

VOLUME 1 SUMMER 2014 CHARTBEAT.COM

# quarterly

CURIOSLY EXPLORING THE DATA SCIENCE OF THE ONLINE WORLD



Matters of time



# What's Inside

## 3 On Engagement and Viewability

WHY HIGH-QUALITY CONTENT MAKES GOOD BUSINESS SENSE

WRITTEN BY JOSH SCHWARTZ

## 4 Data-Driven Web Design

EXAMINING LINKS SIZES, DENSITIES AND CLICK-THROUGHS

WRITTEN BY DAN VALENTE

## 6 The Influence of Tweets

PARSING FIRST-PARTY AND THIRD PARTY TWITTER REFERRALS

WRITTEN BY KRIS HARBOLD

## 8 A Video Voyage

HOW VIEWERS NAVIGATE VIDEOS FROM START TO FINISH

WRITTEN BY JUSTIN MAZUR

# On Engagement and Viewability

WHY HIGH-QUALITY CONTENT MAKES GOOD BUSINESS SENSE

JOSH SCHWARTZ

On March 31, the Media Rating Council (MRC) announced it was lifting its advisory on viewable impressions for display advertising, bringing the industry one step closer to transacting on viewability for the first time.

The point at which publishers are asked to deliver highly viewable campaigns is rapidly approaching. If you haven't already started to develop a strategy to maximize the viewability of your ads, I'd wager that in the next three months, you will.

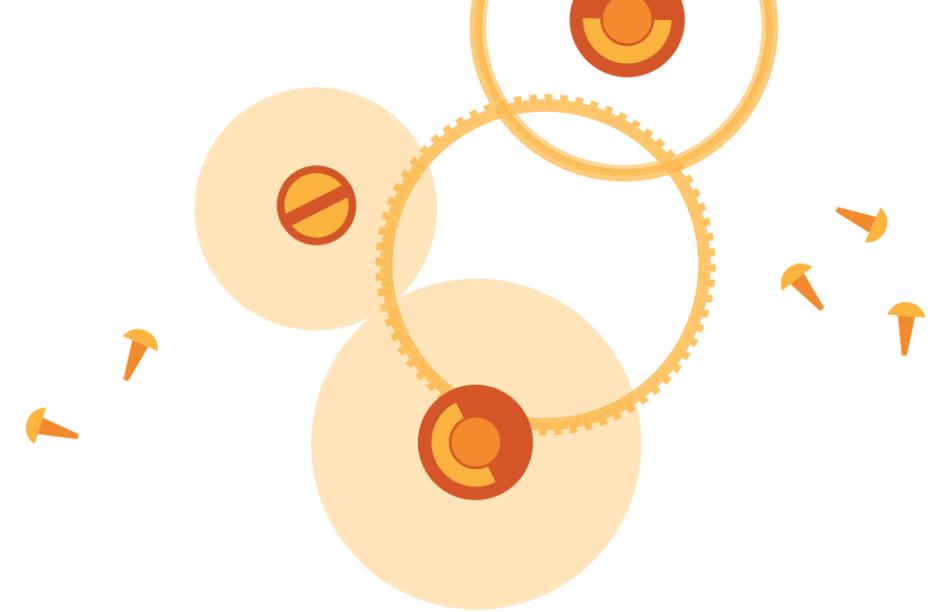
There are many tactics that can be applied to improve your ads' viewability: ensuring fast ad loads; lazy-loading advertisements; and redesigning a website to feature always-in-view units.

One issue has gotten surprisingly little discussion, though: Ads are much more viewable on pages that people actually want to read. Take a look at the figure

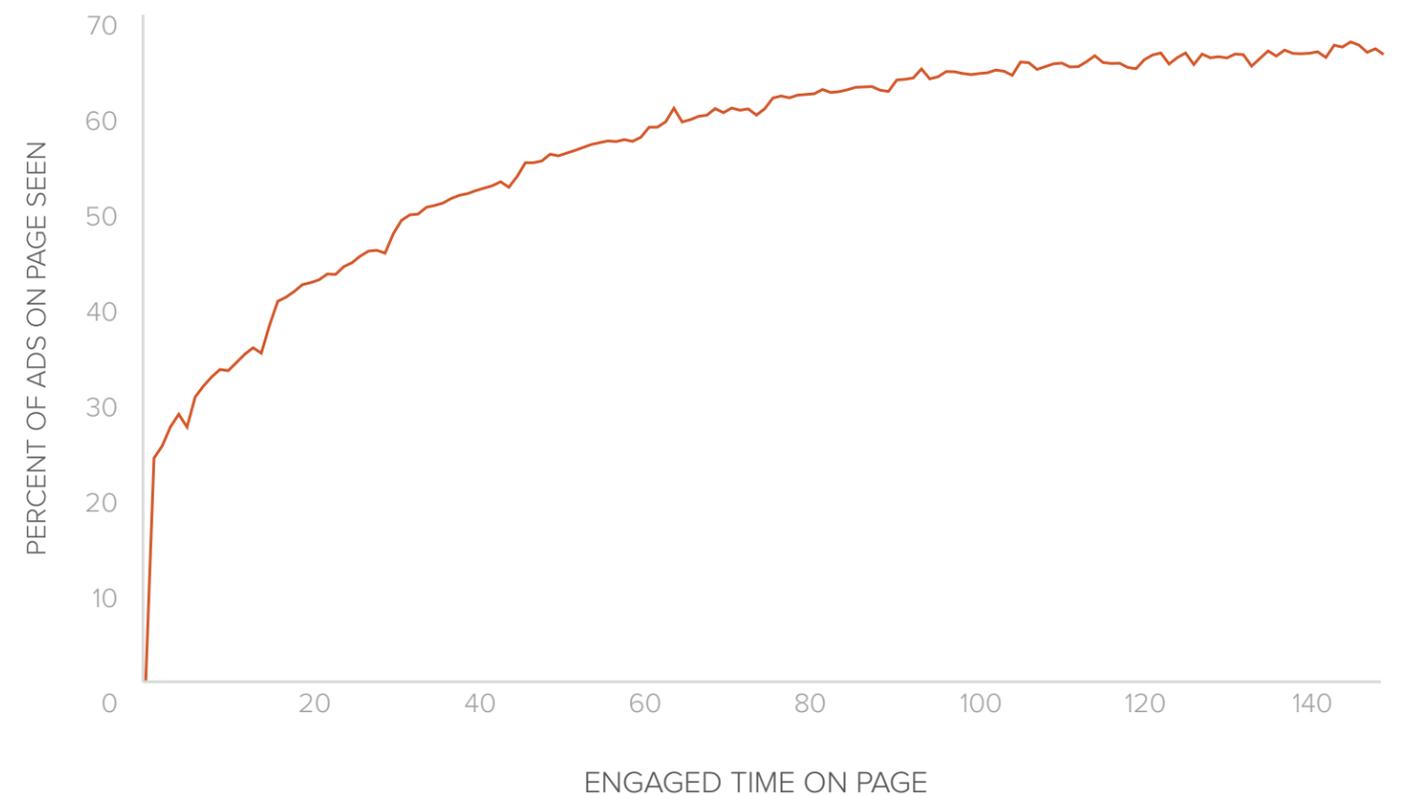
to the right, which was computed across a sample of a billion ad impressions across the month of May 2014.

We see there's a strong relationship between what fraction of ads are seen and how long a person spends reading the page: as Engaged Time increases from 15 seconds to one minute, viewability goes up by over half, from 37% to 57%. Visitors who read for more than 75 seconds see more than 60% of advertisements.

This isn't too surprising. Of course, people who read pages more deeply see more of the ads on the page, but it's still worth taking note. We've argued for years that articles with higher average Engaged Time should be promoted because they represent the articles your audience is most interested in, and—in the days where viewability is more critical than ever—promoting your most deeply read articles makes business sense, too.



THE RELATIONSHIP BETWEEN ENGAGED TIME AND AD VIEWABILITY



# Data-Driven Web Design

EXAMINING LINK SIZES, DENSITIES, AND CLICK-THROUGHS

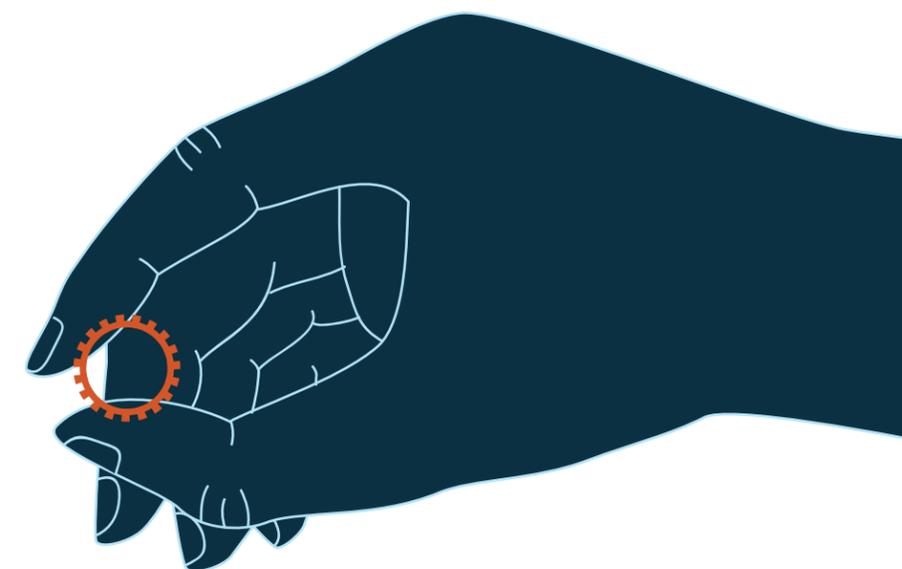
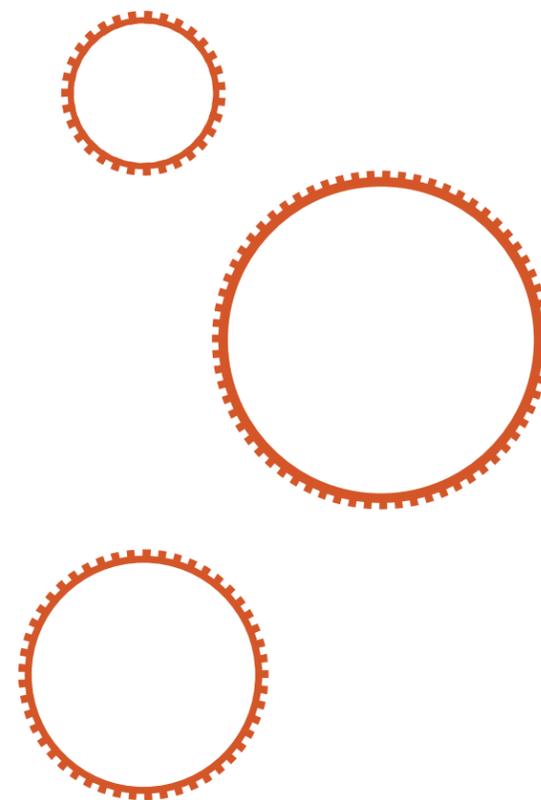
DAN VALENTE

**Many publishers would likely argue that the design of the website is as important for enticing readers to engage with the content as the content itself—humans, unfortunately, do judge books by their covers.**

We wondered if we could use our data to give insight into just how important web design is—a concept we call “data-driven web design.” Are there aspects of a page’s design

that correlate to increased traffic, and even better, increased engagement?

Font sizes and colors, link sizes, link density, interaction, responsiveness: These are elements we can analyze for their ability to draw traffic to content and perhaps even contribute (along, of course, with the content itself) to keeping people there. Do people prefer to read articles surrounded by few links, large fonts, and bright colors? Or, are sparse, simple sites with undecorated text better? For those of us keen on data, could you use these attributes to predict how many people will be drawn to the content?



Understanding how page elements relate to click-throughs is by no means a new idea. For as long as Google AdSense has been around, there have been all kinds of smart people who’ve tried to figure out just how ad size relates to click-through-rates (CTR). But ads and articles are very different beasts. Do the same rules that hold true for ads hold true for articles? Does link size matter? Is it the only thing? Are there even any rules at all?

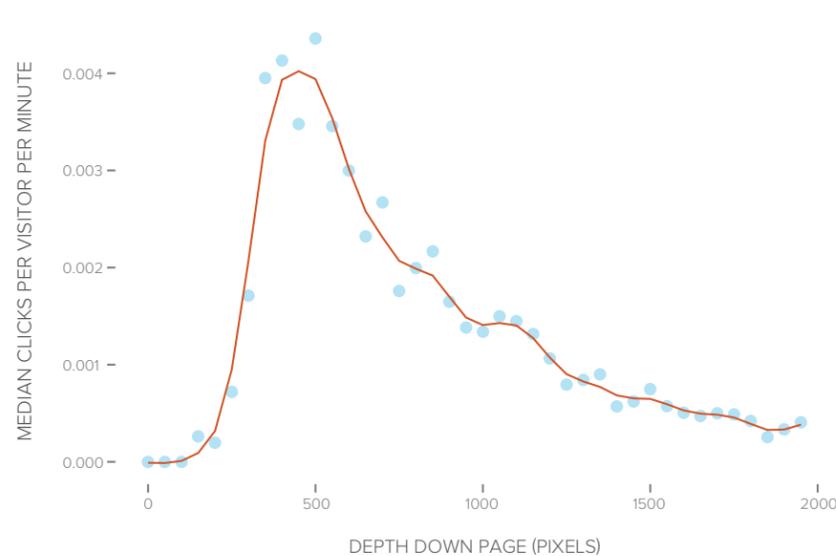
We here at Chartbeat like to focus on engagement, but as a first-pass, we wanted to examine how the almighty click-through relates to the size and distribution of links on a

homepage. We examined a measure of click-through probability, the clicks per minute per active visitor (CPV). The data used in this analysis is the same which powers one of our most popular products, our Heads Up Display (The HUD).

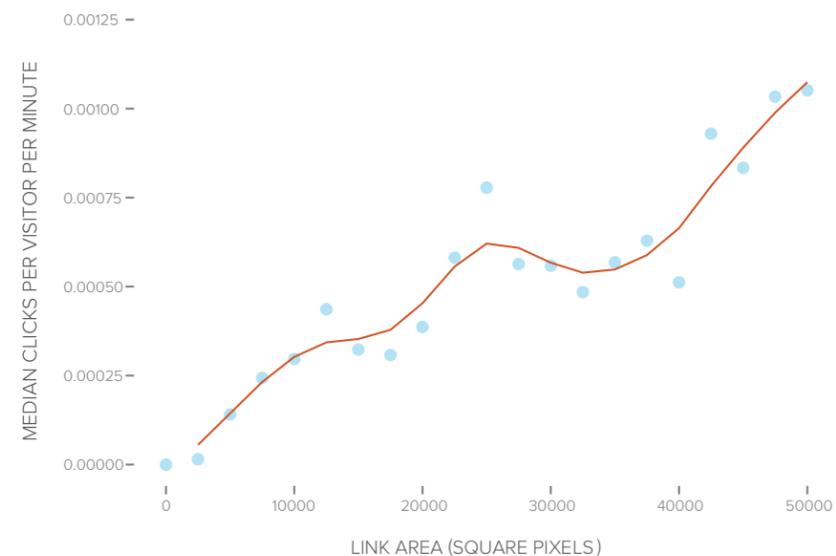
We looked at data from 294 publishing sites during several different times of day across several days to sample a variety of conditions. Much of what we found is not surprising—that is, people click where the design guides them to click. For instance, the majority of clicks happen at page depths of 400 to 600 pixels, where most main content links are located.<sup>1</sup> The other most probable places for clicks are the locations

## In reality, the data is only reflecting a common website design principle – a few large links interspersed with many smaller, closely spaced links.

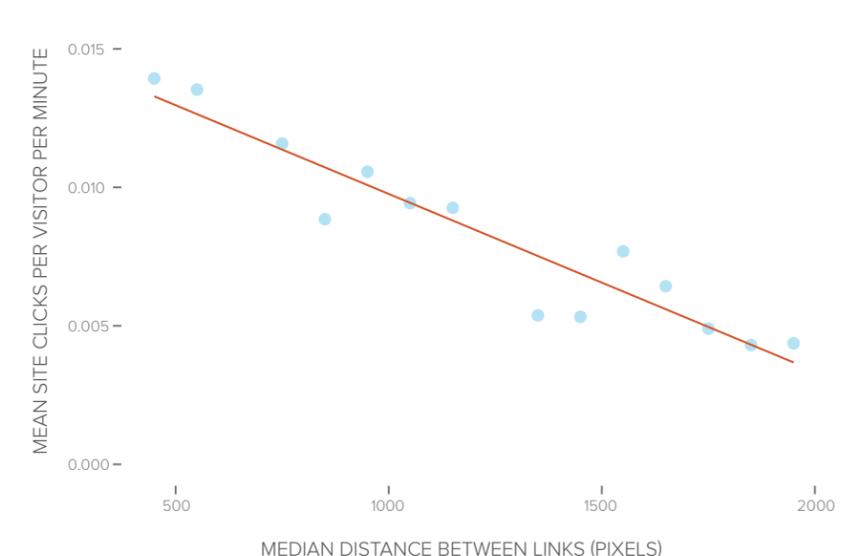
1. WHERE DO VISITORS CLICK?



2. HOW DOES LINK SIZE RELATE TO CLICK-THROUGHS?



3. DOES LINK DENSITY AFFECT CLICK-THROUGHS?



of menus on left and right sides of the page. Nothing surprising here. As far as link sizes go, intuition holds as well: One would expect larger links—which likely represent headline articles—to drive greater traffic. This is certainly true. As a link’s area grows, generally so does the clicks per active visitor. <sup>2</sup>

Larger links correlate with higher click-throughs, but what about link density? For sites with a lot of closely packed links, does this dilute click-through rates? After all, there are only so many concurrent users to split across content. As a proxy for density, we looked at the median distance between links on a site. The data shows that CPVs decrease approximately linearly for links a distance of 450 pixels apart to about 2,000 pixels apart. Sites having more closely spaced links perform about two and a half times better than sites with distant links. It

seems users prefer denser sites. <sup>3</sup>

These two pieces of evidence seem to contradict each other, though, because the distance between large links is necessarily large (assuming, of course, the links aren’t nested!). You might think, “Wait... if I have a lot of large links, I’ll have huge CPV, but they will be spaced far apart, so I’ll have a small CPV!” But, in reality, the data is only reflecting a common website design principle—a few large links interspersed with many smaller, closely spaced links.

In fact, if you ponder these data long enough, it seems that we run into a chicken-and-egg problem. Click-throughs force a tautology. Design forces people to click in certain places, so they do. And we measure this. See why engagement matters?

In any case, the data back up our intuition when it comes to determining how many people will click through to a given piece of content. Given a large enough dataset in which you know where a link is on a page, its height and width, how many people are on the page, and how many are currently engaged with content, you could likely obtain a reasonable prediction for the CPV. And perhaps using this knowledge, one might use such a model to guide the redesign of a website.

We decided to try this (not the site redesign part, the modeling part!). Simple statistical models we have recently built can predict CPV for a link to within 0.007 clicks per min per active visitor for 92% of links. This might seem impressive, but to get a foundation for what this means, only four web-

sites in the set we analyzed have a median CPV greater than this. There is much more work to do until we can really answer the question if design can predict attraction to and engagement with content, but the way forward is promising. Colors, font sizes, responsiveness—the design space is large. These can draw people in, but ultimately, it is the content that will keep people there.

So, the next time you are thinking of undergoing an overhaul or redesign, stare closely at your HUD. Think about link size, link density, and ask yourself what you can do to draw people into that fabulous content.

# The Influence of Tweets

PARSING FIRST-PARTY AND THIRD-PARTY TWITTER REFERRALS

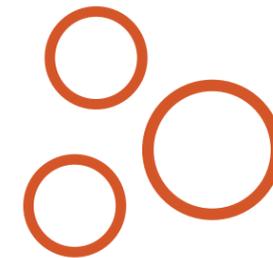
KRIS HARBOLD

**Regardless of your newsroom's size or how many articles you publish every day, chances are you've got a Twitter account.**

What's more, you've likely tried, to greater or lesser success, to leverage the social network for the distribution and promotion of your content. But once your thought-provoking, 140-characters-or-less message is dispatched, what happens next? Will the time and effort you spent pitching your editor pay off? Will you draw in readers who will actively engage with the content? Will you manage to convince readers to

explore additional articles? Could you even convince users to come back over and over again? Or, is all that effort lost in the Twitterverse, drawing in a few readers who come and leave, never to be seen again?

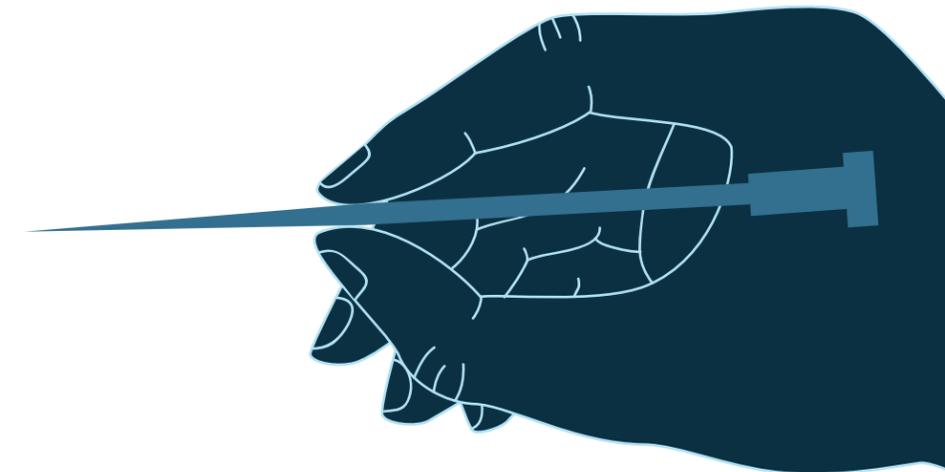
These are just a few of the questions we've been trying to answer here at Chartbeat. But rather than placing all visitors who come from Twitter into a single class and making the assumption that they all behave the same way, we decided to take a deeper dive with an eye toward nuance. We examined the behavior of readers who come from tweets published by content owners (first parties) versus those coming



from independent agents (third parties).

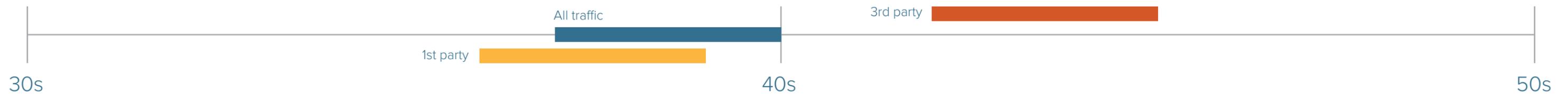
To test our assumption and measure different forms of engagement, we decided on four metrics: average Engaged Time; average number of other pages a viewer visits on a site within two hours of their first visit; percent of users who return after initially visiting; and of those users who return, the number of times on average they will return in the next 30 days. Roughly broken down, this gives us two metrics to look at short-term reader value (engagement and redirects), and two to look at long-term reader value (percent retained and rate of user return).

From previous experience and assumptions we made, it seemed to us that readers coming directly from a content owners' tweet would probably already be a member of that publisher's loyal audience. It would therefore seem logical that these users show qualities similar to that of loyal readers—chiefly that they exhibit a higher than average return rate, and read more pages when visiting. So it wasn't surprising that when observing the percent of users that came back, readers from first-party sources showed returning rates about 15% higher than readers from third parties. During their initial visit, readers coming from Twitter also tend to stick around longer, with first-party consumers reading on average



We found that readers from third parties actually engage with content significantly longer than first-party readers.

### ENGAGEMENT FROM FIRST AND THIRD PARTY TWEETS



three pages during a visit, compared to non-social traffic's one page. The difference between first-party and third-party social consumers, however, does not differ significantly.

The number of times returning users came back however was surprising. Of those users who came back at all, users coming from a first party returned on average 8 to 10 times. A similar user, though, who came from a third party came

back 11 to 13 times. This may suggest that after passing that retention barrier and convincing a reader to come back, the users you receive turn out to be much more valuable, as they return more often, and help bolster your current loyal population.

When looking at the time a reader engaged with a page, we found that readers from third parties actually engaged with

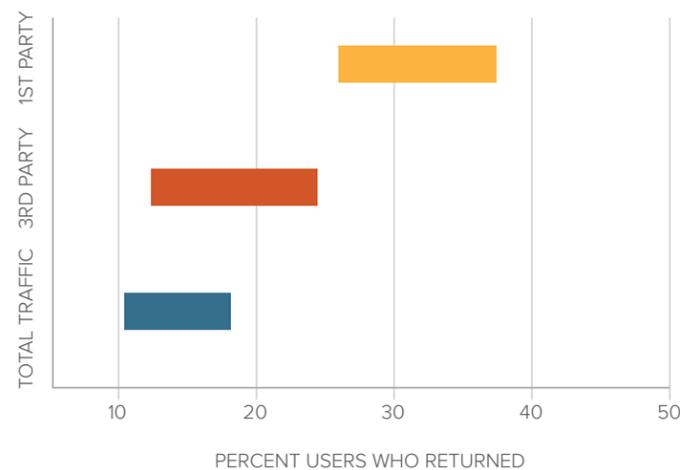
content significantly longer than first-party readers. While first-party consumers engaged with content about the same amount as any other user, regardless of where they came from (averaging between 37 and 39 seconds), third-party readers engaged on average between 42 to 45 seconds<sup>1</sup>. These differences, while seeming small, can lead to practical differences in engagement from a few seconds to nearly a 40% difference in Engaged Time.

Though there are many reasons that these differences may be occurring, one possible conclusion lies in what attracts readers to engage with content. Users who follow publishers on Twitter are apt to know more about the publisher's content; consequently having a greater sense of what type of content they want to read. Loyal readers, as opposed to new readers, may therefore be skimming through content, knowing they will come back later for follow-up stories, or to learn more. Non-loyal readers, however, who are generally the readers coming from third-party tweets, come due to the referral of a friend. These readers may engage deeply with content due to the personal connection with the recommender of that content.

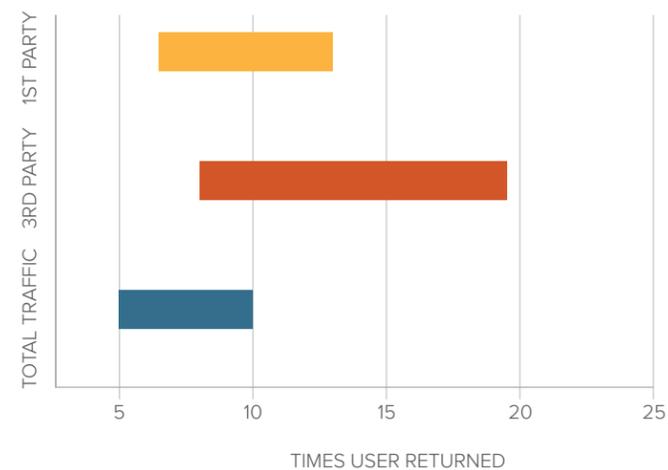
So, to answer the initial questions of whether your tweets actually matter to the health of your site: Of course they do—you already knew that. Your tweets are vital to your loyal audience, and bring in readers who return more often and consume higher quantities of content than readers coming from anywhere else. Don't forget about the importance of making content people want to tweet about, though! Because it turns out people actually listen to the recommendations of their friends, deeply engage in the content before them, and if enamored enough to return in the future, turn into fiercely loyal members of your site's virtual population.

1. Calculated with a p-value of  $p < 0.01$

### PERCENT RETURNING USERS



### VISITS PER RETURNING USER



# A Video Voyage

HOW VIEWERS NAVIGATE VIDEOS  
FROM START TO FINISH

JUSTIN MAZUR

79% OF VISITORS DON'T PRESS PLAY

21% OF  
VISITORS WILL  
PRESS PLAY

When a visitor lands on a page with video content, there's only a 9% chance that they'll watch the entire video. That's right—fewer than 1 in 10 people will watch the average video to its conclusion. But is this as bad as it sounds? Let's dig into the numbers a bit and embark on a video voyage.

To play or not to play, that is the question. And visitors to pages with video content will tend to make up their minds quickly. When a video doesn't start automatically, there is only a 21% chance of us pressing play. In fact, if a video is played over half the time, this video is in the top 25% of all manual start videos on the Internet.

12% OF VIEWERS LEAVE DURING AD

88% OF  
VIEWERS  
WATCH  
THROUGH AD

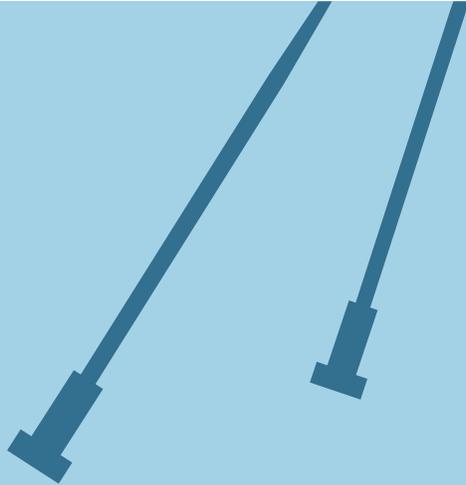
Once the video starts, viewers will often be presented with a pre-roll advertisement. Viewers might find these ads to be annoying, but almost 9 in 10 people will stick around until the actual video content begins. More specifically, once a video with a pre-roll ad has started, there's only a 12% chance that viewers will drop off.

Upon arriving at the video's content, how long will this video keep viewers' attention? It turns out, we are expected to watch at least half of the video. In fact, for videos under 10 minutes, the average viewer watches 73% of the video, while they watch 50% of longer videos. So, if you want to increase audience engagement, focus on getting visitors to just start—because once they begin, they tend to watch the majority of the content.

82% OF VIEWERS LEAVE DURING CONTENT

18% OF VIEWERS  
WATCH TO THE  
END

Last, how likely are viewers to make it all the way through to the video's conclusion? After starting content, there is only an 18% chance that viewers will make it to the very last second of the video player. Sure, very few people make it to the finish line, but at least those that start watching give you a decent chunk of their time.



PRODUCTOUTREACH@CHARTBEAT.COM | 646-218-9333