

Empowering Online Advertisements by Empowering Viewers with the Right to Choose The Relative Effectiveness of Skippable Video Advertisements on YouTube

MAX PASHKEVICH

Google, Inc.
mpashkevich@google.com

SUNDAR DORAI-RAJ

Google, Inc.
sdorairaj@google.com

MELANIE KELLAR

Google, Inc.
kellar@google.com

DAN ZIGMOND

Google, Inc.
djz@google.com

In 2010, YouTube introduced TrueView in-stream advertising—online video advertisements that allowed the user to skip directly to the desired video content after five seconds of viewing. Google sought to compare these “skippable” in-stream advertisements to the conventional (non-skippable) in-stream video advertising formats, using a new advertising effectiveness metric based on the propensity to search for terms related to advertising content. Google’s findings indicated that skippable video advertisements may be as effective on a per-impression basis as traditional video advertisements. In addition, data from randomized experiments showed a strong implied viewer preference for the skippable advertisements. Taken together, these results suggest that formats like TrueView in-stream advertisements can improve the viewing experience for users without sacrificing advertising value for advertisers or content owners.

INTRODUCTION

YouTube is the most popular online video community in the United States and in the world (comScore, 2012; NielsenWire, 2012; DoubleClick, 2011). Founded in February 2005 and purchased by Google in November 2006, YouTube in mid-2012 served more than 4 billion views per day worldwide; at that time, 60 hours of video were uploaded to the site every minute. More than 800 million unique visitors visited YouTube each month (YouTube, 2012).

Online video advertising, still a relatively new and promising concept (Plummer et al., 2007; DoubleClick, 2009), is an essential part of Google’s monetization model for YouTube. In 2012, advertisements (including advertisements from more than 1,000 small advertisers every day) were shown on more than 3 billion YouTube views each week.

YouTube advertisements appear in a variety of contexts:

- as in-stream pre-rolls and overlays in videos that users watch,
- on the right-hand side of the video watch page,
- in search results, and
- on the home page.

One of the dominant video advertising formats on YouTube (and elsewhere online) is the simple in-stream video advertisement, in which a short video—much like a television commercial—is played prior to the user-selected video content. In 2010, YouTube introduced a new variation on this format: TrueView in-stream video ads, in which the user could choose to skip directly to the desired video content after 5 seconds of viewing the advertisement.

With TrueView, advertisers are billed only if the user watches at least 30 seconds (or the complete advertisement, if it is less than 30 seconds long). Marketers have found this format immensely popular; through the middle of 2012, TrueView had

doubled in volume every quarter since it was introduced. As of July 2012, fully 70 percent of in-stream advertisements on YouTube were skippable in this way.

The main motivations behind the new skippable format were

- to improve user satisfaction with the site by reducing the negative consequences of advertisements (such as users leaving the site early), and
- to improve users' overall perception of the site (which would otherwise cause them to return less frequently).

Having an option to skip an advertisement allows users to get to the video content that they intended to watch quickly or to view the in-stream advertisement if it appears to be relevant and interesting.

The viewing options raise a number of questions that can be broadly classified as "user happiness" and "advertising effectiveness"—considerations that usually are seen as inherently conflicting in nature:

- Does making in-stream advertisements skippable indeed reduce negative advertising impacts on users? And, second, how much more attractive is YouTube to users due to the transition to skippable advertisements?
- How does user engagement with in-stream advertisements differ for these two formats? And, second, for advertisers, do their advertisements gain or lose effectiveness when users are permitted to skip them?

In this study, the authors attempt to answer these questions by using data from randomized experiments run on YouTube. They also take a deeper look at the new skippable in-stream format and compare user engagement for advertising impressions that are "viewed" (completed or watched for at least 30 seconds) to those

that are "skipped" (abandoned or skipped early) using observational data from YouTube logs. The authors then summarize their understanding of the new skippable in-stream advertising format and its relative effectiveness with respect to the traditional version.

IMPACT OF SKIPPABLE IN-STREAM ADVERTISEMENTS ON USER SATISFACTION

Randomized Experiment for Advertising Formats Holdback

The vast majority of YouTube advertisements on the watch page fall into one of the following categories:

- in-stream video advertisement (traditional or skippable);
- overlay text or image advertisement that appears in the lower center part of the player while a user is watching a video; or
- mid-page companion unit (MPU) that is a 300 × 250 display advertisement that appears to the right of a video player.

To assess the impact of one or a combination of these advertising formats on YouTube users, the authors ran a multi-arm randomized controlled experiment (Tang et al., 2010) with the control group having all advertising formats enabled ("status quo") and the experimental groups having one, several, or all advertising formats disabled.

For example:

- Users in the "no-ads" experimental group would not get any of the aforementioned formats;
- users in the "overlay-only" group only would see overlays but no in-streams and MPUs; and
- users in the "no MPU" group would only see overlays and in-streams but no MPUs.

Each group in the experiment represented approximately 0.1 percent of YouTube traffic. Users were placed into these groups by selecting random ranges of browser cookies.

A limitation of this approach was that it measured only the impact across all YouTube content when, in fact, there likely was wide variation. Users almost certainly are more tolerant of advertisements on some videos than on others. For the purpose of this article, however, the authors ignored these differences and focused only on the aggregate impact across all YouTube videos watched by each group.

Metrics to Assess User Impact of Advertisements

Total time spent by users in each experiment group watching YouTube videos (briefly referred to as "watch time") was used as a proxy to quantify the negative impact of advertising. Low watch time for a group of users would imply that they had a negative experience at YouTube and tended to leave the site early or return less frequently. On the contrary, higher watch time would suggest that those users were more satisfied with the site.

User Impact of Various Advertising Formats

Removing all three advertising formats resulted in approximately a 5.5-percent gain in watch time in the United States ("no-ads" group; Figure 1). At the same time, just removing in-stream advertisements in the United States resulted in about a 3.8-percent gain ("no in-stream" group), with the other two formats having significantly lower negative user impact ("no-MPU" and "no-overlay groups"). In other words, in-stream video advertisements had the largest negative impact on YouTube users, as has been reported previously (Dorai-Raj et al., 2011).

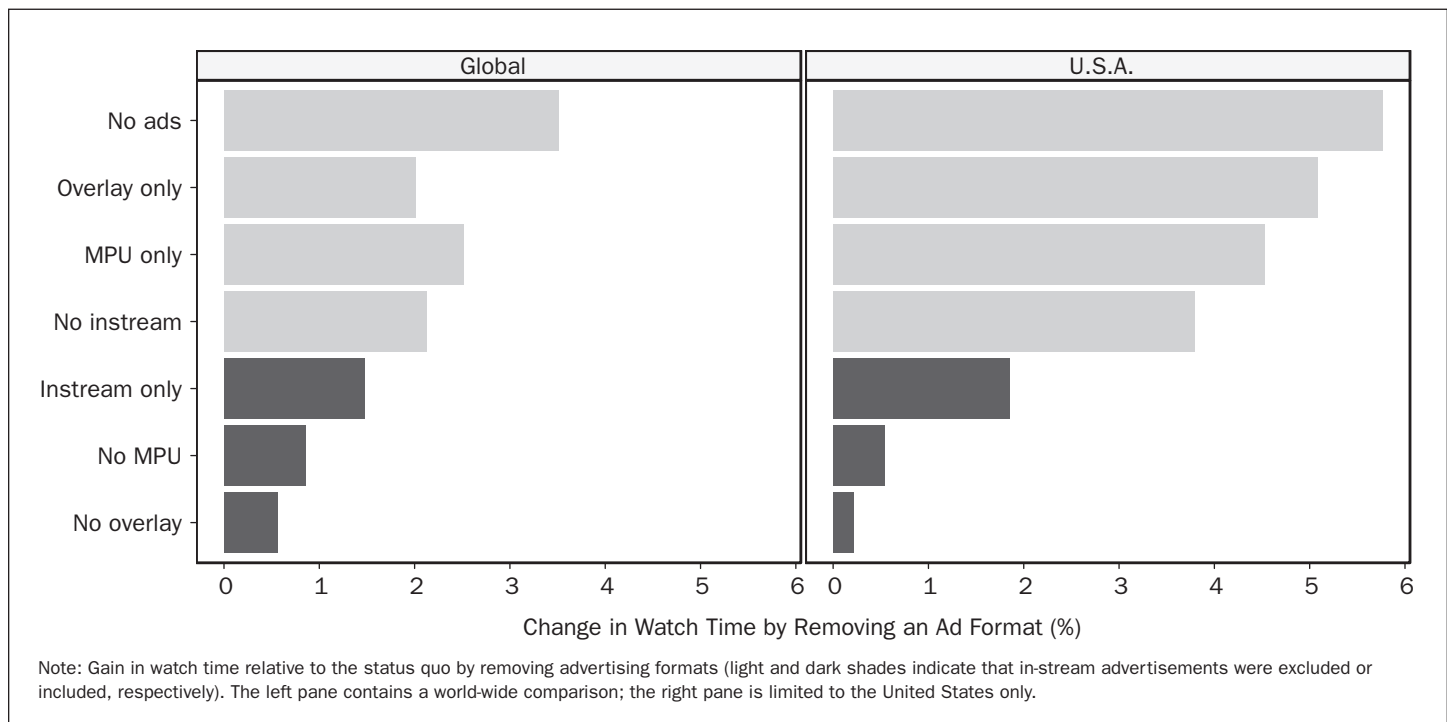


Figure 1 User Impact of Various Advertising Formats

User Perception of Various Advertising Formats

To gather qualitative feedback on user perception of YouTube advertisements, the authors administered an online survey to a subset of U.S. YouTube users in the following experimental groups:

- control,
- no ads,
- no in-stream, and
- no overlay.

More than 2,500 YouTube viewers completed survey responses during 1 week in January 2012.

Respondents were asked to rate their satisfaction with YouTube on a 7-point Likert scale (from “Extremely Dissatisfied” to “Extremely Satisfied”). Though removing all three advertising formats resulted in a 6.9-percent increase in reported user satisfaction among the “no-ads” group, simply removing in-stream advertisements resulted in a 6-percent increase in

user-reported satisfaction (Figure 2). Comparatively, the removal of overlay advertisements had a much smaller negative impact on user-reported satisfaction of YouTube.

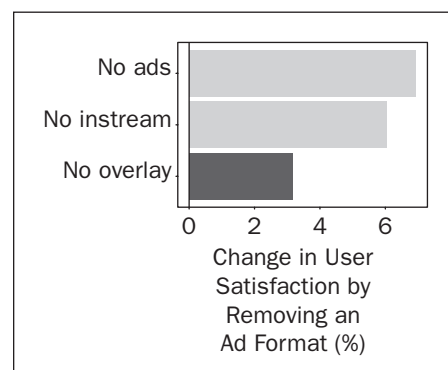


Figure 2 Gain in Reported User Satisfaction Relative to the Status Quo for the Experimental Groups by Removing Advertising Formats

Respondents also were asked to comment on what they liked and disliked most about YouTube advertisements. The authors selected and coded a random sample of 364 free-text responses.

In line with the previous findings, the biggest complaint by respondents was pre-roll in-stream video advertisements, which accounted for 38 percent of all negative feedback. (Respondents were not asked about any specific advertising formats). Further, many respondents specifically mentioned that they disliked non-skippable in-stream advertisements because they value the opportunity to skip uninteresting or irrelevant advertisements.

As one respondent stated, “Usually if an ad is interesting or relevant I will not use the skip feature because I’m genuinely interested in the content of the ad but if it is neither relevant nor interesting, then I don’t feel like I should be forced to watch the duration of a lame ad.”

Another study participant reported

frustration with the delayed access to content in non-skippable advertisements: “The ads that I can’t skip through before watching a video are incredibly irritating. The other ads are tolerable but it’s very frustrating when I’m trying to show a friend a video and we have to wait two minutes just to watch it.”

The findings from this survey provide further evidence that users prefer not to see advertisements in general and in-stream advertisements in particular.

Randomized Experiment to Compare Traditional and Skippable In-Stream Advertisements

The current study also examined whether skippable in-stream advertisements were better tolerated by users.

To compare traditional and skippable in-stream video advertising formats, the authors set up a randomized experiment in which the same creative executions were run in both skippable and non-skippable versions. The experiment had two groups of roughly 50,000 YouTube users, with the first seeing only traditional in-stream video advertisements (skippable advertisements were disabled) and the second seeing the same creative executions but as skippable advertisements (all other in-stream advertisements were disabled).

As a result, each creative execution in the experiment had impressions in both traditional and skippable form; users in the experiment saw either the skippable or non-skippable versions but not both.

Benefits of Making In-Stream Advertisements Skippable

In an analysis of watch time gain relative to the status quo due to removing in-stream advertisements (“no in-stream” group) versus making them all skippable (“no standard in-stream” group), the authors found that completely removing in-stream advertisements resulted in roughly a 3.8-percent

gain in watch time (Figure 3). Simply making all in-streams skippable, however, provided a 1.2-percent gain—a substantial improvement, as turning off all advertisements on the watch page increased the watch time by about 5.5 percent.

The results show that, by allowing a user to skip an in-stream advertisement, YouTube was able to reduce the negative impact of this format by about 30 percent. Thus, making in-stream advertisements skippable, thereby, was shown to significantly improve the YouTube experience, making the site more attractive to users.

In the “status quo” control group, about 70 percent of in-streams were skippable. To further quantify the reduction of the negative impact of in-stream advertisements obtained by making them skippable, the authors considered the following relationships based a simplified assumption that the impact of skippable and non-skippable formats on users was additive:

The authors denoted by “*s*” the penalty (lost watch time in percent relative to the status quo) when all in-streams are skippable and by “*f*” the penalty when they were forced (i.e., not skippable). Then using the aforementioned

numbers, the authors determined the following approximate relationships for the groups with no in-streams and only skippable in-streams:

$$0.7 \cdot s + 0.3 \cdot f = 3.8, \quad s = 1.2$$

Solving this system yielded $f \approx 9.87$ and $f/s \approx 8.2$.

In other words, from the perspective of lost watch time, the skippable in-stream video advertising format was roughly eight times better than the traditional one. And although the derivation of this number was based on simplified assumptions, it demonstrates the dramatic improvement in user experience due to giving a user an option to skip an in-stream advertisement.

USER ENGAGEMENT WITH SKIPPABLE IN-STREAM ADVERTISEMENTS

Engagement Metrics

The authors sought to develop a novel metric for user engagement with a video advertisement by measuring whether the relevancy of YouTube search queries with respect to the advertising content was higher after the advertisement has been viewed.

In other words: If a user searched for the term *music* prior to seeing an automobile advertisement and *cars* after seeing the advertisement, the difference in results can be seen as a sign of user engagement with the advertisement and a rough measure for advertisement effectiveness.

This “search lift” due to an advertisement was calculated for every advertising impression, and the proportion of impressions with a positive search lift was used as a proxy for total user engagement with the ad—a metric the authors have called “follow-on search” (FOS).

A binary indicator of the FOS engagement can be derived from the YouTube logs for each in-stream video advertising impression. Specifically:

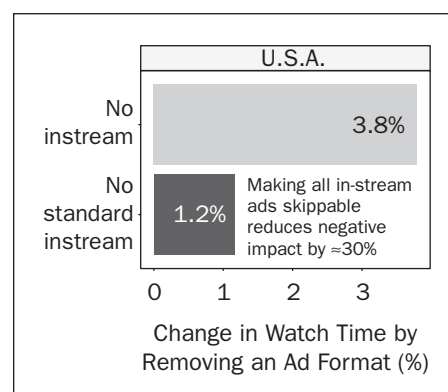


Figure 3 Gain in Watch Time Relative to the Status Quo by Making All In-Stream Advertisements Skippable

Consider a hypothetical user session where the user performed YouTube searches s_1, s_2, \dots, s_k before a particular advertising impression a (traditional or skippable in-stream), and YouTube searches $s_{k+1}, s_{k+2}, \dots, s_n$ after it. Define $w(s)$ to be the set of unique words in a search query s and $k(a)$ be a set of unique keywords describing the video creative played in the advertisement a . Then for the video advertising impression a , define the relevance score of the search query s as the proportion of keywords that were matched by this search query:

$$r(s) = \frac{\|w(s) \cap k(a)\|}{\|k(a)\|},$$

where $\|\cdot\|$ denotes the number of elements in a set.

The authors then computed the maximum relevancy of searches before and after the advertising impression:

$$R_{\text{before}} = \max_{i=1,2,\dots,k} r(s_i), \quad R_{\text{after}} = \max_{i=k+1,\dots,n} r(s_i).$$

Finally, the binary indicator of the search lift due to the video advertising impression a was derived as

$$L(a) = \begin{cases} 1, & R_{\text{after}} > R_{\text{before}} \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

In other words, the authors observed that the impression had a search lift if the maximum relevancy of search queries after the advertisement was higher than the maximum relevancy of searches before it.

FOS score is derived based on the impression-specific lifts (1) in the following manner:

Consider a set of impressions $A = \{a_1, a_2, \dots, a_m\}$ for a specific video advertisement over some period of time (a month, for example). These can be either all impressions of this video advertisement or a specific group such as only billed (or not billed) impressions if this is a skippable in-stream or only watched

to completion (completed) or skipped impressions, etc. The FOS score for the group of impressions A then is defined as the proportion of impressions that had a search lift:

$$S(A) = \frac{\sum_{i=1}^m L(a_i)}{m}.$$

This score can be derived for each video advertisement creative (and potentially its specific subgroups as mentioned above) using the data from the YouTube logs.

The principal advantage of the FOS metrics is that they can be computed mechanically for any number of advertisements without human involvement.

More subjective measures of viewer engagement—such as those based on audience surveys—almost certainly are superior for judging advertising effectiveness, but they are very difficult to conduct at this scale. These analyses involved thousands of distinct advertising creative executions. For example, developing and administering a brand-awareness survey for each one video advertisement would be impractical. Thus, though any individual advertiser would be better served by a more traditional measure, the FOS metrics provide a reasonable proxy and, the authors believe, yield useful aggregate results for large-scale studies of this kind.

Comparing Traditional and Skippable in-Stream Advertising Formats

To compare FOS engagement of the traditional and skippable in-stream advertising formats, the authors used data from the experiment described in the previous section in which the same creative executions were run as traditional (non-skippable) or skippable in-stream ads.

As the same advertising creative executions appeared in both versions, the authors were able to reduce a number of possible biases, including

- natural variability in how engaging users found different creative executions in general,
- which groups were targeted by different advertisers (e.g., younger versus older users), and
- how informative the available creative description might be.

Results obtained by the authors indicate that traditional and skippable in-stream formats do not show significant difference in FOS engagement (Figure 4).

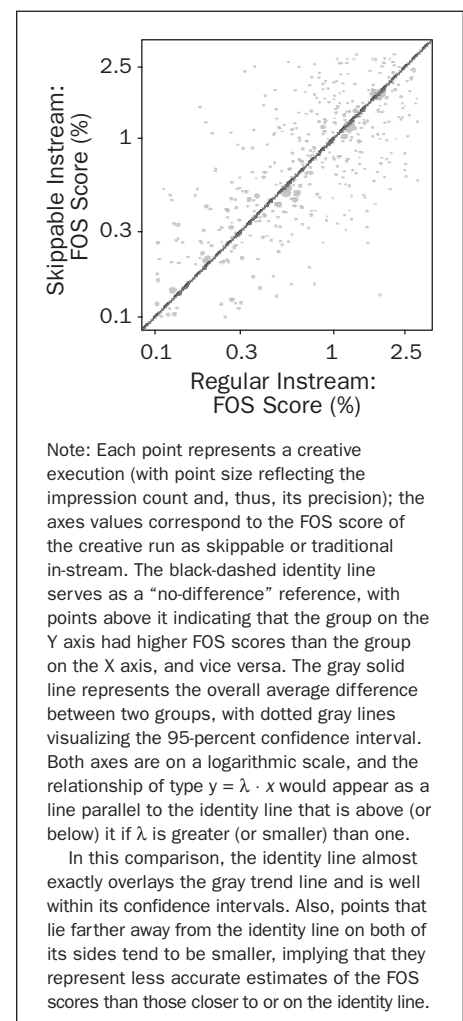


Figure 4 FOS Engagement for Skippable and Traditional in-Streams Does Not Differ Significantly

Comparing Billed and Skipped Impressions for Skippable In-Stream Advertisements

With skippable in-stream advertisements, the advertiser is billed for an impression only if a user either watches the advertisement to its completion or watches at least 30 seconds if the advertisement is longer than that. The authors, in the current study, refer to these impressions as “billed” and to the rest of the impressions (those that were skipped early or abandoned) as “skipped.”

Given this billing model, it is natural to investigate whether billed impressions result in higher effectiveness. In other words, is YouTube charging for the right impressions?

To compare the FOS engagement for the billed and not skippable in-stream impressions, the authors used data for all skippable impressions on non-music content in the United States for 2 months. The non-music content was chosen for comparison because such viewers are known to be more active, making the comparison more accurate (Figure 5).

On non-music content, skippable impressions that become billed views resulted in about 1.2 times higher FOS engagement than those in which a user either abandoned or skipped an advertisement before 30 seconds. This suggests that search behavior on the site is, in fact, influenced by advertising views and that FOS score is, therefore, a meaningful measure of user engagement. It also implies that billed views should be more effective than skipped views due to higher user engagement.

Comparing More and Less Active Completers for Skippable In-Stream

To further validate the proposed follow-on search engagement metrics, the authors compared the FOS engagement for completers of skippable in-stream

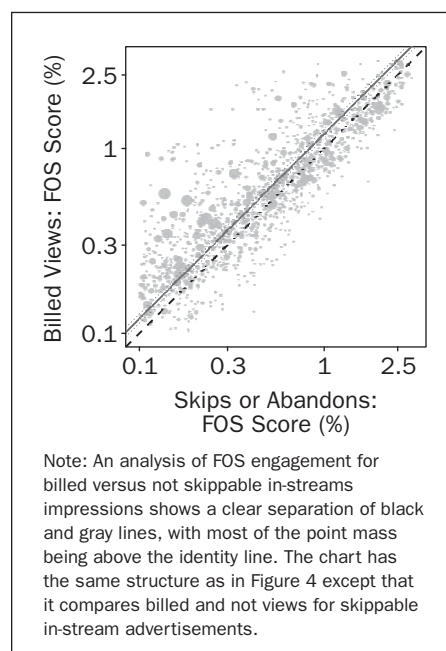


Figure 5 FOS Engagement for Billed Skippable In-Streams Impressions Is Higher than for Skipped or Abandoned Ones

advertisements that had keyboard or mouse activity during the playback after a video advertisement (“more active” completers) to those who did not (“less active”).

User engagement likely falls across a broad spectrum, and users who are more active during video playback (using their mouse or keyboard) may be more engaged during the advertisements. So, it is reasonable to expect that “more active” completers of skippable in-stream advertisements have higher FOS engagement than “less active” ones. The authors’ analysis yielded results consistent with that, with the relative difference being approximately 1.5 (Figure 6). This finding further supports the validity of the proposed engagement metrics—a strong signal consistent with the expected direction, with more active YouTube users more engaged

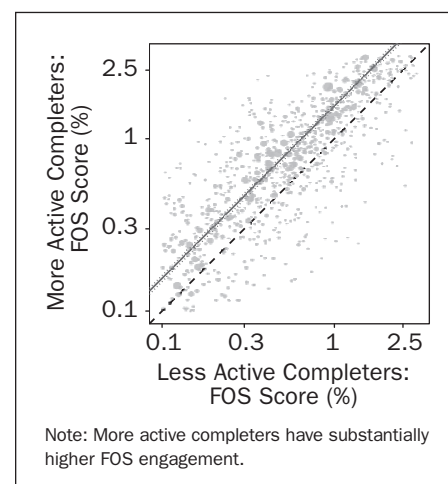


Figure 6 FOS Engagement for More versus Less Active Completers of Skippable In-Streams

with the skippables advertisements that they choose to complete.

DISCUSSION

The results of the current study indicate that YouTube skippable in-stream advertisements deliver effectiveness comparable to the traditional in-stream format while reducing the negative impact of advertising on users.

To arrive at this finding, the authors

- measured time spent watching videos on YouTube when traditional in-stream advertisements were present and when they were replaced by skippable versions;
- compared user-reported satisfaction with the site when in-stream advertisements were enabled and were not;
- analyzed user comments on their perception of advertisements; and
- quantified user engagement with in-stream advertisements using later search queries on the site.

This work has several important limitations. As noted, the negative impact of

advertising on users was computed across all YouTube content, but there still exists wide variation across specific content types and particular channels. This impact also depends on the specific advertising load and will vary from month to month as more (or fewer) advertisements are sold. The aggregate number presented in the current study is unlikely to be representative of the impact of advertisements on any particular YouTube video or channel at any specific time.

Furthermore, the FOS effectiveness metric used for this research was chosen not because YouTube search behavior is a perfect proxy for user engagement or advertising effectiveness. Rather, the authors sought to find a metric that could be applied universally to all advertising campaigns and creative executions without human intervention from the advertisers or YouTube.

Although the FOS metric ensures a valid comparison at a large scale (e.g., across all U.S. traffic), it may be too noisy to use at the level of an individual advertiser in the majority of cases. Traditional survey-based approaches may be more promising here, and YouTube currently is working on developing other advertiser- and campaign-specific effectiveness metrics that will better suit this purpose.

CONCLUSIONS

Online video advertising often is presented as a zero-sum game in which the interests of advertisers and users are inherently at odds. In the current study, the measurements of search activity before and after video advertisements on YouTube suggest that this need not be the case.

In fact, YouTube's TrueView in-stream video advertisements appear to have succeeded in substantially reducing the negative user impacts of online advertising without sacrificing the value of such advertisements to advertisers. This is a substantial accomplishment.

Online video advertising often is presented as a zero-sum game in which the interests of advertisers and users are inherently at odds.

Furthermore, these data imply that online video advertising really does work: Viewing such advertisements affects later user behavior and causes users to pursue relevant search queries in the future.

Finally, giving users the choice to view (or not view) may actually increase this advertising effectiveness by engaging users in the advertising process. In this way, empowering users to choose the advertisements they watch online need not come at the cost of advertiser value but actually appears to serve the interests of advertisers and content owners and users. **JAR**

.....
MAX PASHKEVICH is a quantitative analyst at Google. He holds an MSc in applied mathematics and computer science and a PhD in statistics from the Belarusian State University and did his postdoctoral studies at Stanford University.

.....
SUNDAR DORAI-RAJ is a senior quantitative analyst at Google. His areas of interest include linear models and statistical computing. He has a PhD in statistics from Virginia Tech.

.....
MELANIE KELLAR is a senior user experience researcher at Google. Her research is focused on understanding the online video consumption experience. She has a PhD in computer science from Dalhousie University.

.....
DAN ZIGMOND is an engineering manager at Google, where he directs data warehousing, reporting, and quantitative analysis for YouTube. He holds a BA in computational neuroscience from the University of Pennsylvania.

REFERENCES

- COMSCORE. April 2012. U.S. Online Video Rankings, 2012. Retrieved July 10, 2012 from http://www.comscore.com/Press_Events/Press_Releases/2012/5/comScore_Releases_April_2012_U.S._Online_Video_Rankings
- DORAI-RAJ, S., H. LIU, and D. ZIGMOND. "Advertising and Traffic: Learning from Online Video Data." Presentation at the ARF Audience Measurement conference, 2011.
- DOUBLECLICK. "Research Report: The Brand Value of Rich Media and Video Ads, 2009." Retrieved July 10, 2012, from http://static.googleusercontent.com/external_content/untrusted_dlcp/www.google.com/en/us/doubleclick/pdfs/DoubleClick-06-2009-The-Brand-Value-of-Rich-Media-and-Video-Ads.pdf
- DOUBLECLICK. "Top 1000 sites, 2011." Retrieved July 10, 2012, from <http://www.google.com/adplanner/static/top1000/index.html>,
- NEILSENWIRE. "May 2012—Top U.S. Online Video Sites, 2012." Retrieved July 10, 2012, from http://blog.nielsen.com/nielsenwire/online_mobile/may-2012-top-u-s-online-video-sites/.
- PLUMMER, J., S. D. RAPPAPORT, T. HALL, and R. BAROCCI. *The Online Advertising Playbook: Proven Strategies and Tested Tactics from the Advertising Research Foundation*. Hoboken, NJ: John Wiley & Sons, 2007.
- TANG, D., A. AGARWAL, D. B. O'BRIEN, and M. MEYER. "Overlapping Experiment Infrastructure: More, Better, Faster Experimentation." In *Proceedings of the 16th Conference on Knowledge Discovery and Data Mining*. ACM, Washington, DC, 2010.
- YOUTUBE STATISTICS. 2012. Retrieved June 20, 2012, from http://www.youtube.com/t/press_statistics.