

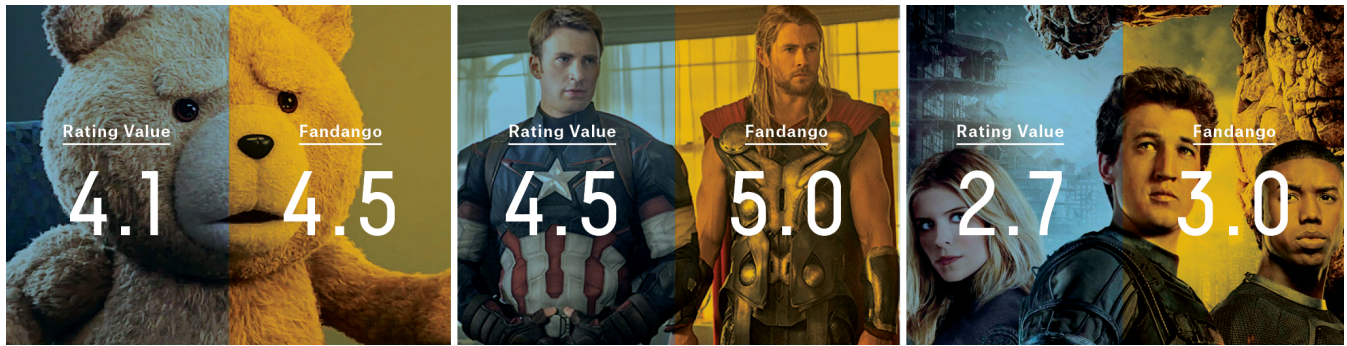
OCT. 15, 2015, AT 9:52 AM

Be Suspicious Of Online Movie Ratings, Especially Fandango's

By Walt Hickey

Filed under Movies

Get the data on GitHub



You were excited for the date: dinner and a movie. Your date picked a restaurant — “It got five stars on Yelp!” — but the movie was up to you. So you checked out what was playing and bought the tickets on [Fandango's website](#). You decided to check out “[Fantastic Four](#),” and even though you hadn't heard great things, Fandango users thought it was good! Over 7,000 people had reviewed it, and it had an average of 3 out of 5 stars. *This is going to be a decent movie.*

It is not a decent movie.

Online movie ratings have become serious business. Hollywood generates something on the order of [\\$10 billion annually at the U.S. box office](#), and online ratings aggregators may hold increasing sway over where that money goes. Sites like Rotten Tomatoes that aggregate movie reviews into one overall rating are [being blamed](#) for poor opening weekends. A single movie critic can't make or break a film anymore, but maybe thousands of critics, professional and amateur together, can.

Several sites have built popular rating systems: Rotten Tomatoes,¹ Metacritic² and IMDb³ each have their own way of aggregating film reviews. And while the sites have different criteria for picking and combining reviews, they have all built systems with similar values: They use the full continuum of their ratings scale,^x

Essentially, it's just as possible for a film to do very poorly on [IMDb](#), [Metacritic](#) and [Rotten Tomatoes](#) as it is to do quite well. If there's a curve, some movies manage to do very badly anyway.

try to maintain consistency,^x

IMDb doesn't go into a ton of detail, but it [uses a weighted average](#) that emphasizes people who review a lot of movies. Metacritic normalizes critic ratings and [transforms them into a metacore](#) through a proprietary but consistent weighted average. And Rotten Tomatoes has [defined thresholds](#) for movies it recommends and discourages.

and attempt to limit deliberate interference in their ratings.^x

IMDb [comes right out and says](#) it calculates a film's score knowing that people are going to try to artificially rate movies high and low. Metacritic draws from a [predetermined blind](#) weighted sample group of critics whom it deems trustworthy. And Rotten Tomatoes has [requirements about what qualifies](#) a person as a trusted critic.

These rating systems aren't perfect, but they're sound enough to be useful.

All that cannot be said of Fandango, a NBCUniversal subsidiary that uses a five-star rating system in which almost no movie gets fewer than three stars, according to a FiveThirtyEight analysis. What's more, as I'm writing this, scores on Fandango.com are skewed even higher because of the weird way Fandango aggregates its users' reviews. And while other sites that gather user reviews are often tangentially connected to the media industry, Fandango has an immediate interest in your desire to see a movie: The company sells tickets directly to consumers.

What started all this? A couple of months ago, a colleague noticed that a bad film had received a decent rating on Fandango and asked me to look into it. When I pulled the data for 510 films on Fandango.com that had tickets on sale this year (you can check out all the data yourself [on GitHub](#)),⁷ something looked off right away: Of the 437 films with at least one review, 98 percent had a 3-star rating or higher and 75 percent had a 4-star rating or higher.

It seemed nearly impossible for a movie to fail by Fandango's standards.

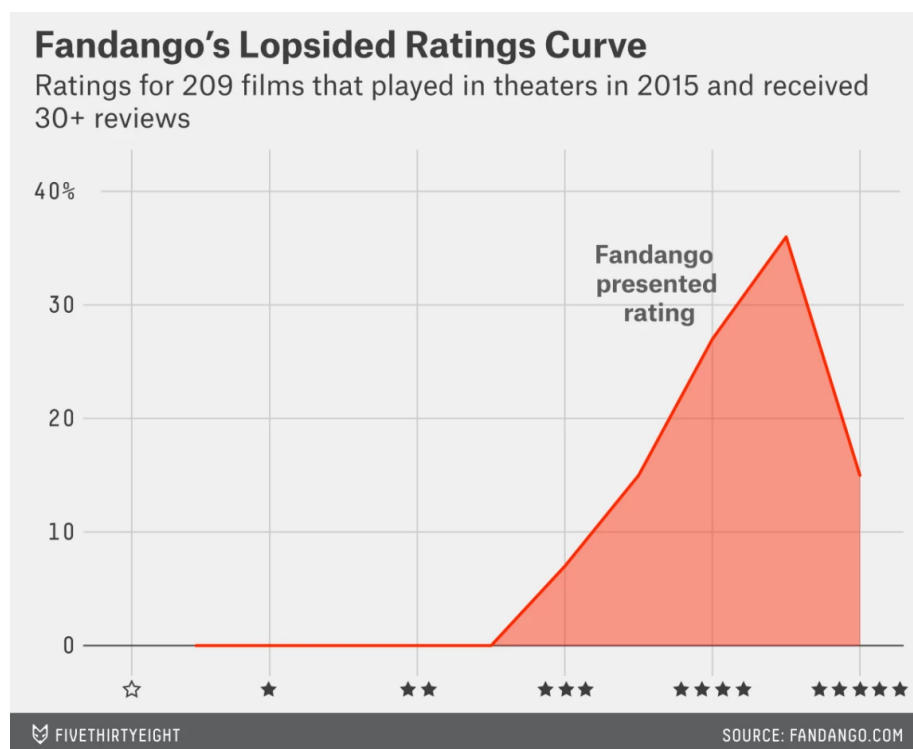
When I focused on movies that had 30^x

To get a decent sample.

or more user reviews,^x

We also excluded three movies that had not been released yet but had still somehow gathered 30 or more reviews on Fandango.com.

none of the 209 films had below a 3-star rating. Seventy-eight percent had a rating of 4 stars or higher.

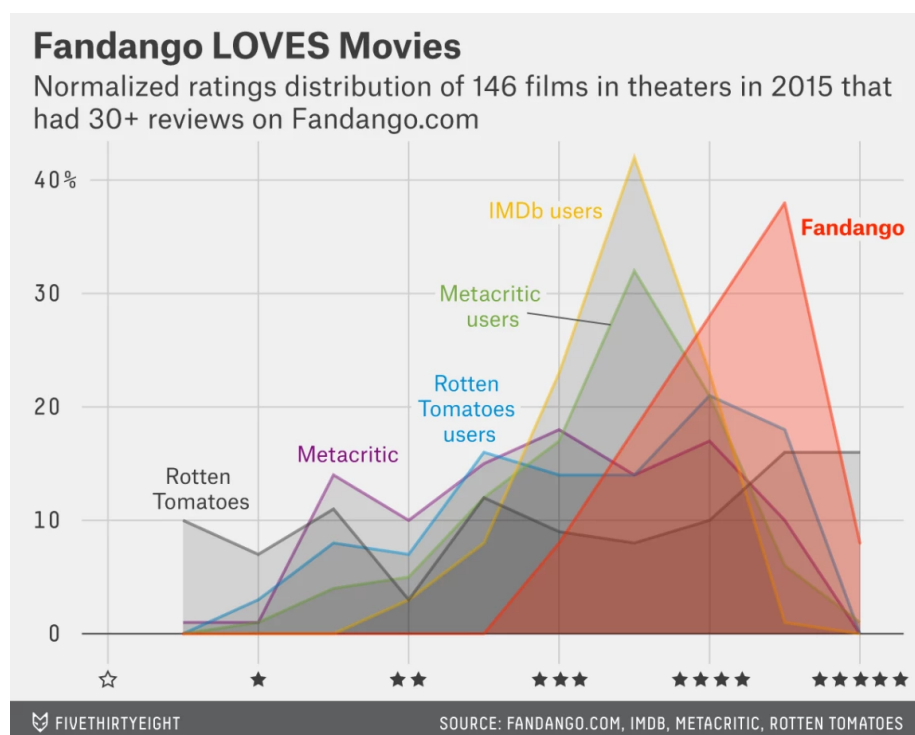


But perhaps the movies were just that good? Maybe we really do live in a society that rates “Mortdecai” as a [3.5-star film](#)?

We don't. The other review sites weren't nearly as charitable. For the 209 films, I pulled IMDb's user rating, Metacritic's aggregate critic rating, Metacritic's user score, the Rotten Tomatoes Tomatometer (critic) score and the Rotten Tomatoes user score. I then normalized these to the five-star rating scale that Fandango uses and rounded it to the nearest half-star.¹⁰

The ratings from IMDb, Metacritic and Rotten Tomatoes were typically in the same ballpark, which makes this finding unsurprising: Fandango's star rating was higher than the IMDb rating 79 percent of the time, the Metacritic aggregate

critic score 77 percent of the time, the Metacritic user score 86 percent of the time, the Rotten Tomatoes critic score 62 percent of the time, and the Rotten Tomatoes user score 74 percent of the time.



There are all sorts of reasons that scores might be higher on a site like Fandango compared with competitors; after all, if you ask people about a movie after they've paid \$15 for it and devoted a couple of hours of their life to it, maybe they'll have a more favorable opinion of the work. Maybe the profoundly rightward shift in Fandango's bell curve is just a moviegoer's version of Stockholm syndrome.

Still, this is a deeply flawed rating system. It's not clear why so few movies earn less than 3 stars, and Fandango didn't offer any explanation. "As we have not analyzed other sites' user ratings systems and we do not have access to their customers' profile and engagement behavior, it is unfair for us to speculate how our ratings may or may not differ from theirs," Fandango said in an emailed statement.

So for all intents and purposes, Fandango is using a 3 to 5 star scale. And that's not the only thing wrong with its ratings. I found an issue with the methodology Fandango uses to average user ratings on its website: Fandango never rounds the average down.

On a given film's page on Fandango's website, its aggregate user rating is displayed in one spot: the stars next to the film's poster, above the area that provides showtimes. The stars are expressed on a five-point scale at half-star increments. Beneath the star ratings, Fandango lists the number of reviews the film has received.

But when you pull the HTML source of a page on Fandango's website, there's more information. Take "[Ted 2](#)." When I pulled data for it on Monday, the film had 4.5 stars from 6,568 reviews.



You can see that information on the HTML backend of the page; the “AggregateRating” schema says “Ted 2” had 6,568 ratings, a maximum score of five stars and a minimum score of 0 stars. That all makes sense.

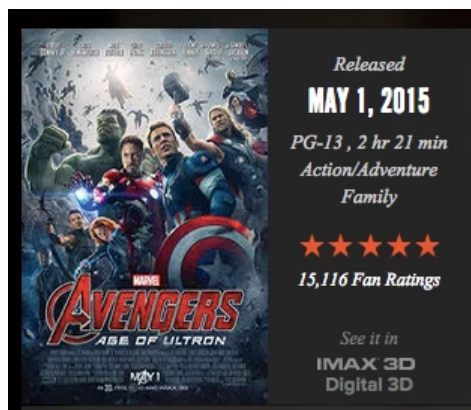
Here’s the thing, though: According to the code for the page, “Ted 2” had a “ratingValue” of only 4.1 stars.

```
<span itemprop="trailer" itemscope itemtype="http://schema.org/VideoObject">
  <meta itemprop="name" content="Ted 2">
  <meta itemprop="url" content="http://www.fandango.com/movie-trailer/ted2-trailer/169836">
  <meta itemprop="image" content="http://images.fandango.com/z98.3/imagerelay/121/87/video.fandango.com%2Ffandan
1600x1200_2651781474_gen.jpg/image.jpg/redesign/static/img/noxSquare.jpg">
  <meta itemprop="duration" content="PT2M58S">
  <meta itemprop="contentUrl" content="http://video.fandango.com/NewFandangoMovies/357335-1/481/958/357335-1-750
  <meta itemprop="embedUrl" content="http://www.fandango.com/fplayer/Player.aspx?mpxId=&genre=Comedy&w=620&title
smb=user&mid=169836&h=3494&t=R">
  <meta itemprop="description" content="Seth MacFarlane, Mark Wahlberg, Amanda Seyfried, Liam Neeson, Morgan Fre
  <meta itemprop="thumbnailUrl" content="http://video.fandango.com/FandangoMovies/357335-1_1600x1200_2651781474_
  <meta itemprop="uploadDate" content="2015-02-03" />
</span>

<span itemprop="aggregateRating" itemscope itemtype="http://schema.org/AggregateRating">
  <meta itemprop="ratingValue" content="4.1">
  <meta itemprop="ratingCount" content="6568">
  <meta itemprop="bestRating" content="5" />
  <meta itemprop="worstRating" content="0" />
</span>
</span>
```

In a normal rounding system, a site would round to the nearest half-star — up or down. In the case of “Ted 2,” then, we’d expect the rating to be rounded down to 4 stars. But Fandango rounded the “ratingValue” up. I pulled the number of stars listed on the page of each film in our sample of 437 (with at least one user review), as well as the ratingValue listed on the page’s source. And I found that Fandango doesn’t round a rating down when we’d mathematically expect that (it appears Fandango does round correctly on its mobile app — more on this in a moment).

There are even more extreme cases than that of “Ted 2.” Take “Avengers: Age of Ultron.” When I pulled data for that on Monday, the film had 5 stars from 15,116 reviews.



But according to the code for the page, “Avengers: Age of Ultron” had a “ratingValue” of only 4.5 stars, meaning that it gained a full half-star from rounding.

```

<span itemprop="aggregateRating" itemscope itemtype="http://schema.org/AggregateRating">
  <meta itemprop="ratingValue" content="4.5">
  <meta itemprop="ratingCount" content="15116">
  <meta itemprop="bestRating" content="5" />
  <meta itemprop="worstRating" content="0" />
</span>
</span>

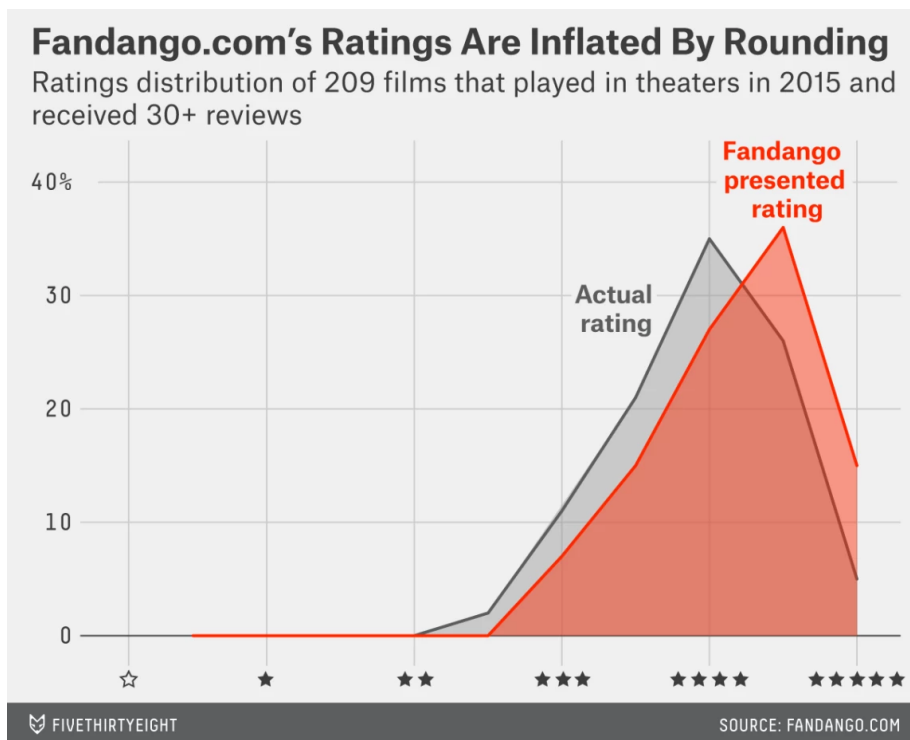
```

So what kind of effect did this have across the board?

Here's a breakdown of how Fandango.com's system handled the rounding for each of the 437 films in our sample:

- On 109 occasions, about a quarter of the time, the ratingValue was the same as the number of stars presented. This means that a movie's average rating was already at a half-star and no rounding occurred.
- On 148 occasions, about 34 percent of the time, Fandango rounded as you would expect — rounding up 0.1 or 0.2 stars,¹¹ like one would round a 3.9 or 3.8 to a 4.
- On 142 occasions (including for “Ted 2”), 32 percent of the time, Fandango added 0.3 or 0.4 stars to the rating,¹² when one would normally round down, juicing up the score by a half-star. Think of it this way: That's the equivalent of saying your SAT score¹³ was about 100 to 120 points higher than it actually was.
- On 37 occasions, about 8 percent of the time, Fandango's rounding system added a half-star to the film's rating. It's not clear why this happened — why “Avengers: Age of Ultron” would have its 4.5 ratingValue rounded up to 5 stars — but it happened about 1 in 12 times. It may be that Fandango is rounding at the second decimal place — e.g., 4.51 to 5. But again, it's not clear; the “ratingValue” in the HTML code is only shown to the first decimal place.
- On one occasion, a film was rounded up by an entire star, from a 4 to a 5.

The cases above include movies with very few reviews; the average rating for these movies is more likely to fall on a whole or half star, which doesn't require rounding. Returning to the 209 films that had 30 or more user reviews on Fandango.com, the average movie gained 0.25 stars from this rounding. Using a normal system, that average should be close to 0.



When I initially asked Fandango about its rounding practice, public relations coordinator Alison Ver Meulen said this in an email: “We always display stars rounded up to the nearest half star. So for example 3.6 stars would show up as 4

on our site.”

However, after further back and forth, the company described the rounding disparity — by which, for example, 4.1 is rounded to 4.5 — as a bug in the system rather than a general practice. “There appears to be a software glitch on our site with the rounding logic of our five star rating system, as it rounds up to the *next highest* half star instead of the nearest half star,” the company said in an emailed statement.

Fandango told us that it plans to fix the rounding algorithm on its website “as soon as possible.”

Fandango also said that “the rounding logic is accurately displayed on our mobile apps.” And that appears to be true; I checked several films that had raised red flags on the company’s website and found that their scores were accurately represented on Fandango’s iOS app. Still, the star-based scores on the app skew just as high as on the website.

Fandango.com’s rounding methodology, even if it was just an innocent bug, is a good example of why you should be skeptical of online movie ratings, especially from companies selling you tickets. If this kind of bug can survive unnoticed on the website of a major American ticket seller for who knows how long, there’s no reason a similar bug — or another issue we’re missing — couldn’t be on any other site we’re using to figure out if something is good or not.

And the Federal Trade Commission, which protects consumers from deceptive and anti-competitive business practices, pays attention to the [use of ratings and endorsements](#) to promote products. “User ratings would be material to consumers, so they have to be truthful and non-misleading,” said Mary Engle, who directs the FTC’s division related to advertising practices. Engle couldn’t comment on any specific company not already under investigation, but said there is an expectation that companies that present user ratings do so accurately. “We know that nowadays user reviews are very important, whether it’s a movie, a vacation purchase, electronics, whatever,” she said. “You go online to see what other consumers are saying. And so we’re looking at issues where those reviews aren’t what they purport to be.”

What’s The Point: Walt Hickey on the world of online reviews and the wisdom of the crowd.

Subscribe to all the [FiveThirtyEight podcasts here](#).

All of this matters more to movie studios now than it did in the past.

“If you look over the last 20 years, the release strategy used to be much more based around a movie playing for a long time, perhaps releasing regionally and building word-of-mouth around the country,” said box office analyst Bruce Nash, who operates [The-Numbers.com](#), which tracks box office data. “Today, it’s much more focused on getting into theaters opening weekend and hitting as hard as you can with the opening. For that, I think the reviews can have more of an effect.”

Looking at it this way, the idea of a studio inventing a critic to promote its films, as one [was accused of](#) in the early 2000s, starts to seem reasonable.

Fandango might be an extreme case, but its problems are indicative of the limitations of online movie ratings generally. Freelance film critic [Ben Kenigsberg](#) said, “They’re a useful shorthand or heuristic for readers, but I think they’re kind of a tongue-in-cheek way of looking at movies, and I think they should be taken as such.”

When it comes to critical aggregators like Metacritic and Rotten Tomatoes, the act of boiling down a nation’s worth of critics to a number does have an inaccurate air of finality to it. “I like both sites,” said [Todd VanDerWerff](#), culture editor at Vox.com. “But I feel like they have created a sense that there’s an answer to whether a movie is good or bad when really that’s a very personal question.”

He added: “Because it *looks* like math, we have it in our head that it’s somehow objectively true, but in reality, it’s all based on subjective experience.”

Footnotes

1. Owned by [Flixster](#), which sells movie tickets and is a Warner Bros. subsidiary.
2. Owned by [CBS Interactive](#).
3. Owned by Amazon.
4. Essentially, it’s just as possible for a film to do very poorly on [IMDb](#), [Metacritic](#) and [Rotten Tomatoes](#) as it is to do quite well. If there’s a curve, some movies manage to do very badly anyway.
5. IMDb doesn’t go into a ton of detail, but it [uses a weighted average](#) that emphasizes people who review a lot of movies. Metacritic normalizes critic ratings and [transforms them into a metacore](#) through a proprietary but consistent weighted average. And Rotten Tomatoes has [defined thresholds](#) for movies it recommends and discourages.
6. IMDb [comes right out and says](#) it calculates a film’s score knowing that people are going to try to artificially rate movies high and low. Metacritic draws from [a predetermined blind](#) weighted sample group of critics whom it deems trustworthy. And Rotten Tomatoes has [requirements about what qualifies](#) a person as a trusted critic.
7. Specifically, any film that was on Fandango’s “[movies in theaters](#)” page on Aug. 24 — as well as those that were there on April 6 or June 29, according to the Internet Archive’s Wayback machine. This allowed me to get a random sample of films.
8. To get a decent sample.
9. We also excluded three movies that had not been released yet but had still somehow gathered 30 or more reviews on Fandango.com.
10. Not all the movies sold by Fandango had IMDb or Metacritic scores — streamed live events and foreign films in particular were missed — but the final set was made up of 146 films with complete data.
11. 66 instances of rounding up 0.1 stars, 82 occasions of rounding up 0.2 stars.
12. 72 instances of rounding up 0.3 stars, 70 occasions of rounding up 0.4 stars.
13. Out of 1600.