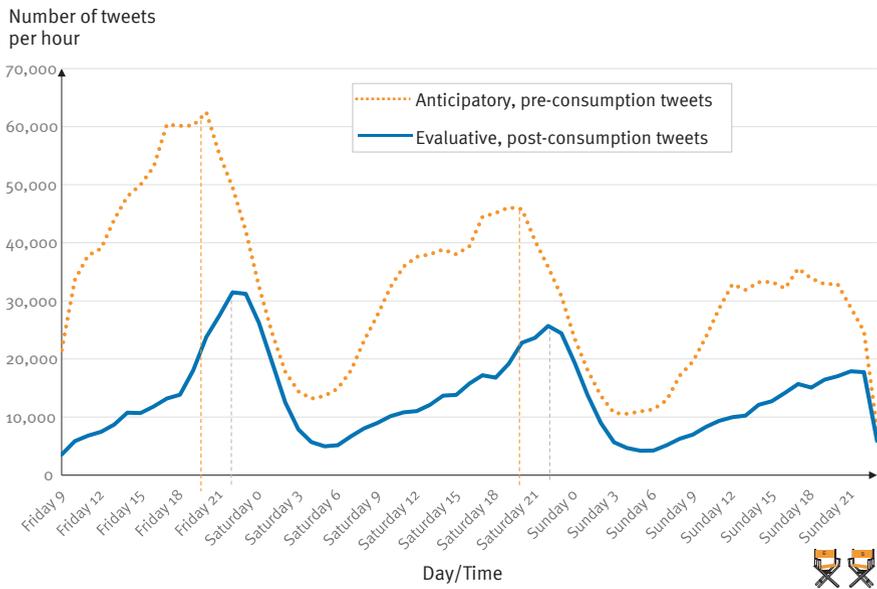
Number of tweets
per hour

Fig. 12.2 When people tweet about movies during their opening weekend
Note: Reprinted with minor modifications with permission by the Journal of the Academy of Marketing
Science/Springer from Hennig-Thurau et al. (2015, p. 382).

or increase on Saturday and Sunday)[326] with (1) the share of moviegoers who
tweeted positively valenced comments about a film after having seen a film
on its first Friday and (2) the share of opening-day moviegoers who tweeted
negatively about their experience; we also included the sheer volume of open-
ing-day SWOM and controlled for multiple other "success factors."[327]

Our results show that the "Twitter effect" does indeed exist. Interestingly,
it is much stronger for *negative* tweets that discourage consumers from
going to see a film on Saturday and Sunday. The parameter for positive
tweets (which are sent *much* more frequently than are negative ones) is only
one-tenth of the one for negative tweets and not statistically significant.
Simulations show that consequences of negative SWOM can be quite sub-
stantial: if the share of negative tweets was 24% for all films (the maximum
among the films in our data set), the average opening weekend box office

---

[326]Specifically, we measured the share of opening weekend revenues a film accrued on the Saturday and
Sunday that followed its opening day (Friday).

[327]Those controls included a movie's ad spending and the pre-release buzz it received. Besides, the
results remained largely the same when we used the sales as our dependent variable, but such a model
carries some additional econometric challenges.

per film would have been 15% lower—the equivalent of $0.5 million. And the *absence* of negative tweets for all films would have increased the average opening weekend box office by more than 4%, or $1 million per film. A film such as the 2010 remake of horror classic Nightmare on Elm Street would have lost almost $5 million of its $15.7 million first-weekend revenues if it had a 24% negative-tweets share, but it would have gained about $3 million if negative tweets were completely absent on its opening day.

Whereas the typology of WOM types helps us to understand their respective effects, differences in WOM effects can even exist *within* a single kind of WOM across platforms. This is also what Chevalier and Mayzlin (2006) find in their study of EWOM effects on book sales: EWOM at bn.com had a far weaker effect than EWOM at Amazon.com. A one-star increase at bn.com corresponded with a 16% improvement versus the 52% the scholars estimated for Amazon, and changes in the percentages of 5-star and 1-star ratings are also clearly less strong. So we recommend entertainment managers to pick their WOM data source very carefully.

Not All Word-of-Mouth Givers are Equal, Too…

There is evidence that, for WOM to have an impact, it also matters *who* shares thoughts and feelings about an entertainment product—and what receivers know about him or her.

With regard to personal communication via TWOM, the idea that some consumers are "opinion leaders" that have a stronger influence on others is a central element of Katz and Lazarsfeld's (1955) "two-step flow" theory of communication. Accordingly, opinion leaders are more exposed to media, and their acquired expertise serves as the basis for their influence on less-informed consumers. Other scholars have stressed the critical role of product-specific motivations and involvement for such opinion leadership (such as for cars or video games). Whereas the concept's origin was in politics, Katz and Lazarsfeld themselves provided evidence that opinion leaders exist in different areas of life—including movie-going!

As EWOM usually does not involve a personal connection between giver and receiver, what makes someone an opinion leader in this context? Meiseberg (2016) finds in her large book data set that a review by a "Top 500 reviewer," a status given by Amazon, matters, above and beyond the "normal" WOM valence that exists for a book. Her results do not take into account the valence of the Top reviewer's rating, though, only its mere existence.

In addition, Forman et al. (2008) show that in an environment characterized by anonymity, it can make that EWOM giver's judgment more influential on the receiver's subsequent choices when a receiver learns more about the EWOM giver's expertise and, in turn, builds a sense of closeness with him or her. The scholars learn this from a panel regression of monthly sales ranks of 768 books (which had been a bestseller at Amazon in 2005–2006), which indicates that the more the author of a review at Amazon.com reveals about his or her identity, the more influential the reviews are. The coefficient for such "identity disclosure" suggests that if the percentage of WOM givers for a book who disclose their identity increases by 50%, this would translate into a 15% improvement in sales rank.

The Timing of Word of Mouth

The role of WOM for success also varies over an entertainment product's life cycle. We have reported that the elasticity of EWOM valence for video games is 0.47 in our study (Marchand et al. 2016)—but this effect is actually an average over the whole first 9 weeks of our investigation. A look at distinct time windows shows that the elasticity can differ quite substantially over time. For the EWOM that is articulated 7–9 weeks after a game's release, the valence elasticity is as high as 1.11. Specifically, our "rolling windows" approach, in which we estimated effects for different windows of three subsequent weeks, shows that the effect of EWOM valence *grows* in importance over a product's life cycle. The EWOM parameter only becomes significant about seven weeks after the release.

Why these differences over time? The reliability of WOM valence grows over time with the number of reviews, resulting in more stable averages which offer more powerful insights for consumers. The long-term availability of digital WOM can play to the late strength of this type of information; as time goes on, the availability of alternate information sources (advertising, professional reviews, charts, and buzz) becomes clearly limited, increasing the *relative* importance of EWOM. But we also need to keep in mind that many more copies of a product are sold in the earlier windows, so that higher elasticities do not necessarily mean higher absolute sales volumes as a result of WOM, but only higher *changes* in sales.

Another study that compares the role of EWOM at different points in time is the one by Schmidt-Stölting et al. (2011). In their large-scale analysis of determinants of book sales, the scholars find a smaller effect of EWOM valence for hardcover titles than for paperback versions of the same books

(which are usually published several months after the hardcover versions). This result again stresses the role of information availability and the "costs" that accrue for consumers who search for them at a given point in time: because more people have already read a book by the time it is released in paperback, more (and more reliable) WOM should be available about a book compared to when it exists only in hardcover format.

When Consumers Have Different Views: The Role of Word-of-Mouth Variance

Finally, when consumers hear about a new entertainment product's quality from many others, the assessments can differ widely across people. Such variation is not captured by the average WOM valence: for example, a mean score of 4 stars can be the result of ten 4-star ratings, of five 5-star and five 3-star ratings, or of six 5-star, two 4-star, and two 1-star ratings. *Entertainment Science* scholars have also looked whether such variance of WOM valence is informative on its own for consumers, and whether high variance is perceived by consumers as positive or negative information.

In their study of book sales, Jabr and Zheng (2014) find that the variance of Amazon ratings is influential, but in a negative way, with higher variance *reducing* a book's success potential.[328] Sun's (2012) results for books confirm an influence of WOM variance, but also suggest that a more complicated mechanism is at place. Adding an interaction term between WOM valence and variance shows that variance can act in a negative, but also a positive way, depending on a book's WOM valence. If a book's average WOM valence is highly positive, then high variance is not helpful; instead, it discourages consumers from buying a book, as dissonant voices confuse rather than stir interest in a book.

But if the WOM valence for a book is low, a higher level of variance across consumers' Amazon ratings is a positive thing, resulting in more sales. In this constellation, variance might suggest that there is something appealing "hidden" in the book which some reviewers do not see. In Sun's data set, the threshold value is a rating of 4.1, which means that 35% of the books in her data would benefit from more variance—but the majority would suffer from it. Eventually, Karniuchina's (2011) movie-related results point to another potentially positive

---

[328]Because the independent variables are neither log-transformed nor standardized in Jabr and Zheng's study, their effects cannot be compared.

effect of variance in ratings: variance can stimulate discussions among consumers, contributing to debates and higher interest and buzz.

In summary, we have seen that informed cascades, via WOM, can have a quite substantive impact on the success of entertainment products, and we have laid out the conditions under which it matters more and those under which WOM is less influential. Let us shift the focus now to "uninformed," or "action-based" cascades, where consumers infer quality levels from the actions, rather than the articulated judgments, of other consumers.

Uninformed Cascades: The Power of Herds

"'Eat sh#t, a hundred billion flies can't be wrong,' the old graffito used to say. 'Follow Stephen, two million tweeters can't be wrong,' I say."
  —*Novelist* Stephen Fry *(2010)*

Even when we, as consumers of entertainment, are not actively talking with others or reading what they have written, we are often still observing what they are doing. What we observe informs our own decision making through what psychologists have named "observational learning" or "social learning" (e.g., Bandura 1977). Through this specific learning mechanism, an action of one person can initiate the actions of others who observe it, who, in turn, affect others, setting into motion so-called "action-based" cascades. Because the information garnered from observation is less rich compared to what can be learned from personal communication, another term for such cascades is "uninformed" cascades. Economists have compared such behavior to that of herds ("herding"), in which members instinctively follow the actions of other herd members without deeper reflection of the logic or wisdom underlying these actions (e.g., Bikhchandani et al. 1998).

In this section, we study two different, but related types of uninformed cascades that are relevant for the success of entertainment products. The first type happens when the adoption of a product by other consumers serves as a signal of the product's quality; here the success of the new product is made visible to others via bestseller or Top Ten lists. This "social proof" increases the desirability of the product to others (through mechanisms we will discuss below) and, thereby, initiate self-enhancing "success-breeds-success" effects. The second type of uninformed cascade takes place even earlier: when the "buzz" about a forthcoming entertainment product is interpreted as a quality signal, enticing others to join the "buzz train" and to eagerly anticipate adopting the product as soon as it is available.

Post-Release Action-Based Cascades: When Entertainment Success Breeds Entertainment Success

Some Words (and Numbers) on the Mechanisms at Work

The information that a product is popular or successful influences its success via two psychological routes. The first mechanism is that popularity biases consumers' perception of the product's quality. Remember that judging the quality of entertainment is notoriously difficult for consumers for several reasons, including the lack of search qualities, and the hedonic character of entertainment which puts emphasis on subjective, holistic aesthetic and artistic achievements rather than objective functional attributes. And artistic taste standards imply that a judgment depends also on one's own, idiosyncratic cultural capital, making the judgment task even more challenging.

Because making a quality judgment is so complicated, an entertainment product's popularity sends a clear and straightforward signal to consumers that is also easy to access, as the media is biased toward hits. The popularity provides "social proof": others must have made the complex choice decision before, so following them should be a good way to get quality. And Lynn et al. (2016) show that the influence of popularity on quality perceptions is not limited to our consumption choices, but even remains *after* we have experienced an entertainment product. When they use experiments to study the reactions of several thousand consumers to downloads of unknown "indie rock" songs, the scholars find that information about a song's popularity (the number of previous downloads by others) impacts the quality rating that consumers give to a song. More popular songs are rated systematically higher by consumers, even though the scholars control for the "inherent quality" of a song, as measured via other consumers' judgments.

But popularity does not affect all of our quality judgments equally: whereas popularity has very little impact on songs with high-quality ratings (i.e., 4 or higher on a "1-hate to 5-love" scale), higher popularity (i.e., ten additional downloads) increases consumers' ratings of a low-quality song (2 or lower on the 1–5 scale) quite remarkably (by 0.25-points). As entertainment consumers, we seem to be confident in our own judgment when we like something, but much less so when we *don't* like it.

The second mechanism through which popularity influences success is a social one: people often want to join the bandwagon. Just as sports fans flock to a winning team (Bayern Munich anyone?!), if an entertainment

product is successful, consumers enjoy becoming part of the movement and do not want to miss out. We discuss this "bandwagon effect" in more detail in the context of pre-release cascades based on the buzz for an upcoming entertainment product.

We will now take a closer look at popularity cascades. We will start with an analysis of the impact that popularity information can have on product success within a channel or market and the circumstances that influence this impact. Because entertainment often involves multiple sequential channels and is marketed globally, we then take an inter-channel and inter-market perspective, investigating how popularity in one channel or market can spill over to others. And we will look at moral (and immoral) ways entertainment managers can attempt to harvest the success-breeds-success effect.

The Impact of Popularity on Entertainment Success

Much of what we know today about success-breeds-success effects in entertainment stems from two experiments ran by Salganik et al. (2006). In their inspiring study, the scholars created a number of "artificial markets" in which more than 14,000 consumers were invited to download music (48 songs from 18 bands). In the first experiment, participants were randomly assigned either to a market in which they were provided only with the names of the songs and the bands (the "independent" scenario) or to one of eight markets in which they were also shown the number of previous downloads by other consumers for a song (the "social influence" scenario).

In either scenario, participants could listen to songs and then decide whether or not to download them (for free). All songs started with zero downloads and were presented in random order throughout the experiment. The second experiment was similar—the main difference was that popularity information was presented more prominently; in the "social influence" scenario, the *order* in which songs were presented now reflected the number of their previous downloads.

In essence, the results of their experiments show conclusively that popularity can have a substantial influence on consumers' entertainment choices. Popular songs were *more* popular than less popular songs in all experimental conditions in which downloading information was available, compared to the "independent" scenario (in which no information on a song's popularity

existed). And when the downloading information was presented more prominently, the difference in popularity was even higher.[329]

Whereas the download ranks of the songs in the independent, or purely "quality-based" market correlate with those from the social influence markets, they did so far from perfectly—the songs' "performance" took a life on its own based on their early popularity. This is what Panel A of Fig. 12.3 shows by plotting each song's "independent" performance against its performance when popularity information was available. Panel B in the figure shows that the performance dispersion was clearly higher when popularity information was presented in a more prominent way.[330] In other words, the popularity information added unpredictability regarding a song's performance.

The correlational pattern in the figure also stresses that popularity does not fully substitute for the role of a product's quality, but that both coexist and co-determine an entertainment product's economic fate. As Salganik
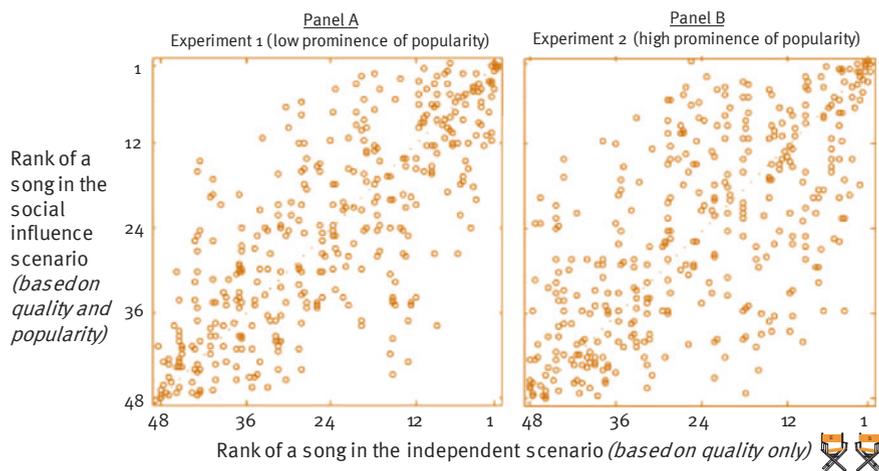

Fig. 12.3 Performance of songs in "artificial music markets"

Notes: Reprinted with minor modifications with permission by the American Association for the Advancement of Science from Salganik et al. (2006, p. 855). The ranks are derived from the number of downloads a song has received in the respective artificial market.

---

[329]These differences between popular and less-popular songs are statistically reflected in Gini coefficients, which measure the equality between the "market shares" of often and less often downloaded songs. In the first experiment, the Gini coefficient for the independent scenario was 0.25 versus an average of 0.34 for the social influence scenarios (a difference of about 36%). In the second experiment, the Gini coefficients were 0.19 versus 0.50 (+163%, or 4.5 times as high).

[330]The scholars do not report correlations.

et al. argue: "the 'best' songs never do very badly, and the 'worst' songs never do extremely well, but almost any other result is possible" (p. 855).

Keep in mind that this result is based on a lab experiment in which popularity solely develops based on consumers' experiences with the product's quality. Would things look different if popularity develops *before* consumers can actually experience a product? In a follow-up study using the same "artificial market" approach and the same 48 songs, Salganik and Watts (2008) gathered more insights about such popularity that is not based on the actual quality of entertainment. This time they manipulated the songs' *initial* popularity and looked how this "fake" popularity affected the performance of songs among consumers. They found that it indeed also impacts success—low-quality songs became hits solely because of their reported popularity, and the downloads of high-quality songs would suffer greatly due to a lack of initial popularity. We get back to this issue in our next section on pre-release buzz.

There is some positive news though for those who believe in the power of "great entertainment." In Salganik and Watts' experiment, the very best songs in the data set recovered from their low popularity over time. Still, they ended up with fewer downloads compared to when no fake popularity information was provided. In other words, high-quality songs can make up a lack of initial popularity, but only to a certain degree. Further, let us keep in mind that entertainment products usually do not have much time to recover, given the competitive nature of the markets and short product life cycles.

Other scholars have provided additional evidence for the role of popularity using actual sales instead of lab data. Market data certainly add realism, but also carry problems: it is tough outside the lab to demonstrate conclusively that the empirical links from early to later popularity are truly causal (instead of simply reflecting the underlying reasons why a product is successful in the first place). In the music context, Bhattacharjee et al. (2007) analyze albums' survival on the charts, controlling for a number of other variables (such as the artist's star status, the producing label, and the release time). They find that a less-successful debut is associated with a shorter time on the charts—each additional rank at entry alters survival by about 2%. But be aware, however, that the authors do not control for factors like advertising that occurs before release, and we assume that an album's chart entry will often be influenced by these factors (instead of being exogenous).

Similar concerns apply to two studies on success-breeds-success effects for books. Clement et al. (2008) study how the rank of a book in Germany's *Spiegel* bestseller list, and three threshold dummy variables (Top 10, Top

20, and Top 50) affect accumulated weekly sales of 609 novels published in hardcover format in Germany over the course of the 12 months following release. The scholars run OLS and panel regressions in which they also control for a large number of other factors (star status of the author, sequel, the book's quality as rated by readers, and its price—but not advertising). Their results again support the existence of success-breeds-success effect. But they find that the effect is highest in the first week after a book's release (when a higher rank translates into about 4% higher sales) and then deteriorates quickly. After week 4, the specific rank of a book no longer matters—instead, it is whether a book was among the Top 10 or Top 20 in the previous week that affects sales.

We learn that whereas consumers seem to be very attuned to a book's success when it first comes out, and the information is often heightened by media attention and advertising, things change to a more generic perspective, over time. Then it just matters whether a no-longer-new book is a hit or not. This insight is consistent with what Alan Sorensen (2007) reports for an analysis of how being on the New York Times bestseller list affects the weekly sales of 800 books released in 2001 or 2002. In his case, a book's initial appearance on the list adds about 8% in terms of sales, but no significant effect exists in the following weeks.

Logic suggests that popularity should also matter for movies and games. And our own research lends at least tentative empirical support: via structural equation modeling, we find that the long-term theatrical success of a film is highly related to its opening performance, even when controlling for the studio's efforts to produce, promote, and distribute an appealing product (Hennig-Thurau et al. 2006a). For the 331 films in our data set, the results suggest that the short-term success essentially absorbs all marketing measures when it comes to the performance in the following weeks, except for the film's quality itself.

With the pinball marketing environment in place now, we expect the power of success-based cascades to increase. As former Amazon Studios chair Roy Price argued: "If you have one of the top five or ten shows in the marketplace, it means your show is more valuable because it drives conversations and it drive subscriptions" (quoted in Littleton and Holloway 2017). In other words, in the digital era, information about hits will travel farther and faster among consumers, which should let firms profit more than before from their hits.[331]

---

[331]This effect is also reflected in today's shortened diffusion patterns.

We have so far restricted our analysis of success-breeds-success effects to a single channel or market. But entertainment products are usually released in more than one channel or market, and popularity-related information can flow between them. This is what we will look at now.

Success-Breeds-Success Between Channels and Markets

A specific area in which popularity matters is when the success of an entertainment product in one channel (such as movie theaters) or market (such as North America) influences consumers' decision making in a different channel (think DVDs) or market (think Germany). *Entertainment Science* provides some insights into this issue as well. Keep in mind the limitations of real-word data for analyzing success-breeds-success effects—they also apply to most of the studies we mention in the following sections.

Inter-Channel Success-Breeds-Success

Regarding movies, we show for a cross-sectional data set of 331 movies released in theaters and on video during 1999–2001 that the link between theatrical success and video rental success is quite substantial (Hennig-Thurau et al. 2006b). When analyzing a complex structural model with partial least squares, both a film's opening weekend and its performance in theaters in the following weeks are strongly linked with video rental success. The path coefficients are about 0.35 (on a scale from −1 to 1), on par with the impact of advertising and higher than those of all other "success drivers."

Whereas our standardized parameters do not offer elasticities, those from Jozefowicz et al. (2008) do. For a smaller data set that is biased toward theater hits (93 of the 100 highest grossing films in 2001) and few controls, they run two OLS regressions with VHS and DVD revenues as dependent variables. The scholars find that a 10% higher box office corresponds with 5–6% increases in rental revenues for the two home channels—those are large success-breeds-success effects, but with diminishing returns.

Finally, Jordi McKenzie (2010) studied how success at the Australian theatrical box office is linked with DVD sales for a data set of 760 films (from 2004 to 2007). Consistent with our own evidence on rental success, he finds a *very* strong positive correlation between the two channels of 0.88. While correlations are strong across genres and age ratings, they are highest for the biggest films in the data set and for films for which the time gap is neither very short nor very long—success information needs time to

spread, but is forgotten after a while. From a simple OLS regression that uses only box office and its squared term as regressors, he derives that the success-breeds-success relationship appears to be nonlinear—the most successful films in theaters experience a disproportionally high bump in DVD sales, compared to less successful films.

Similar inter-channel popularity effects have been reported for other entertainment products. Papies and van Heerde (2015) find that record sales trigger concert revenues, with the effect being part of a feedback loop, as concert tickets also spur higher record sales. In line with our expectation that the digital environment will facilitate such cascades, the scholars show that this effect only exists in recent times with extensive broadband Internet diffusion.

And for books, Sumiko Asai (2015) studies the role of previous hardcover sales on paperback sales for 254 newly released novels which appear on the Japanese Top 200 paperback chart from 2010 to 2013. Using a GMM approach in which she estimates sales and price separately (and controls for several factors—but not advertising and quality), Asai finds an elasticity for hardcover sales of 0.11—on average, a 10% increase in a novel's hardcover sales is linked to 1% higher paperback sales. Schmidt-Stölting et al. (2011) suggest that the role of success in previous channels might be even higher when books which were not hits are also considered. For their large data set of German books, a SUR analysis shows an success-breeds-success effect that is five times as high as the one reported by Asai.

Inter-Market Success-Breeds-Success

Until recently, entertainment media products were often released in their domestic market first and only later in foreign markets. Today, a global "day-and-date" release is nothing unusual for major movies, albums, games, and novels, despite the additional work that is often needed for foreign launches (such as translations, dubbing, etc.). A major advantage of a global release is the synergies in building buzz that producers hope will spread across countries via the Internet. But a potential drawback of this new approach might be the lack of success-breeds-success effects: when a product is released simultaneously in its home market and abroad, there is little time for action-based cascades to develop. What does the evidence say?

Two studies on movies provide evidence that success-breeds-success effects do exist across markets, but these studies also point to problems that occur if the delay between markets is too long. The first study is the often-cited one by Elberse and Eliashberg (2003), in which the two scholars use 3SLS

to estimate the effects of a film's success in North America on its foreign reception in France, Germany, Spain, and/or the U.K. For between 127 and 140 films, they find a direct and positive impact of the North American performance on foreign moviegoers in all four European countries. The direct impact of a 10% higher home market success ranges from an increase of 1.6% (in Germany) to as much as 9% (in the U.K.) in their data set. For France, Germany, Spain, the *total* success-breeds-success effect is even higher, as the North American performance significantly impacts the "supply side." Specifically, theater owners in the European markets show films that were more successful in the U.S. on more screens, which also increases the films' success potential with European audiences.

Elberse and Eliashberg's results also teach us that the strength of this success-breeds-success effect is moderated by the time that passes between the two market entries. An interaction between the films' North American box office and the time lag between market entries is strong and negative for the three continental European countries, such that a larger time lag reduces the impact of a film's success in North America for its foreign release. When we analyzed the box-office results of 231 North American films in both their home market and in German theaters (all released between 1998 and 2001), we found the same decay effect for action-based cascades: whereas the correlation between box-office results in the U.S. and Germany was a very high 0.87 for films released in Germany *within 3 months* after their home release, it was only half as strong for films released abroad 6 months or more after the home market release (Hennig-Thurau et al. 2003).

But let us note that both of these studies assumed a linear effect of time—the effect size weakens continuously with every additional month that passes after the home release and across the whole time period. This assumption leaves no room for potential *non-linear* patterns, such as allowing some time for inter-market action-based cascades to build. When we revisited the 1998–2001 data for this book, we find indications of such non-linearity. Take a look at Fig. 12.4: the shared variance between film's North American and German performance follows an inverse U-shape pattern over time, with the highest amount of variance explanation existing not for films which are released simultaneously in both countries, but for films released in Germany 2–3 months after their release in North America. Only then do the success-breeds-success effects start to decline.

These insights must be treated with care, as the published studies do not consider the potential endogeneity of studios' international release decisions. But it seems reasonable that managers systematically market their flops differently than their hits when it comes to international release decisions. But

Percentage of German box
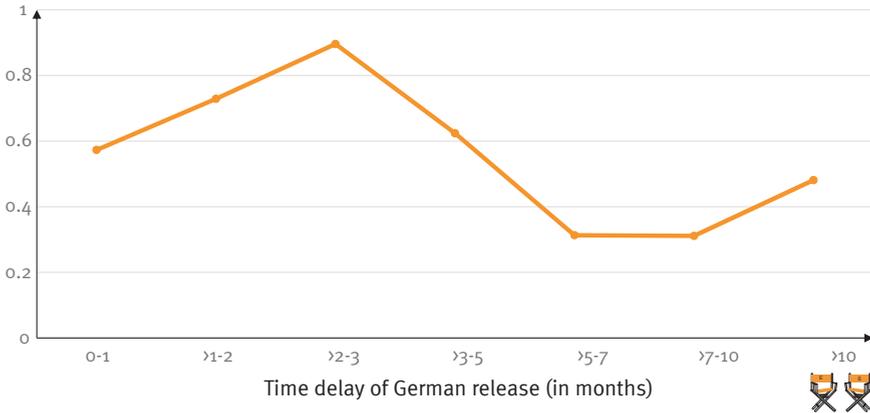office explained by North
American box office

Fig. 12.4 Relationship between films' North American and German box office
Note: Authors' own illustration based on data for 231 U.S. films, as described in Hennig-Thurau et al. (2003).

based on what we see in the data, we are tempted to conclude that delaying the foreign release of what is expected to be a flop in the home market might not be the worst thing: launching at the normally optimal time may enhance the likelihood that the flop information spills over and makes the product dead-on-arrival in the foreign market. But as release decisions and contracts for entertainment products are usually made months in advance, making effective use of inter-market success-breeds-success effects will work best if the producer is able to forecast a new product's success early on—which links our work in this section to the forecasting models we discussed in the context of innovation decisions.

As with all aspects of "earned" communication, managing success-breeds-success effects is a challenge for entertainment managers because of the low controllability of these channels. So, let us take a look at what managers can do (in addition to timing their global releases accordingly) to facilitate such cascades.

Managerial (Mis-)Use of Post-Release Uninformed Cascades

Our findings about the importance and mechanisms of popularity-based action-based cascades for entertainment products carry a number of managerial

implications. Some of these fall into the category of "smart" decisions, whereas we consider others to be outright immoral.

But let's start with those we assign to the "smart decisions" category. When Steven Spielberg's movie Jaws was approaching its theatrical release in the summer of 1975, Universal's chairman Lew Wasserman made a tough choice. Although his marketing team had initially arranged for release in a then-unprecedented number of 900 theaters in North America, because of the high demand that theater owners were sensing, Mr. Wasserman dropped more than 400 of those.[332] In a time when box office statistics were not omnipresent and immediate as they are today, Mr. Wasserman's vision was to create long waiting lines which would generate massive news coverage and, in turn, trigger action-based cascades (Shone 2004).

Although it is hard to assert causality in a single case, we know that his decision certainly did not hurt the film, which ran all summer and became the first in history to break the $100 million barrier. In more general terms, Mr. Wasserman's logic was to create images of high demand through an artificial shortening of supply. This is an approach that Hollywood still practices these days when it follows a "limited release" strategy, implying that the showing of a film in carefully selected venues (such as arthouse theaters in art-centric areas) will create high demand signals that are then picked up by the media and, in turn, set off cascades, attracting audiences in other parts of the country and, ideally, the world. And the media today indeed helps by highlighting high "per-screen" averages for limited-release films that pack selected theaters.

Other approaches to exploit the power of success-breeds-success effects might be considered equally clever, but of more questionable morality. A first one was reportedly practiced by management "gurus" Michael Treacy and Fred Wiersema in 1995, when they secretly purchased 50,000 copies of their own book "The Discipline of Market Leaders." They bought them from the stores that are influential for the New York Times bestseller list, helping the book to flourish on the charts, despite mediocre reviews (Bikhchandani et al. 1998). The book then "sold well enough to continue as a bestseller without further demand intervention by the authors" (p. 151).

Such behavior appears to be quite common for novels and music today; it might not be technically illegal, but it certainly generates a fake popularity

---

[332]See also our more detailed discussion of Jaws as the first incarnation of the blockbuster concept in the integrated entertainment marketing chapter.

that cannot be detected by consumers, who are misled by the numbers. Music manager Tom Silverman, in 2010, argued that music labels have long "hyped the charts" by buying their own songs, and are still doing it. The economics of the digital economy make such behavior even less costly; now the labels only have to cover the 30% share of the digital retailer, such as Apple/iTunes, but not the production costs for the disks. "[I]f they buy 50,000 songs, we're talking $50,000 less 70%, so it would cost about $15,000. For $15,000 in a week, they can buy 50,000 more song downloads, which could drive the record up three or four positions on the chart" (Silverman, quoted in van Buskirk 2010). This success, Mr. Silverman argues in line with the logic of action-based cascades we discussed on the previous pages, makes potential consumers assume that a song is of high quality, which then triggers the "actual" success.

A related, and no less dubious, approach is to tell audiences that a product is a bigger success than it actually is, an approach referred to as "overreporting." Whereas doing so for actual sales would be outright cheating, Malhotra and Helmer (2012) offer empirical evidence that some film studios systematically practice overreporting in publishing box office *estimates*; these are then used by the media for their success-related coverage and Monday morning charts.

Malhotra and Helmer analyze such estimates in combination with actual box office numbers for all 1,000+ wide-released films from 2003 to 2010, finding that overreporting differs strongly between studios.[333] They find that overreporting is particularly strong when the incentives for the studios are the highest—in the release week (versus later), when cascades can be stimulated most effectively, and when competition is high. And if the difference to the second-placed film is less than $10 million by Saturday evening, overreporting for Sunday is highest. Based on these results, the scholars conclude that overestimation is a fact-of-life in the film industry, and that it "is highly unlikely to be due to chance" (Malhotra and Helmer 2012, p. 1411).

---

[333]In their data set, Malhotra and Helmer find overreporting to be highest for Sony (+10.3% on average) and Summit (+7.9%, now part of Lionsgate), and lowest for Fox, Disney, and DreamWorks (4%, 3.6%, and 1.2%, respectively). Of course, these differences *might* also be attributed to the studios' differing monitoring skills…

Pre-Release Action-Based Cascades: Buzz

> "Failure to create the right buzz beforehand [means] less anticipation and ultimately fewer ticket purchases. … Greater emphasis has been placed on social media in the hopes of generating the right buzz that will serve to build anticipation, pushing the film into a 'must-see' status."
> —Freedman (2015)

Action-based cascades can also be observed *before* the release of a new entertainment product. At this point in time, no information about the actual performance of a product can exist—but consumers can interpret the "buzz" that they observe in the market for a new product as an indicator of the overall anticipation for the product. According to this logic, the observed anticipation itself serves as a signal of the forthcoming product's attractiveness and quality.[334] Building buzz has become a major element in the distribution of new entertainment products, a cornerstone of the now-dominant blockbuster strategy. The idea behind such efforts to stimulate buzz is that pre-release buzz cascades translate into opening success, which might then trigger post-release action-based cascades, as discussed on the previous pages.

Because one cannot measure (let alone manage) something for what no clear understanding exists, let us first clarify what we mean when we say "buzz," a term that is often used in a "happy-go-lucky" way by consumers, journalists, and managers alike. After offering a scholarly definition of the concept, we take a closer look into what *Entertainment Science* scholars have unearthed regarding the drivers of buzz, the concept's link with entertainment success, and finally its mediating role on the road to product success.

So, What Exactly is "Buzz"?

People inside and outside of entertainment have been using the term "buzz" in a pretty superficial way. As a result, we rarely talk about the same thing when discussing buzz, something that has affected the many ways buzz is measured. It has also hindered scientific progress on the topic: there is just no way for *Entertainment Science* scholars to explore a concept and its role for managing entertainment without a common understanding of it. So, what is a useful way to define buzz?

To understand what constitutes buzz, we asked those whose responses to it determine any impact buzz can have on entertainment success: the

---

[334]See our initial discussion of the buzz concept in the context of (social) consumer behavior.

consumers. Through several depth interviews and focus groups with consumers, we found that new product buzz describes the "aggregation of observable expressions of anticipation by consumers for a forthcoming product" (Houston et al. 2018, p. 349). Let us highlight four core characteristics of buzz that we learned from our exploration:

Buzz is an aggregate-level concept. Buzz is not about what any single person does or says, but it is about what is happening with many consumers at an overall, "macro" level—the totality of things, if you want, not just its individual parts.

Buzz has an anticipatory nature. Buzz is forward-looking, referring to something "to come," not looking back at what has already happened. It is also a positive concept: at least in the context of entertainment, buzz is fueled by people's hopes, not their criticism. "Negative buzz" exists only if consumers are afraid that their hopes will not be met by a new product. Will the AVATAR sequels match the stellar-high standard of the original film? Those studies which include a measure for pre-release *sentiment* of consumer communication find it to be uncorrelated with later success (e.g., Gopinath et al. 2013).

Buzz is multi-behavioral. Consumers' perceptions of the buzz for a product are fed not only by what is said and written about a new product, but also by two other kinds of observable behaviors: people's searches for the new product (as observed in metrics such as IMDb's MovieMeter) and also their "participative" actions (e.g., trailer views on YouTube, the scavenger hunts that took place as part of THE DARK KNIGHT alternate reality game, the re-watching of an earlier movie to prepare for an upcoming sequel, etc.).[335] Figure 12.5 illustrates buzz' multi-behavioral nature and provides examples for each type of behavior.

Buzz is two-dimensional. The amount of buzz, as reflected in "most-popular" rankings on Twitter and other websites, certainly marks a key type of information when consumers judge the buzz for a product. But in addition, buzz involves a more qualitative dimension: *who* is anticipating the new product, in addition to how many? This is the "pervasiveness" dimension of buzz: it tells us to what degree the collective anticipation for a new product is *spread across consumer segments* (see also Fischer et al. 2017). Are only the fan boys excited? Or has the excitement exceeded the niche and reached the mainstream?

---

[335]In their introspective study, Wohlfeil and Whelan (2008) describe in detail how the anticipation for watching the then-forthcoming movie PRIDE & PREJUDICE shaped the first author's behavior, including participative actions such as buying a new version of the book (which featured the poster as book cover) and buying newspapers which featured articles about the film.
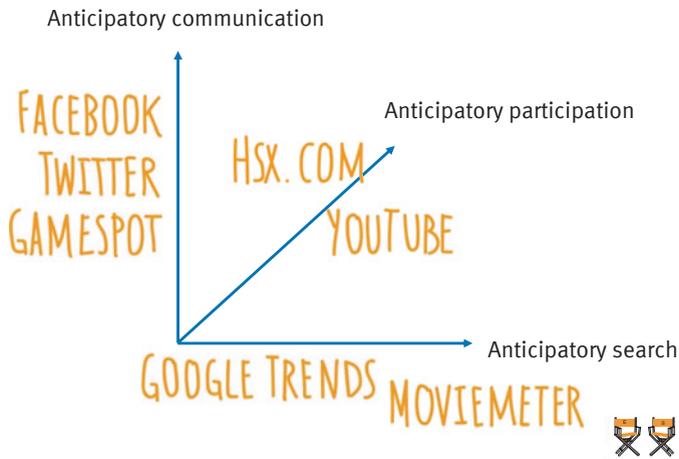
Fig. 12.5 Three general kinds of buzz behaviors (and examples)
Notes: Authors' own illustration based on Houston et al. (2018). Brands are trademarked. With graphical contributions by Studio Tense.

For consumers' evaluation of a coming attraction, niche buzz and mainstream buzz are often separate things.

This two-dimensional character of buzz provides an explanation why some films such as KICK-ASS and SCOTT PILGRIM VS. THE WORLD flopped at the box office, despite having been strongly hyped by Comic-Con attendees, whereas other films that were surrounded by similar buzz levels at Comic-Con became huge box office hits (such as AVATAR and IRON MAN).[336] The buzz for the latter films came from fans *and* mainstream consumers (i.e., high pervasiveness), but the former films excited *only* the niche. What did Kaye (2012) say about SCOTT PILGRIM? "Only nerds like movies about nerds; [leading actor] Michael Cera is not a leading man to the rest of the world."

What do we gain from such an elaborate conceptualization of buzz? Generally speaking, the better someone is able to capture the essence of a concept that influences a product's performance, the more accurate success predictions can be. In the next section, we will first investigate what empirical studies tell us about the link between buzz and success in general, before getting back to the question whether it matters how we conceptualize buzz (and how we measure it).

---

[336]IRON MAN director Jon Favreau explicitly attributed the success of his film to the roaring response of fans at Comic-Con in 2006, telling attendants "It all started here. Nobody cared before you did" (quoted in Horn 2009).

How Buzz is Linked to Product Success

The Link Between Buzz Volume and Product Success

Despite the differences in empirical measures, studies consistently show that buzz is linked to entertainment success. But like with post-release action-based cascades, establishing the causal nature of this link, which is so essential for determining its role in the entertainment marketing mix, is far from trivial.

When Google analysts Panaligan and Chen (2013) reported that they were able to explain about 70% of a movie's opening weekend box office with the search volume for a movie title via Google in the week prior to its release (and that adding the number of theaters and brand status variables increases the variance explained to 92%), they could be sure to get wide attention among entertainment managers. And even more impressive was their proclamation that the search volume for a movie's *trailer* four weeks before release explained 62% of the opening weekend success, 94% when paired with release timing and brand status variables. These result sound too good to be true, and they might be. The authors report very little about the data and method they use, and the same is true for their selection criteria for the small sample of 99 films they analyze—which, as every researcher knows all too well, can have a *huge* impact on the results. We don't know how much of the box office explanation can be attributed to search buzz: the analyses, and particularly the "search-only" ones, suffer from a massive omitted variables bias.[337] And for their "joint" analyses, the authors do not even reveal the share of success that is explained by search buzz, above and beyond the other variables.

Let us thus turn to more rigorous approaches to shed light on the link between buzz and success. In their study of 681 games (all released in 2009 and 2010), Xiong and Bharadwaj (2014) look at how pre-release communication buzz (on a variety of blogs and forums) and search buzz (from Google Trends) affect the games' opening week sales. The scholars conduct separate regressions with robust standard errors for the two buzz behaviors; they also include a large set of controls, including whether a game is a sequel, advertising spending, its genres, and the release timing. They estimate an elasticity of 0.46 for communication buzz (i.e., a 10% higher communication buzz

---

[337]See the opening chapter in this book for a short introduction to the omitted variables bias.

transfers into 4.5% higher opening week sales)—an effect which is almost nine times higher than the effect of pre-release advertising in their data set.

For movies, results also point to a strong role of buzz. Gopinath et al. (2013), in their investigation of 75 movies, report a strong elasticity of 0.27 for pre-release communication buzz (articulated on blogs) on a film's opening weekend, and Liu (2006), in his OLS analysis of 40 movies, finds an even higher elasticity of 0.59 for pre-release communication on the Yahoo Movies site, which might be somewhat inflated, however, because of a smaller set of controls. Whereas Kim and Hanssens (2017) do not report elasticities, they use search buzz three weeks prior to movie release, along with advertising and distribution efforts, to predict the opening weekend box office of 41 movies. They find that prediction is about 8% (or $6 million) more accurate for the model that includes search buzz versus a model without search buzz.

We, too, looked at movies when we predict the opening weekend box office of 254 movies which were widely released in North America from 2010 and 2011 with partial least squares analysis. We used a multi-behavioral buzz measure that, in line with our conceptualization above, combines information from Twitter for communication, Google for search, and Facebook likes for participation in addition to large set of other "success drivers." We find that this model explains 25% more of the variance in the movies' opening weekend box office than a model that is exactly the same except that it contains no buzz information (Houston et al. 2018).

Scholars have also compiled evidence that buzz matters for TV show episodes, although the effect found in scholarly studies is less radical in this context, probably because episodes, being part of an ongoing series, are less "innovative" compared to a new movie or game. Studying the weekly ratings for the episodes of 30 TV shows when aired in North America between 2008 and 2012, Liu et al. (2016) use a panel GMM estimation approach and find that the number of pre-broadcast tweets increases the explanation of a show episode's rating by 0.5 percentage points, compared to a model that uses only previous ratings for a show.

Finally, only limited evidence exists for the power of buzz for new *music*. The little we know comes from Buli and Hu (2015), employees of music analytics company Next Big Sound, who reported that first-week album sales correlate at 0.70 with views of an artist's Wikipedia page (a measure of search buzz) in the week before release and at 0.25 with YouTube video views (a kind of participation buzz). The authors report Granger causality tests to address the causal nature of these parameters, but there is simply not enough transparence regarding the methods used to judge their findings.

The results we have assembled here give us a solid understanding that buzz is indeed influential for entertainment success; the size of the buzz effects reported by scholars also puts those named by commercial analysts into perspective and gives us an idea of their inflated nature. But the section also shows that quite different measures of buzz are used. Let us now see if we can shed some more light on differences between those measures and the corresponding roles of the buzz behaviors and the pervasiveness dimension of buzz, in addition to buzz' sheer volume.

Differences Between Buzz Measures: Where Behaviors,
Pervasiveness, and Content Matter

To understand the success impact of buzz better, studies that compare alternative specifications of buzz can be informative. In their games study, Xiong and Bharadwaj (2014) find a large difference between communication buzz and search buzz. In a separate analysis to their model in which they use communication buzz, they estimate an elasticity of search buzz (via Google) on opening week sales of 0.05, or about just one-tenth of the one by communication buzz.

Results for movies show a different pattern, though. When Kim and Hanssens (2017) report prediction results for their data set of 41 movies for an alternative model which contains communication buzz (via Google blogs), instead of search buzz, they find the search measure to be more effective than the communication measure. And when we estimate alternative models to our multi-behavioral buzz specification (which combines Twitter, Google Trends, and Facebook likes), we also find quite substantial differences between individual buzz behaviors (Houston et al. 2018). When studied in isolation, search buzz explains the most (i.e., has the highest R^2) and also has, along with Twitter, the lowest prediction error (i.e., the smallest MAPE). The number of Facebook likes for a movie, our measure of participation buzz, is clearly less effective in both regards, a finding that is consistent with our results in Kupfer et al. (2018). But none of the buzz measures can, when used in isolation, match the performance of the multi-behavioral buzz specification, which explains, on average, almost 8 percentage points more of movies' opening weekend success.

We also compared the three buzz measures, that are all used by all kinds of consumer segments and thus reflect high-pervasiveness buzz, with low-pervasiveness, niche measures (the number of posts on the fan forum

joblo.com for communication buzz, the number of MovieMeter searches on IMDb for search buzz, and the number of page edits by enthusiasts of a film's Wikipedia entry as participation buzz). Using these low-pervasiveness measures for each buzz behavior reduces the explained variance of movie success by 13 percentage points. In other words, at least for our data set of wide-released movies, pervasiveness matters for buzz, and mainstream buzz plays a bigger role than niche buzz.

Finally, Liu et al. (2016) show that the buzz *content* can also matter for the effect that communication buzz has on success. In their analysis of TV series ratings, an in-depth coding of what people chat about via Twitter with regard to a show increases the explanation of the show's rating by 15 percentage points—up to 90%. When buzz features strong doses of positive emotions such as "excited" and "love," the scholars find it to exert a particularly strong effect on success.

In essence, not all buzz is equal, just as we concluded in the case of WOM. Different buzz behaviors produce different insights and suggest different decisions. And although the results for different buzz behaviors and measures vary somewhat between studies and criteria, among the most important insight is that a multi-behavioral buzz measure that captures high pervasiveness is most informative. Understanding this is essential for picking the "right" measures when tracking, and stimulating, the buzz for an entertainment product.

Buzz Cascades and Thresholds

The studies we have covered so far implicitly assume that the volume of buzz has a linear effect on success. But this assumption somewhat conflicts with an argument we made in the consumer behavior chapter of this book. Buzz sends a signal of collective interest to people (the "We-desire") that triggers a cascade, and that at least some consumers only join the cultural bandwagon once they feel that buzz is substantial enough to deserve their interest. In other words, after the buzz for a product has passed a certain threshold, a cascading, non-linear effect should take place.

And studies indeed provide evidence for the existence of such buzz cascades. The results of Xiong and Bharadwaj (2014), for buzz communication about games, show that buzz creates more buzz on its own over the 180 days prior to a game's release, with this self-enhancing effect growing over time. And Kim and Hanssens' (2017) VAR analyses of movies and

games point in the same direction; for both forms of entertainment, they find positive "self-elasticities" for both communication buzz (for which they find the effect to become smaller over time) and search buzz (effect becomes stronger).[338] They also show that buzz cascades can exist also between different buzz behaviors; in their data set, communication and search buzz mutually enhance each other in both entertainment contexts, with the effect of communication buzz on search buzz being more than four times as strong as the effect of advertising.

Whereas we now have evidence of buzz cascades' existence, none of these studies has addressed the existence of buzz thresholds that need to be exceeded to set the self-enhancing effect in motion. But Soderstrom et al. (2016) have looked for such thresholds empirically, using buzz data for 309 movies released from 1999 to 2001. Their buzz data come from a professional survey firm that contacts about 300 consumers in the week preceding a particular movie's release, asking consumers whether their friends were talking about the film. In this analysis, the measure of (communication) buzz is the percentage of consumers for which this was the case. To analyze how such buzz is linked with movies' opening weekend success, the scholars combine OLS regressions with a "spline" method, in which separate functions are estimated for different parts of the data set and then the slopes for the parts are compared.

Soderstrom et al.'s findings provide further support for the relevance of (communication) buzz, as adding the buzz measure to a list of controls (such as distribution and stars—though not advertising) increases the explained variance from 44% to 72%. With regard to the cascading effect, their results suggest that a buzz threshold indeed exists, and that once it is passed, the effect of buzz is 3.6 times stronger than before. And where is the threshold located? For the set of movies analyzed by Soderstrom et al., it is when more than 21% of consumers sense the buzz for an upcoming film; this "cut-off" is illustrated in Fig. 12.6. We cannot say whether this insight can be transferred to other data sets and entertainment products, and the authors' findings come with certain limitations. But we learn that for entertainment managers, investments in buzz pay off much more in success once a critical buzz level is passed.

---

[338]Kim and Hanssens find the self-enhancing effect of "buzz on buzz" to become smaller over time for communication buzz, whereas for search buzz the effect becomes stronger over time until the release.
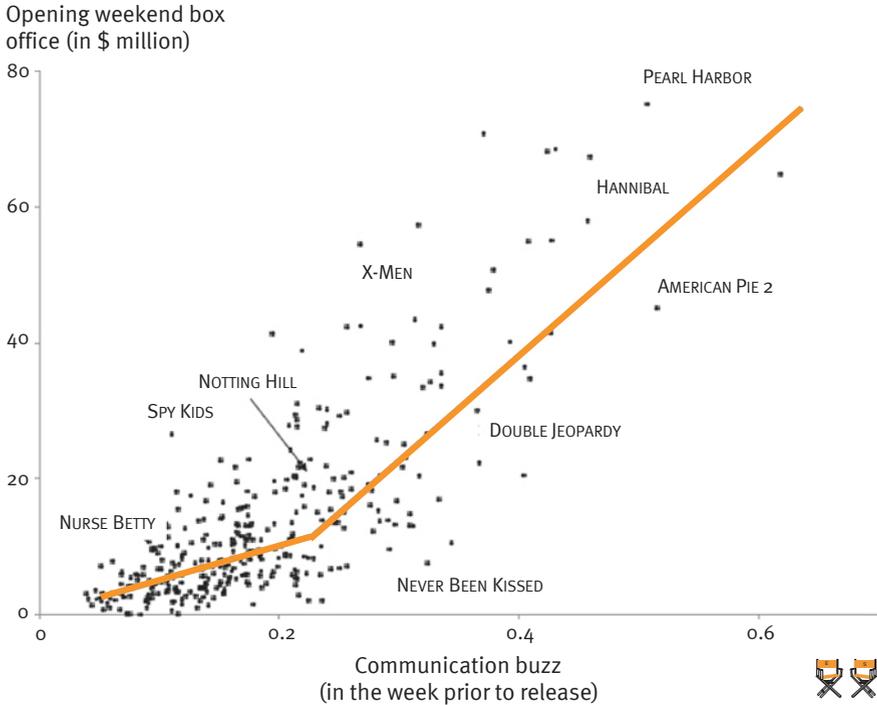
Opening weekend box
office (in $ million)

Fig. 12.6 A buzz threshold

Note: Reprinted with some modifications from Soderstrom et al. (2016, p. 925), which is distributed under the terms of the Creative Commons Attribution 4.0 International License (https://goo.gl/n6FkkT).

Buzz Patterns Over Time

Buzz builds over time, and scholars have shown that, beyond its amount at a given point in time and its pervasiveness across groups, there is also value in understanding its *pattern of development*, over time. Foutz and Jank (2010) use functional shape analysis to determine patterns of how the "stock prices" of movies develop over time on the Hollywood Stock Exchange, which, they argue, mirrors the "hype" for a movie as perceived by those consumers who trade movie stocks on HSX. From a data set of 262 films released in North America in 2003–2005, they extract four unique buzz patterns from the movies' stock prices over the 52 weeks prior to their releases: a "higher trading average," a "steeper upward trend," a "last moment hype," and an "early preannouncement hype" pattern.

The two scholars find that all four patterns are linked to higher opening weekend success in their data set. And they show that the patterns might also be used by managers for early forecasting: whereas the average

prediction error for a model based on movie characteristics and advertising is about 90% until ten weeks prior to release, Foutz and Jank calculate that the addition of buzz pattern information reduces this error substantially to 35% at 40 weeks before release, and to 20% at ten weeks before release.

Xiong and Bharadwaj (2014) reveal similar insights for their data set of games when they also identify four main functional shapes for pre-release communication buzz, as expressed in blogs and forums before a game becomes available. When the scholars link the shapes with opening week sales in a separate regression model, two pattern are associated with higher sales, both of which had also been found to be influential by Foutz and Jank: the "higher trading average" pattern and "steeper upward trend" pattern. Xiong and Bharadwaj also repeat their analysis with buzz *search* data, finding similar results.

And What Drives Buzz?

Now we have shown what buzz is and that it matters economically. For managing buzz successfully, understanding the factors that influence it is crucial. From the studies that explored variables that drive buzz (or its facets), two major categories of drivers have emerged: the manager's *marketing actions* and the *quality* of the product that the buzz is about. Let us note that most existing studies look at the amount/volume dimension of buzz, but very little is known what determines whether buzz is niche or mainstream.

Marketing actions. Advertising has been shown to be a main factor when it comes to triggering buzz. In Xiong and Bharadwaj's (2014) study of video games, the authors use a panel regression approach to understand what drives buzz communication during the 180 days prior to a game's release. They find that daily *online* advertising has the expected effect on buzz, but it varies over time, with its effect size peaking around 90 days prior to release (but remaining influential until release).

The pattern for TV advertising is similar, but the peak at −90 days is stronger (and the effect weaker closer to release). These patterns might be influenced by current industry practices—a bias that Kim and Hanssens' (2017) longitudinal study of the link between advertising and buzz in a games context counters by using a VAR modeling approach. For a data set of 66 games from 2013 and 2014, they estimate a cumulative advertising elasticity of 0.09 for communication buzz (on blogs) and of 0.06 for search buzz—10% higher advertising spending corresponds with 0.9 and 0.6% higher buzz, respectively. The pattern of results for a data set of 137 movies

from around 2009 looks similar, but the effects of advertising on buzz are about twice as high for movies than for games.

Other marketing variables have also been found to influence buzz, with different kinds of brands playing a prominent role. Xiong and Bharadwaj (2014) find that games with a major publisher brand (such as EA) experience more buzz; the publisher brand is particularly effective in the early stages, but becomes small and insignificant about 90 days before the game's release. Craig et al. (2015), in a regression analysis of a sample of 62 wide-released films (from 2008 to 2009), find that buzz (they combine communication and trailer views as a form of participation buzz) is, on average, higher for sequel brands than original films.[339] And Divakaran et al. (2017) find that stars, as human brands, also boost communication buzz (the number of user comments on Fandango, the pre-order fan website about their films) for a data set of 373 movies from 2009 to 2010 which they analyze with partial least squares. Related, Karniouchina's (2011) results suggest that consumers' *chatter* about stars spills over to the chatter about the films in which they appear; she estimates separate equations for the search intensity for films and stars with 3SLS.

Buzz also differs between certain entertainment genres. Xiong and Bharadwaj find much higher buzz for sports games (versus other game genres played on the Xbox360); for movies, Craig et al. find action and horror films to have more buzz, on average, and Liu (2006) reports this to be the case also for action movies and adventure movies. Liu also finds that movies with a rating less restrictive than R produce an average of 150% more communication buzz (on Yahoo Movies).

Quality. Fewer studies have investigated the role of product quality on buzz, but their results suggest that higher quality translates into more buzz, in general. Xiong and Bharadwaj, who measure quality by experts' ratings, find that product quality has an effect on communication buzz, which grows over time, and only loses its impact in the week before the release. In our own partial least squares (PLS) analysis of 254 movies, we also find quality, again measured as experts' quality perceptions, to be positively linked to our multi-behavioral buzz measure (Houston et al. 2018).

But how strong is the (relative) impact of quality on buzz compared to those for the marketing actions we have identified as buzz drivers before?

---

[339]Craig et al.'s result, and those of Divakaran et al. (2017), also suggest a buzz-driving role of the production budget of entertainment products, but here the same concerns apply that we mentioned in our discussion of the production budget's role on product success. Neither study included advertising, whose impact might have been appropriated by the budget variable.

In our study, we also compare the effect size of movie quality on buzz to the effect size of a multi-faceted marketing actions variable (which comprises a film's budget, rating, star, and sequel character). The PLS results show that marketing actions and quality together explain a sizable chunk of buzz (about 35%), and that marketing actions account for a larger part: its impact on buzz is 3.4 times as high as the impact of quality.

The Mediating Role of Buzz

Our analysis on the previous pages has shown that buzz influences the success of entertainment products, but is also affected itself by various marketing actions and the quality of the product–factors that we know influence the market performance of entertainment products on their own.[340] Thus, buzz should not be treated as an exogenous construct, but instead as a mediator between a product's quality and other marketing actions, on one hand, and the product's success on the other (Houston et al. 2018). Figure 12.7 illustrates this mediating role by putting buzz between quality/other marketing actions and success.

Our empirical analyses of this structural model inform us that buzz serves as a *partial*, not a full mediator—both the product's quality and the firm's other marketing efforts still influence consumers directly, aside from buzz. But a substantial share of their impact is transmitted through buzz—in our study, 42% of the effect of quality on movies' opening box office is explained via buzz, and 39% of the effect of other marketing efforts is moderated by the buzz these efforts are able to stimulate (see also Divakaran et al. 2017 for a similar argument).

This mediating role has far-reaching implications for the management of entertainment: because buzz is acknowledged as an important success driver, producers increasingly look for marketing actions that help build buzz. A look at the buzz determinants tells us that this logic favors certain types of products over others. Freedman (2015) states, with regard to the film industry, that producers' attempts to "adapt to [the critical role of buzz] are revealed in the types of films they choose to produce and the manner in which they have amped up their marketing efforts to generate buzz."

In contrast, those products that have lower buzz potential are considered inferior: "It is still possible to start with unfamiliar characters, but only if they can properly be transformed through the magic of social media"

---

[340]We discuss the role of quality for product success in Chapter 7 of this book, which is dedicated to the entertainment experience.
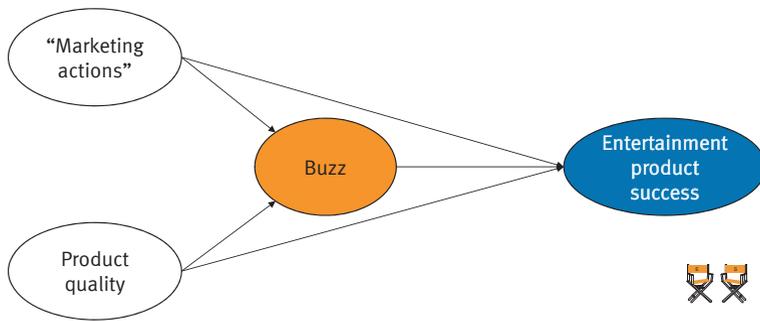
Fig. 12.7 The mediating role of buzz
Note: Authors' own illustration based on ideas in Houston et al. (2018).

(Freedman 2015). We continue this discussion and critically evaluate the industry's adaptation of the powers of buzz in the shape of the "pre-sales" approach to entertainment in the final chapter of this book.

But before we do that, let us inspect other kinds of "earned" media. We begin with what might be considered the most obvious infusion of data analytics in entertainment communication: automated personalized recommendations of entertainment products.

Automated Personalized Recommendations

> "One day we hope to get so good at suggestions that we're able to show you exactly the right film or TV show for your mood when you turn on Netflix."
> —*Netflix-CEO* Reed Hastings *(quoted in The Economist 2017)*

Digitalization has broadened the concept of word of mouth by making the entertainment experiences of other consumers available for us on websites and via social media. It has further facilitated action-based cascades by letting us access success information about entertainment products and the buzz that exists for them in unprecedented ways.

But digital technologies also provide entertainment firms access to new "big" databases and powerful hardware that can process such data and advanced algorithms with enormous speed. The combination of both enables firms to offer personal product recommendations based on the experiences of other consumers who we have never met, and probably will never meet. This is what automated recommender systems do: they apply algorithms to large data sets to generate individualized product recommendations based

on the choices of other consumers with similar preferences to our own. Recommenders today exist for all entertainment products that we feature in this book: Pandora and Spotify use them for recommending songs, Netflix and moviepilot.de for guiding us to movies, Gamefly and GameStop develop personalized recommendations for games, and Barnes and Noble suggest which books we should read next. And Amazon uses recommenders for all these products (and for everything else).

Although recommender systems could be considered as a variation of WOM, we discuss them at this point of the book for a reason: unlike WOM and also unlike buzz, recommender systems require the activities of third parties beyond consumers and producers. It is usually an intermediary (such as a retailer like Amazon), a streaming service (such as Spotify or Netflix), or a web platform (such as moviepilot.de or movielense.org) that generates personalized recommendations.[341] For them, recommenders can constitute important strategic resources, or even a "mission-critical technology" (Jannach et al. 2016). Because data technology is highly dynamic, keeping a competitive edge can be quite costly; Netflix is reported to have spent over $150 million as of 2015 just to improve their existing recommendations (*tickld* 2016).

This intermediary involvement is what sets recommender systems apart from the kinds of "earned" media we have discussed so far. Intermediaries, as recommendation providers, blend consumer information with their own skills and ideas. And the latter is what links recommendations with the "earned" media types we will discuss in the following sections, which also involve information from parties other than consumers, namely reviews by critics and awards from (industry) experts.

In the following, we look into the different kinds of recommenders that exist and their respective algorithmic logic; we then discuss selective issues beyond algorithms that also affect the effectiveness of recommenders. Please note that recommendation systems have developed into an academic

---

[341]One might argue that, with Amazon and Netflix (and, to a certain degree, also Spotify) now also producing entertainment products themselves, their recommender systems fall more into the category of "owned media" than "earned media." However, as their recommenders cover a wide range of products from various sources, with their own productions being only a small fraction of potential recommendations, we consider their recommender systems more part of their value creation role as platform providers than as producers of content. But this discussion raises an interesting issue: at least theoretically, recommender systems could be "owned media," depending on the range of products which are considered. In practice, as we point out in our discussion on the following pages, the value generated by recommenders depends on the (large) number of product alternatives—which presents a serious limitations for the effective use of recommenders as owned media.

discipline on their own, with conferences being held and extensive articles and books being published on the issue. Our discussion does not intend to capture the technical details of this vibrant exploration, but aims to introduce our readers to the fundamental issues of the debate.

How Recommender Systems Work: A Look at Their Algorithmic Logic

The basic idea behind all automated recommender systems is that consumers have individual preferences; thus, person-specific recommendations of entertainment products are more valuable for a consumer than general, average ones. At the same time, the approaches to determine individualized product recommendations differ strongly between systems: whereas collaborative filtering methods use the *judgements or behavior of other consumers* with regard to other products for providing a consumer with suggestions to meet his or her idiosyncratic preferences, content-based recommendations believe the solution for doing so is in the *product and its attributes*.

All recommender systems face a set of challenges against which they are judged. One is the "cold-start" problem: how to deal with new products (for which no consumer ratings yet exist) and new users of the system (who have not rated any products themselves). A second basic problem is referred to as the "serendipity" problem: as captured by the sensations-familiarity framework, consumers value more of the same (the familiarity component of the framework), but they also look for new impressions (the sensations component), i.e., the "serendipity" moment.

Recommenders have a strong bias toward familiar attractions; they generally struggle to provide users with fresh sensations. Some people, such as veteran film journalist Peter Bart, consider this a given and a basis for their aversion against technology: "I wanna try new things!" (Bart 2017 on why he does not like recommenders). As we will show below, different recommender systems address these problems to varying degrees, a fact that has motivated scholars to develop recommendation models that combine the advantages of different systems—so-called "hybrid" recommenders.

Collaborative Filtering: The Matrix Completion Challenge

Collaborative recommenders are all in the "matrix-completion business." Imagine a matrix that includes the ratings of multiple products by multiple

users in which not all products have been rated by all users. A recommender algorithm's focal goal in this case is to predict how users would rate those products which they have not yet rated. In other words, how much would a consumer like a movie he has not seen yet?

Figure 12.8 shows such a matrix, which consists of four consumers, or users (Claudia, Mark, Nancy, and Thorsten), and five movies, or items (Notting Hill, Terminator 2, Pirates of the Caribbean, The Good the Bad and the Ugly alias GBU, and Love Actually). The rating scale ranges from 1 = "Don't like at all" to 5 = "Like very much." The scale itself does not matter—it could also be from 1 to 10 or binary (0/1). The question marks in the matrix indicate that a user has not yet seen a film, such as Mark has not yet seen GBU. Would he like the film?

The answer depends on the recommender approach used. We will now take a quick look at the two main approaches for generating personalized recommendations via collaborative filtering: user-to-user filtering and item-to-item filtering. Let us note that the "matrix-completion-view-of-the-world" also carries its own definition of what defines a "good" recommender: if you intentionally leave out some of the ratings in the matrix when training the algorithm, the "goodness-of-fit" of the recommender is measured by how accurately the algorithm can predict those left-out ratings.

User-to-User Collaborative Filtering

The logic of user-to-user filtering is pretty straightforward: find other users whose ratings of products are similar to a target user (i.e., have a similar taste) and then predict to what degree this user will like a product he or she has not yet experienced, based on its rating by those "taste neighbors" (e.g., Ekstrand et al. 2010). Doing so requires three major decisions: how the similarity between users is determined, how many neighbors should be considered for the recommendations, and how similarities and ratings are transformed into recommendations.

	Notting Hill	Terminator 2	GBU	Pirates …	Love Actually
Claudia	5	3	?	4	5
Mark	?	5	?	4	1
Nancy	3	3	1	?	2
Thorsten	5	4	5	1	4

Fig. 12.8 A simple matrix of four consumers' ratings of five films
Notes: Authors' own illustration. Brands are trademarked.

Extensive research has studied potential alternative approaches for user-to-user collaborative filtering (for an overview, see Ekstrand et al. 2010). One popular way is to use the correlation between users as a measure of their respective similarity and to use this similarity measure as a weight when calculating the average of a film's ratings by other users. This approach resembles how consumers integrate the different attributes of a product into an overall attitude rating, according to attitude theory (e.g., Fishbein and Ajzen 1975). And how many neighbors should be considered? Whereas considering too few neighbors is suboptimal because of the artifacts it tends to produce, including too many usually adds too much noise. Empirical research thus suggests that between 20 and 50 "taste neighbors" is a good compromise.

Let us exemplify this by predicting (albeit with fewer-than-ideal neighbors) how much Mark would like GBU, using the data from Fig. 12.8. In this case, only two other users have rated GBU who have rated other films that Mark (M) has also rated. These two can be considered as Mark's taste neighbors: his correlation with Nancy (N) is a perfect 1.0, whereas the one with Thorsten (T) is -0.28. We adjust the rating for Mark's general tendency to judge films positively, as expressed in his average rating $\bar{r}_M$ across all films in the data set. The corresponding equations look as follows (with $r =$ rating and $s =$ similarity/correlation):

$$p_{Mark,GBU} = \bar{r}_M + \frac{s(M,N) \times \left(r_{N,GBU} - \bar{r}_N\right) + s(M,T) \times \left(r_{T,GBU} - \bar{r}_T\right)}{|s(M,N)| + |s(M,T)|}$$

$$p_{Mark,GBU} = 3.33 + \frac{1.0 \times (1 - 2.25) + (-0.28 \times (5 - 3.8))}{1.0 + 0.28}$$

$$p_{Mark,GBU} = 2.09$$

In essence, because Mark's movie taste is (much) closer to that of Nancy (who does not like GBU), the user-to-user filtering algorithm does not recommend GBU to Mark, even though Thorsten considers it a masterpiece. Keep in mind that this is a *very* simple illustrative example—correlations are only stable when many more shared observations exist. In reality, reliable results simply cannot be generated with only two neighbors.

User-to-user recommenders have potential for serendipity in that any surprise discovery made by neighbors has a chance to spread to other users via the algorithm. But even using a large database of users and items, the approach is troubled by the cold-start problem: if new users have not rated a sufficient number of items, the system will not be able to produce any meaningful recommendations for them. However, in practice, an even

bigger problem for user-to-user filtering is that it is enormously resource-intensive. Calculating the similarities between the users in a large database requires a large amount of computing time (e.g., for a database with only 100,000 users, 100,000 * 999,999/2 = 49,999,950,000; almost 50 billion similarities have to be calculated). As similarities change with *every* single product rating by a user, this strongly limits the usability of this type of system in many entertainment settings, despite the ever-growing power of computing hardware.

Item-to-Item Collaborative Filtering

A practical alternative for collaborative filtering is item-to-item recommenders—they are used by many retailers and websites, including Amazon and Spotify. Item-to-item, or item-based, recommenders use similarities between rating patterns of *products* (or items), rather than between individual users (Ekstrand et al. 2010). If two items are both liked by one group of users and both disliked by a second group, then the items are considered to be similar for consumers preferences. It is then expected that a user who likes one item but does not know the other will also like the latter (because of its preference-related similarity to the first item).

Based on the illustrative data set from Fig. 12.8, we can calculate product similarities between our five example films. How is this done in such an approach? A robust and often-used measure is cosine similarity, but let us use Pearson correlations once more for simplicity. This results in the item-to-item similarity matrix that is shown in Fig. 12.9.[342] The matrix illustrates that, in our miniature data set, users who like NOTTING HILL also tend to

	NOTTING HILL	TERMINATOR 2	PIRATES …	LOVE ACTUALLY
NOTTING HILL	1	.50	*cbc*	.95
TERMINATOR 2		1	.00	-.57
PIRATES …			1	-.28
LOVE ACTUALLY				1

Fig. 12.9 Similarity matrix for four movies based on rating patterns
Notes: Authors' own illustration. Values are pairwise Pearson correlations derived from data in Fig. 12.8 on p. 633. cbc = The correlation cannot be calculated because of a constant value for one film. Brands are trademarked.

---

[342]Please note that we have to leave GBU out because there are not sufficient observations for it.

like Love Actually, whereas fans of Terminator 2 tend to dislike the latter film.

Let us now use the correlations between products/items to predict how much Mark would like Notting Hill (= NH), another movie he has not seen; just like we did above with the user-to-user filtering for GBU. We can use the available similarity information and his ratings of two movies, namely Terminator 2 (T2) and Love Actually (LA). This leads us to the following equations (the notation remains the same as above):

$$p_{Mark,NH} = \frac{s(NH,T2) \times r_{Mark,T2} + s(NH,LA) \times r_{Mark,LA}}{|s(NH,T2)| + |s(NH,LA)|}$$

$$p_{Mark,NH} = \frac{0.50 \times 5 + 0.95 \times 1}{0.50 + 0.95}$$

$$p_{Mark,NH} = 2.38$$

Item-to-item filtering would thus not recommend Notting Hill to Mark, mainly as a result of him not liking Love Actually—the movie has a very similar rating pattern as Notting Hill by the other users in the database. The recommendation level is still higher for Notting Hill than his rating for Love Actually—this is because Notting Hill also shows a decent rating overlap with Terminator 2, a film which Mark admires.

A major advantage of item-to-item over user-to-user filtering is that it requires much less calculation time and effort. In theory, when a user changes or adds a rating, it might affect the similarity between items, but in practice, at least in constellations with large user and rating numbers relative to the number of products, similarities between products will be relatively stable. The cold-start problem is also at least partly resolved here; a new user can get solid recommendations after having rated very few products himself, because information about the similarity between products has been generated solely from other users' ratings. (In practice, the system can recommend "similar" products once the user names, or clicks on, one product he or she likes.)

However, the cold-start problem still exists for new *products*, because their similarities with other products are yet to be determined. One approach to overcome this issue is to use so-called "content-based" recommendations, alone or in combination with collaborative approaches in hybrid models—we discuss this in the next section. And whereas user-to-user recommenders provide room for serendipity (because of surprise selections of individual

"neighbors"), this is much less the case for item-to-item recommenders, which average ratings between products across consumers.

In summary, collaborative filtering approaches are potentially powerful. They align nicely with consumer behavior theories and can make use of rich insights buried deep in large databases. A major stream of recommender research has been dedicated to improve recommendations by reducing overlap between items as well as users, identifying underlying, latent item dimensions and user segments. Such attempts are often associated with the concept of matrix factorization (singular value decomposition, in particular), a method which was popularized by its use in the algorithm than won the "Netflix Prize," a $1 million award given to the team that increased the prediction accuracy of the former DVD-rental service's recommender algorithm by at least 10% (Gower 2014; see Koren et al. 2009 for a general overview).

In any case, all collaborative filtering approaches aim to complete matrices, and the fewer empty cells a recommender matrix has, the more information can be used for generating user-to-user or item-to-item recommendations. The desire to "fill in the matrix" was a driving force behind Netflix's switch from metric rating data to only asking subscribers to provide "thumps up/down" information. The firm presumes the expected increase in the number of ratings (because of the ease-of-use of the new scale) will overcome the loss of detail associated with switching to a less-granular scale (Goode 2017).

Let us now look at the strength and weaknesses of another type of automated recommender systems which trade information on user preferences against product expertise: content-based recommendations.

Content-Based Recommendations

Content-based recommendation systems share their basic logic with item-to-item collaborative recommenders: they try to suggest new products to a user based on the products he has liked before. But in contrast to item-based recommenders, which identify similar products via the overlap in user ratings, content-based recommenders use the inherent attributes of a product, not other consumers' subjective judgments, as the source of similarity. Products whose attributes match those of products liked by a consumer are recommended, while those with different attributes are not.

The approach combines three key steps (Lops et al. 2011). First, the content of products has to be analyzed and the product categorized accordingly. This is crucial because the attributes used for categorization are the ones that define a product's fit with a consumer's preferences. Second, the system needs to learn about a consumer's preferences regarding the attributes. This is usually done based on the consumer's previous behaviors, expressed either explicitly when rating products or implicitly based on the kinds of products that the consumer has purchased or rented (or at least searched for online). The third step then is matching a product's attributes and the consumer's preferences derived in the previous step to enable the actual provision of recommendations.

Content-Based Recommenders: An Example

To illustrate how content-based recommenders work, let us return one more time to our small exemplary data set of five movies. We have now coded the movies with regard to four attributes: their romantic appeal, level of action, average consumer rating, and runtime. We then transformed each score into a 1–5 scale; Fig. 12.10 shows the results for the five movies.

We once more use Pearson correlations to calculate the similarity between movies (although other criteria could also be used), this time based on their attributes. As we did in the case of item-to-item recommenders, we then combine similarity information for two films which Mark has seen and his ratings for them to determine his predicted rating for Notting Hill. For this purpose, we choose Love Actually (the most similar to Notting Hill, with an almost perfect correlation of 0.99) and Pirates of the Caribbean (PIR; the most dissimilar film to Notting Hill—the correlation between

	Action content	Romantic appeal	Consumer rating	Runtime
Notting Hill	1	5	3.50	3.85
Terminator 2	5	1	4.25	4.26
GBU	4	1	4.45	5.00
Pirates …	5	2	4.00	4.44
Love Actually	1	5	3.85	4.19

Fig. 12.10 Matrix with attribute ratings for five movies
Notes: Authors' own illustration. Values are our own codings of the five movies, transformed into a 1–5 scale (where 1 is the lowest and 5 the highest level of the respective attribute). Brands are trademarked.

the two films is a strongly negative −0.82). Using the same equation as we did for item-to-item recommenders results in the following[343]:

$$p_{Mark,NH} = \frac{s(NH,LA) \times r_{Mark,LA} + s(NH,PIR) \times r_{Mark,PIR}}{|s(NH,LA)| + |s(NH,PIR)|}$$

$$p_{Mark,NH} = \frac{0.99 \times 1 + 0.82 \times 2}{0.99 + 0.82}$$

$$p_{Mark,NH} = 1.45$$

This approach results in a predicted rating of 1.45, which (again) suggests that recommending NOTTING HILL to Mark might not be a great idea.

Challenges for Content-Based Recommenders:
The Critical Role of Attributes

Our approach might look intuitive; it shows that content-based recommenders largely circumvent the "cold-start" problem that recommenders usually deal with. As with item-to-item recommenders, only few consumer reactions are needed to find similar products to the ones a consumer prefers. And, as similarities between products are based on expert codings (rather than consumer reactions) in this case, no cold-start issues hamper the integration of new products. In addition, content-based recommendations can be transparent, because the reasons for suggesting a product can be explicitly provided, something that is hardly possible with collaborative filtering, where the only explanation can be that an item is recommended because either "people like you" have rated it highly or because you like "similar" products.

But our example also points to potential challenges for content-based recommendations, most of which have to do with the selection of attributes and their integration. The usefulness of content-based recommenders is strongly influenced by the attributes which are used to determine similarity among items: Even if you agree that the attributes we selected in our example are relevant for consumers' preferences, it seems similarly problematic to believe that this set of attributes is comprehensive: we discuss *many* others in this book that influence consumers' decision making, but leave them out here.

---

[343]Please note that to account for the negative correlation between Notting HILL and PIRATES OF THE CARIBBEAN, we transformed Mark's evaluation for the latter film from 4 (the second *best* category) to 2 (the second *worst*). By doing so, we assume that a strong *dis*similarity between items has the inverted effect of a strong similarity (i.e., "if a product is highly dissimilar to one I like, I will not like it").

The closer chosen attributes link with consumers' preferences, the more powerful recommendations will be. In our example, we use highly aggregated genre labels such as romantic appeal and action content. Doing so results in predictable ratings: If a consumer has liked one romance movie, he or she will receive recommendations for other films from the same genre. That is why in practice, content-based recommenders often struggle with the serendipity problem, leaning toward satiation—users' lust for sensation is not adequately addressed by them.

But aggregated attributes not only bore us, but they also often simply miss the point why consumers like a particular piece of entertainment. A lot of people like Clint Eastwood's western UNFORGIVEN, not because of a general preference for westerns, but for its star. So recommending a western with John Wayne would only disappoint them. Some of those who like the film because of its star might not be content with other films Mr. Eastwood has directed and appeared in, such as BRONCO BILLY and MILLION DOLLAR BABY. And others may like the film because of its subtle connections to the Italo westerns directed by Sergio Leone and Mr. Eastwood's personae of the Man with No Name in films like A FISTFUL OF DOLLARS—whereas these consumers reject most other westerns and other films by Mr. Eastwood.

In addition to such stable preferences, contingencies also influence what we like (or don't like) at a certain moment: maybe you are among those who can work most effectively while listening to low-key classical film music? Whether a streaming service's content-based recommender can offer a fitting program depends on its ability to separate "low key" soundtracks from others which might be popular, but distract us from our work—take John Williams' spectacular INDIANA JONES theme as an example. In other words, the crucial question is what a streaming service understands when we tell it that we "like" a song we just heard—the value of the information that the service "will play more song like this one" depends on how well it understands *why* we like the song.

To address the issue of fine-grained, contingent preferences, companies such as Netflix and Pandora have invested enormous amounts into a highly differentiated coding of their products. Netflix internally has called its approach "Quantum Theory," which encompasses almost 77,000 unique labels of "micro-genres" for their content. Employees who assign tags to films and series receive a 36-page training document (see Madrigal 2014—a fascinating and highly recommended read, by the way). The firm considers conventional genres, in the words of its vice-president Todd Yelling, to be "just wrappers." It is the micro-genres which are intended to recommend films that are similar to the ones a user loves.

In some ways, this approach even aims to *add* serendipity to content recommenders: Netflix wants to "break these pre-conceived [genre] notions

and make it easier for [users] to find stories they'll love, even in seemingly unlikely places" (quoted in *Netflix* 2017). Similarly, music streaming service Pandora has musical experts analyze each song using up to 450 distinct musical facets—their very own "Music Genome Project" (Lasar 2011). By going into this level of detail, the firms hope to minimize the limitations of content-based recommendations while harvesting their strengths.

Let us end this section by naming some further challenges for content-based recommenders. How did we determine the values of the attributes for a given movie in our example? Whereas "objective" information exists for some attributes (such as the films' runtime), this is not the case for others, such as consumer ratings (which we took from the IMDb, but which may differ on other websites) or genre: how romantic is Pirates of the Caribbean *really*? We gave it a 2 on our 1-to-5 romance scale; but what would *you* rate it? It seems also questionable to assume that attributes are of equal importance, and that this attribute importance is the same across consumers—an issue that can be addressed by adding weighting parameters.

Hybrid Approaches: The Best of All Recommender Worlds?

Scholars have also tried to integrate the different recommender approaches, combining their respective strengths while minimizing their limitations. Collaborative filtering suffers, to varying degrees, from the cold-start problem of handling new users and items and the lack of detailed explanation. And content-based recommenders do not make use of the information that is available about the other users in a database and also often have problems in offering fresh sensations.

Popular ways to combine the different approaches include the following (see Burke 2002 for more details):

- *Weighted hybrid recommenders*. The ratings of multiple recommenders for an item are combined via a weighting routine, producing a weighted average recommendation.
- *Switching hybrid recommenders*. The ratings provided to the user come from different recommender systems, which are chosen depending on the characteristics of the situation (e.g., the number of existing ratings for an item or by a user). A challenge here is to determine the "right" switching criteria.
- *Mixed hybrid recommenders*. The user is provided with multiple ratings for a product generated by different recommenders. Such an approach carries the risk of overloading the user with "unnecessary" information. A variation of this approach is to use multiple recommenders, but to

present their respective results not jointly for a specific product, but separately for each recommender (first the top recommendations by content-based recommenders, than by item-to-item filtering etc.). This allows the consumer to decide which recommender to use. Such an approach has been employed by Netflix (Gomez-Uribe and Hunt 2015).

- *Sequential hybrid recommenders*. Different ways have been suggested to combine several recommenders in a sequential order. For example, the output of the initial recommender (such as a content-based approach) is used as input for a second recommender (such as a collaborative filtering approach).

Are such hybrid treatments of recommendations effective? Several studies point to the potential gains of hybrid recommenders; for example, in Marx et al. (2010), we combine content-based and user-to-user collaborative filtering using switching criteria, which increases the fit of the pure content-based recommender by 16% and the fit of the collaborative filter by 10%. But those who offer recommenders always have to weigh the benefits in effectiveness against the higher complexity, longer processing times, and higher costs of such approaches, which limits their usefulness in many practical settings.

So far, we have focused on the algorithms that create recommendations, the backbone of automated recommender systems, if you will. In the following section, we will show that the value of such recommendations for both consumers and those who offer them is influenced also by several other factors.

Recommenders are Way More Than Algorithms: Beyond Matrix Completion

> "Predicting movie ratings accurately is just one aspect of [Netflix's] world-class recommender system."
>   —Amatriain *and* Basilico *(2012)*

Algorithms are essential for providing powerful recommendations, and the clear majority of scholarly research on recommenders has focused on the development of "better" algorithms, i.e., those that predict user preferences with minimal error. But maximum accuracy at the matrix completion task is far from the only thing that matters for recommenders to be successful (Jannach et al. 2016).

Although it paid a million dollars for it, Netflix has never implemented the algorithm that won the Netflix Prize. The reason is that the firm sees

higher gains from other parts of the recommender challenge: "the days when stars were the focus of recommendations [at Netflix] have long passed" (Gomez-Uribe and Hunt 2015).[344] Recommender research and practice have recently turned toward other aspects of recommenders and their implementation at the user interface. Let us name some of those issues that recommender providers in the entertainment world also have to deal with— and the relevant knowledge that scholars have gathered.[345]

Contextualization

Whether we like an entertainment product depends not only on the product itself, as recommenders usually imply. Instead, our liking is influenced by the context in which we use a product. Adomavicius et al. (2011) provide sound arguments that the following specific context dimensions are of particular importance for the enjoyment we derive from an entertainment product—and that taking them into account when making product recommendations can strongly increase the recommendations' usefulness.

Physical context. Our consumption patterns differ with time and place. For example, the music we like to hear when we wake up usually differs from what we like during the day or when exercising. A consumer who generally likes a classical composition might find its recommendation inadequate in the morning hours. Also, whereas a consumer might love epics from David Lean or Paul Thomas Anderson for long-weekend nights, he might prefer much shorter programs for weekday evenings. We know that enjoying the four hour-long Lawrence of Arabia on a Wednesday evening could be a treat, but that we might pay a high price for the lack of sleep during the Thursday morning's work meeting.

Historical context. Any evaluation of an entertainment product is part of a sequence of consumption acts; there is always a previous novel as there is always a next song. The sequence of events influences a consumer's reaction to a particular product, such as when the similarity of previous experiences triggers satiation with a product. After having experienced five sequels that

---

[344]Let us add that another reason for Netflix's decision to not implement the "best" algorithm was that the firm changed its business model shortly afterward, from renting DVDs to streaming films and shows. Whereas recommenders have remained crucial for the firm, the new business model introduced a different usage context for recommenders; the now immediate link between recommendations and consumption made other aspects more relevant than prediction accuracy, some of which we discuss in this section.

[345]Jannach et al. (2016) provide a highly informative overview about these non-algorithmic challenges.

satisfy a consumer's desire for familiarity, he or she might have a stronger appetite than before for new sensations from more-original content and would be thankful for being offered corresponding recommendations. Also, dissatisfaction with previous products can lower the expectation level for the next one. The fact that the second and third Rambo films were such disappointments might provide an explanation for the relatively positive reception of the fourth entry in the Rambo series.

Emotional context. Managing our mood and emotions is often a motive for consuming entertainment, so the value of recommendations will also vary with our emotional needs and desires at a given point in time. Anger requires a different soundtrack than happiness, sadness demands other films or books than does being in love. Our entertainment choices differ in response to certain emotional states, but standard recommenders have no room for our emotions.

Social context. Are we going to the movies alone or with a friend? Earlier in this book, we discussed how social context influences our decision-making process. Recommenders would benefit from accounting for social context instead of assuming we always act in isolation. Recommending The Salvation, a Danish homage to Leone's westerns *dall Italia*, to a dedicated fan of the latter films makes sense in general—but much less so if the person is going on a date with someone who is "allergic" to subtitles and westerns.

Recommendation scholars have suggested considering such context information when designing recommenders, calling for context-aware recommender systems ("CARS"; Adomavicius et al. 2011). How can such contextualization be achieved? One approach that has been suggested is to use contextual filtering: in every situation, recommendations are generated only based on similar situations; all other ratings are filtered out and ignored (i.e., pre-filtering). For example, for someone who wants to watch a movie on Saturday with a group, only those ratings collected for the weekend and/or for group consumption are used for recommending films. Alternatively, results can be filtered post hoc based on context criteria (i.e., post-filtering).

A second approach for "CARS" is contextual modeling. Here, the recommendation function includes contextual information by using statistical techniques, such as decision trees and hierarchical Markov chains. Netflix makes use of this approach, estimating Markov chains. According to Neil Hunt, the firm's former Chief Product Officer, Netflix records "what you have seen first, then what next and what after next. Then we compare this with other people's behavior and calculate 'transition probabilities'" (quoted in Brodnig 2015). Both approaches require extensive data collection, as every cell in the matrix needs to be tied to the specific context in which it

was entered. But subscription services such as Netflix and Spotify have access to such data, as they know the timing of when a consumer watched a movie or listened to a song which he then rated.

Also, content-tagging can help (is a song sad, or angry, or happy?), which can then be linked with consumers' moods in a given situation via content-based recommenders, if consumers are willing to reveal their mood. Spotify, for example, has asked listeners for their current mood (Jannach et al. 2016). Another way to make use of content tags is to let listeners self-select mood-adequate titles by offering "context playlists," as Amazon Music does with lists such as "in love" and "relaxed breakfast." Doing so also enables a content-based recommender to address the consumer's desire for "low-key" film music in a working situation; a challenge we have mentioned earlier.

For addressing the social context of a recommendation, scholars have proposed group recommenders that combine the preference data of the different members of a group (O'Connor et al. 2001). Our own suggestion is a two-step approach, where in the first step the individual preferences of each group member are used to generate predicted ratings of potential product alternatives, which are then, in step 2, transformed into collective ratings for the group as a whole, essentially using a weighing mechanism (Hennig-Thurau et al. 2012a).

Does such an approach offer "better" recommendations than a standard one? We tested this approach with two lab experiments with 460 consumers who had to actually watch a recommended film; the recommendations were created with a user-to-user collaborative filtering recommender based on about 4.8 million ratings from a popular German movie website. The results provide evidence that a group recommender can indeed outperform a "single" recommender: the group's satisfaction with the recommended film (the average of both group members) was more than one point higher for the group-recommender condition (on a 0–10 point scale) than when the group watched the film recommended to the "agent" (who made the choice for the group) in isolation. Not surprisingly, most of this effect came from the "partner's" satisfaction with the recommended film.[346]

---

[346]In the described scenario, the group members had to watch the film recommended by the recommender system. When we allowed the group members to choose freely either to follow the recommendation or to "overrule" it, the group recommender still outperformed the single recommender in terms of group satisfaction with the film. But this effect held only when the group members liked each other and the agent followed the recommender and chose a compromise film over one that maximized his own preferences.

Design and Interaction

Scholars have also highlighted the importance of the usability of a recommender system, as determined by the design of the interface (Jannach et al. 2016). This affects the input of customer information, but also the presentation of recommendations.

Regarding input, the core challenge is to let the users of the system enter ratings easily and intuitively. In the case of "passive" data which derives the consumer's preferences from his or her usage behavior (does a user watch the whole movie? Does he or she listen to the whole song?), no action is required from the consumer, although *a lot* of artificial intelligence is needed to transfer behaviors into preferences. For "active" data, the rating scale matters, as does the input mechanism that must account for consumers' need for convenience (as an example, see Netflix's recent scale change from 5-star to binary; Goode 2017). In this regard, voice-based systems, such as Amazon's Echo device, might provide new ways to increase the usability of recommender systems.

With regard to the presentation of recommendations, it is mostly about how intuitive the results are presented to the user by the system. To address the varying preferences of its consumers for the use of recommenders, Netflix retired it single-recommender approach and substituted it with a collection of different algorithms "all serving different use cases" (Gomez-Uribe and Hunt 2015).

Trust

Many consumers are skeptical about artificial intelligence-powered solutions, so a major hurdle for every recommender system is to earn the trust of its users. Trust involves competence *and* benevolence (e.g., Sirdeshmukh et al. 2002), and both are relevant here. Are the recommended products truly the best for the user, i.e., is the recommender algorithm smart enough to find alternatives that deserve his or her time? One way to demonstrate a recommender's competence is to provide explanations for the recommended titles (e.g., "Titanic is recommended to you because you like big-budget Hollywood movies and films with Leonardo Di Caprio"; see Marx et al. 2010 on the empirical power of such explanations).

And is the platform provider really acting in my best interest, or are recommendations given in a way that primarily benefits the platform? With its increasingly high spending for original productions, Netflix has begun to allocate a disproportional share of its home screen to those films and series,

even when recommendation scores are rather low (Bishop 2017). Such a practice carries the risk of threatening consumers' trust in the benevolence of the service's recommender system.

How "Good" are Recommender Systems?

We have seen that the other kinds of "earned" media that provide consumers with information about a product's quality actually influence the consumers' decisions. How about recommender systems? Do they offer consumers insights that they find valuable? Let us take a look at whether the usage of recommenders helps consumers to make "better" decisions and the role of recommendations for such decision making and product success.

A rare study that compares the predictive performance of recommenders to situations in which consumers have no access to them is Krishnan et al. (2008). In this study, the scholars compared the accuracy of film ratings for 14 consumers by a recommender system with the predictions of 50 human raters (who had access to each consumer's past ratings). The recommender, a item-to-item collaborative model of the MovieLense recommender that incorporated 15 million ratings at the time of the study, predicted ratings with a 19% lower average error (MAE) than the human raters—but the human raters' predictions were biased by a number of outlier predictions, and the predictions of the recommender system were closer to actual ratings for less than half of the 14 consumers.

In our own exploration of the performance of group recommenders, we compared the predictions of our user-to-user collaborative filtering approach with a condition in which the "agent" (i.e., the group member who made the choice) lacked access to a recommender system (Hennig-Thurau et al. 2012a). In the latter condition, the agent had to choose a movie based solely on movies' titles, countries of origin, main genre, and a thumbnail poster (with no access to the Internet or word of mouth). We learned that the chosen (and watched) movie was more positively evaluated by the group when group members liked each other and when the agent held a positive attitude toward recommender systems, in general. But for a standard (single) recommender, we found no difference in movie satisfaction: a result which illustrates that the power of recommenders should not be taken for granted across contexts and conditions.

Meiseberg (2016), in her study of book success, did not look at the quality of recommendations or users' satisfaction with recommended products, but their practical impact on decision making: does a recommendation for a product on Amazon's website influence the product's sales at Amazon?

In her analyses of a large book sample (in which she controlled also for a large number of other "success drivers"), she used the sales of all books in which a certain book title was recommended as measure of recommendation intensity for the title. Her results demonstrate that, at least in the context of her study, product recommendations do indeed trigger book sales—and do so to a remarkable degree: correlations are higher than those for word-of-mouth and prices, and comparable to the book title appearing on TV. The recommender variable has the strongest effect for the Top 20 selling books in the sample, followed by the lowest 20% quantile.

This impact of recommenders on product choices is also reflected in Netflix's disclosure that only about 20% of the hours consumers use the service result from searches on the site, while about 80% are inspired by popularity rankings (see our section on "success-breeds-success" effects elsewhere in this chapter) or recommenders (Gomez-Uribe and Hunt 2015). Among those 80%, the share of personalized recommendations that result in the consumer watching a film or series (the "take-rate") is, for attractive titles, up to almost four times higher than for recommendations based on popularity. The two Netflix managers Carlos Gomez-Uribe and Neil Hunt also claim that personalized recommendations at Netflix lead to "lower subscription cancellations rates" (p. 13), but they do not provide empirical details or evidence.

Although recommender systems are operated by a platform, they still use data that are "earned" from consumers, just like word of mouth and action-based cascades. In the final sections of our discussion of entertainment communication, we will now look at the role of other stakeholders of entertainment, namely the professional critics and industry peers (who hand out awards) whose judgments might also influence the success of an entertainment product.

Professional Reviews

"I heard that some studio insiders want to hold off critic screenings until opening day or cancel them all together."
—Anthony D'Alessandro *(2017) on film managers' reaction to the disappointing opening weekend results for* Baywatch *and the fifth* Pirates of the Caribbean *movie*

"At his very best, a critic is a cheerleader for films that need support."
—*Film critic* A. A. Dowd *(2015)*

One of the most obvious consequences of entertainment's difference from other products is that media outlets employ journalists to write about them.

It is entertainment's cultural character that makes it so media-worthy, granting films, books, music, and video games space on the pages and websites of prominent newspapers and magazines. From this "media-worthiness" of entertainment, the question arises regarding what role the judgments of professional reviewers play for the commercial success of the products they review.

Whereas this might be an age-old question, it is still highly controversial. Entertainment producers often blame critics' negative comments for the failure of their products—see the introductory quote by Mr. D'Alessandro above. The second quote by Mr. Dowd suggests that critics themselves are much more skeptical regarding the impact of their own work. And they have been that skeptical for quite a while: when we asked Peter Körte, a leading German film critic, what he thinks about the influence of his own writings on audiences, he quoted what French reviewer André Bazin wrote about his occupation in the 1950s: "Film criticism is when you stand on a bridge and spit into a river" (Körte 2009, p. 194).

In this section, we turn to *Entertainment Science* scholars to learn about whether critics influence consumers and, eventually, product success. As with many other questions we have addressed in this book, finding a credible answer econometrically is far from trivial, requiring care and expertise. Critics' judgments, like consumers' word of mouth, overlap with an entertainment product's "true" artistic qualities, so that correlations must not be confused with causal effects. But scholars offer ambitious attempts to isolate such causal effects (i.e., critics as "influencers") from spurious ones (critics as "predictors"), whose main insights we will summarize here. We will then again go beyond "average effects," shedding light on moderators and related issues.

Review(er) Effects on Consumers: The "Influencer Versus Predictor" Controversy

Expert reviews are a source of third-party evaluative information about the quality of entertainment products. Traditionally such information could be accessed by consumers only via newspapers and magazines on a review-by-review basis, and the review literally disappeared the day after it was published. But the Internet has changed the availability and format of reviews. The Internet has created an archive of reviews which, if they are not locked behind a publication's pay wall, can be accessed at any time, independent of their publication date. The Internet has also given us aggregator websites

and services such as Rotten Tomatoes (containing reviews of movies and TV shows, founded in 1998) and the more selective Metacritic (for movies, TV shows, games, and music; launched in 1999) that provide consumers with links to individual reviews as well as summary signals, such as mean scores and the number of reviews. Take the movie ALIEN: COVENANT: at the time of this writing, Rotten Tomatoes reports a mean rating of 6.3 (out of 10) and a "Fresh" score of 68% based on 324 different reviews, whereas Metacritic has calculated a weighted average of 65 (out of 100) from 52 reviews. Both aggregators are owned by entertainment conglomerates; Rotten Tomatoes is a joint venture of Universal/Comcast and AT&T/Warner, and Metacritic is part of CBS/National Amusements.

Professional critics have been called "institutional gatekeepers" because they screen the entertainment products that are offered by producers, "winnowing them into a much smaller number of select goods from which everyday consumers then choose" (Hsu 2006a, p. 468). But do expert reviews actually *influence* consumers, and to what extent? When scholars have run regression models with entertainment product success as dependent variable, they often include a product's average review score from critics, as calculated by one of the aggregators.

The estimated parameter in these models is usually positive and significant, which demonstrates that the quality assessments of professional critics go along with a product's commercial performance on average.[347] However, as we have stressed throughout this book, associations are not necessarily causal, and providing evidence that reviewers' tastes are *associated* with what consumers buy should not be confused with evidence for a causal effect by the reviews themselves.

Initial Insights: The Eliashberg and Shugan Study

The first scholars who dived deeply into the issue of review effects were Eliashberg and Shugan (1997): they coined the terms "influencer effect" (for a causal link between critics' reviews and a movie's box office performance) and "predictor effect" (for spurious correlations in which the statistical relationship is not causal, but in which positive reviews *indicate* a film's success because the

---

[347]Examples of studies that find professional reviews to be positively linked to entertainment product success as part of a regression-type analysis, but do not address the more complex issues we discuss below in their empirical approach, are Clement et al. (2014) and Elberse and Eliashberg (2003) for films, Marchand (2016) and Cox and Kaimann (2015) for games, Clement et al. (2007a) and Schmidt-Stölting et al. (2011) for books, and Lee et al. (2003) for music.

product quality underlying the positive reviews would also drive success).[348] They conducted the first empirical attempt to separate those effects by explaining the weekly North American box office of 56 films released in 1991–1992 (i.e., in the pre-Internet era) with the percentage of positive and negative reviews via regressions. Eliashberg and Shugan found that the regression parameters for positive and for negative reviews are only significant from week 5 onward after a movie's release, but insignificant in the first four weeks.

They considered this finding as support for the dominance of the "predictor effect," because a causal "influencer effect" would be reflected in higher parameters in the earlier weeks (when review information is still fresh, as most reviews are published around the release of a film). But there are alternative explanations which Eliashberg and Shugan's approach does not rule out: think of early audiences often being fan boys who might find reviews less informative than later moviegoers. Also, quality information, the core of any critical review, might have needed more time back then to spread among consumers (as we have shown in our discussion of the timing of word of mouth earlier in this chapter). Basuroy et al. (2003) offer another explanation: maybe the small data set is to be blamed. When they replicated Eliashberg and Shugan's approach for a larger set of films from the same time period, they found significant review parameters for *all* weeks.

What We Have Learned About Review Effects Since Then

From the foundation of this seminal work, other scholars have developed alternative, more sophisticated empirical designs and used even larger data sets in attempts to isolate reviews' spurious "predictor effect" based on movie quality from a truly causal impact.

Also with movies as product category, Reinstein and Snyder (2004) use a "difference-in-differences" approach to circumvent the problem of a spurious correlation. To make use of the *timing* of a review, the authors studied whether films that received a positive review during their opening weekend experienced a higher opening box office than other films (those that were reviewed later and/or received a negative review). The scholars' measure of professional ratings was a single review source: the ratings by prominent

---

[348]See our discussion of the link between product quality and commercial success in entertainment in Chapter 7 about the "entertainment experience."

U.S. reviewers Gene Siskel and Roger Ebert in their television show Siskel & Ebert/At the Movies. Their data set encompasses 609 films released in North American theaters during the long period of time (from 1982 to 1999) that Mr. Siskel and Mr. Ebert's show was on air.

Using an OLS regression (in which they control for distribution size and use movie guide book author Leonard Maltin's much later published ratings as a proxy for the films' "true" artistic quality), they find that films that got "Two thumps up!" from the reviewers during their opening weekend generated an average of 28% higher box office over their first three days than later- and/or negatively reviewed films, despite the fact that the TV show was only aired on the Saturday of a film's opening weekend.[349] Reinstein and Snyder's results indicate that the review effect is non-linear: if only one of the two reviewers gave a positive rating, the effect drops by more than half and becomes insignificant.

One limitation of their study, beyond considering only a single, most-popular, no-longer-existing review source, is that movies of higher quality had a higher chance of being reviewed on their opening weekend by Mr. Siskel and Mr. Ebert, which could have blurred the causal and spurious effects.[350] Whereas Reinstein and Snyder try to reduce the effect distortion by including the "true" quality variable in their regression, we used a different approach when studying the issue ourselves. In an investigation of all 1,370 fictional films that were released in North America from 1998 to 2006 (and generated at least $1 million at the box office), we isolated the unique part of the reviewers' ratings that was not the same as the consumers' quality perceptions (Hennig-Thurau et al. 2012b).

We did this by running a separate ("auxiliary") regression in which we explained the average reviewer rating of a movie (from Metacritic) with the movie's consumer ratings (from Netflix and Yahoo). We then used the residual (the part of the Metascore *not* explained by consumer ratings) to explain movies' box office, along with a large number of "success drivers" including advertising and buzz. In one regression, the opening weekend box

---

[349]The regression also included one variable that measured whether a film got a positive judgment from at least one of the two reviewers on its release weekend. Let us note that the confidence level for the finding was only 90%, not the usual 95%; this also applies to most of the subsample-related findings in Reinstein and Snyder's study we report later in this section.

[350]The higher "inherent" quality of the early reviews could have erroneously boosted the review effect. The authors empirically show that such a quality effect indeed exists in their data set in a separate probit regression.

office served as the dependent variable; in the other, the dependent variable was the box office generated in the weeks *after* the one in which the movie was released. We find that the unique part of professional reviews indeed impacts product success, but only does so later on in a movie's theatrical run. Specifically, the effect of reviews is not significant for the opening weekend, but positive and significant for the later box office.

And how strong is the latter effect? In our data, a ten-unit increase in professional reviews (e.g., from a Metascore of rating of 60 to 70) corresponds with a 15% increase in box office after the opening week. When including a squared rating parameter in the regression, our findings also corroborate the non-linear nature of review effects suggested by Reinstein and Snyder's results. Figure 12.11 shows that highly positive reviews provide a disproportionate benefit to a film's financial performance.

Whereas all major reviews for movies are published at the same time (around a film's release date), reviews for other entertainment products such as books tend to be published over a longer period. Sorensen and Rasmussen (2004) make use of this characteristic by linking book reviews by the New York Times with sales changes in the weeks following the review's publication.
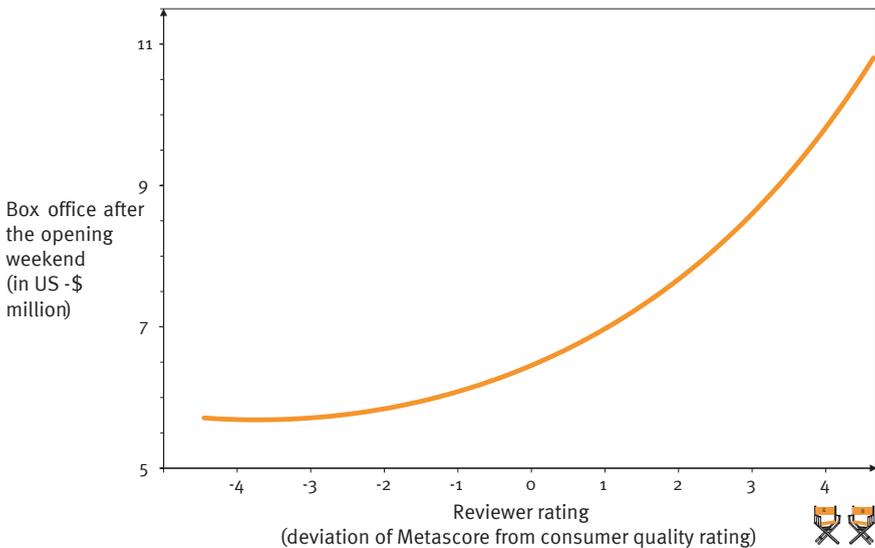

Fig. 12.11 The non-linear relationship between professional reviews and the box office after the opening weekend

Notes: Reprinted with minor modifications with permission by the Journal of Cultural Economics/ Springer from Hennig-Thurau et al. (2012, p. 272). We transformed box office values into raw format from their logarithmized form (which was used in the estimations).

Investigating weekly sales numbers for 175 hardcover books for 2001–2003, the scholars find, via a panel regression in which they control for a book's sales in a previous week (but not for other book variables, such as star author status), that a positive book review in the New York Times causes an enormous 63% increase in next-week sales in the U.S.

Altogether, the skepticism of critics regarding any impact of their writing, as expressed at the beginning of this section by A. A. Dowd and Peter Körte, appears inappropriate. In contrast, entertainment producers' concerns seem justified: at least on average, the ratings of an entertainment product by professional critics have the power to influence its success.[351]

Before we now dig deeper and look beyond average effects, let us address one related issue: can there also be, above and beyond the causal effect of review valence we have looked at here, a publicity (or awareness) effect of professional reviews, in a "all-reviews-are-good-reviews-even-bad-ones" sense? Sorensen and Rasmussen (2004) indeed find that, at least for the set of books they examine, *negative* reviews also have a *positive effect* on sales, which they attribute to the awareness such mentioning raises. But this positive effect is only half as strong than the valence effect and can thus only partially mitigate the negative consequences of being negatively reviewed. For movies, studies which find such a "publicity effect" do not control for the "general" popularity of a film (which exists because of advertising and/or buzz). When we do so, we don't find any success-enhancing effect of the number of reviews written about a film (Hennig-Thurau et al. 2012b).[352]

---

[351]Let us add that this conclusion is also consistent with insights from lab experiments, in which scholars expose consumers to reviews and measure how that impacts them; for example, Wyatt and Badger, as early as 1984, reported a between-subjects design in which consumers who were shown a negative review had a 30% lower interest in seeing the reviewed movie. Whereas such experiments illustrate the value potential of reviews, their contribution is limited by their artificial nature. Media managers are interested in learning about the impact of a specific piece of information in a world in which many other sources of information compete for consumers' attention, whereas lab studies mostly account for only a subset of those alternative sources. In Wyatt and Badger's (1984) study, the professional review was the *sole* information source; in a later study, the authors provided the participants also with information about a film's plot and its stars in a within-subjects experiment with about 200 consumers (Wyatt and Badger 1990). Here they found a 12% increase in interest in a film when shown a positive review, and a 7% decrease when shown a negative one. Separately, the results by Chen et al. (2012) suggest that review effects are not limited to (post-opening weekend) revenues, but can also escalate to the producing studio's stock price—their event study of the professional reviews of 220 movies from 2005/2006 finds, on average, "abnormal returns" of the studio's stock of 12% for above-average reviews and of −24% for below-average reviews. In other words, investors react to critical reviews, and do so stronger to bad reviews than to good reviews.

[352]Please also note that several reviewed books are of a rather "niche" kind, with small advertising budgets, so that awareness effects appear rather more feasible for them than for mainstream-targeted movies or other forms of entertainment (which can be expected to have a higher awareness "baseline.")

Toward an Even Richer Understanding of Professional Reviews: Moderators and Mediators of Their Impact

Scholars have also made steps toward a finer-grained understanding of how reviews affect product-level success for entertainment products, not just averages. Entertainment managers and journalists often argue that products differ in their degree of "review-proneness," and we can provide some statistical evidence regarding what kind of products fall into that category. Beyond product characteristics, studies also shed light on consumers, macro-economic conditions, and the effects of reviews on *distributors*.

Product Factors as Moderators

Scholars have analyzed differences in review-proneness for commercial (mainstream-targeted) versus independent (niche-targeted) products, as well as between genres.

Regarding commercial versus independent products,[353] Gemser et al. (2007) argue that audiences of independent products may focus more on the product's artistic achievements (which are what reviews are mainly about), while commercial products will be judged mainly based on other, less artistically ambitious criteria.[354] Also, commercial products, with their attached stars and strong advertising, offer consumers more quality signals (or quasi-search attributes) than independent products do, which should further limit the importance of reviews for them.

Empirical results confirm such a logic. When Reinstein and Snyder (2004) split their data set of films into wide-released and narrow-released ones (in the study of TV reviews by critics Roger Ebert and Gene Siskel), they find that the review effect exists only for movies that were released narrowly. For these films, a "Two thumps up!" is associated with almost 45% higher opening weekend revenues, whereas *no* effect is found for wide releases. In our own study of movie reviews, we also see that professional reviews affect the box office of narrow releases more strongly (Hennig-Thurau et al. 2012b), and Lampel and Shamsie (2000), for a data set of 409 films from 1991 and 1992 which they analyze with OLS regression, report a

---

[353]We discuss these fundamental product types in our chapter on entertainment business models.

[354]See also our discussion of how the different levels of taste overlap between professional reviewers and mainstream versus niche audiences in the chapter on entertainment product characteristics.

negative interaction effect between a review score and its distribution size on the total North American box office.

In addition, we also find differences for other factors that vary between commercial and independent products: reviews are less influential for films that feature stars, are sequels, and have strong advertising and high buzz. Finally, our results also show that the effect of reviews on later box office is lower for films that open strongly—we assume that early success stimulates action-based cascades, which then crowd out the information offered by professional reviews.

Scholars have also compiled evidence in the context of movies that some genres are affected more strongly by professional reviews than others. Both Reinstein and Snyder (2004) and our own work (in Hennig-Thurau et al. 2012b) support the industry's assumption that dramas are more review-dependent than other genres: Reinstein and Snyder found a 90% boost caused by "Two thumps up!" reviews for this genre. We assume that audiences trust reviewers more when it comes to judging quality as a matter of dialogue than of spectacle; Reinstein and Snyder find no effect of reviews for action films. We find that comedies benefit from a high Metascore even more so than do dramas. Why could this be? Humor is a highly idiosyncratic matter, with the benefits consumers derive from watching a comedy being highly uncertain, so it's the professional reviewers to the rescue for many of us. But not all reviewers provide equally valuable guidance when it comes to deciding what kind of comedy is fun: Reinstein and Snyder do not find a comedy's opening success to be impacted by Mr. Ebert's and Mr. Siskel's ratings.

These insights are a solid start for determining what characterizes a "review-prone" entertainment product. But the product variables studied so far by scholars still encompass quite high levels of heterogeneity—there are all kinds of commercial products, as there are all kinds of dramas. When the movie CHILD 44, a dramatic thriller set in a dystopian Stalin-led Russia, performed poorly at the box office, industry analyst D'Alessandro (2015b) blamed its negative reviews, arguing that "[w]hen you've got spectacularly grim subject matter, you need strong critical support." Are there indeed certain stories and (sub-)genres which depend on reviewers in particular? And which ones? Hopefully, future *Entertainment Science* studies will shed additional light on the matter.

Consumer Factors as Moderators

Different entertainment products appeal to different consumer segments, which might vary in their sensitivity toward professional reviews. One

approach to study such consumer factors with secondary data is to code products by their target groups. We have done this for our data set of 1,370 films, using tags by the (no longer available) movie recommender website Jinni.com. The results show that films targeted at families are systematically less impacted by professional reviewers' ratings than films with other target groups; this effect is found only after the opening week.

More detailed light on the issue is shed by a team of Canadian scholars who collected consumer-level data, again in the context of movies. d'Astous and Colbert (2002) surveyed a convenience sample of 120 Canadian student moviegoers and tested a number of hypotheses regarding their usage of professional movie reviews when making movie-going decisions. The scholars' OLS regression results suggest that professional reviews are used more by consumers who (a) have high knowledge about film, (b) are, in general, more susceptible to the social influence of others, and also (c) have lower levels of self-esteem. And the results of a replication study for which the scholars collected data from about 450 consumers in Austria, Colombia, and Italy (d'Astous et al. 2005), propose that these findings can be generalized beyond Canada.[355]

Macro Factors as Moderators

Are there more general factors, beyond characteristics of products and consumers, that alter the impact of professional reviews on the success of entertainment products? Dhar and Weinberg (2015) use a data set that covers almost 1,700 movies over a long time frame—between 1983 and 2009—and test whether the macro-economic conditions at a given time influence the role that professional reviews have for film success.

Using a consumer sentiment index as a measure of the macro-economic conditions and estimating results via a GMM regression,[356] the scholars interpret their results as support for their expectation that consumers are more sensitive toward professional reviews in times of economic crisis (when spending decisions are made more carefully). In contrast, in times in which consumer sentiment is high, they find that people tend to be less critical when making entertainment choices.

---

[355]For their Canadian sample, d'Astous and Colbert (2002) also report a *negative* impact of film involvement on the usage of professional reviews. We suspect that this might be an artifact though based on the concept's overlap with film knowledge—the effect of involvement is *positive* in both settings in which knowledge was not also included in the regression model.

[356]The scholars run separate equations for box office—the "demand side"—and the number of theaters in which movies are shown—the "supply side."

But Dhar and Weinberg also offer a disclaimer: the small size of the inter-action parameter they estimate suggests that the change in consumer behavior in less-favorable times is hardly of managerial concern, at least when it comes to the role of reviewers. That is different for the "main effect" of the economic condition; we return to this issue when we investigate the timing of the release of an entertainment product in our distribution chapter.

Distributors as Mediators

Finally, there is initial empirical evidence that professional reviews appear to be read not only by consumers, but also by those in the entertainment value chain who decide which products are offered to consumers, such as movie theater owners. In a study of more than 165,000 exhibition decisions for 788 films (specifically, weekly theater-level choices by theater owners in Quebec, Canada, spanning 2002 to 2011), Legoux et al. (2015) find that professional reviews influence the theater owners' programming decisions.

The scholars use the number of weeks a movie is shown in a theater as the dependent variable in their analysis and estimate a discrete-time survival model in which they control for numerous variables (including a film's performance in the previous week). They find that movies which earned excellent ratings from professional critics had, on average, a 37–66% higher probability of still being shown in the theater in the following week, compared to less positively reviewed films. This effect becomes more pronounced over the life cycle of a film; it is also higher for more successful films.

Now let us conclude our discussion of the effects of professional reviews by taking a look at their relevance for consumers' *quality perceptions*. We stressed earlier in this book that judging the quality of entertainment is not an easy task for consumers—not only prior to experiencing entertainment, but also during and even afterward. We have also shown that consumers' evaluations of entertainment are influenced by its popularity, and there is similar evidence for an effect of professional reviews on consumers' *post*-experience judgments of entertainment products.

This evidence comes from an experiment that Wyatt and Badger (1984) ran. The scholars showed 89 U.S. students professional reviews (that the scholars had manipulated to be positive, negative, or mixed) for the comedy movie THE NATIONAL HEALTH. The researchers found that, on a 77-point multi-attribute liking scale, the average rating for those who had read a positive review before watching the film was 59—a number that was significantly higher than the average liking score for those students who had read a mixed review (51) or a negative review (49).

Remember that the actual film watched was exactly the same in all conditions; only the review read prior to watching the film was different. Overall, in an analysis of variance, the review accounted for a remarkable 11% of the consumers' movie ratings. So it looks that reviewers have an even stronger overall impact than what we have reported in this section so far.

Managerial (Mis-)Use of Professional Reviews

"The producing team of BIG DADDY has produced another winner!"
—*Fictitious film critic* Dave Manning *about the film* THE ANIMAL *(Hoaxes.org 2014)*

Professional reviews are, per our definition, "earned" media. But as with other kinds of "earned" media, that does not prevent entertainment managers from trying to influence the role reviews play for the success of their products. There are two major ways to do so: by influencing the availability of reviews and by combining reviewers' assessments with other, more controllable forms of communication (namely paid and owned media).

A manager's influence on the availability of professional reviews is generally limited as a result of the freedom of the press in a given country. But entertainment managers can decide *when* to let critics access their newest creations. In practice, when producers provide critics access to a new product before its release (a long tradition in the industry for movies, games, songs, or books), they now often contractually prevent reviewers from publishing their reviews prior to a certain date (often the release date or close to it)—establishing an "embargo" (e.g., Brew 2016). Doing so can help prevent reviews from tampering with the producer's own buzz-building plan or to avoid the leaking of unwanted information about the product and its storyline.

A more extreme way to impact the timing of professional reviews is by not providing critics access to the product *at all* prior to its public release. Historically, whereas movie studios have occasionally prevented pre-release access for the reasons we have named above,[357] doing so has been mainly reserved for their *worst* creations in an attempt to prevent negative

---

[357]For his thriller PSYCHO, Alfred Hitchcock famously not only did not allow press screenings, but also bought copies of the underlying novel to prevent the massive twist to be leaked. And the producers of the eccentric SNAKES ON A PLANE movie also did not offer screenings mainly to add to the "mystery buzz" surrounding it which was considered so essential for the film (Golder 2015).

information from hampering opening week sales (for illustrious examples, see Golder 2015). But there is a catch to this approach: the mere act of withholding a product from critics has become a quality signal itself; products for which no pre-release access is provided are almost never worth their price. For example, when the movie Hansel & Gretel: Witch Hunters was not shown to critics, Rotten Tomatoes informed its users (and potential moviegoers) that they'd "love to tell you more about this one, but it doesn't screen for critics until later in the week, which is never a good sign" (*TV Tropes* 2013).

To multiply the power of positive professional reviews, managers often use them as ingredients in their paid and owned media efforts, such as by including snippets from the reviews on posters and in trailers; we discuss the effectiveness of doing so in the context of "good advertising." However, managers sometimes take some liberty with this approach, using the reviewers' critique out of context. The advertisement for the film 27 Dresses quoted Entertainment Weekly's review of the film: "Katherine Heigl glows!" But those three words were only the first part of a *photo caption* that continued: "…but 27 Dresses' formulaic romantic comedy stumbles on the way to the altar" (Bialik 2009). A. A. Dowd (2015), the A.V. Club's film editor, dedicated a whole article to a misleading quote on the DVD package for the film Accidental Love. His article's headline left not much room for doubt: "No, I didn't call your shitty movie a "comedic masterstroke." In these cases and several others, the full quote suggests a very different judgment than the one implied by the advertisement.

Probably the most extreme case of malpractice in the (mis)use of professional reviews was when around 2000, a marketing manager at Sony who had grown up in Ridgefield, Connecticut, created Dave Manning. As a reviewer for The Ridgefield Press, Mr. Manning "wrote" euphoric reviews for four Sony films, quotes from which were featured in print ads (*Hoaxes.org* 2014)—our introductory quote for this section provides an example of Mr. Manning's virtuosity. After a Newsweek reporter found about the true (i.e., nonexistent) nature of Mr. Manning, Sony first argued that their action was justified by "free speech," but in the end agreed to pay $1.5 million to consumers who saw the films in theaters to settle a class-action lawsuit (Phipps 2005).

The downside risk of such an approach would be enormous today; consumers would certainly detect such a farce quickly and soundly destroy it via social media's pinball mechanism. And be aware that *Entertainment Science* teaches us that there is not even an upside potential, as being endorsed by an

unknown critic has only very limited potential to drive audiences toward a new product. Remember that Rao et al. (2017) found that it is only blurbs by "top reviewers" (a status that certainly would *not* apply to an unknown critic from an unknown newspaper) that make a difference in advertising effectiveness.

This book has highlighted several far more promising ways to communicate the attractions of a new entertainment product to its target audiences. Regarding professional reviewers, producing a "great" product can still be considered the best assurance. Doing so can have one additional potential benefit: let us now end our discussion of entertainment communication by taking a look at the success potential of winning *awards*.

Awards as Recognitions of Excellence in Entertainment

> "The [Oscar] is important in order to bring people to the movie theater. That's the only principle meaning of any award."
> —*Actor* Javier Bardem *(quoted in Likesuccess.com 2017)*

Some Essentials of Entertainment Awards

The artistic nature of entertainment products is not only reflected in the ubiquitous presence of professional reviews, but also in the existence of awards and the attention that our society devotes to the ceremonies in which the awards are handed out. Awards for entertainment are granted by various institutions specific to each type of entertainment product. Although each award follows its unique rules, the logic is almost always the same: to honor artistic quality and achievements.[358]

For films, the by far most prominent recognitions of artistic excellence are the Oscars awarded by the Academy of Motion Picture Arts and Sciences (AMPAS, an elite group of film creatives and producers); other noteworthy movie awards include the Golden Globes, the Cannes International Film Festival's Palme d'Or, and the British BAFTA awards. For fiction books,

---

[358]A second type of entertainment award is based on the *commercial* performance of a product (e.g., Gold or Diamond record). In economic terms, such awards function as a visualization of a product's commercial success, similar to charts.

there's nothing bigger than the Nobel Prize in Literature handed out for writers by an expert jury of the Swedish Academy; other widely publicized honors for authors include the Pulitzer Prize, the Man Booker award, and the Newbery Medal (which is given to the most distinguished contribution to children's literature).

For music, the Grammy awards (handed out by the Recording Academy, a peer group of music professionals) stand out; other well-known music prizes include the American Music awards and the MTV Music awards. Game awards include the Game Developers Choice awards (as decided by members of the Game Developers Conference, a peer group of video game developers) and the Game Awards. The most renowned awards for TV productions are the Emmys (with key prizes decided by the Television Academy, another group of production and distribution peers) and, as for movies, the Golden Globes.

Awards and Success: What *Entertainment Science* Can Tell Us About Their Link

Why Determining the Commercial Impact of Awards is Quite a Challenge

The high TV ratings for award shows clearly indicate that we, the people, are interested in what entertainment product experts think are of high quality. When Beck's album MORNING PHASE won the Grammy for best album in February 2015, search volume on Google increased 100 times, to a level that was five times higher than the previous maximum search volume for the album. And when Mario Vargas Llosa was awarded the Nobel Prize for his writing in 2010, people searched ten times more frequently for his 1970s novel AUNT JULIA AND THE SCRIPTWRITER than in the previous months.

However, we also see that search effects vary greatly between products and they are often short term, dropping back to previous levels shortly afterward. And search means consumer interest, but does not equate with commercial success. Learning about the impact of awards on commercial success is a complicated matter though: just like word of mouth, awards can provide information about a product's quality that can be part of an informed cascade among consumers. But as with professional reviews, awards overlap with quality (*and* with professional reviews), which makes measuring the true *causal* contributions of awards a challenging endeavor. A basic

requirement for quantifying the effect of awards is to include these other quality variables—otherwise, awards just pick up their share erroneously, and its value becomes inflated.[359]

But determining such a true causal effect is even more complicated in the case of awards, because entertainment products are usually honored months *after* their market release. Thus, in conjunction with entertainment's short life cycles, *reverse* causality is a serious concern: an Oscar given out in February to a movie released in October just cannot impact moviegoers *before* February. In other words, timing is particularly crucial when estimating statistical models for awards.

And there is just one more thing: because of awards' artistic focus, they are usually only given to products which fall into our category of artistic, independent entertainment, but are rarely bestowed upon commercial products. The task thus is not to compare the Oscar-winning, low-budget film MOONLIGHT with Marvel's THE AVENGERS when it comes to the economic success of an Oscar win, but to compare MOONLIGHT to a similar low-budget film that did *not* win the award. When mixing both categories in a joint model, results tend to be biased, with the relatively lesser commercial performance of the Oscar winner being falsely attributed to the award.[360]

Monetizing the Oscars and Other Learnings

So, what then do we know regarding the financial value of awards? The most authoritative study on the matter is by Nelson et al. (2001), who investigate the effect that Oscar wins and nominations in three categories (best picture, best leading actor/actress, best supporting actor/actress) have on a film's box office. Their data set comprises all 131 films that received an Oscar nomination in one of the three categories between 1978 and 1987. The scholars pair the films with a set of another 131 films that received *no* nominations. By comparing apples (Oscar-nominated films) to apples (*similar* films,

---

[359]Take for example the historic study by Smith and Smith (1986), who, running an OLS regressions to explain the revenues of 600 films released between the 1940 and 1980), report a much higher Oscar parameter than we do for a more recent set of films (Hennig-Thurau et al. 2006b). Whereas Smith and Smith do not include *any* quality controls, we control for CinemaScore ratings *and* IMDb ratings from consumers.

[360]We assume that not addressing this "apples-to-oranges" problem (essentially a selection/endogeneity problem) is the reason why Luan and Sudhir (2010) find a *negative* effect of Oscar nominations on DVD sales, and maybe also why we do not find any significant impact of Oscar wins and nominations on video rental success (Hennig-Thurau et al. 2006b).

except that they garnered no nomination), not to oranges (films which got nominated, but differ systematically), the idea is to avoid the treatment (or "selection") bias we have argued to exist in the context of awards above (i.e., the MOONLIGHT vs. AVENGERS comparison) and which we have already discussed with regard to other success drivers in previous parts of this book.[361]

Nelson and his colleagues then used weekly box-office data to determine, with a fixed-effects regression model, the effect of the award categories on weekly box-office results. They combine this information with the results of a survival function in which they estimate the average number of weeks a film is shown in theaters (assuming that awarded films are shown longer—a supply-sided effect).[362] They find that all three award categories extend a film's run, but only best picture and lead actor categories influence weekly revenues. Using the estimated parameters in a simulation in which they compare an average film that received no Oscar nominations to an equal one that garnered nominations (and wins), their results show that films can benefit from awards (and, to a lesser degree, nominations). But the award effect differs between award categories and also by the time at which the film was released: across categories, in their analysis, a fourth quarter release generates seven times more value from an Oscar than a first quarter release.

In dollar amounts, the "Oscar effect" that Nelson et al. find is quite substantial, as we report in Fig. 12.12: the scholars estimate that a best picture win increases a film's revenues by an average of $16 million (in 1987 value—which translates into $35 million in 2017 currency) for a film that is released in the fourth quarter. While they treat the award categories as

---

[361]We discuss such treatment/selection bias in the context of entertainment branding. There's a caveat, however, with regard to Nelson et al.'s approach: their sole selection criterion is to pick a film released in the same week as a nominated film, which does not warrant the removal of potential differences in terms of other "success drivers" (think of genres, budgets, or—of particular importance—movie quality). Clement et al. (2007b) conducted a closer investigation of the differences between those films that received an Oscar nomination and those which did not; among others, their results provide evidence that the quality of Oscar nominated films as judged by critics is clearly higher than that for non-nominated films. As a consequence, the results from Nelson et al. we report in this section might still involve a systematic bias between films that were recognized with awards and those that were not; they should thus be considered as "conservative" estimates that mark the lower end of the spectrum.

[362]Our observations of the marketplace suggests that showing a film longer might not be the only supply-sided effect. Theaters also often increase a film's *availability* as a result of Oscar nominations and wins. For example, when THE SHAPE OF WATER received 13 nominations in 2018, its number of theaters more than doubled (from 853 to 1,854, or +117%), hand in hand with a 171% increase in box office compared to the pre-nomination weekend. Nelson et al. did not divide their results into supply-sided and demand-sided effects, so we can only report this as anecdotal evidence.
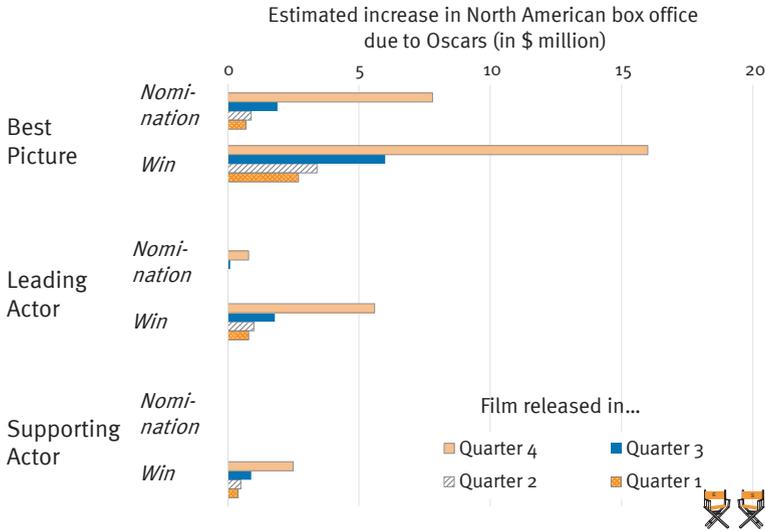
Estimated increase in North American box office
due to Oscars (in $ million)

Fig. 12.12 Monetary effects of Oscar wins and nominations
Notes: Authors' own illustration based on results reported by Nelson et al. (2001). The estimates are box office effects which are adjusted for probability of survival. All numbers are in 1987 currency; multiplying them by factor 2.2 approximates their 2017 value.

independent, we expect that negative interactions exist between them—the incremental value of an additional best actor Oscar should be less than if it is the *only* win for a film.

When evaluating these effects, we have to keep in mind that the scholars only look at the success in North American theaters. Movie producers now earn strong returns in channels other than theaters,[363] and we suspect that awards help to gain revenues in them too, either directly or via inter-channel action-based cascades.

With the Oscars being by far the most prominent award in the field of film, to what degree do such effects also exist for other awards? Gemser et al. (2008) study the role of 13 different awards in the financial performance of films, using a data set of around 300 award-winning films that were released between 1997 and 2001 and were still shown in theaters when the award was announced. The scholars run separate OLS regressions for the second and fourth week after the announcement, estimating the effect that awards

---

[363]For a more detailed discussion of distribution channels for entertainment, see our distribution chapter.

have on both box-office results and on distribution (the number of screens on which a film is shown). They control for the box office and screens prior to the award, as well as several other success drivers (including advertising spending); they do not correct for any potential treatment bias, though. Gemser et al.'s results confirm that the number of total awards won by a film (across categories), and also the number of *different* awards it has won, enhances box-office results and the number of screens allocated to the film. Thus, the economic potential of awards should not be limited to an Oscar win.

Interestingly, the scholars find that *who* picks the winner also matters for any award's impact on audiences: across the different awards, those awards given by experts (such as the Golden Globes) are significantly more influential for audiences and distributors than are awards chosen by industry peers or those picked by consumers. At least in this regard, the Oscars certainly constitute an exception to the (statistical) rule.

We have to admit that we know clearly less about the effects of awards in other fields of entertainment. Future research on music, book, and game awards would help us to better understand their impacts and could serve as the foundation for managerial efforts toward winning an award, which we discuss in the following sub-section. But before we get to this next issue, let us mention one additional thing we indeed do know from scholarly work: the benefits of awards are not limited to *financial* gains.

In a study of the life paths of Oscar winning actors, Redelmeier and Singh (2001) find that Oscar winners live almost four years longer, on average (80 versus 76 years). They derive this insight by comparing the life expectancies of 1,649 Oscar winners and nominees with those of non-awarded same-sex cast members in the same films; birth demographics do not affect the finding. Whereas winning an Oscar more than once tends to add even more life expectancy, being nominated (but ending up not being the winner) does not grant any advantage over those who were not nominated at all. So, whereas an Oscar can't make us immortal, it tends to let us live longer. Why is this? Because, using the authors' words, "success confers a survival advantage" (Redelmeier and Singh 2001, p. 960).

Managerial (Mis-)Use of Awards

We have seen that awards, as a special kind of "earned" media coverage, can grant commercial advantages for entertainment products. As with professional reviews, entertainment managers have developed a number of practices to reach for these advantages: by influencing the chances to win and, if successful, by leveraging those wins.

When it comes to influencing winning probabilities, managers use practices that span the marketing mix. One aspect involves release timing: Nelson et al. (2001) provide evidence that award-winning films that were released (in theaters) later in the year have a *much* higher monetization potential—simply because a later release date increases the chances that a film is still showing in theaters when it receives the award. But, as we discuss in the next chapter, distribution timing is a complex matter, and time periods that are optimal for award potential might be less attractive with regard to other criteria, such as consumer demand and competition. Managers should include award effects in their distribution considerations but balance them with other factors, particularly as winning an award is rare and highly uncertain.

A second popular managerial practice is to dedicate advertising budgets to an award candidate. Very often, studios design campaigns that are targeted at those who choose the award winners. These days, Hollywood's film studios together spend about $150 million for their Oscar campaigns (Kirkham 2015). Follows (2015), in a detailed analysis of such Oscar campaigns, reports that studios spend $5 to $8 million per candidate film; in certain conditions, spending can be twice that amount (e.g., the Weinsteins spent $15 million for Shakespeare in Love's successful Oscar campaign, which equates with almost $23 million in 2017 value). Most of the money flows into targeted advertisements (characterized by "For Your Consideration" notes), followed by producing and sending "screener" versions of their films to members of the AMPAS, along with hosting theatrical screenings. In the film business, more than half of the campaign budget is spent prior to the announcement of the nominations; the remaining budget is then concentrated on those films that managed to achieve a nomination, trying to convert them from nominees into winners (Follows 2015).

Is such spending justified? Putting the campaign costs in relation to Nelson et al.'s revenue estimates can provide a tentative answer. Assuming that an Oscar-winning film released in the fourth quarter returns an additional $20 million to the studio from theaters and, a rough estimate, two times that number from other channels over its lifetime (plus higher brand value that might be harvested by producing sequels or selling the rights),[364]

---

[364]Among Oscar winning films for which sequels (or prequels) were made are Rocky, The French Connection, In the Heat of the Night, The Silence of the Lambs, and The Lord of the Rings. And The, Godfather of course. We don't argue that the sequels to these films were made *because* of their Oscar wins, but we assume that, based on the value of the Oscar we have shown in this section, they will have benefited from their predecessors' wins. Why have not more award winners been turned into franchises then? We speculate that this might have to do with Oscar voters having shown a preference for drama and sad endings (think: Titanic!), which affect brand value.

spending up to \$20 million on an Oscar campaign would certainly be justified. But, of course, often a film does *not* win. Thus, a portfolio perspective is adequate: over time, a studio must balance its spending in a way that is justified by the number of winners and nominees. The work of some *Entertainment Science* scholars who have tried to demystify the logic behind Oscar nominations and wins might be of help here (e.g., Krauss et al. 2008; Pardoe and Simonton 2008).

Finally, once an entertainment product has managed to win an award, how can this achievement be commercialized? Here, the same basic rules apply as for professional reviews: the nominations and wins are printed on new posters, featured on web pages, added to new trailers, and noted on the packaging of a product (e.g., on the album or DVD cover). And beyond communication, because awards are quality signals (specifically, "substitute cues"), they can also be used as justification for releasing special editions of a product (such as Universal's "Oscar Edition" of films on DVD) as part of a versioning approach. Some movies that had already ended their theatrical run when an award was announced *return* to theaters as a "re-release;" the producer hopes to trigger the interest of those consumers who ignored the film in its original run.

Concluding Comments

We began this chapter by reviewing what insights *Entertainment Science* scholars have compiled regarding the most commonly considered form of earned communication, word of mouth. We distinguished between three types of word of mouth, traditional, social media, and other electronic word of mouth, which today impact entertainment product sales, being more than substitutes for each other. In the weeks and months following an entertainment product's release, the valence of word of mouth influences sales more than advertising and many other "success drivers." This strong effect results from word of mouth being a "substitute cue" for a consumer and involve the sharing of actual people's actual experiences with an entertainment product; we thus refer to the spread of such communication as "informed cascades."

But *un*informed information cascades are also quite influential. We reported that in addition to the desirability that is signaled by high chart rankings ("The top romantic comedy in America!"), consumers' feverish anticipation expressed in pre-release buzz provides a powerful signal; both

drive the success of new entertainment products, though at different points in time. Furthermore, automated personal recommender systems process information about consumers' liking of certain products into an information source that is considered as valuable by many consumers. We portrayed the basic approaches that offer recommendations (collaborative filtering, content-based recommenders, and hybrid approaches) and discussed their respective strengths (and limitations).

Whereas recommenders combine consumer data with the actions of those who offer them, we also looked at information from other stakeholders of an entertainment firm that needs to be "earned." Reviewing the empirical evidence, we concluded that professional reviewers do indeed have an impact on product sales, and at least some industry awards such as the Oscar do so, too. We demonstrated how their effect can be monetized from different award categories, offering insights the factors that determine what can be earned by getting nominated or even winning.

This concludes our analysis of entertainment communication. We now turn to distribution: if consumers cannot act on their desire for an entertainment product because it is unavailable or too troublesome to acquire, the product's success will be hindered. In entertainment, the main distribution challenges include finding the right time to release a product, to orchestrate the many distribution channels that exist for entertainment in our digital times, and to deal with illegal competition that takes the form of a pirated version of an entertainment producer's own product.

References

Adomavicius, G., Bamshad, M., Francesco, R., & Tuzhilin, A. (2011). Context-aware recommender systems. *AI Magazine, 32*.

Amatriain, X., & Basilico, J. (2012). Netflix recommendations: Beyond the 5 stars (Part 2). *The Netflix Tech Blog*, June 20, https://goo.gl/N6QNK9.

Arndt, J. (1967). Role of product-related conversations in the diffusion of a new product. *Journal of Marketing, 4*, 291–295.

Asai, S. (2015). Determinants of demand and price for best-selling novels in paperback in Japan. *Journal of Cultural Economics, 40*, 375–392.

Austin, B. A. (1989). *Immediate seating: A look at movie audiences*. California: Wadsworth Pub. Co.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, N.J.: Prentice-Hall.

Bart, P. (2017). Peter Bart: Amazon raises bet on movie business, but rivals still baffled about long-term strategy. *Deadline*, March 10, https://goo.gl/t46Vvj.

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

Berger, J. (2014). Word of mouth and interpersonal communication: A review and directions for future research. *Journal of Consumer Psychology, 24*, 586–607.

Berger, J., & Schwartz, E. M. (2011). What drives immediate and ongoing word of mouth? *Journal of Marketing Research, 48*, 869–880.

Berger, J., & Milkman, K. L. (2012). What makes online content viral? *Journal of Marketing Research, 49*, 192–205.

Bhattacharjee, S., Gopal, R. D., Lertwachara, K., Marsden, J. R., & Telang, R. (2007). The effect of digital sharing technologies on music markets: A survival analysis of albums on ranking charts. *Management Science, 53*, 1359–1374.

Bialik, C. (2009). The best worst blurbs of 2008. *Gelf Magazine*, January 2, https://goo.gl/gvDHhB.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of Economic Perspectives, 12*, 151–170.

Bishop, B. (2017). How Netflix is trying to rewrite movie marketing with Bright. *The Verge*, December 19, https://goo.gl/VnDjM7.

Böll, H. (1963). *Ansichten eines Clowns*. Cologne: Kiepenheuer & Witsch.

Brew, S. (2016). Movie embargoes: What are they, and why do they matter? *Den of Geek*, July 15, https://goo.gl/Za6igw.

Brodnig, I. (2015). Netflix-Produktchef Neil Hunt: 'Ich weiß das alles über Sie'. *profil*, December 11, https://goo.gl/67ACyu.

Buli, L., & Hu, V. (2015). Data science and the music industry: What social media has to do with record sales. *hypebot*, December 6, https://goo.gl/tET39C.

Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction, 12*, 331–370.

Chen, Y., Liu, Y., & Zhang, J. (2012). When do third-party product reviews affect firm value and what can firms do? The case of media critics and professional movie reviews. *Journal of Marketing, 76*, 116–134.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research, 43*, 345–354.

Chintagunta, P., Gopinath, S., & Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science, 29*, 944–957.

Clement, M., Proppe, D., & Rott, A. (2007a). Do critics make bestsellers? Opinion leaders and the success of books. *Journal of Media Economics, 20*, 77–105.

Clement, M., Christensen, B., Albers, S., & Guldner, S. (2007b). Was bringt ein Oscar im Filmgeschäft? Eine empirische Analyse unter Berücksichtigung des Selektionseffekts. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung, 59*, 198–220.

Clement, M., Hille, A., Lucke, B., Schmidt-Stölting, C., & Sambeth, F. (2008). Der Einfluss von Rankings auf den Absatz – Eine empirische Analyse der Wirkung von Bestsellerlisten und Rangpositionen auf den Erfolg von Büchern. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung, 60*, 746–777.

Clement, M., Wu, S., & Fischer, M. (2014). Empirical generalization of demand and supply dynamics for movies. *International Journal of Research in Marketing, 31*, 207–223.

Corliss, R. (2009). Box-office weekend: Brüno a one-day wonder? *Time Magazine*, July 13, https://goo.gl/wYf5v3.

Cox, J., & Kaimann, D. (2015). How do reviews from professional critics interact with other signals of product quality? Evidence from the video game industry. *Journal of Consumer Behavior, 14,* 366–377.

Craig, C. S., Greene, W. H., & Versaci, A. (2015a). E-word of mouth: Early predictor of audience engagement—How pre-release 'E-WOM' drives box-office outcomes of movies. *Journal of Advertising Research, 55*, 62–72.

D'Alessandro, A. (2015b). Tom Hardy Soviet drama 'Child 44' bombs at box office: What the hell happened? *Deadline*, May 1, https://goo.gl/HFaSmh.

D'Alessandro, A. (2017). How 'Pirates' & 'Baywatch' are casualties of summer franchise fatigue at the domestic B.O. *Deadline*, May 29, https://goo.gl/WZJjVP.

D'Astous, A., & Colbert, F. (2002). Moviegoers' consultation of critical reviews: Psychological antecedents and consequences. *International Journal of Arts Management, 5*, 24–35.

D'Astous, A., Carù, A., Koll, O., & Sigué, S. P. (2005). Moviegoers' consultation of film reviews in the search for information: A multi-country study. *International Journal of Arts Management, 7*, 32–45.

De Vany, A., & Lee, C. (2001). Quality signals in information cascades and the dynamics of the distribution of motion picture box office revenues. *Journal of Economic Dynamics & Control, 25*, 593–614.

Dewan, S., & Ramaprasad, J. (2012). Music blogging, online sampling, and the long tail. *Information Systems Research, 23*, 1056–1067.

Dhar, T., & Weinberg, C. B. (2015). Measurement of interactions in non-linear marketing models: The effect of critics' ratings and consumer sentiment on movie demand. *International Journal of Research in Marketing, 33*, 392–408.

Divakaran, P. K. P., Palmer, A., Alsted Søndergaard, H., & Matkovskyy, R. (2017). Pre-launch prediction of market performance for short lifecycle products using online community data. *Journal of Interactive Marketing, 38*, 12–28.

Dowd, A. A. (2015). No, I didn't call your shitty movie a "comedic masterstroke". *A.V. Club*, July 27, https://goo.gl/s3FWnY.

Duan, W., Bin, G., & Whinston, A. B. (2008). The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry. *Journal of Retailing, 84*, 233–242.

Ekstrand, M. D., Riedl, J. T., & Konstan, J. A. (2010). Collaborative filtering recommender systems. *Foundations and Trends in Human-Computer Interaction, 4*, 81–173.

Elberse, A., & Eliashberg, J. (2003). Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures. *Marketing Science, 22*, 329–354.

Eliashberg, J., & Shugan, S. M. (1997). Film critics: Influencers or predictors? *Journal of Marketing, 61*, 68–78.

Fischer, S. F., Hammerschmidt, M., & Weiger, W. H. (2017). Signals from the echoverse – the informational value of brand buzz dispersion. *Proceedings of Winter AMA Conference, 28*.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley. [An online version of the book can be found at Mr. Ajzen's website at https://goo.gl/Re6HGP].

Follows, S. (2015). How much do Hollywood campaigns for an Oscar cost? January 12, https://goo.gl/ke5nx4.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research, 19*, 291–313.

Foutz, N. Z., & Jank, W. (2010). Prerelease demand forecasting for motion pictures using functional shape analysis of virtual stock markets. *Marketing Science, 29*, 568–579.

Freedman, N. (2015). How social media is changing Hollywood. *Digital America*, April 6, https://goo.gl/ufned5.

Fry, Stephen (2010). Two million reasons to be cheerful, November 30, https://goo.gl/8Ffq5w.

Gemser, G., Van Oostrum, M., & Leenders, M. A. A. M. (2007). The impact of film reviews on the box office performance of art house versus mainstream motion pictures. *Journal of Cultural Economics, 31*, 43–63.

Gemser, G., Leenders, Mark A. A. M., & Wijnberg, N. M. (2008). Why some awards are more effective signals of quality than others: A study of movie awards. *Journal of Management, 34*, 25–54.

Golder, D. (2015). 14 movies that were 'Not Screened For Critics', August 5, https://goo.gl/2WkDhb.

Gomez-Uribe, C. A., & Hunt, N. (2015). The Netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems, 6*, 13–19.

Goode, L. (2017). Netflix is ditching five-star ratings in favor of a thumbs-up. *The Verge*, March 16, https://goo.gl/wRWgxY.

Gopinath, S., Chintagunta, P. K., & Venkataraman, S. (2013). Blogs, advertising, and local-market movie box office performance. *Management Science, 59*, 2635–2654.

Gower, S. (2014). Netflix Prize and SVD. Working Paper.

Hennig-Thurau, T., Walsh, G., & Bode, M. (2003). Exporting media products: Understanding the success and failure of Hollywood movies in Germany. Working Paper, Bauhaus-University of Weimar.

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). What motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing, 18*, 38–52.

Hennig-Thurau, T., Houston, M. B., & Sridhar, S. (2006a). Can good marketing carry a bad product? Evidence from the motion picture industry. *Marketing Letters, 17*, 205–219.

Hennig-Thurau, T., Houston, M. B., & Walsh, G. (2006b). The differing roles of success drivers across sequential channels: An application to the motion picture industry. *Journal of the Academy of Marketing Science, 34*, 559–575.

Hennig-Thurau, T., Marchand, A., & Marx, P. (2012a). Can automated group recommender systems help consumers make better choices? *Journal of Marketing, 76*, 89–109.

Hennig-Thurau, T., Marchand, A., & Hiller, B. (2012b). The relationship between reviewer judgments and motion picture success: Re-analysis and extension. *Journal of Cultural Economics, 36*, 249–283.

Hennig-Thurau, T., Wiertz, C., & Feldhaus, F. (2015). Does Twitter matter? The impact of microblogging word of mouth on consumers' adoption of new movies. *Journal of the Academy of Marketing Science, 43*, 375–394.

Hoaxes.org (2014). Dave Manning. https://goo.gl/tNadxL.

Horn, J. (2009). ComicCon's buzzmakers. *Los Angeles Times*, July 27, https://goo.gl/uAXFGT.

Houston, M. B., Kupfer, A., Hennig-Thurau, T., & Spann, M. (2018). Pre-release new product consumer buzz. *Journal of the Academy of Marketing Science, 46*, 338–360.

Hsu, G. (2006a). Evaluative schemas and the attention of critics in the US film industry. *Industrial & Corporate Change, 15*, 467–496.

Jabr, W., & Zhiqiang, Z. (Eric) (2014). Know yourself and know your enemy: An analysis of firm recommendations and consumer reviews in a competitive environment. *MIS Quarterly, 38*, 635–654.

Jain, K., & Srinivasan, N. (1991). An empirical assessment of multiple operationalizations of involvement. *Advances in Consumer Research, 17*, 594–602.

Jannach, D., Resnick, P., Tuzhilin, A., & Zanker, M. (2016). Recommender systems–beyond matrix completion. *Communications of the ACM, 59*, 94–102.

Jozefowicz, J., Kelley, J., & Brewer, S. (2008). New release: An empirical analysis of VHS/DVD rental success. *Atlantic Economic Journal, 36*, 139–151.

Karniouchina, E. V. (2011). Impact of star and movie buzz on motion picture distribution and box office revenue. *International Journal of Research in Marketing, 28*, 62–74.

Katz, E., & Lazarsfeld, P. F. (1955). *Personal influence: The part played by people in the flow of mass communication*. New York: The Free Press.

Kaye, D. (2012). 4 films that failed to live up to their blockbuster Comic-Con buzz. *Syfy Wire*, December 14, https://goo.gl/baBFuu.

Kim, H., & Hanssens. D. M. (2017). Advertising and word-of-mouth effects on pre-launch consumer interest and initial sales of experience products. *Journal of Interactive Marketing, 37*, 57–74.

Kirkham, E. (2015). Hollywood spends $150 million on Oscar campaigns each year. *The Philadelphia Inquirer*, January 15, https://goo.gl/D7d8s7.

Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer, 42*, 30–37.

Körte, P. (2009). Filmkritik ist, wenn man von der Brücke in den Fluss spuckt. In T. Hennig-Thurau & V. Henning (Eds.), *Guru Talk – Die deutsche Filmindustrie im 21. Jahrhundert* (pp. 192–197). Marburg: Schüren.

Krauss, J., Nann, S., Simon, D., Fischbach, K., & Gloor, P. (2008). Predicting movie success and Academy Awards through sentiment and social network analysis. In *Proceedings of 16th European Conference on Information Systems* (pp. 2026–2037). Galway, Ireland.

Krishnan, V., Narayanashetty, P. K., Nathan, M., Davies, R. T., & Konstan, J. A. (2008). Who predicts better?—Results from an online study comparing humans and an online recommender system. In *Proceedings of the 2008 ACM Conference on Recommender Systems* (pp. 211–218). New York: ACM.

Kupfer, A.-K., Pähler vor der Holte, N., Kübler, R., & Hennig-Thurau, T. (2018). The role of the partner brand's social media power in brand alliances. *Journal of Marketing, 82*, 25–44.

Lampel, J., & Shamsie, J. (2000). Critical push: Strategies for creating momentum in the motion picture industry. *Journal of Management, 26*, 233–257.

Lasar, M. (2011). Digging into Pandora's music genome with musicologist Nolan Gasser. *Ars Technica*, January 13, https://goo.gl/yWmG8M.

Lee, J., Boatwright, P., & Kamakura, W. A. (2003). A Bayesian model for prelaunch sales forecasting of recorded music. *Management Science, 49*, 179–196.

Legoux, R., Larocque, D., Laporte, S., Belmati, S., & Boquet, T. (2015). The effect of critical reviews on exhibitors' decisions: Do reviews affect the survival of a movie on screen? *International Journal of Research in Marketing, 33*, 357–374.

Li, X., & Hitt, L. M. (2008). Self-selection and information role of online product reviews. *Information Systems Research, 19*, 456–474.

Likesuccess.com (2017). The award is important in order to bring people to the movie theater. That's the only principle meaning of any award. https://goo.gl/qpYdkm.

Littleton, C., & Holloway, D. (2017). Jeff Bezos mandates programming shift at Amazon Studios. *Variety*, September 8, https://goo.gl/GFtV37.

Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing, 70*, 74–89.

Liu, X., Singh, P. V., & Srinivasan, K. (2016). A structured analysis of unstructured big data by leveraging cloud computing. *Marketing Science, 35*, 363–388.

Lops, P., de Gemmis, M., & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender systems handbook* (pp. 73–104). New York: Springer.

Lovett, M. J., & Staelin, R. (2016). The role of paid, earned, and owned media in building entertainment brands: Reminding, informing, and enhancing enjoyment. *Marketing Science, 35*, 142–157.

Luan, Y. J., & Sudhir, K. (2010). Forecasting marketing-mix responsiveness for new products. *Journal of Marketing Research, 47*, 444–457.

Lynn, F. B., Walker, M. H., & Peterson, C. (2016). Is popular more likeable? Choice status by intrinsic appeal in an experimental music market. *Social Psychology Quarterly, 79*, 168–180.

Madrigal, A. C. (2014). How Netflix reverse engineered Hollywood. *The Atlantic*, January 2, https://goo.gl/4frKk3.

Malhotra, N., & Helmer, E. (2012). Inflation in weekend box office estimates. *Applied Economics Letters, 19*, 1411–1415.

Marchand, A. (2016). The power of an installed base to combat lifecycle decline: The case of video games. *International Journal of Research in Marketing, 33*, 140–154.

Marchand, A., Hennig-Thurau, T., & Wiertz, C. (2016). Not all digital word of mouth is created equal: Understanding the respective impact of consumer reviews and microblogs on new product success. *International Journal of Research in Marketing, 34*, 336–354.

Marx, P., Hennig-Thurau, T., & Marchand, A. (2010). Increasing consumers' understanding of recommender results: A preference-based hybrid algorithm with strong explanatory power. In *Proceedings of the fourth ACM Conference on Recommender Systems* (pp. 297–300). New York: ACM.

McKenzie, J. (2010). How do theatrical box office revenues affect DVD retail sales? Australian empirical evidence. *Journal of Cultural Economics, 34*, 159–179.

Meiseberg, B. (2016). The effectiveness of e-tailers' communication practices in stimulating sales of niche versus popular products. *Journal of Retailing, 92*, 319–332.

Nelson, R. A., Donihue, M. R., Waldman, D. M., & Wheaton, G. (2001). What's an Oscar worth? *Economic Inquiry, 39*, 1–16.

Netflix (2017). Decoding the defenders: Netflix unveils the gateway shows that lead to a heroic binge. Press Release, August 22, https://goo.gl/2nmGZf.

Niraj, R., & Singh, J. (2015). Impact of user-generated and professional critics reviews on Bollywood movie success. *Australasian Marketing Journal, 23*, 179–187.

O'Connor, M., Cosley, D., Konstan, J. A., & Riedl, J. (2001). PolyLens: A recommender system for groups of users. In *Proceedings of the 7th European Conference on Computer-Supported Cooperative Work* (pp. 199–218). Dodrecht: Kluwer.

Panaligan, R., & Chen, A. (2013). Quantifying movie magic with Google Search. White Paper, Google.

Papies, D., & van Heerde, H. (2015). How the Internet has changed the marketing of entertainment goods: The case of the music industry. Working Paper, Marketing Science Institute.

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.

Phipps, K. (2005). The ghost of David Manning will have his revenge. *AV Club*, August 3, https://goo.gl/CCSWCW.

Rao, V. R., Ravid, A. S., Gretz, R.T., Chen, J., & Basuroy, S. (2017). The impact of advertising content on movie revenues. *Marketing Letters, 28*, 341–355.

Redelmeier, D. A., & Singh, S. M. (2001). Survival in Academy Award-winning actors and actresses. *Annals of Internal Medicine, 134*, 955–962.

Reinstein, D. A., & Snyder, C. M. (2004). The influence of expert reviews on consumer demand for experience goods: A case study of movie critics. *The Journal of Industrial Economics, 53*, 27–51.

Rosario, A. B., Sotgiu, F., De Valck, K., & Bijmolt, T. H. A. (2016). The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. *Journal of Research in Marketing, 53*, 297–318.

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science, 311*, 854–856.

Salganik, M. J., & Watts, D. J. (2008). Leading the herd astray: An experimental study of self-fulfilling prophecies in an artificial cultural market. *Social Psychology Quarterly, 71*, 338–355.

Schmidt-Stölting, C., Blömeke, E., & Clement, M. (2011). Success drivers of fiction books: An empirical analysis of hardcover and paperback editions in Germany. *Journal of Media Economics, 24*, 24–47.

Shone, T. (2004). *Blockbuster*. New York: Free Press.

Smith, S. P., & Smith, V. K. (1986). Successful movies: A preliminary empirical analysis. *Applied Economics, 18*, 501–507.

Soderstrom, S. B., Uzzi, B., Rucker, D. D., Fowler, J. H., & Diermeier, D. (2016). Timing matters: How social influence affects adoption pre- and post-product release. *Sociological Science, 3*, 915–939.

Sorensen, A. T. (2007). Bestseller lists and product variety: The case of book sales. Working Paper, Stanford GSB Research Paper No. 1878.

Sorensen, A. T., & Rasmussen, S. J. (2004). Is any publicity good publicity? A note on the impact of book reviews. Working Paper, Stanford University.

Sirdeshmukh, D., Singh, J., & Sabol, B. (2002). Consumer trust, value, and loyalty in relational exchanges. *Journal of Marketing, 66*, 15–37.

Strause, J. (2016). How 'My Big Fat Greek Wedding' became an indie phenomenon. *Hollywood Reporter*, March 25, https://goo.gl/RV4prv.

Sun, M. (2012). How does the variance of product ratings matter? *Management Science, 58*, 696–707.

The Economist (2017). How to devise the perfect recommendation algorithm. https://goo.gl/VouerJ.

tickld (2016). 31 fascinating things most people don't know about Netflix. https://goo.gl/8pNw7X.

TV Tropes (2013). Not screened for critics. *TV Tropes*, https://goo.gl/QxQiCs.

Wohlfeil, M., & Whelan, S. (2008). Confessions of a movie-fan: Introspection into a consumer's experiential consumption of 'Pride and Prejudice'. *Proceedings of European ACR Conference* (pp. 137–143).

Wyatt, R. O., & Badger, D. P. (1984). How reviews affect interest in and evaluation of films. *Journalism Quarterly, 61*, 874–878.

Wyatt, R. O., & Badger, D. P. (1990). Effects of information and evaluation in film criticism. *Journalism Quarterly, 67*, 359–368.

Xiong, G., & Bharadwaj, S. (2014). Prerelease buzz evolution patterns and new product performance. *Marketing Science, 33*, 401–421.

You, Y., Vadakkepatt, G. G., & Joshi, A. M. (2015). A meta-analysis of electronic word-of-mouth elasticity. *Journal of Marketing, 79*, 19–39.

Zhu, F., & Zhang, X. M. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing, 74*, 133–148.