Acceptable Ads Committee

In-view ad refresh survey

Measuring ad-blocking users' perceptions of in-view ad refresh

Date: December 2022

Performed by: Dynata and eyeo GmbH

Commissioned by: The Acceptable Ads Committee

Table of Contents

List of Figures

List of Tables

About the Acceptable Ads Committee

Over 250MM online users worldwide have chosen to exclusively receive Acceptable Ads. Established in 2017, the Acceptable Ads Committee (”AAC”) is a non-profit organization whose objective is to protect the user experience, while simultaneously providing publishers and content creators with meaningful monetization opportunities. The AAC does so by conducting independent research into the acceptability of various ad formats and codifying the results in the Standard for what constitutes an Acceptable Ad. One of the few advertising bodies that provides a voice for internet users, the AAC aims to maintain a sustainable open internet by balancing the needs of its stakeholder representatives, including users, publishers, advertisers, ad tech companies, and digital rights organizations.

Abstract

The Acceptable Ads Standard ensures that ads are not intrusive to the users' web experience. To suggest changes to the Standard, the Acceptable Ads Committee (AAC) is required to base their recommendations on data about users’ perceptions of new types of ads or on evolving perceptions of existing types of ads.

The present study investigates the sentiment of ad-blocking users towards websites featuring an ad refresh implementation. Ad refresh refers to a web page changing (or ‘refreshing’) ad content within an already-served ad unit without the page itself necessarily reloading. To analyze sentiments towards ad refresh, a single ad refreshing ad was implemented on one of two mock webpages: a scrollable single page in the form of a long article; and, a non-scrollable single page in the form of an online game. To explore potential implementations of ad refresh, different ad format sizes, placements and ad refresh time-based triggers were tested. These ad refresh implementations were compared with the same websites showing static ads in compliance with the latest Acceptable Ads Standard. Participants were asked to rate the ad refreshing ad as well as their overall ad experience on the webpage. The survey was distributed to 7,427

ad-blocking participants from the US, Germany, and France. Results showed that certain implementations of ad refresh are acceptable to users. Specifically, when an ad refreshing ad is placed adjacent to the primary content, the following ad formats and refresh trigger rates were rated positively by users:

Scrollable single-page (e.g. Article experience)

- 300x250 banner ad (or any smaller size) in the side rail, with an ad refresh time-based trigger rate of 30 seconds or longer.
- 160x600 banner ad (or any smaller size) in the side rail, with an ad refresh time-based trigger rate of 30 seconds or longer.

Non-scrollable single- page (e.g. Game experience)

- 300x250 banner ad (or any smaller size) in the side rail, with an ad refresh time-based trigger rate of 30 or 60 seconds, provided that there are no other ads in the viewport.
- 160x600 banner ad (or any smaller size) in the side rail, with an ad refresh time-based trigger rate of 60 seconds or longer, provided that there are no other ads in the viewport

The following cases for a long-dwell (non/minimal-scroll) webpage show an inconclusive picture:

- 300x250 banner ad (or any smaller size) in the side rail, with an ad refresh time-based trigger rate of 90 seconds, provided that there are no other ads in the viewport.
- 160x600 banner ad with an ad refresh time-based trigger rate of 30 seconds.
- 300x250 banner ad for 30, 60 seconds when a top-banner ad is present.
- 728x90 top banner ad refresh in the presence of an ad to the side of the primary content.

In contrast to previous studies, socio-demographic factors did not play a significant role on the overall ratings of the different advertisements. Only participants from Germany perceived the ads in general more negatively compared to participants from the US and France. Additionally, the general attitude towards online advertising is again a strong predictor of the ad ratings: the more negative the attitude the worse the ratings.

1. Introduction

The Acceptable Ads Committee (AAC) commissioned a study to explore user perceptions of in-view ad refresh. The AAC is an independent, non-profit organization responsible for setting the standards for what constitutes an Acceptable Ad – an ad that respects user experience while delivering value to publishers and advertisers. The AAC is composed of three equally represented coalitions: for-profit (advertisers, ad tech companies, publishers, and content creators), expert (user and creative agents, researchers, specialists in online advertising) and user (digital rights organizations and ad-blocking users) groups. The role of the AAC is to set ad-filtering ad standards online, but unlike other institutions or industry bodies that advocate for a better advertising experience for general users, the AAC focuses primarily on the experience of ad-filtering users. These users do not block all ads but are willing to accept certain ad formats or placements if their overall user experience is not negatively affected.

The Acceptable Ads Standard currently prohibits in-view ad refresh, described in the section "What is not considered an acceptable ad?" as: "ads that visibly reload new ads if there is no change to the Primary Content[1]". A concern is that refreshing ad units can negatively affect user experience (e.g., increased bandwidth, decreased page performance, negative visual experience or disrupted attention) and may be less valuable to advertisers owing to concerns about the viewability of ad units (Fennell, 2019; Jatain, 2020; Southern, 2020b). Certain implementations of ad refresh, such as site-initiated ad refresh, are currently not recommended by some companies or ad servers for this reason (e.g., Google AdSense; setupad, 2021). However, research on the negative effects of ad refresh on user experience is limited, and it is unclear whether certain implementations of ad refresh would be deemed acceptable to ad-filtering users. The aim of the present study was to evaluate how users employing an ad blocker rate these ads as well as their overall website experience when ad refreshing ad units are included in the Acceptable Ads ad experience.

[1] The 'Primary Content' is defined as (based on the WHATWG Specifications description of the HTML element): Content that is directly related to, or expands upon the central topic of a document or the central functionality of an application.

2. Literature Review

2.1. Role of in-view ad refreshing units in the advertising industry

Ad refresh is the automatic serving or changing of advertising content, either directly prompted by a user (user initiated) or set by a website (site initiated; Interactive Advertising Bureau, 2017). In order to guarantee that the ad refresh is visible to the study's participants, we test only in-view sticky ads and ads that are permanently viewable owing to the website's layout. Ad refresh is generally recommended for websites with long-dwell times, high engagement, low bounce rates, and where the viewability of an ad unit can be isolated, for instance, on service-based websites, gaming sites, blogs, online media sites, or pages with infinite scroll (Jatain, 2020; PubGalaxy, 2021; setupad, 2021; Southern, 2020a).

There are three main types of in-view ad refresh implementations:

- ***User action***: an ad refresh occurs based on a user-guided action (e.g., clicking a "read more" button to reveal previously hidden content, scrolling, or the user reaches the end of an article).
- ***Event action***: the ad refresh occurs when a publisher-defined event is performed (e.g., refresh or update of webpage content)
- ***Time***: ad refresh occurs according to a pre-specified time interval (e.g., 30 seconds)

These implementations can be executed individually, in combination, or in addition to other page view metrics (e.g., viewability). For instance, time and event-based triggers can be combined such that an ad only reloads when users are active and the ad has been in-view for a pre-specified time (Graham, 2021; PubGalaxy, 2021; Quentel-Reme, 2020). Incorporating viewability into ad refresh implementations can help overcome potential negative effects of ad refresh inventory on CPMs, for instance by increasing advertiser's confidence in the value of the publisher's inventory (PubGalaxy, 2021, Southern, 2020). Recommendations to publishers include using longer time intervals to guarantee better viewability, and not replacing an ad with an ad of a different format

(e.g., ad format size) or replacing an ad in a way that increases Cumulative Shift Layout (the stability of the website layout) (Jatain, 2021; setupad, 2021).

However, industry acceptance of ad refresh implementations is not universal, and companies and ad servers can follow different policies or establish different guidelines for ad refresh (Google Ad Manager Help, 2022, Interactive Advertising Bureau, 2017; PubGalaxy, 2021, setupad, 2021). In general, transparency in sharing ad performance metrics and clearly defined metrics are advised to facilitate exchange between publishers and advertisers and to ensure there are clear definitions and criteria for implementation and cost evaluation (Expósito-Ventura et al., 2021).

2.1.1. Ad refresh from the perspective of publishers and content providers

The appeal of ad refresh to publishers and content providers that it increases the number of ad impressions on a page and therefore correspondingly increases the amount of inventory publishers have available to sell (Southern, 2020a). However, implementing ad refresh can come with certain trade-offs, such as decreasing viewability and click-through-rates. This can have a negative impact on page RPM (advertising revenue per 1000 page views), and impact traffic if implementations slow down page performance and damage the user experience (Fennell, 2019; Quentel-Reme, 2020; Southern, 2020a). For instance, Opti Digital found that despite an increase in the volume of impressions, global auto-refreshing of ads could have a negative impact on page RPM as webpages that included auto-refresh resulted in lower click-through rates, viewability rates and eCPM impressions (Quentel-Reme, 2020).

2.1.2. Ad refresh from the perspective of advertisers

Despite Google creating standards for ad refresh on its exchange in 2016, practices are generally not standardized and not all vendors follow recommendations about declaring ad refresh inventory (e.g., that they are selling a refreshing ad, or the time trigger the refresh is based on; Fennell, 2019; Southern, 2020). Further, advertisers often do not know which impression they are buying (e.g., the first, second or other; Fennell, 2019). Accordingly, ad refreshing inventory tends to be valued less by

advertisers who may offer lower bids for each successive refresh owing to concerns about viewability (Jatain, 2020). However, a recent academic paper evaluating the click-through-rate (CTR) for mobile ad refreshing units found that, in contrast to reports from industry data, the CTR for subsequent ads in a session can be higher as variety draws attention to the ad (Rafieian & Yoganarasimhan, 2021). Nevertheless, certain ad refresh implementations can be more appealing to advertisers, such as employing longer time intervals and/or imposing minimum viewability thresholds (PubGalaxy, 2021; setupad, 2021).

2.2. Role of in-view ad refreshing units for users

Some of the above-noted recommendations for ad refresh do take into consideration the effect of ad refresh implementations on user experience. For instance, implementations that refresh ads of differing sizes are considered to negatively affect the user experience as the page jumps around when the ad refreshes (Southern, 2020a). Similarly, ad refresh implementations increase bandwidth usage compared to non-refreshing ads which can have negative impacts on users with limited or low bandwidth (Jatain, 2020; Publift, 2022). Pairing user-action or event-based triggers with time and viewability thresholds is also thought to result in a better user experience (PubGalaxy, 2021). However, it is not clear how these assumptions have been tested with users (e.g., A/B testing with different implementations on websites).

Only one study has examined the effect of ad refresh on actual user experience by evaluating perceptions of refreshing ad units. In 2016, the Coalition for Better Ads (CBA) included two ad refresh formats in their evaluation of consumer perceptions of desktop and mobile ads (Coalition for Better Ads, 2016). A standard 300x250 image ad appeared in-line with the content of a short article and was replaced with another ad every 15 or 30 seconds. Out of 40 different ad formats tested in the study, the 30-second ad refresh was ranked 5/40 for the most favored experience while the 15-second ad refresh was ranked 15/40. These results suggest a somewhat favorable view of ad refreshing units relative to other ad formats tested in the study (e.g., inline, sticky, and prestitial ads). Testing ad refresh ad units in combination with other ads on a webpage (e.g., multiple

ads), and measuring participant engagement (e.g., time spent on a webpage, awareness of ad refresh) will help extend these findings to establish whether ad refresh could be incorporated into Acceptable Ads.

To our knowledge, there have been no additional studies evaluating the effects of ad refresh on user experience from the perspective of the users themselves (i.e., there are studies evaluating optimal ad refresh rates for different publisher metrics, such as click-through-rates, or user responsiveness to ads, Rafieian & Yoganarasimhan, 2021; Truong, 2016). As such, the present study would be the first study to evaluate the disruptiveness of ad refresh on user experience from the perspective of users who have an ad blocker on their devices. Further, the study aims to take a comprehensive approach by testing multiple ad refresh time triggers, for different ad format sizes and placements, and on different web experiences (e.g., scrollable single-page and non-scrollable single-page sites).

2.3. Background research and input from AAC and industry representatives

To narrow the scope of the ad refresh project, most common use-cases for ad refreshing units and the types of web experiences where these should be tested were identified. A review of academic and industry literature on the use of ad refreshing units was conducted. Additionally, several entities participating in the Acceptable Ads program were also asked to share insights on their use of in-view ad refresh. Two publisher representatives and an advertising representative from AAC were consulted to provide feedback on the survey design, too.

Based on this research and feedback, the study was designed to focus on the following ad refresh type and web experiences:

- **Desktop**: Focus on desktop (vs. mobile or other device) implementations of ad refresh.

- **Viewability**: The ad refreshing ad unit would be viewable at the time it reloaded new advertising content.
- **Single ad refresh unit**: Only one ad refreshing ad unit would be tested on a webpage. Multiple ad refreshing units would require extensive technical implementation and lead to an exponential increase in the number of study conditions that would need to be tested.
- **Ad refresh units that reload ads of the same dimension**: Reloading ads of different dimensions can be more disruptive to users as they change the visual layout of the webpage in addition to the ad unit itself (Southern, 2020b). For these as well as technical and budgetary reasons, testing different ad size/dimension reloading units was not feasible.
- **Web experience**: There are multiple potential use-cases for web experiences to display ad refresh. Two use-cases were selected as indicative of common experiences:
    1. A webpage where the user is engaged with content on a single-page that is consumed by scrolling down the page (e.g., reading a long news article or blog post). **This experience is referred to as the "Article experience"**.
    2. A webpage where the user is engaged with content on a single-page for some time without the need to scroll, for instance, owing to active engagement with content or where content updates passively (e.g., a gaming website or gallery page, or a weather/stock reporting webpage, respectively). Of these, a game experience would ensure the user was actively engaging with the page content for at least a few minutes to be able to view ad refreshing units. **This experience is referred to as the "Game experience"**.

Further, the scope was narrowed to focus on the following ad refresh implementations:

- **ad refresh time-based triggers**: The trigger for ad refresh was time-based and consisted of the following time intervals: 30 seconds, 60 seconds, and 90 seconds. The minimum time was chosen based on Google's requirement which does not allow any ad refresh time shorter than 30 seconds. Ad refresh may also be implemented in conjunction with a viewability threshold (e.g., that a portion

of the ad unit needs to be within the user's viewpoint to be counted as part of the time trigger) or a user-initiated action or trigger (e.g., scroll). However, it was not feasible to technically implement these additional criteria within the study design. Nevertheless, the web experiences for the study were designed in such a way to guarantee the ad refreshing unit would be viewable at the time it reloaded new advertising content (e.g., sticky-ad unit or a non-scrollable webpage) and efforts were taken to motivate active engagement with the webpage. As such, both the time trigger and viewability threshold would be satisfied.

- **Placement of ad refresh unit**: Ad refresh placements would be selected such that they would be in-view for the user at the time of the ad refresh. Thus, a sticky right-rail ad unit would be viewable for a scrollable single-page (Article experience), or a right rail or top banner ad unit would be viewable for a webpage without scroll (Game experience).
- **Ad format size**: Ad format sizes were selected based on common ad format sizes while meeting the Acceptable Ads Standard for above the fold and below the fold size restrictions: 728x90 for a top banner ad and either a 160x600 or 300x250 for the right rail ad (see also section 3.2 for additional context).

2.4. Additional design considerations for web environments

2.4.1. Realistic Acceptable Ads web experiences

In previous research commissioned by the AAC, studies have been designed to test ad formats in isolation: web experiences are created to test a single ad unit on a page without additional ad units present.

The present study sought to test more realistic ad experiences for Acceptable Ads users and evaluate the *addition* of an ad refreshing unit to the ad experiences currently allowed under the Acceptable Ads Standard.

2.4.2. Engagement with web experience

Engagement with the web experience would be crucial to ensure the ad refresh was viewable to the participant at the time it reloaded new content. Two strategies to promote engagement with the webpage for the study were incorporated.

First, to increase the probability that the content would be of interest to participants, participants had to choose between one of two different Article or Game experiences to complete the study. To decide on content that would be appealing to a broad audience, a pilot study was conducted where a small sample of participants (n=299) from different age groups were asked to rank their interest in reading five different Article titles and playing four different Games. The two top-ranking Articles/Games were selected for the present study.

Second, prior to viewing the web experiences, participants were told that it was very important that they engage with the webpage and that they would be asked three comprehension questions following the experience. A minimum viewing time was set at 150 seconds so that participants were not able to move forward in the study until this threshold had been reached.

3. Study methodology and survey design

3.1. Participant Recruitment

Representatives of the online consumer population were recruited by a third-party provider 'Dynata' using their online recruitment platform. Respondents were recruited from the US, Germany and France, as these three countries are the three biggest ad-blocking markets. Respondents who were 18 years and older were invited to participate in an online study. Respondents were asked an initial screener question to determine whether they were using ad blocking software and only those who met this criterion were invited to complete the study. Quota sampling was applied such that the study sought to recruit an equal representation of gender (1:1 male-to-female ratio) and proportions of participants from different age groups representative of typical ad

blocker usage age groups. Participants were required to complete the survey on desktop computers.

3.2. Study Design

The study sought to explore ad refresh implementations that varied according to: (a) ad placement, (b) ad format size and (c) ad refresh time-based trigger. However, as the two web experiences differed in the ad units that could be shown while meeting the Acceptable Ads Standard size requirements, the design of the ad experiences were tailored to the Article and Game webpages.

3.2.1. Article experience

The Article experience was designed as a scrollable single-page article containing a combination of image, text and gallery content. Under the current Acceptable Ads Standard, a scrollable single-page article could contain multiple ad units as long as they occupy no more than 15% of the viewable area above the fold and 25% below the fold. Given that the study required the ad refreshing ad unit to be viewable at the time it reloaded new ad content, a sticky right rail ad was selected. That is, the right rail ad would always be viewable as the user scrolled through the Article content.

Fitting within the Acceptable Ads size requirements, the top banner ad (728x90) and a right rail ad (300x250) are among the most common ad formats presented above the fold that would meet the size criteria when shown in combination, and common right rail ad sizes (300x250 or 160x600) would meet the criteria when presented below the fold with an in-content ad (728x90 or 840x150).

A realistic Acceptable Ads Standard Article experience without an ad refreshing unit was created as a comparison for other study conditions. The Acceptable Ads Standard Article experience included:

- Top banner ad: 728x90 [shown *above* the fold]
- Sticky right rail ad: 160x600 [shown *below* the fold]

- In-content ad: 728x90 or 840x150 [shown *below* the fold]

An illustration of the study conditions is provided below in Figure 1.

The remaining study conditions were created by varying the following criteria:

- Ad refresh time-based trigger: 30 seconds, 60 seconds, 90 seconds
- Right rail ad format size: 160x600 or 300x250
- In-content ad: present or absent

The inclusion of the in-content ad placement as a test condition served to evaluate ad refresh in cases where in-content ads were not incorporated in the webpage primary content. It would also enable an evaluation of whether ad experiences that included an in-content ad were associated with similar or more negative user experience than those without an in-content ad.

Therefore, the study included an additional 12 Article experiences (ad refresh time-based trigger [*30 seconds, 60 seconds, 90 seconds*] x right rail ad format size [*160x600, 350x200*] x in-content ad [*present vs. absent*]). An illustration of the study conditions is provided below in Figure 2.

Figure 1. The Acceptable Ads Standard control Article experience study condition (no ad refresh). Ads are highlighted with the blue boxes.

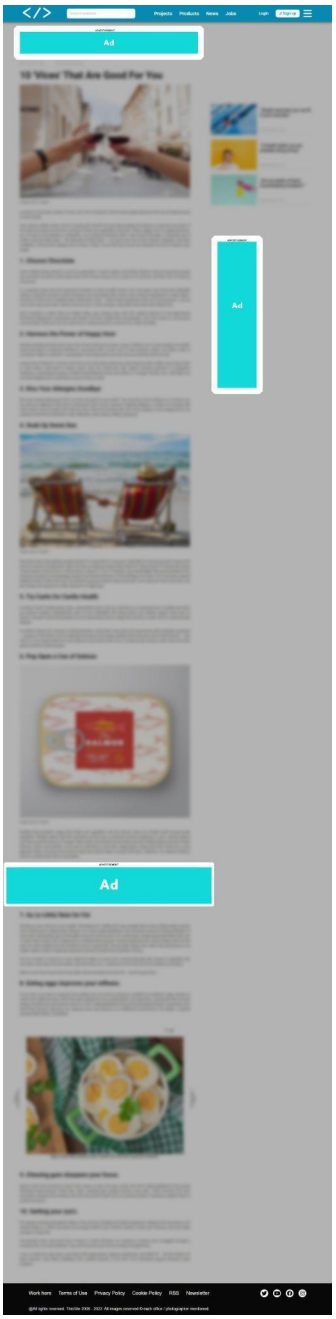
Acceptable Ads Standard experience with a 782x90 in-content ad	*Acceptable Ads Standard experience with a 840x150 in-content ad*
	

Figure 2. Study conditions showing web experiences with and without in-content ad present, and with different ad format sizes for the sticky right rail ad refresh. For each of these conditions, the ad refresh time-based trigger was set at either 30, 60 or 90 seconds. Ads are highlighted with the blue boxes.

	In-content ad: **present**		In-content ad: **absent**	
	In-content ad: **728x90**	In-content ad: **840x150**		
Sticky Right Rail: **160x600**				*Compare right rail ad format sizes: 160x600 vs. 300x250*

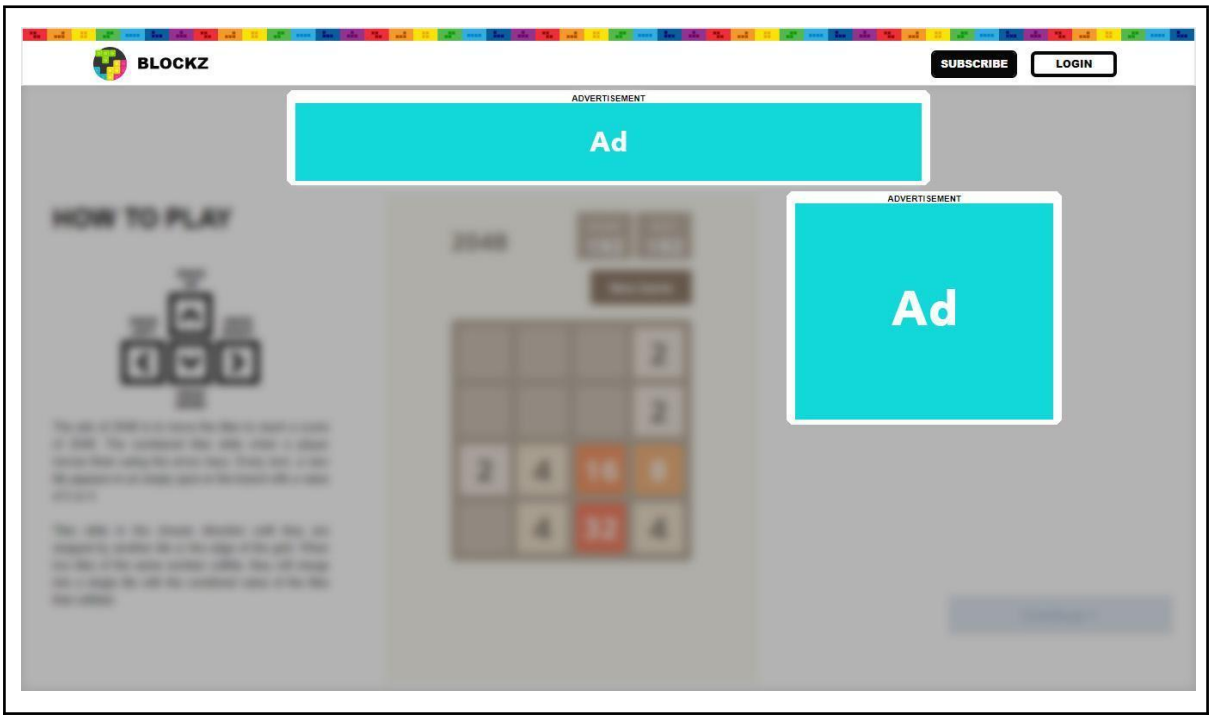
Sticky Right Rail: **300x250**			
	Compare in-content ad format sizes: 728x90 vs. 840x150		
	Compare ad experiences with vs. without in-content ads		

Note: Comparing web experiences in the left and right panels allows for the evaluation of ad experiences with and without the inclusion of in-content ads. Comparing web experiences in the lower and upper panels allows for the evaluation of different ad refresh formats sizes.

3.2.2. Game experience

The game experience was designed as a long-dwell non-scrollable page where the primary purpose is for the user to engage in content presented within a single viewport; as such, the webpage would need to ensure ads occupied no more than 15% of the viewable area (above the fold size requirement). The Acceptable Ads control ad experience contained a top banner ad (728x90) and right rail ad format size (300x250) alongside the game. An illustration of the study conditions is provided below in Figure 3.

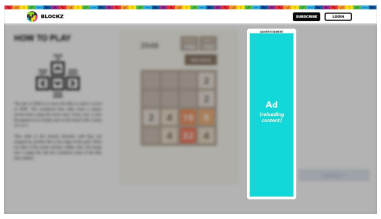
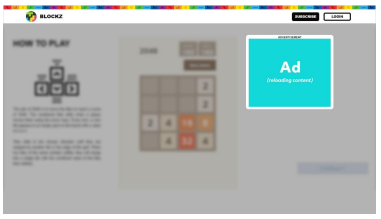
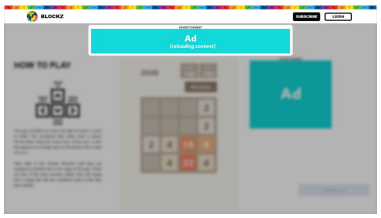
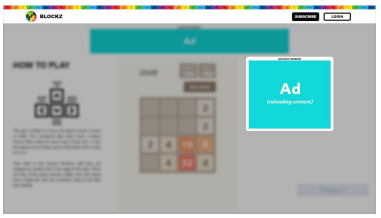
Figure 3. The Acceptable Ads Standard control Game experience study condition (no ad refresh). Ads are highlighted with the blue boxes.



Further, as the user would see all ads displayed within the viewport at all times, ad experiences with different ad refresh placements (top banner vs. right rail ad) as well as different right rail ad format sizes (160x600 and 300x250) were created. Specifically: (1) when a top banner (728x90) and right rail ad (300x250) were both shown on the page, the refreshing ad could be shown in either placement; (2) if no top banner ad was shown, a larger right rail ad refreshing unit could be tested (160x600).

Thus, the study included an additional 12 Game experiences (ad refresh time-based trigger [*30s, 60s, 90s*] x right rail ad format size [*160x600, 350x200*] when the top-banner ad *is not present*; and ad refresh time-based trigger [*30s, 60s, 90s*] x ad refresh placement [*top-banner vs. right rail ad*] when the top-banner ad *is present*). An illustration of the study conditions is provided below in Figure 4.

Figure 4. Study conditions showing ad experiences for the Game experience with and without a top-banner ad present, different right rail ad format sizes, and ad refreshing ad placements. For each of these conditions, the ad refresh time-based trigger was set at either 30, 60 or 90 seconds.

<table>
<tr><td></td><td colspan="2">Compare ad experiences for different right rail ad format sizes (160x600 vs 300x250)</td></tr>
<tr><td></td><td>Right rail ad refresh:
160x600</td><td>Right rail ad refresh:
300x250</td></tr>
<tr><td>Top banner ad:
Absent</td><td>  </td><td>  </td></tr>
<tr><td></td><td colspan="2">Compare ad refresh ad placements (top banner vs. right rail ad refresh)</td></tr>
<tr><td></td><td>Top banner ad refresh:
728x90</td><td>Right rail ad refresh:
300x250</td></tr>
<tr><td>Top banner ad:
Present</td><td>  </td><td>  </td></tr>
</table>

Note: Comparing web experiences in the upper panel allows for the evaluation of ad refreshes for different ad format sizes in cases where the webpage does not include a top-banner ad. Comparing web experiences in the lower panel allows for the evaluation of different ad refresh placements on webpages containing both a top-banner and right rail ad.

3.2.3. Ad creatives

A set of ad creatives of different sizes were developed for display on the web experiences. The content for each ad creative was designed to be a unisex product that would be equally appealing across genders. The brands used were mock brands, ensuring that the brands' effect could be isolated from any existing or known brands. An overview of the ad creatives designed for each ad unit and refresh location can be found in Appendix 1.

3.2.4. Study preview

The following link provides access to the original survey:

https://survey-d.yoursurveynow.com/survey/selfserve/53c/2211256?list=0&ot=DumSID&C=2&W=1&decLang=english

Please note that the link will remain accessible for three months (until March 2023) and that assignment to any of the testing experiences is random (the link can be accessed multiple times in order to view different testing experiences).

3.3. Study Measures and Procedure

After consenting to participate in the study, participants completed the study questionnaire (the questionnaire can be found in Appendix 2). Participants provided basic demographic information (i.e., age, gender, highest level of education and income category) before being randomly allocated to complete one of the Article or Game experiences. Participants were informed that they would be presented with a choice between two Articles/Games and that they would be required to engage with the webpage for a minimum of two and a half minutes before being able to continue with the survey. Participants were told that it was important that they engage with the webpage as they would be asked three questions about its content.

For the Article experience, participants could choose between two topics:

- "10 'Vices' That Are Good For you"
- "17 Little Cooking Habits You Should Actually Ditch ASAP"

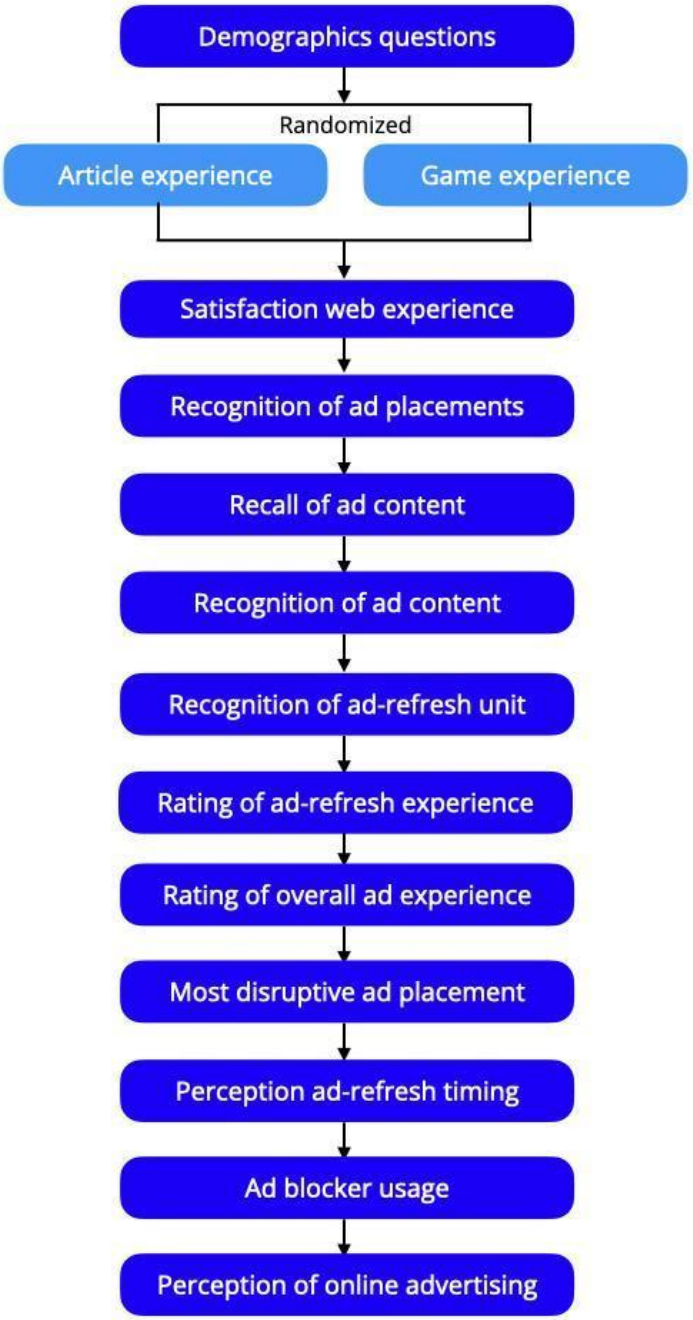
For the Game experience, participants could choose between:

- Tetris-like a puzzle game
- 2048-like a puzzle game

After viewing the webpage, participants rated their satisfaction with the overall web experience and completed the three comprehension questions. Participants were asked to indicate the ads that they recognized appearing on different locations of the webpage, to list the content of the ads they recalled, and to complete the ad recognition task. Participants were also asked to rate the disruptiveness, annoyance, enjoyableness and intrusiveness of the ad refreshing unit[2]. All participants then rated their overall ad experience – that is, their experience with all ads shown on the webpage – using the same rating scales and identified the ad that was most disruptive to their web experience. Participants concluded the study by reporting their perception of acceptable ad refresh timing, their perception of the most disruptive ad (see Appendix 3 for results), their current ad blocker usage, and by rating their overall attitudes towards online advertising (based on Redondo & Aznar's questionnaire; 2018). The study flow is shown in Figure 5.

[2] Participants who indicated that they did not recognize ads appearing on the webpage were told that they would be shown different ads that may have appeared on the webpage prior to completing the ad recognition task and ad refresh ratings.

Figure 5. Study flow



3.4. User testing and soft launch of survey

Unmoderated user tests were performed with six participants who completed either the Article experience or the Game experience. The results of these user tests were used to evaluate the comprehensibility of the survey questions and identify any issues with the survey flow or wording. Feedback from user tests were integrated into the survey. As a

second step, the English survey was launched to a subset of 714 participants (353 completed the Article experience and 361 completed the Game experience) to check the implementation and programming of the survey, and to ensure participants were randomized to the different conditions as intended.

3.5. Criteria for determining acceptability

Participants rated the disruptiveness of a) the ad refreshing ad unit itself ("How disruptive was the reloading of new ad content to your experience?") as well as the disruptiveness of b) the ad refreshing ad unit as part of the overall ad experience ("How disruptive were ALL the shown ads to your overall experience?"). Both ratings are essential for evaluating the disruptiveness of in-view ad refresh for the Acceptable Ads Standard, as ad refreshing ad units will rarely be implemented in isolation and need to be evaluated in terms of how they affect the overall ad experience. As such, only refreshing ad units that meet the 35% acceptability criteria for both disruptiveness ratings should be considered further for their acceptability.

4. Sample Description

4.1. Demographics

The study recruited a total of 7,427 participants. Participants who demonstrated scrolling behavior that indicated they did not engage with the webpage (e.g., scrolled immediately to the end of the Article web experience and waited until they could continue the study) or who stated that they did not use an ad blocker when asked which ad blocker they used (n=6) were filtered out of the data a priori. A total of 7,134 participants were included in the analysis: USA (n=2,518), Germany (n=2,083) and France (n=2,533). As the age distribution of the sample differed slightly from the internet using population (see Table 1), the data was weighted to better represent the internet using population. The survey aimed for an even gender distribution, which was achieved with a male-to-female ratio of 1.

Table 1. Socio-demographic characteristics of the sample relative for each country and in reference to the internet using population.

	USA n (%)	**Germany** n (%)	**France** n (%)	**Age distribution of the internet using population**
Age category				
18-24	468 (18.6%)	223 (10.7%)	473 (18.7%)	18%
25-34	799 (31.7%)	589 (28.3%)	748 (29.5%)	32%
35-44	441 (17.5%)	457 (21.9%)	501 (19.8%)	19%
45-54	361 (14.3%)	362 (17.4%)	370 (14.6%)	14%
55-64	260 (10.3%)	268 (12.9%)	258 (10.2%)	10%
65+	189 (7.5%)	184 (8.8%)	183 (7.2%)	7%
Gender				**Male-to-Female ratio of one**
Male	1216 (48.3%)	1144 (54.9%)	1259 (49.7%)	50%
Female	1261 (50.1%)	932 (44.7%)	1265 (49.9%)	50%
Other	31 (1.2%)	7 (0.3%)	8 (0.3%)	-
Prefer not to say	10 (0.4%)	NA	1 (0%)	-
Education[a]				
Low	44 (1.7%)	301 (14.5%)	726 (28.7%)	
Middle	2010 (79.8%)	959 (46%)	1199 (47.3%)	
High	434 (17.2%)	819 (39.3%)	583 (23%)	

Prefer not to say	30 (1.2%)	4 (0.2%)	25 (1%)	
Income[b]				
Low	754 (29.9%)	888 (42.6%)	673 (26.6%)	
Middle	1129 (44.8%)	899 (43.2%)	1505 (59.4%)	
High	545 (21.6%)	188 (9%)	227 (9%)	
Prefer not to say	90 (3.6%)	108 (5.2%)	128 (5.1%)	
Perceptions of online advertisement[c]	**Median** *[95% confidence interval]*	**Median** *[95% confidence interval]*	**Median** *[95% confidence interval]*	
18-44 years	3 [3.0, 3.1]	3.1 [3.1, 3.2]	3.2 [3.2, 3.3]	
45 years +	3.3 [3.3, 3.4]	3.5 [3.4, 3.6]	3.7 [3.6, 3.7]	

Notes: [a]According to International Standard Classification of Education (ISCED): Low includes completion of lower secondary education, Medium refers to upper secondary education through to post-secondary non-tertiary education, and High refers to education inclusive and beyond short-cycle tertiary education. [b]Low income is defined as an annual gross income lower than 75% of the disposable income per capita of each country. Middle income is defined as an annual gross income between 75% and 200% of the disposable income per capita of each country. High income is defined as an annual gross income higher than 200% of the disposable income per capita of each country. Each income was converted using the purchasing power parity (source: The World Bank, 2021). [c]Responses to 11 questions assessing participant's perceptions of online advertising on a 5-point scale with lower scores representing more positive attitudes. See Appendix 3 for further details.

4.2. Ad blocker usage

Participants were asked to indicate the main ad blocker they are currently using. The most common ad blocker was Adblock, followed by Adblock Plus, the simultaneous use of Adblock Plus and Adblock, and AdGuard. Differences between the countries are negligible. The distribution of ad blockers can be found in Table 2.

Table 2. Ad blocker usage by participants across countries

Ad blocker	USA	Germany	France
Adblock	1093 (43.2%)	883 (35.1%)	752 (36.1%)
Adblock Plus	459 (18.1%)	453 (18%)	403 (19.3%)
AdGuard	62 (2.4%)	300 (11.9%)	108 (5.2%)
Other	66 (2.6%)	203 (8.1%)	142 (6.8%)
Adblock Plus+Adblock	202 (8%)	105 (4.2%)	136 (6.5%)
uBlock	116 (4.6%)	186 (7.4%)	150 (7.2%)
Brave	102 (4%)	90 (3.6%)	71 (3.4%)

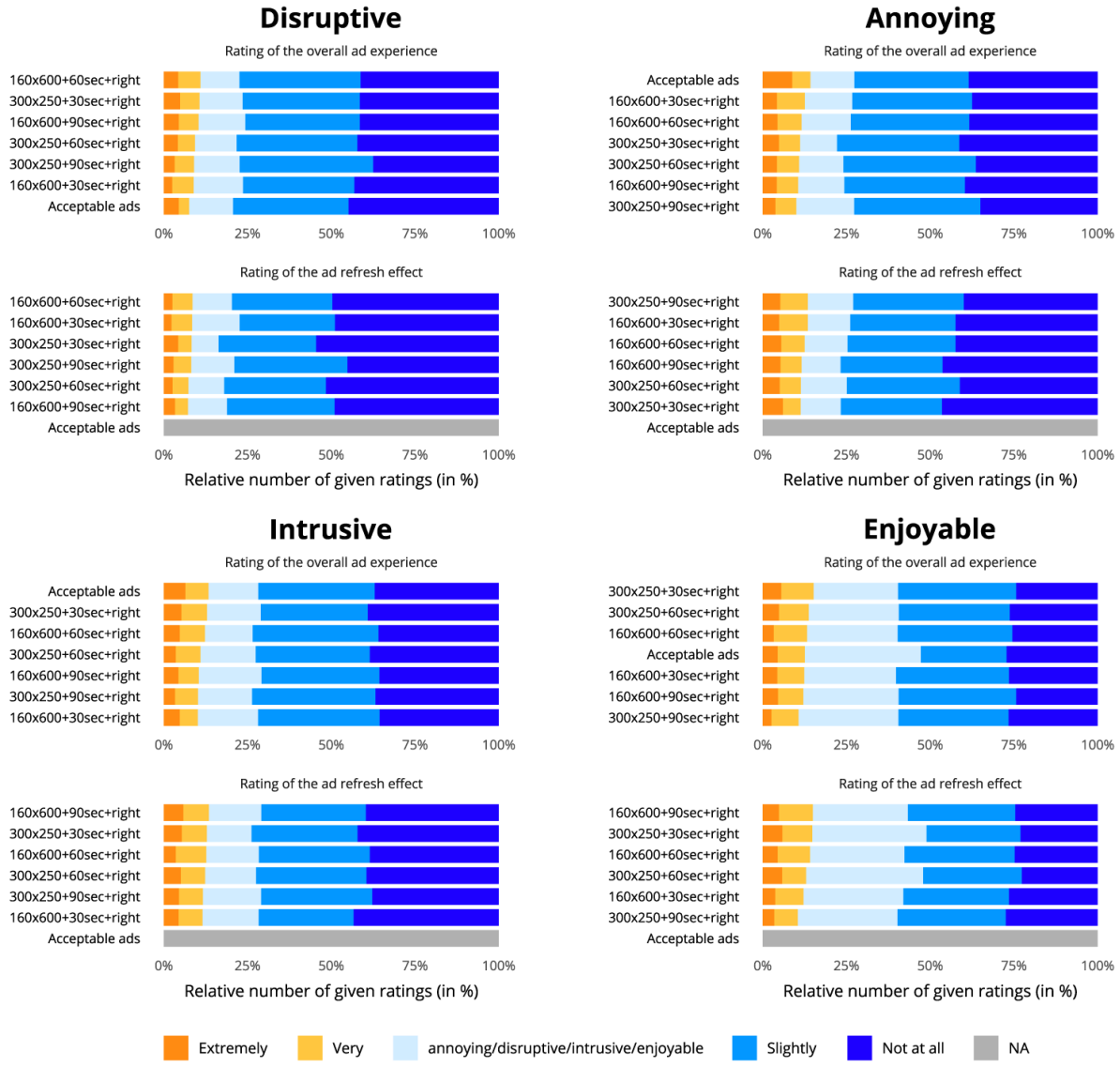
5. Study Engagement

As stated in the introduction of the report, in order to elicit reliable ratings, an important goal of the ad refresh study was to create realistic web experiences that were engaging for the participants. In general, participants spent more than the set minimum amount of time on the webpage, demonstrated expected scrolling behavior for the Article experience, and the majority answered 2 out of 3 of the comprehension questions correctly. Around a third of participants in the Article experience and half of the participants in the Game experience did not recall seeing any ads. When both the right rail ad and the top banner ad were shown, the right rail ad was more likely to be recognized. Further details about how participants engaged with the web experience and with the ad experience can be found in Appendix 3.

6. Individual ad ratings

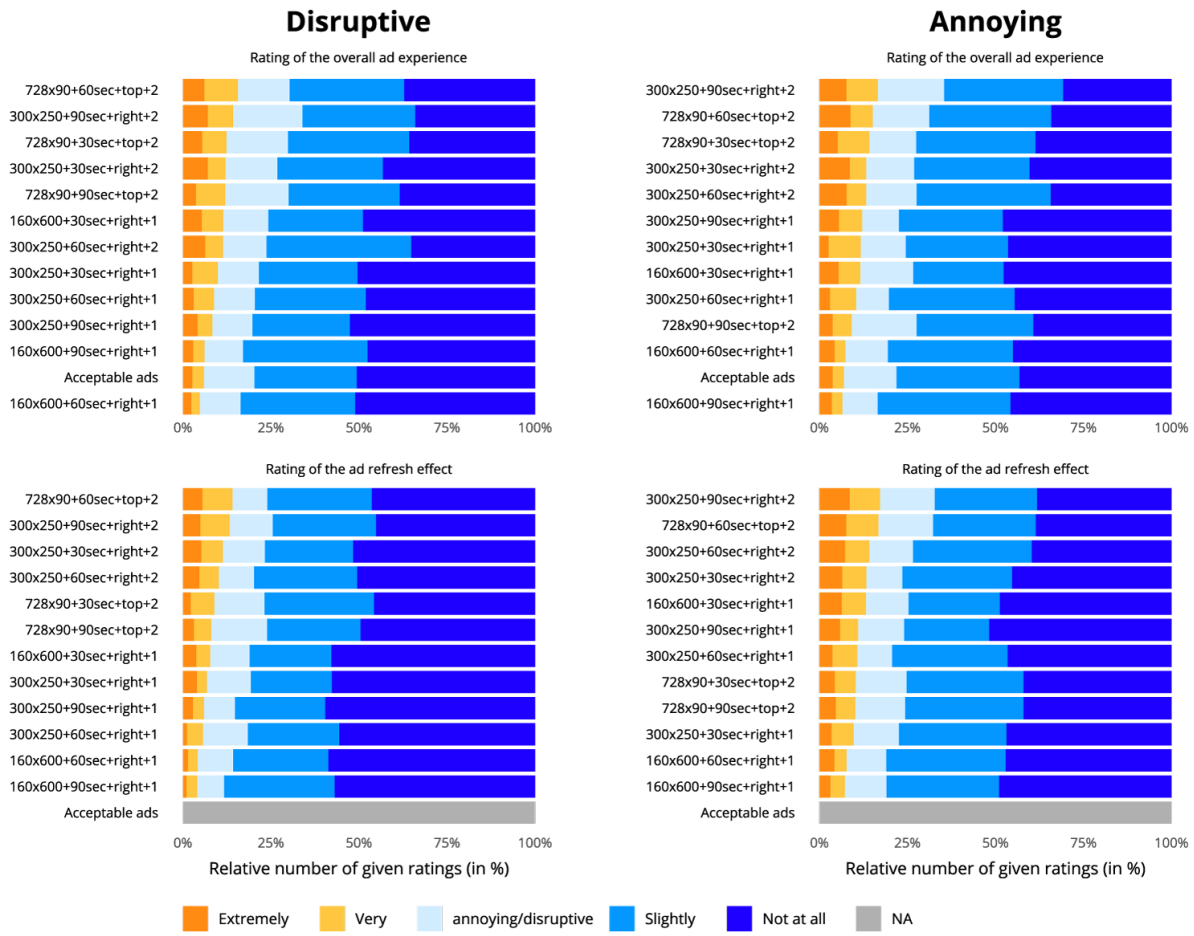
To evaluate the disruptiveness of the ad refreshing ad unit to a web experience, participants rated a) the ad refreshing ad unit itself and b) the ad refreshing ad unit as part of the overall ad experience in terms of how disruptive, annoying, intrusive and enjoyable each experience was. Figure 6 and 7 present the ratings for each web experience for the Article and Game experiences, respectively. Responses were weighted to reflect the age and gender of the internet-using population. The labels for the ad experience conditions can be interpreted as follows: refreshing ad format size+ ad refresh time-based trigger rate+ad refresh position. For example, in the Article experience, the ad experience "*300x250+30sec+right*" represents the web experience with a refreshing ad format size 300x250, an ad refresh time-based trigger rate of 30 seconds, and a positioning on the website's right rail (as the presence of an in-content ad had minimal effects on ratings, the results for these conditions are combined with those where no in-content ad was shown). In the Game experience, if the web experience contained one (right rail ad only) or two ads (both the right rail and top banner ad), the number is appended to the end of the label. For example, "728x90+60sec+top+2" represents the web experience with a refreshing 728x90 top banner ad, an ad refresh time-based trigger of 60 seconds, and a 300x250 static side rail ad.

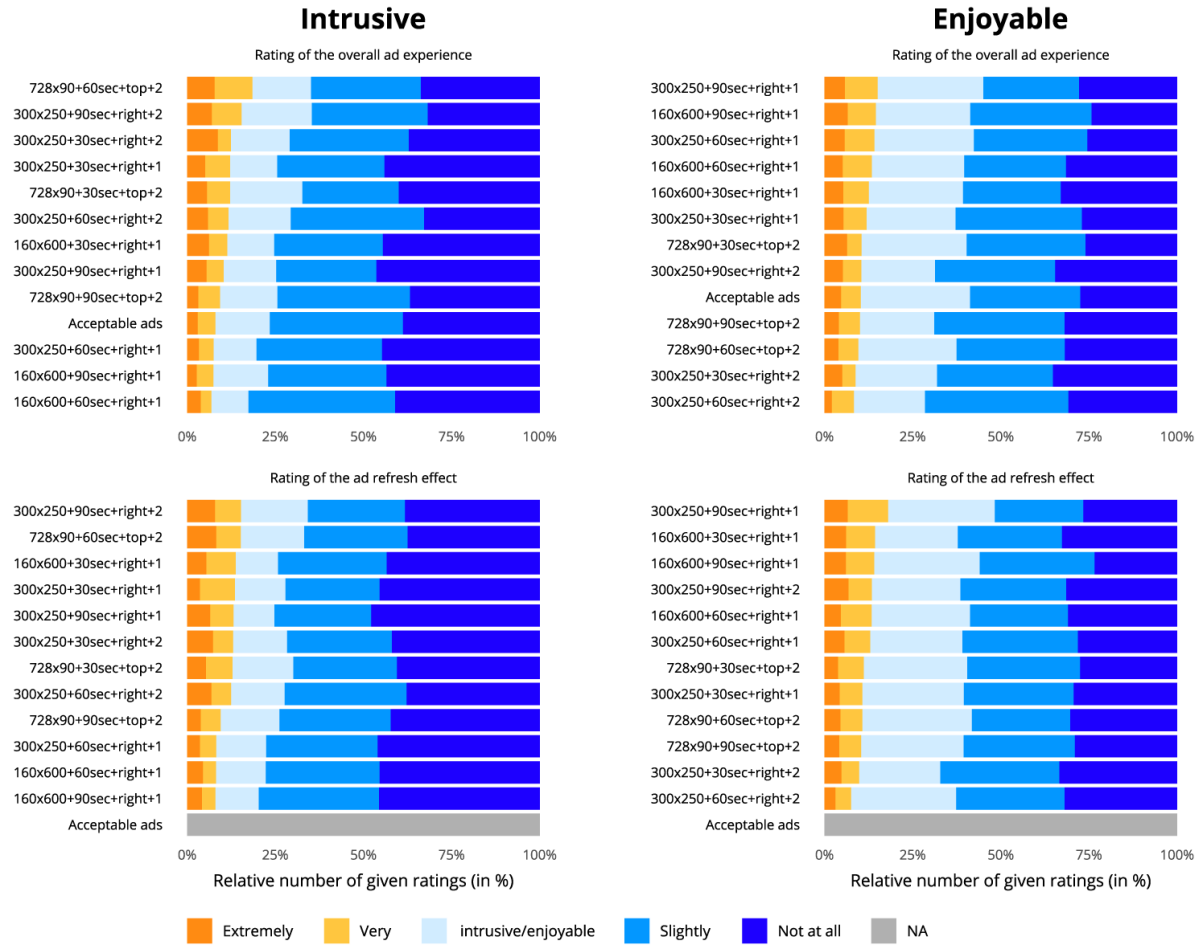
Figure 6: Ratings of the disruptiveness, annoyance, intrusiveness and enjoyment of the overall ad experience and the ad refresh experience for the *Article* experience.



Note: Labels for the ad experiences refer to the ad refresh format size, ad refresh time-based trigger rate and the ad position. Ratings for the overall ad experience refer to the question: "How disruptive were ALL the shown ads to your overall experience?". Ratings for the ad refresh effect refer to the question: "How disruptive was the reloading of new ad content to your experience?" "Acceptable ads" refers to the control experience where no ads are refreshed. Accordingly, there is no data for "Rating of the ad refresh effect".

Figure 7: Ratings of the disruptiveness, annoyance, intrusiveness and enjoyment of the overall ad experience and the ad refresh experience for the *Game* experience.





Note: Labels for the ad experiences refer to the ad refresh format size, ad refresh time-based trigger rate, ad position and whether the experience included one ad (right rail) or two ads (top banner and right rail). Ratings for the overall ad experience refer to the question: "How disruptive were ALL the shown ads to your overall experience?". Ratings for the ad refresh effect refer to the question: "How disruptive was the reloading of new ad content to your experience?" "Acceptable ads" refers to the control experience where no ads are refreshed. Accordingly, there is no data for "Rating of the ad refresh effect".

Overall, participants' ratings in the Article experience did not differ substantially across the different ad experiences. Participants rated all ad experiences similarly and at a similar level to the Acceptable Ads experience. For the Game experience, the data indicated an increase in disruptiveness when the ad experience included two ads (top banner and right rail) as opposed to one (right rail), suggesting an ad clutter effect. The pattern is evident in ratings of the ad refresh experience and the overall ad experience. When only one ad was visible, there is a general pattern that longer ad refresh time-based triggers received slightly better ratings. One can also see that there is a tendency for ads with a 30 second refresh rate to receive more negative ratings than

ads with either a 60 or 90 second refresh rate. Around 75% of participants rated the Acceptable Ads experience as not at all or slightly annoying/disruptive/intrusive.

7. Impact of demographics on the ad's ratings

To explore whether there were any factors that were associated with a higher probability of rating an experience as disruptive, logistic regression models were run on the ad refresh and the overall ad experience disruptiveness ratings. In these models, the webpage experience factors (e.g., whether the experience included an in-content ad, the size of the reloading ad, time spent on the webpage, etc.) as well as socio-demographic factors (e.g., generation, gender, attitude towards online advertising) were included. Factors that reflected participant's engagement with the web experience and the study (e.g., if the participant could correctly answer 2/3 of the comprehension questions, or if the participant stated that they did not recall seeing any ads) were also included. Socio-demographic categories with few participants were removed from the data prior to the analysis (i.e., gender categories "other" (n=126) and "prefer not to say" (n=30) and education "prefer not to say" (n=162) each representing <.01% of the sample). The model's results can be found in Tables 4 and 5.

Table 3 - Estimated odds-ratios from the logistic regression results for an ad being rated as disruptive (extremely, very or disruptive) vs. not disruptive (slightly or not at all disruptive) in the Article web experience.

The reference is a male US citizen aged 45 years or older who was classified as being in the low education and low income categories who had an ad experience without an in-content ad and a 160x600 sticky right rail ad that refreshed every 30 seconds and spent the average amount of time on the webpage. An interaction between the ad refresh time-based trigger and the time on the webpage was included in the model to capture the effect of being exposed to a different number of ads on a webpage the shorter the ad refresh time-based trigger. An odds ratio can be interpreted as follows: the odds that a person rates an experience as disruptive in Germany is 1.31 times as high as the odds for a person in the USA – that is, the odds that a person would rate the ad experience as disruptive is 31% higher if that person comes from Germany as opposed to the USA. A predictor is considered to have

an effect if its confidence interval does not contain 1 (equal odds). This significance is marked with an asterisk based on a 95% confidence interval.

	Rating of reloading ad effect	**Rating of overall ad experience**
Predictors	*Odds Ratios*	*Odds Ratios*
Webpage experience factors		
In-content ad: Present	0.87	1.02
Reloading ad size: 300x250	0.86	0.92
ad refresh time-based trigger: 60 seconds	0.95	0.92
ad refresh time-based trigger: 90 seconds	1.02	0.99
Time on webpage[a]	1.06	1.07
Interaction: ad refresh time-based trigger: 60 seconds x time on webpage	0.82	0.88
Interaction: ad refresh time-based trigger: 90 seconds x time on webpage	0.93	1.00
Socio-demographic factors		
Generation: Younger	0.81*	1.18
Gender: Female	1.04	0.81*
Education: Middle	1.10	1.24
Education: High	1.48*	1.52*
Income: Middle	1.16	1.04
Income: High	1.19	1.18
Income: Prefer not to say	0.99	0.96
Country: Germany	1.31*	1.01
Country: France	0.77*	0.73*
Perception of online advertising	1.73*	2.32*
Engagement with web experience		
Did not correctly answer 2/3 comprehension questions	1.74*	1.76*

Participant did not seen any ads	0.87	0.73*

Note: [a]Time on the webpage was centered on the average time, which means that the odds ratio can be interpreted as follows: when a participant spent one minute longer than the average time on the webpage, the odds are 6% higher that they rate the reloading ad experience as disruptive. [b]Higher scores = more negative attitudes to online advertising.

For the Article experience, no web experience factors were associated with being more likely to rate an ad experience as disruptive. Notably, the presence of an in-content ad did not result in participants being more likely to rate an ad experience as disruptive. However, the odds of rating a refreshing ad as disruptive were higher for individuals from Germany and for participants with a high level of education, and for those participants who had more negative attitudes towards online advertising. Conversely, the odds were lower for younger participants and individuals from France. When rating the overall ad experience, females and participants from France were less likely to rate the experience as disruptive while participants with a high level of education were more likely to rate the experience as disruptive. Participants who did not engage strongly with the webpage and could not answer at least 2/3 of the comprehension questions correctly were more likely to rate both the refreshing ad and the overall ad experience as disruptive. Further, participants who did not recall seeing any ads were less likely to rate the overall ad experience as disruptive.

Table 4 - Estimated odds-ratios from the logistic regression results for an ad being rated as disruptive (extremely, very or disruptive) vs. not disruptive (slightly or not at all disruptive) in the Game web experience.

The reference is a male US citizen aged 45 years or older who was classified as being in the low education and low income categories who had an ad experience showing only a 160x600 right rail ad that refreshed every 30 seconds and spent the average amount of time on the webpage. An interaction between the ad refresh time-based trigger and time on the webpage was included in the model to capture the effect of being exposed to a different number of ads on a webpage the shorter the ad refresh time-based trigger. An odds ratio can be interpreted as follows: the odds that a person rates an experience as disruptive if they are shown both a top banner and right rail ad is 1.42 times as high as the odds for a person who is only shown a right rail ad – that is, the odds that a person would rate the ad experience as disruptive is 42% higher if that person is shown two versus one ad placement. A predictor is

considered to have an effect if its confidence interval does not contain 1 (equal odds). This significance is marked with an asterisk based on a 95% confidence interval.

	Rating of reloading ad effect	**Rating of overall ad experience**
Predictors	*Odds Ratios*	*Odds Ratios*
Webpage experience factors		
Reloading ad size: 300x250	1.21	1.08
Reloading ad size: 728x90	1.27	1.21
Top banner + right rail ad shown	1.42*	1.56*
ad refresh time-based trigger: 60 seconds	0.86	0.82
ad refresh time-based trigger: 90 seconds	0.87	0.97
Time on webpage[a]	0.82*	0.84*
Interaction: ad refresh time-based trigger: 60 seconds x time on webpage	1.22	1.10
Interaction: ad refresh time-based trigger: 90 seconds x time on webpage	1.05	1.00
Socio-demographic factors		
Generation: Younger	1.11	1.22*
Gender: Female	0.84*	0.85
Education: Middle	0.96	0.97
Education: High	1.07	0.85
Income: Middle	0.97	0.99
Income: High	1.08	1.24
Income: Prefer not to say	0.85	0.92
Country: Germany	1.29*	1.27*
Country: France	0.93	0.81
Perception of online advertising	1.69*	2.16*
Engagement with web experience		

Did not correctly answer 2/3 comprehension questions	1.58*	1.63*
Participant did not seen any ads	0.81*	0.81*

Note: [a]Time on the webpage was centered on the average time, which means that the odds ratio can be interpreted as follows: when a participant spent one minute longer than the average time on the webpage, the odds are 18% lower that they rate the reloading ad experience as disruptive. [b]Higher scores = more negative attitudes to online advertising.

For the Game experience, being shown both a top banner and right rail ad on a webpage was associated with greater odds of rating the refreshing ad and overall ad experience as disruptive relative to being shown only the right rail ad. Spending more time on the webpage was associated with being less likely to rate the refreshing ad and overall ad experience as disruptive. German participants and participants who had more negative attitudes towards online advertising were more likely to rate both the refreshing ad and overall ad experience as disruptive (relative to USA participants), females were less likely to rate the refreshing ad as disruptive relative to males while younger participants were more likely than older participants to rate the overall ad experience as disruptive. Participants who could not answer at least 2/3 of the comprehension questions correctly were more likely to rate the refreshing ad and overall ad experiences as disruptive whereas participants who did not recall seeing any ads were less likely to rate these as disruptive.

In summary, there were no specific web experience factors associated with a greater likelihood of rating a refreshing ad or overall ad experience as disruptive for the Article experience. However, for the Game experience, having two as opposed to one ad placement on the webpage increased the odds of rating the experience as disruptive. While spending a greater amount of time on the webpage was associated with a lower likelihood of rating the experience as disruptive, it cannot be concluded from the data whether this was because people who were more disrupted by the experience tended to spend less time on the webpage or whether the experience became less disruptive the more participants engaged with the webpage. Similarly, those who appeared to engage with the web experience more (e.g., answered two out of three comprehension questions correctly) or claimed to have not noticed any ads were less likely to rate the

refreshing ad and overall ad experience as disruptive. It is not clear whether increased engagement in a web experience may result in individuals being less likely to react negatively or even to notice the ads, even with ad refresh.

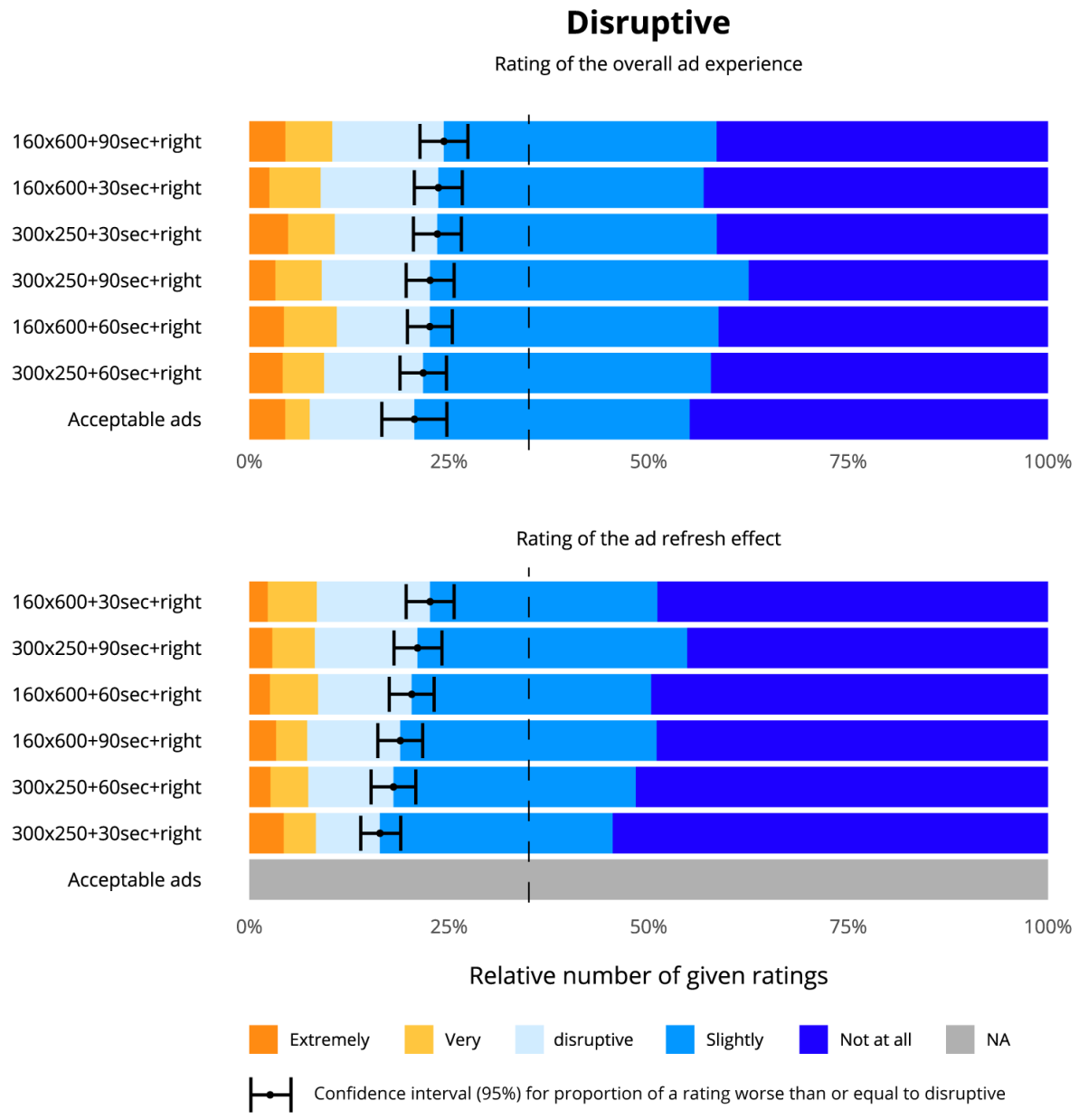
8. Acceptability

To be able to determine the acceptability of an ad format, the Acceptable Ads Committee is required to determine the annoyance level of the tested ad format. Only if an advertisement format fulfills the requirement of being “equivalent to 35 on the ‘Level of Disruption’ scale” can an ad type be taken into consideration for addition to the Acceptable Ads Standard. The level of disruption is demonstrated in Figures 8 and 9 along with the 95% confidence interval.

In the Article experience, all the ads would fulfill the AAC criterion of acceptability in terms of ratings for the ad refreshing ad as well as the overall ad experience. In the Game experience, some of the ad experiences would not be acceptable according to overall ad experience rating: nearly all ad refresh experiences where another ad was visible (that is, there was a top banner and right rail ad present) were seen as exceeding or bordering on the disruptiveness threshold. These results suggest that there is a negative ad clutter effect such that an ad refreshing ad unit is considered disruptive when there is another ad present in the viewport.

Figure 8: Disruptive ratings for all tested ad formats for the *Article* experience, with the 95% confidence interval of the proportion of ads being disruptive, very disruptive, or extremely disruptive.

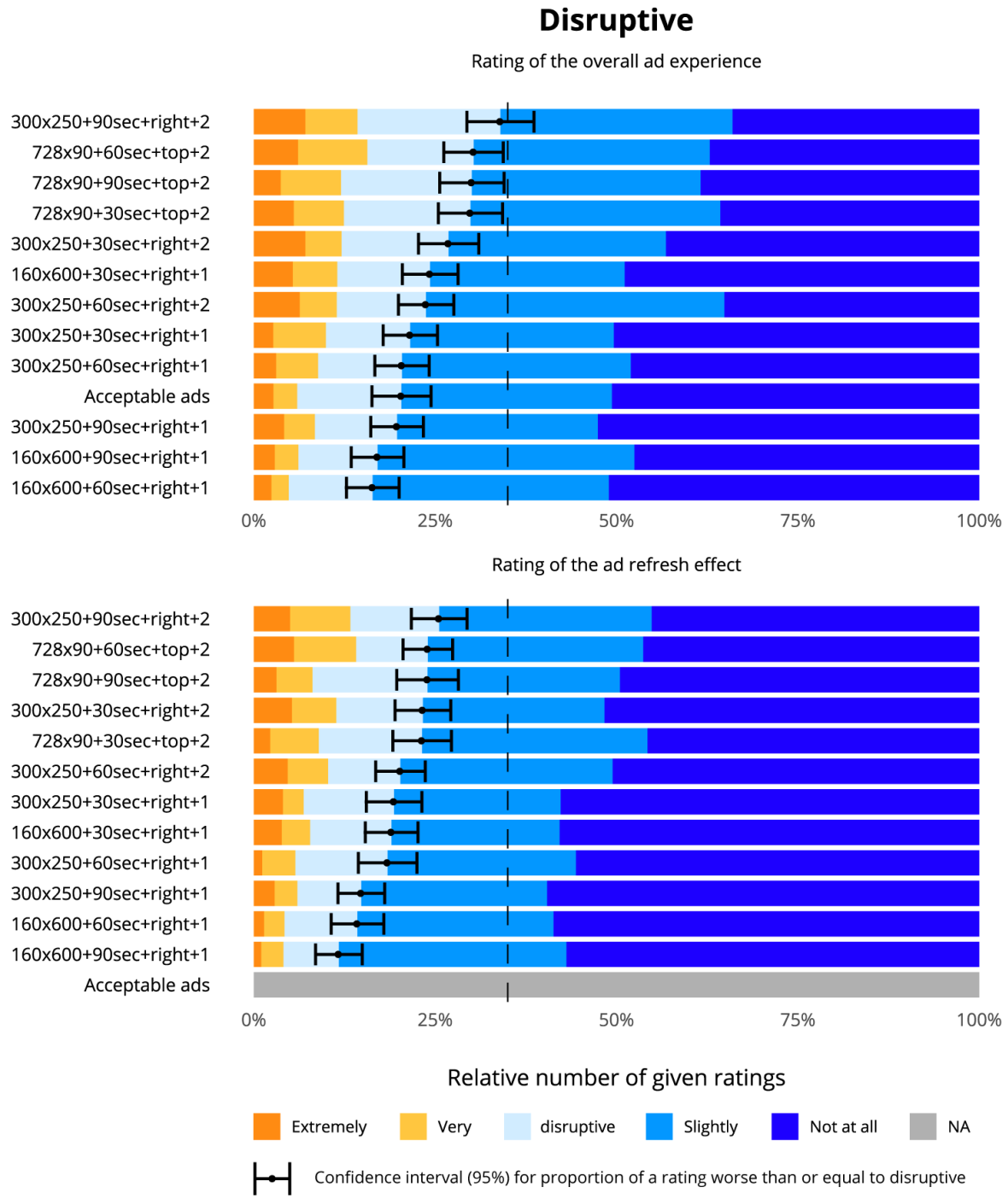
The black dashed line indicates the 35%-threshold stated in the AAC bylaws. Disruptiveness is measured for the overall ad experience and the ad refresh effect itself.



Note: Labels for the ad experiences refer to the ad refresh format size, ad refresh time-based trigger rate, ad position. Ratings for the overall ad experience refer to the question: "How disruptive were ALL the shown ads to your overall experience?". Ratings for the ad refresh effect refer to the question: "How disruptive was the reloading of new ad content to your experience?"

Figure 9: Disruptive ratings for all tested ad formats for the *Game* experience, with the 95% confidence interval of the proportion of ads being disruptive, very disruptive, or extremely disruptive.

The black dashed line indicates the 35%-threshold stated in the AAC bylaws. Disruptiveness is measured for the overall ad experience and the ad refresh effect itself.



Note: Labels for the ad experiences refer to the ad refresh format size, ad refresh time-based trigger rate, ad position and whether the experience included one ad (right rail) or two ads (top banner and right rail). Ratings for the overall ad experience refer to the question: "How disruptive were ALL the shown ads to your overall

experience?". Ratings for the ad refresh effect refer to the question: "How disruptive was the reloading of new ad content to your experience?"

The analysis demonstrates that the following ad formats can be taken into further consideration for the **Article** experience:

- 300 x 250 (or smaller) banner ads, with a time-based refresh rate of 30 seconds or longer.
- 160 x 600 (or smaller) banner ads, with a time-based refresh rate of 30 seconds or longer.

For the **Game** experience the following ad formats can be taken into further consideration:

- 300 x 250 (or smaller) banner ads, with a time-based refresh rate of 30 seconds or longer, provided there are no other ads on the page
- 300 x 250 (or smaller) banner ads, with a time-based refresh rate of 30 seconds or 60 seconds, if there is another ad on the page
- 160 x 600 (or smaller) banner ads, with a time-based refresh rate of either 30 seconds or longer, provided there are no other ads on the page
- Refreshing ads in the top banner of the website, with caution. In all three conditions with the 728x90 top banner ad refresh, disruptiveness ratings were very close to the disruptiveness threshold. With a confidence level of 99%, all three ad experiences would cross the disruptiveness threshold, suggesting less confidence that these ad experiences are firmly below the threshold.

The findings for the Game experience were mixed. In the case that a 300x250 right rail ad refresh is shown alongside a static top banner ad, a 90 second ad refresh time-based trigger exceeded the disruptiveness threshold whereas the 30 and 60 second ad refresh time-based trigger rates did not. The following section investigates potential explanations for this finding. It also explores the overall robustness of results across different subgroups.

8.1. Consistency and robustness of results

Two additional analyses were conducted to check the consistency and robustness of results: (1) exploring whether some of the participant characteristics associated with disruptiveness ratings as identified in Chapter 7 suggest potential subgroup differences in acceptable formats and (2) exploring whether patterns were consistent when taking into account more data from the ratings of the ad experiences.

Firstly, as the logistic regression model in Chapter 7 found that generation, gender, country and education level influenced the odds of rating an ad as disruptive, the results were checked for consistency and robustness across different subgroups[3]. For generation and gender, no additional ad formats would be considered less acceptable for the Article or Game experience when comparing ratings for the different subgroups. High education participants tended to rate ad experiences more negatively according to both the overall ad experience and ad-refresh ratings such that all ad experiences would exceed the disruptiveness threshold for the Article experience. When comparing ratings across countries, participants from Germany also found ad experiences more disruptive in general. Although there was no real pattern for the Article experience, the 160x600 and 300x250 with a 30 second ad-refresh trigger (both without a top-banner ad present) exceeded the disruptiveness threshold for the Game experience.

Taken together, the general pattern of results does not change when comparing these subgroups. The disruptiveness ratings for all ad experiences in the Article experience were below the threshold and the ad experiences with multiple ads in the Game experience continued to exceed the disruptiveness threshold. As the 160x600 and 300x250 ad formats with a refresh rate of 30 seconds were rated by some subgroups as unacceptable, one could caution their inclusion in the Acceptable Ads Standard.

[3] Potential differences between users of products that offer ad-filtering (ABP/Adblock) and users of other ad-blocking software were also explored. The general pattern for disruptiveness ratings did not change, with the exception that the 160x600 ad format exceeded the disruptiveness threshold in the Game experience. However, one should be cautious interpreting these results as the majority of participants used ABP/Adblock and the proportion of other users rating each ad experience condition was relatively small.

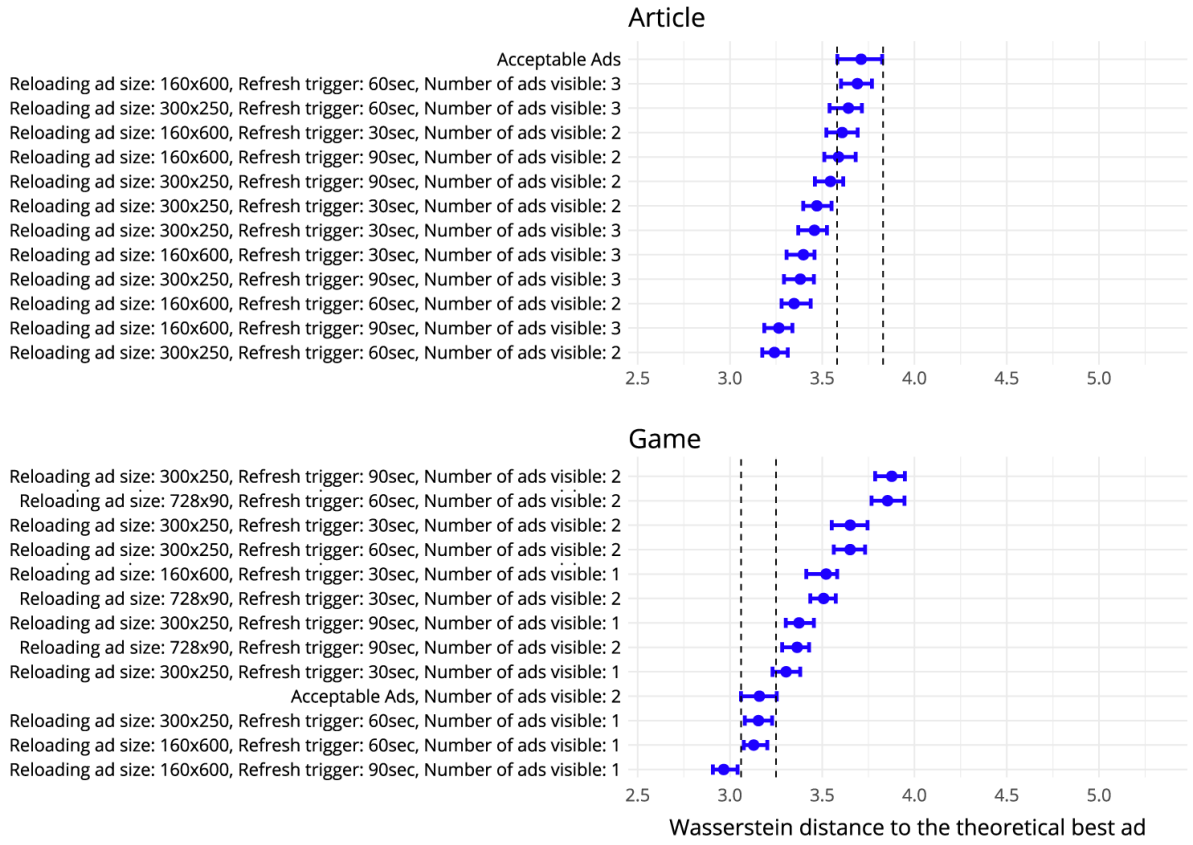
Second, consistent with the analysis informing the video ads and in-content ads studies, the ad experiences were compared to a theoretical "best ad" experience using a combination of the perceived disruptiveness, intrusiveness, annoyance and enjoyment ratings. The theoretical best case would be an ad experience that receives the lowest possible negative (= not at all disruptive/intrusive/annoying) and highest possible positive ratings on each of the scales (= extremely enjoyable) for the overall ad experience and for the ad refresh effect itself. That is, the theoretical ad experience would be not at all disruptive, annoying, or intrusive and extremely enjoyable.

The distribution of each total score forms an empirical cumulative density function (ECDF). To understand how different these ECDFs are, a Wasserstein metric is used to measure the distance between the distribution of ratings for a given ad experience and the distribution of ratings for the theoretical best case ad experience. Thus, the lesser the distance between the ad experience and the theoretical best ad, the better the ad experience's performance. This metric can be used to evaluate the distance between an ad experience and the theoretical best ad and compare it to the distance between the Acceptable Ads experience and the theoretical best ad. If the ad experience is within the bounds or below the distance for the Acceptable Ads experience – which is the ad experience users are currently experiencing online and that meets the acceptability criteria in the current study – one can be confident that such an ad experience is acceptable to users even when taking into account more comprehensive data from all disruptive, intrusive, annoying and enjoyment ratings.

Figure 10 presents the results for the Article and Game experiences separately. As the Acceptable Ads experience does not include a refreshing ad, the Wasserstein distance is calculated based on the ratings of the overall ad experience only. These ratings tended to be more negative than ratings for the ad-refreshing ad, which may explain why it appears that the Acceptable Ads experience was rated more negatively than the other ad experiences, particularly in the Article experience.

Figure 10: Wasserstein distance metric for the ECDF between ad types and the theoretical best ad with a 95% confidence interval

Comparison of all ad experiences for each of the different web experiences. The dashed lines mark the confidence intervals for the Acceptable Ads experience in the Article and Game web experiences. Ad experiences to the right of the dashed lines are those that are rated worse (a greater distance to the theoretical best ad experience) than the Acceptable Ads experience.



Note: As the Acceptable Ads experience does not include a refreshing ad, the Wasserstein distance is calculated based on the ratings of the overall ad experience only. As evident in Figures 8 and 9, ratings for the overall ad experience tended to be more negative than ratings for the ad-refreshing ad. This is most likely because, when rating for the overall ad experience, participants rated their experience seeing multiple ads. For this reason, the Acceptable Ads experience appears to be rated more negatively than the other ad experiences.

Figure 10 shows that, compared to the ratings for the Acceptable Ads experiences, all ad experiences in the Article experience would be considered as acceptable as their distance to the theoretical best ad is either lower or not statistically different from the distance of the Acceptable Ads experience to the theoretical best ad. These results are consistent with the analysis based on the disruptiveness ratings. In Figure 10, the Acceptable Ads experience appears to be rated more negatively than the other ad

experiences. However, this is likely because the data is only based on the "overall ad experience" ratings ("How disruptive were ALL the shown ads to your overall experience?") where participants rate their experience seeing all the ads in their web experience (e.g., the Acceptable Ads experience included a static top banner (728x90), static sticky right rail (160x600) and an in-content ad (either 840x150 or 728x90)). As such these ratings tended to be more negative than ratings for the "ad-refreshing ad" ("How disruptive was the reloading of new ad content to your experience?"; compare the two ratings in Figures 8 and 9) as it required participants to rate their experience viewing the multiple ads presented to them in their ad experience. As the Acceptable Ads experience did not include a refreshing ad, participants did not answer the question on their perceptions of the ad-refreshing ad experience, meaning the results reflect only the more negative ratings of their overall, multiple-ad experience.

For the **Article** experience, any tested ad format fulfills the criterion that its Wasserstein distance is smaller than the distance of the Acceptable Ads to the theoretical best ad. Therefore, all ad experiences could be considered acceptable.

For the **Game** experience, consistent with the analysis based on the disruptiveness ratings, all ad experiences showing more than one ad are statistically different from and exceed the threshold based on the Acceptable Ads experience. This differs slightly from the analysis based on the disruptiveness ratings where ad experiences showing a 300x250 ad refresh with a 30 second or 60 second ad refresh time-based trigger rate and a top banner ad did not exceed the disruptiveness threshold. However, as can be seen in Figure 7, these ad experiences received more negative intrusive and annoying ratings. In addition, the 160x600 ad format with a 30 second ad refresh time-based trigger is statistically different from the Acceptable Ads experience, a finding that is consistent with the subgroup analysis above. One anomaly that was found in the analysis of disruptiveness ratings also persists here: the ad experience with a 300x250 ad format with a 90 second ad refresh time-based trigger and no top-banner ad is rated as statistically worse than the Acceptable Ads experience whereas the 30 second and 60 second ad refresh time-based triggers were not. This finding is counterintuitive and one

could argue that if a lower ad refresh time-based trigger is considered acceptable, it is logical that longer ad refresh time-based triggers could be considered acceptable as well. A comparison of the disruptiveness and Wasserstein distance analyses for the ad experiences in the Game experience are presented in Table 5.

Table 5: Comparison of the acceptability of ad experiences according to the disruptiveness ratings and Wasserstein distance analyses for the *Game* experience.

Ad format	**Criterion:** 35% disruptiveness threshold	**Criterion:** Wasserstein distance	Comment
160x600, 30 seconds refresh rate, no other ad visible	✅	❌	The results for this ad format are inconsistent across analyses. However, given that this ad experience was rated as exceeding the disruptiveness threshold for some subgroups (e.g., German participants), one could caution its inclusion as an Acceptable Ad.
160x600, 60 seconds refresh rate, no other ad visible	✅	✅	
160x600, 90 seconds refresh rate, no other ad visible	✅	✅	
300x250, 30 seconds refresh rate, no other ad visible	✅	✅	The findings that the 30 and 60 seconds refresh rates were found to be acceptable across both analyses, one could argue that the 90 seconds refresh rate should also be acceptable. The more negative ratings for the 90 second refresh rate are counterintuitive.
300x250, 60 seconds refresh rate, no other ad visible	✅	✅	
300x250, 90 seconds refresh rate, no other ad visible	✅	❌	
300x250, 30 seconds refresh rate, static top banner visible	✅	❌	Findings for the 300x250 ad format for 30, 60, and 90 seconds are inconsistent across analyses. However, given the general pattern across analyses that ad
300x250, 60 seconds refresh rate, static top	✅	❌	

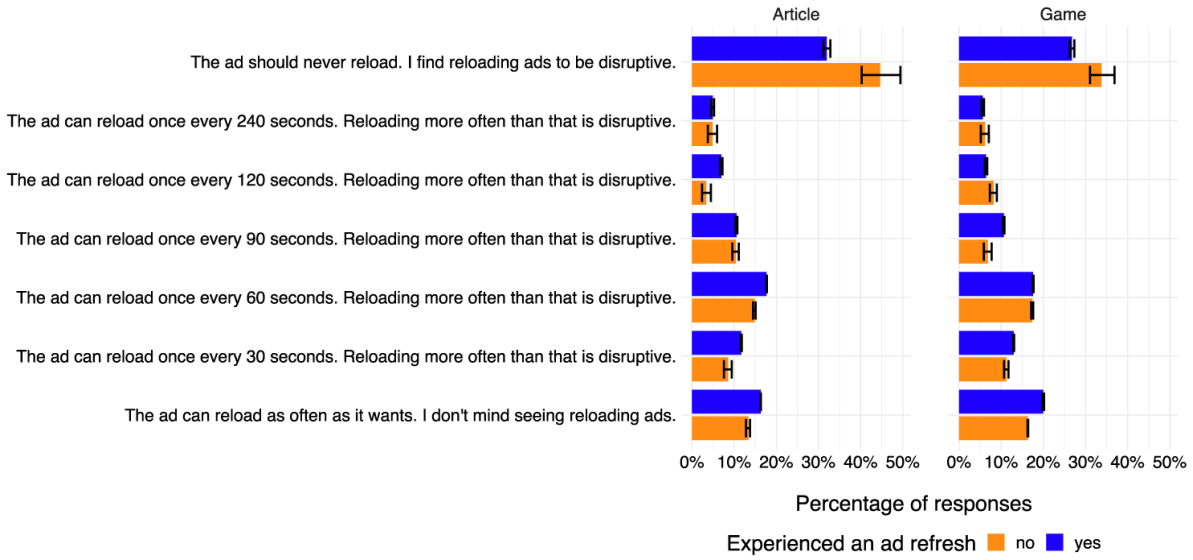
banner visible			experiences that showed multiple ads were rated more negatively, one could consider all of these ad experiences as not acceptable.
300x250, 90 seconds refresh rate, static top banner visible	❌	❌	
728x90, 30 seconds refresh rate, static right rail visible	✅*	❌	*The disruptiveness ratings for the 728x90 top banner ad refresh were very close to the disruptiveness threshold. With a confidence level of 99%, all three ad experiences would cross the disruptiveness threshold, suggesting less confidence that these ad experiences are firmly below the threshold.
728x90, 60 seconds refresh rate, static right rail visible	✅*	❌	
728x90, 90 seconds refresh rate, static right rail visible	✅*	❌	

8.2. Participant's perceptions of acceptable ad refresh time-based trigger rates

At the conclusion of the study, participants were asked how often an ad could refresh (i.e. a new ad shown in the same spot) before it becomes annoying or disruptive. The majority of participants (70%) reported that at least some ad refresh time-based trigger would not be considered annoying or disruptive, with 12% stating an ad refresh lower than 30 seconds and 18% stating that an ad refresh lower than 60 seconds would be disruptive (see Figure 11). Around one fifth (18%) reported that any type of ad refresh would not be annoying or disruptive regardless of how often it occurs. Interestingly, participants who *did not* experience an ad refresh in their web experience (the Acceptable Ads experiences), were more likely to state that they did not want to see any ad refreshes (39%) as compared to those participants who experienced an ad refresh (29%).

Figure 11: Proportion of answers to the question: "In your opinion, how often can an ad refresh (i.e. a new ad shown in the same spot) before it becomes annoying/disruptive?"

With 95% confidence interval



8.3. Comparison with previous survey results

As all previous AAC surveys were designed to examine how disruptive, annoying, intrusive and enjoyable different ad experiences were, the data from the present study could be combined with the data from the 2020 video advertisement survey and the 2021 in-content ad survey to create an overall ranking of all previously tested ad formats. Again, the Wasserstein distance relative to the theoretical best ad was calculated in order to compare all ad experiences tested in these three surveys (as described in section 8.1). Although all AAC surveys used the same rating scale for rating a tested ad format (e.g.: "How disruptive was the ad to your experience?"), the current study also included ratings for the overall ad experience (e.g.: "How disruptive were ALL the shown ads to your overall experience?"). As such, the ratings used in the Wasserstein distance analysis reflect both ratings for ad experiences from the present study. The results are presented in Figure 12.

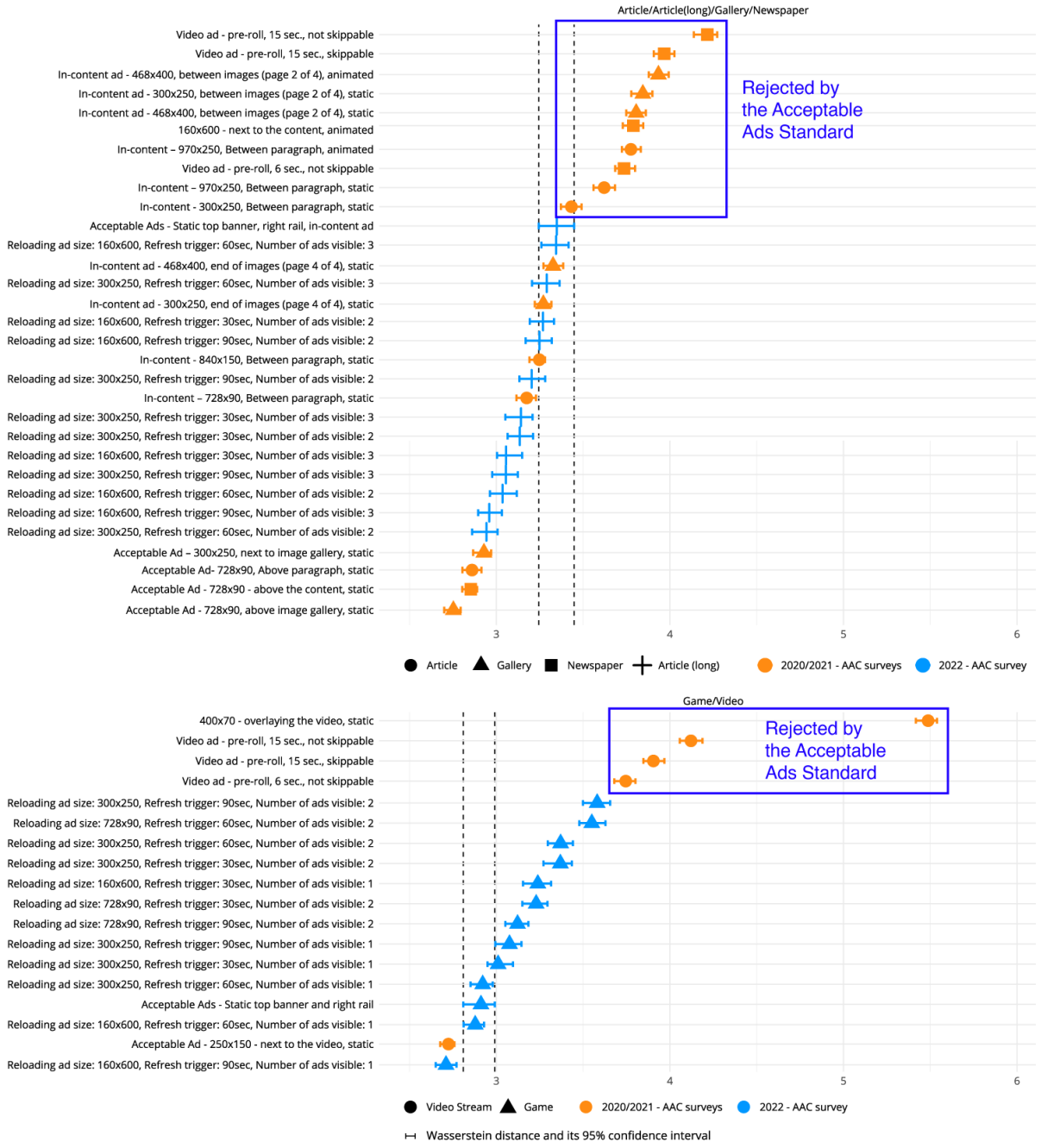
To differentiate the Article experience in the 2021 survey from the Article experience in the 2022 survey, the latter is named “Article (long)”, as the article length was significantly longer in the current 2022 web experience.

Figure 12 shows that for the long Article experience (scrollable single-page) all refreshing ad formats fall into the cluster of ads that were previously determined to be acceptable by the AAC.

For the Game experience (no-scroll website) all ad refreshing experiences where no other ad was visible in the viewport clearly fall in the area of previously accepted ad formats. The only discrepancy is the 160x600 right rail ad with a 30 seconds refresh trigger which is rated within the same area as the ad refresh experiences where two ads were visible in the viewport.

Figure 12: Wasserstein distance metric for the ECDF between ad types and the theoretical best ad

Comparison of ad experiences in the current study (long Article and Game) to ad experiences tested in the 2020 AAC Video advertisement study (newspaper and video stream) and the 2021 AAC in-article and in-gallery ads survey (article and gallery). The dashed lines mark the confidence intervals for the Acceptable Ads experience in the (long) Article and Game web experiences.



Note: In the 2020 and 2021 survey, all ad experiences contained only a single ad, whereas in the current 2022 survey, all Article experiences and most of the Game experiences contained multiple ads (Acceptable Ads plus a refreshing ad). In the 2022 survey participants not only rated the ad refreshing ad ("How disruptive/intrusive/

annoying/enjoyable was the reloading ad to your experience?") but also their overall ad experience ("How disruptive/intrusive/annoying/enjoyable were ALL the shown ads to your overall experience?").

9. Conclusion

In summary, the present study explored how disruptive ad-blocking users perceive a single ad refreshing ad when included in an Acceptable Ads experience on a scrollable single-page (Article) and non-scrollable single-page (Game) experience. The study examined different ad format sizes, placements, and ad refresh time-based trigger rates and evaluated users' perceptions of the ad refresh experience and the overall ad experience. Strengths of the study include focusing on realistic web experiences where the ad refreshing ad appeared in combination with other static Acceptable Ad formats – including the recently added in-content ad format – and ensuring participants engaged with the web experience. These strengths increase the reliability of and confidence in the results of the study.

In general, there is some consistency in the results across the Article and Game experience. Specifically, **when the refreshing ad is located to the side of the primary content,** there is support for the acceptability of the following ad formats and ad refresh time-based trigger rates:

Scrollable single-page (Article experience):

- 300x250 banner ad (or any smaller size) in the side rail, with an ad refresh time-based trigger rate of 30 seconds or longer
- 160x600 banner ad (or any smaller size) in the side rail, with an ad refresh time-based trigger rate of 30 seconds or longer

Non-scrollable single-page (Game experience):

- 300x250 banner ad (or any smaller size) in the side rail, with an ad refresh time-based trigger rate of 30 or 60 seconds, provided that there are no other ads in the viewport

- 160x600 banner ad (or any smaller size) in the side rail, with an ad refresh time-based trigger rate of 60 seconds or longer, provided that there are no other ads in the viewport

There is mixed support for the acceptability of the following cases with respect to the Game experience:

- 300x250 or any smaller ad format with an ad refresh time-based trigger rate of 90 seconds when no other ad is visible in the viewport. However, given the findings that the 30 and 60 seconds refresh rates were found to be acceptable across both analyses, one could argue that the 90 seconds refresh rate should also be acceptable. The more negative ratings for the 90 second refresh rate are counterintuitive.
- 160x600 ad format with an ad refresh time-based trigger rate of 30 seconds. This format was rated by some subgroups as exceeding the disruptiveness threshold and was rated as statistically more negative relative to the Acceptable Ad experience in the Wasserstein distance analysis.
- 300x250 ad format for 30, 60 seconds when a top-banner ad is present. Results are inconsistent across analyses. However, given the general pattern across analyses that ad experiences that showed multiple ads were rated more negatively, one could consider all of these ad experiences as not acceptable.
- 728x90 top banner ad refresh in the presence of an ad to the side of the primary content. The disruptiveness ratings for this ad experience bordered on the disruptiveness threshold and would cross the threshold if a higher confidence level of 99% is used. These ad experiences were also statistically more negative relative to the Acceptable Ad experience in the Wasserstein distance analysis and there was a general pattern that when ad refreshes occurred in the presence of another ad format, ratings were more negative in general.

It is not possible to ascertain from the study whether a top banner ad refresh would be rated as acceptable if it was the only ad format shown in the viewport at any given time,

as the study did not test this scenario explicitly. However, given that Game experiences in which the ad refresh occurred in the top banner were consistently rated as more disruptive than those with a right-rail ad refresh, one would caution applying the results of the study to allow top banner ad refresh.

In addition to the results on the disruptiveness of ad-refreshing ads, the study also provides additional support for the introduction of in-content ads to the Acceptable Ads Standard as web experiences containing an in-content ad were not rated as more disruptive than those that did not contain an in-content ad. A caveat to this finding is that the study only incorporated a single in-content ad within a scrollable single-page and, when asked directly, participants stated that the in-content ad was the most disruptive ad format even when their web experience did not include an in-content ad.

In line with the Acceptable Ads Committee bylaws, the committee's decision and potential changes to the Standard based on this report will be published and opened for further user feedback for the period of a month.

10. Limitations and Implications of Study Findings

10.1. Study limitations

Despite efforts to evaluate the effect of introducing refreshing ad units using more realistic and engaging web experiences, the study environment remains artificial and may not adequately reflect a user's actual internet viewing experience. For instance, it is possible that users are exposed to more ads on a webpage than what was presented in the study. Further, the disruptiveness of ad experiences may accumulate over the course of a web browsing experience. Nevertheless, the study aimed to incorporate common ad format sizes and placements that maximized the ad space available under the current Acceptable Ads Standard to make the browsing experience as realistic as possible for participants.

The study tried to encourage engagement with the web experience by providing participants with a choice between two articles or games, informing participants that we would ask them comprehension questions about their experience, and setting a minimum time threshold to spend on the webpage. While some indicators suggest participants did engage with the web experience (more than two thirds answered comprehension items correctly, and participants spent more than one minute longer than the minimum time threshold on the webpage, on average), it is not possible to know how engaging the experiences were for participants and how this affected ratings.

Further, as the study was only conducted with participants from three markets (United States of America, France and Germany), it is not possible to know whether the results would generalize to all other global markets where Acceptable Ads are used.

10.2. Practical implications of study findings

The study focused on incorporating a single refreshing ad unit into the Acceptable Ads experience. Based on the design of the study, the following points should help inform practical implementations of the study results:

1. A single ad refresh per viewport. The study focused on the minimum possible adaptation where a single ad refreshing ad unit appeared in an Acceptable Ads experience to determine if users were in principle opposed to in-view ad refresh. Given the general finding that ad refresh was more disruptive when shown in combination with a static ad, it is highly likely that multiple ad refreshing ads would be more disruptive to users.
2. Ad units that reload an ad of the same size/dimensions. Ad refreshing units that reload new ads of different sizes/dimensions were not tested and are likely to be more disruptive for users as these refreshes may lead to a more noticeable visual change on the webpage.
3. Ad refresh of in-content ads. Ad refresh of in-content ads was not tested in the study and it is unlikely that the results will translate to in-content ads.
4. Alternative ad refresh time-based triggers. User-action or event-based ad refresh triggers (e.g., clicks, scroll) were not tested in the study. Nevertheless, if these ad

refresh triggers are implemented in addition to an acceptable time trigger and/or viewability threshold, these alternative ad refresh time-based triggers should not be prohibited.

5. Ad format sizes. It could be assumed that the results of the study would generalize to any ad format sizes that were *not* tested in the current study but are *smaller* in height and width dimensions than any of the tested ad format sizes.

Glossary

Above the fold	The portion of the webpage that is visible without scrolling.
Acceptable Ads size requirements:	All ads that are visible in the browser window when the page first loads (i.e. above the Fold) must not collectively occupy more than 15% of the visible portion of the webpage. If placed lower on the page (i.e. below the Fold), ads must not collectively occupy more than 25% of the visible portion of the webpage.
Ad experience	The specific combination of ads shown on the web experience.
Below the fold	The part of a webpage that can't be seen without scrolling down.
Web experience	A scrollable single-page (Article experience) or non-scrollable single-page (Game experience).

Bibliography

Coalition for Better Ads. (2016, April). *An Experimental Methodology to Rank N Ad Experiences by Consumers' Perceptions—Coalition for Better Ads*. https://www.betterads.org/research/rankingpaper/

Expósito-Ventura, M., Ruipérez-Valiente, J. A., Parra-Arnau, J., & Forné, J. (2021). A Survey of the Role of Viewability Within the Online Advertising Ecosystem. *IEEE Access*, *9*, 134593–134610. https://doi.org/10.1109/ACCESS.2021.3115979

Fennell, D. (2019, May 8). *A Suggestion For Ad Refreshing*. AdExchanger. https://www.adexchanger.com/the-sell-sider/a-suggestion-for-ad-refreshing/

Google Ad Manager Help. (2022). *Declare inventory that refreshes*. https://support.google.com/admanager/answer/6286179?hl=en#zippy=%2Ctypes-of-refresh-triggers

Graham, K. (2021, February 25). *How Does Google Measure Ad Viewability in 2022*. MonetizeMore. https://www.monetizemore.com/how-google-measure-ad-viewability/

Interactive Advertising Bureau. (2017). *Desktop Display Impression Measurement Guidelines (version 7.1)*. Interactive Advertising Bureau Technology Laboratory Media Rating Council. https://www.iab.com/wp-content/uploads/2017/11/Desktop-Display-Impression-Measurement-Guidelines-US-MMTF-Final-v1.1.pdf

Jatain, V. (2021, January 28). *Why ad viewability matters and how to improve it*. https://blockthrough.com/blog/why-ad-viewability-matters-and-how-to-improve-it/

Kuisma, J., Simola, J., Uusitalo, L., & Öörni, A. (2010). The Effects of Animation and Format on the Perception and Memory of Online Advertising. *Journal of Interactive Marketing*, *24*, 269–282. https://doi.org/10.1016/j.intmar.2010.07.002

PubGalaxy. (2021). *Ad refreshing with Google Ad Manager and Ad Sense*. PubGalaxy. https://www.pubgalaxy.com/ad-refreshing-with-google-ad-manager-and-ad-sense/

Publift. (2022, April 24). *What is Ad Refresh and How Does It Work? - Publift*. https://www.publift.com/blog/what-is-ad-refresh-and-how-does-it-work

Quentel-Reme, M. (2020, February 17). Are auto-refresh ads profitable? *OptiDigital*. https://www.optidigital.com/en/are-auto-refresh-ads-profitable/

Rafieian, O., & Yoganarasimhan, H. (2021). Variety Effects in Mobile Advertising. *Journal of Marketing Research*, 00222437211056090. https://doi.org/10.1177/00222437211056090

Redondo, I., & Aznar, G. (2018). To use or not to use ad blockers? The roles of knowledge of ad blockers and attitude toward online advertising. *Telematics and Informatics*, *35*(6), 1607–1616. https://doi.org/10.1016/j.tele.2018.04.008

setupad. (2021, October 21). What is Ad Refresh and Is It Necessary for Your Website? | Setupad. *Setupad.Com*. https://setupad.com/blog/what-is-ad-refresh/

Southern, L. (2020a, January 8). 'It can be a money-printing machine': The incredible revival of auto-refreshing ads. *Digiday*. https://digiday.com/media/can-money-printing-machine-incredible-revival-auto-refreshing-ads/

Southern, L. (2020b, July 1). Ad refreshing is the publisher ad tactic both ad buyers and readers dislike—And is on the rise in tough times. *Digiday*. https://digiday.com/media/buyers-dont-like-it-readers-dont-like-it-but-ad-refresh-is-on-the-rise-as-publishers-seek-revenue/

Truong, V. N. X. (2016). Optimizing mobile advertising using ad refresh interval. *2016 International Conference on Electronics, Information, and Communications (ICEIC)*, 1–4. https://doi.org/10.1109/ELINFOCOM.2016.7562948

Appendix 1

Table A.1. Ad creatives developed for the web experiences

Ad unit	Ad dimensions	Mock ad brand	Advertised product
Static top-banner ad	728x90	Gym-Beast	Workout/Gym center
Static right rail ad	300x250 160x600	Jasmine	Tea
Static in-content ad	728x90 840x150	Mansion Coffee	Coffee
Refresh (1)	728x90 top banner 300x250 right-rail 160x600 right-rail	Socks Box	Socks
Refresh (2)		Stylophone	Phone cases
Refresh (3)		Chocochino	Chocolates
Refresh (4)		Voxanon	Clothes
Refresh (5)		Home sweet home	Interior items
Refresh (6)		Getgadget	Tech gadgets
Refresh (7)		gogogo	Ride sharing service
Refresh (8)		Listener	Audio books
Refresh (9)		Travel guru	Travel agency
Refresh (10)		Natural Juice	Orange

Appendix 2 - Questionnaire

Question	If equal for both experiences	Article Experience	Game experience
	Consent page		
Q1	What is your gender?		
Answer code			
a	Male		
b	Female		
c	Other		
d	Prefer not to say		
Q2	How old are you?		
	____ years		
Q3	**Please indicate the highest level of education you have completed:**		
Answer code			
a	No formal schooling		
b	Less than high school		
c	High school diploma		
d	Technical certificate		
e	Bachelor's		
f	Master's		
g	Doctorate		
h	Prefer not to say		
Q4	**Please provide your current gross annual income in USD:**		
Answer code			
a	Under $25,000		
b	$25,000 - $34,999		
c	$35,000 - $49,999		

d	$50,000 - $74,999		
e	$75,000 - $99,999		
f	$100,000 - $149,999		
g	$150,000 - $199,999		
h	$200,000 and over		
i	Prefer not to say		
Q5	**What types of technologies do you currently own or use? Please select all that apply:**		
Answer code			
a	Home assistant (Google Home, Amazon echo, etc.)		
b	Smart watch (Apply Watch, Pebble etc.)		
c	Ad-blocking software (Adblock Plus, AdBlock, uBlock Origin, etc.)		
d	VPN (virtual private network)		
e	Streaming service (Netflix, Hulu, etc.)		
f	AI-powered political content blocker		
INFO1	**Task description**		
		On the following page we will present you with a set of web articles. You can select one you would like to read. Please read the article carefully, as we will ask you some questions about what you have read and seen on the following pages. You will not be able to return to the article once you have left the page.	On the following page we will present you with a set of online games. You can choose one by yourself. Please play the game carefully, as we will ask you some questions about what you have seen on the following pages. You will not be able to return to the game once you have left the page.

		It is important to us that you take the time to read the article and explore the webpage. There will be a **minimum time limit of 2.5 minutes** before you can click ahead in the survey.	It is important to us that you take the time to play the game and explore the webpage. There will be a **minimum time limit of 2.5 minutes** before you can click ahead in the survey.
Q6		Content selection	Game selection
Answer code		Please select one of the following articles you would like to read:	Please select one of the following games you would like to play. We provided you a short description and a screen recording of the games to give you a better idea of the games. Watching the screen recording is not mandatory.
a		10 'Vices' That Are Good For You - A square of chocolate, a glass of wine, and a bit of sunshine: Find out about guilty pleasures that can actually be good for your health.	**Tetris:** The aim of Tetris is simple; you bring down blocks from the top of the screen. You can move the blocks around, either left to right, and/or you can rotate them. The blocks fall at a certain rate, but you can make them fall faster (by pressing the downward arrow) if you're sure of your positioning. Your objective is to get all the blocks to fill all the empty space in a line at the bottom of the screen; whenever you do this, you'll find that the blocks vanish and you get awarded some points. Your game is over if your pieces reach the top of the screen. Tetris has a cunning means of adding challenge. It becomes faster (e.g. the blocks fall faster) the better you do.

b		17 Little Cooking Habits You Should Actually Ditch ASAP - Cooking is both rewarding and challenging. Whether you're just starting out or have been honing your kitchen skills for a while, there are likely some mistakes you might still be making.	**2048:** 2048 is played on a plain 4×4 grid, with numbered tiles that slide when a player moves them using the four arrow keys. Every turn, a new tile randomly appears in an empty spot on the board with a value of 2 or 4. Tiles slide as far as possible in the chosen direction until they are stopped by either another tile or the edge of the grid. If two tiles of the same number collide while moving, they will merge into a tile with the total value of the two tiles that collided. The goal is to reach the highest possible tile by continually merging tiles of the same value.
Q7	Overall Web Experience		
	We would now like to ask you a few questions about your overall experience viewing the webpage.		
Answer code	**How satisfied were you with this OVERALL browsing experience?**		
5	Very satisfied		
4	Slightly satisfied		
3	Neutral		
2	Slightly dissatisfied		
1	Very dissatisfied		
INFO2		**Website Comprehension**	**Game Comprehension**
		We would like to ask you a few questions about the article you just read.	We would like to ask you a few questions about the game you just played.

Answer code		IF THE PARTICIPANT HAS CHOSEN THE "Vices" ARTICLE:	IF THE PARTICIPANT HAS CHOSEN THE "Tetris" GAME:
Q8.1.1		**Which of the following statements is true?**	**What of the following elements did not appear in the game?**
a		Carbs are good when eaten in the morning to give your body more time to digest	[Picture of a Tetris Block]
b		Carbs are good when you need to lower your insulin level	[Picture of a false Tetris Block]
c		Carbs are good when in form of simple carbohydrates	[Picture of a Tetris Block]
d		Carbs are good when eaten in the form of complex carbohydrates such as root vegetables and whole grains over refined	[Picture of a Tetris Block]
Q8.1.2		Happy Hour is okay when “light and moderate”. What is considered “light to moderate” consumption?	**How many individual blocks is the T-shaped element composed of?**
a		One alcoholic drink per day	3
b		Good rule for women: maximum 4 drinks per week	4
c		Less than one alcoholic drink per day.	5
d		Woman can have 4 and men can have 3 alcoholic drinks per week	6
Q8.1.3		Which of the following vices was not recommended in the article?	**In the game you just played, which of the following colors were not used for any of the blocks?**
a		Kiss Your Allergies Goodbye	Yellow
b		Soak Up Some Sun	Blue

c		Harness the Power of Happy Hour	Grey
d		Indulge in a little Netflix & Chill	Green
		IF THE PARTICIPANT HAS CHOSEN THE "Habit" ARTICLE:	IF THE PARTICIPANT HAS CHOSEN THE "2048" GAME:
Q8.2.1		**How often is it recommended to hone your kitchen knives?**	**Which of the following strategies is not a good strategy to achieve a tile with a value of 2048?**
a		Once per month	Combine if possible tiles with high values first
b		Once per week	Always move your tiles up
c		Twice per week	Keep your tiles tidy
d		Once every three months	Isolate tiles with high value at the edges of the playing field
Q8.2.2		**According to the article, why should you not always use a nonstick pan to cook everything?**	**Which of the following statements is true?**
a		They are not suitable for using metal utensils.	You lose the game if a tile has reached the value 13.
b		They get too hot too fast	You can rotate the tiles.
c		They can’t get as hot as other pans	With every movement, tiles with the value 2 or 4 randomly appear on the screen.
d		Their heat distribution is always even	The tiles do not change their color.
Q8.2.3		**Is it a good idea to use the pasta cooking water for further cooking?**	**Which of the following statements is true?**
a		Yes because the cooking water is filled with wheat that will help bind the pasta and sauce together	The game gets faster the longer you play.

b		Yes because the cooking water is filled with starch that will help bind the pasta and sauce together	You are able to take back a made move.
c		No, it is usually too salty which is also why rinsing the pasta is recommended.	You are only able to merge tiles of the same value.
d		Yes, because the cooking water is filled with loosened proteins which act as a thickening agent.	When you merge two tiles the new tile is the multiplication of both tiles' values.
Q9	**Next we would like to ask you about your experiences with ads that may have appeared in different locations on the webpage.**		
Answer code	**Please indicate if you recognized ads appearing in each of the following locations on the webpage:**		
a		I have seen an ad at the top of the webpage	I have seen an ad at the top of the webpage
b		I have seen an ad on the right side of the webpage	I have seen an ad on the right side of the webpage
c		I have seen an ad placed in the middle of the article text	
d		I cannot remember seeing any ad	I cannot remember seeing any ad
Q10	**Ad Experiences**		
	In this next section, we would like to ask you about your experiences with any ads you may have seen on the webpage.		
Answer code	**Please indicate what ads you can recall (one ad per field). Please use simple keywords to**		

	describe the ad(s) (e.g., the content or brand of the ad):		
a	1st ad		
b	2nd ad		
c	3rd ad		
INFO3	**Recognition of Ads**		
	Next we would like to show you some different ads that may have appeared on the webpage. For each ad, please indicate whether you recognize the ad and answer the following questions about each ad experience.		
	Please look at the two ads below. Please indicate whether you saw one, both, or neither of these ads on the webpage you viewed.		
INFO3.1	**Recognition of Ads**		
	You have indicated in the previous question that you did not recognize ads appearing on the website. Nevertheless, we would now like to show you some different ads that may have appeared on the webpage. For each ad, please indicate whether you recognize the ad and answer the following questions about each ad experience.		

	Please look at the two ads below. Please indicate whether you saw one, both, or neither of these ads on the webpage you viewed		
Q11.1	**Pair A:**		
Answer code	Picture of Ad A / Picture of Ad B		
a	I saw ad A		
b	I saw ad B		
c	I saw both ad A and ad B		
d	I saw neither ad		
Q11.2	Pair B:		
Answer code	Picture of Ad A / Picture of Ad B		
a	I saw ad A		
b	I saw ad B		
c	I saw both ad A and ad B		
d	I saw neither ad		
Q11.3	Pair C:		
Answer code	Picture of Ad A / Picture of Ad B		
a	I saw ad A		
b	I saw ad B		
c	I saw both ad A and ad B		
d	I saw neither ad		
Q13	**You indicated that you saw an ad on the right side of the webpage.**		
Answer code	**What type of ad did you notice in this location?**		
a	An ad with either an image or text or both (but not animated content), where the ad content did not change		

b	An ad with either an image or text or both (but not animated content), where the ad content did change (i.e. new ads were loaded in the same spot)		
c	An ad with animated content		
d	I noticed an ad but I don't recall what it looked like		
Q13.1	**You indicated that you saw an ad on the top of the webpage.**		
Answer code	**What type of ad did you notice in this location?**		
a	An ad with either an image or text or both (but not animated content), where the ad content did not change		
b	An ad with either an image or text or both (but not animated content), where the ad content did change (i.e. new ads were loaded in the same spot)		
c	An ad with animated content		
d	I noticed an ad but I don't recall what it looked like		
INFO4	**For each of the following, please rate your ad experience.** **As a reminder, a stylized screenshot of the actual web experience you were presented with is shown. There were ads that reloaded new ad content according to a set timer. The reloading ad position is highlighted in the screenshot.** **You can enlarge the screenshot by clicking on it.**		

Q14.1	**How disruptive was the reloading of new ad content to your experience?**		
Answer code			
5	Extremely disruptive		
4	Very disruptive		
3	Disruptive		
2	Slightly disruptive		
1	Not at all disruptive		
Q14.2	**How enjoyable was the reloading of new ad content to your experience?**		
Answer code			
1	Extremely enjoyable		
2	Very enjoyable		
3	Enjoyable		
4	Slightly enjoyable		
5	Not at all enjoyable		
Q14.3	**How annoying did you find the reloading of new ad content to your experience?**		
Answer code			
5	Extremely annoying		
4	Very annoying		
3	Annoying		
2	Slightly annoying		
1	Not at all annoying		
Q14.4	**How intrusive did you find the reloading of new ad content to your experience?**		
Answer code			
5	Extremely intrusive		
4	Very intrusive		

3	Intrusive		
2	Slightly intrusive		
1	Not at all intrusive		
INFO5	**Ad experience ratings for overall ad experience**		
	Now we would like to ask you to rate your experience with all the ads you have seen while visiting the webpage. **Please consider your OVERALL experience of the ads and their locations on the webpage.** **As a reminder, a stylized screenshot of the actual web experience you were presented with is shown. The screenshot highlights all ads. You can enlarge the screenshot by clicking on it.**		
Q15.1	**How disruptive were ALL the shown ads to your overall experience?**		
Answer code			
5	Extremely disruptive		
4	Very disruptive		
3	Disruptive		
2	Slightly disruptive		
1	Not at all disruptive		
Q15.2	**How enjoyable were ALL the shown ads to your overall experience?**		
Answer code			
1	Extremely enjoyable		
2	Very enjoyable		
3	Enjoyable		

4	Slightly enjoyable		
5	Not at all enjoyable		
Q15.3	**How annoying did you find ALL the shown ads to your overall experience?**		
Answer code			
5	Extremely annoying		
4	Very annoying		
3	Annoying		
2	Slightly annoying		
1	Not at all annoying		
Q15.4	**How intrusive did you find ALL the shown ads to your overall experience?**		
Answer code			
5	Extremely intrusive		
4	Very intrusive		
3	Intrusive		
2	Slightly intrusive		
1	Not at all intrusive		
Q15.5	Please choose the ad that was most disruptive to your web experience.		
Answer code			
a		ad at the top of the webpage	ad at the top of the webpage
b		ad at the right hand side of the webpage	ad at the right hand side of the webpage
c		ad placed in the text of the webpage	
INFO6	**Now we would like to ask you a few questions about your views on online advertising.**		

Q16	**In your opinion, how often can an ad refresh (i.e. a new ad shown in the same spot) before it becomes annoying or disruptive?**		
Answer code			
a	The ad can reload as often as it wants. I don't mind seeing reloading ads.		
b	The ad can reload once every 30 seconds. Reloading more often than that is disruptive.		
c	The ad can reload once every 60 seconds. Reloading more often than that is disruptive.		
d	The ad can reload once every 90 seconds. Reloading more often than that is disruptive.		
e	The ad can reload once every 120 seconds. Reloading more often than that is disruptive.		
f	The ad can reload once every 240 seconds. Reloading more often than that is disruptive.		
g	The ad should never reload. I find reloading ads to be disruptive.		
Q16.1	**Some ads on a webpage "reload" - that is, after some period of time passes, the original ad disappears and is replaced by another ad in the same spot. The ads that you saw earlier loaded a new ad every [x] seconds.** **In your opinion, what is an acceptable time period before**		

	an ad should be allowed to reload?		
Answer code			
a	The ad can reload as often as it wants. I don't mind seeing reloading ads.		
b	The ad can reload once every 30 seconds. Reloading more often than that is disruptive.		
c	The ad can reload once every 60 seconds. Reloading more often than that is disruptive.		
d	The ad can reload once every 90 seconds. Reloading more often than that is disruptive.		
e	The ad can reload once every 120 seconds. Reloading more often than that is disruptive.		
f	The ad can reload once every 240 seconds. Reloading more often than that is disruptive.		
g	The ad should never reload. I find reloading ads to be disruptive.		
Q16.2	**Please briefly explain why you chose this answer:**		
	[OPEN TEXT FIELD]		
Q17	**What is the main ad blocker you are currently using? Please select all that apply:**		
Answer code			
a	Ghostery		
b	ublock Origin		
c	Brave		
d	AdBlock		
e	AdGuard		

f	Adblock Plus		
g	None of the above (please specify): _____		
Q18	**Blocking ads can improve your browsing experience, but it can also significantly impact publishers and advertisers. Which of the following do you agree will result from using an ad blocker that blocks all ads?**		
Answer code			
a	Users may not have access to certain content or users may be required to pay for certain content		
b	There will be less independent content available on the internet (i.e. more sponsored content)		
c	Publishers will not be paid or will be paid less, for the content that they create		
d	Content will not be equally available to all internet users		
e	I am not aware of any negative effects of ad blocking		
INFO7	**For the following statements, please indicate how much you agree or disagree with each statement:**		
Q19.1	**Generally, I consider Internet advertising to be a good thing.**		
Answer code			
5	Completely disagree		
4	Disagree		
3	Neither agree nor disagree		
2	Agree		

1	Completely agree		
Q19.2	**I appreciate seeing advertising messages on the Internet.**		
Answer code			
5	Completely disagree		
4	Disagree		
3	Neither agree nor disagree		
2	Agree		
1	Completely agree		
Q19.3	**Internet advertising supports free access to content.**		
Answer code			
5	Completely disagree		
4	Disagree		
3	Neither agree nor disagree		
2	Agree		
1	Completely agree		
Q19.3	**Online advertisements promote competition, which benefits consumers.**		
Answer code			
5	Completely disagree		
4	Disagree		
3	Neither agree nor disagree		
2	Agree		
1	Completely agree		
Q19.5	**Online advertisements support content creators.**		
Answer code			
5	Completely disagree		
4	Disagree		
3	Neither agree nor disagree		

2	Agree		
1	Completely agree		
Q19.6	**There are too many advertisements on the Internet.**		
Answer code			
5	Completely disagree		
4	Disagree		
3	Neither agree nor disagree		
2	Agree		
1	Completely agree		
Q19.7	**Internet advertisements intrude on the content I am accessing.**		
Answer code			
5	Completely disagree		
4	Disagree		
3	Neither agree nor disagree		
2	Agree		
1	Completely agree		
Q19.8	**Online advertising disrupts my activity on the Internet.**		
Answer code			
5	Completely disagree		
4	Disagree		
3	Neither agree nor disagree		
2	Agree		
1	Completely agree		
Q19.9	**Consumers may obtain reliable information through Internet advertising.**		
Answer code			
5	Completely disagree		
4	Disagree		

3	Neither agree nor disagree		
2	Agree		
1	Completely agree		
Q19.10	**Viewing online advertisements is a pleasant experience for me.**		
Answer code			
5	Completely disagree		
4	Disagree		
3	Neither agree nor disagree		
2	Agree		
1	Completely agree		
Q19.11	**Sometimes I take pleasure in thinking about what I saw or heard in online ads.**		
Answer code			
5	Completely disagree		
4	Disagree		
3	Neither agree nor disagree		
2	Agree		
1	Completely agree		
INFO7	**Thanks for participating in this survey. Your answers are of high value for us.**		

Appendix 3 - Additional data summaries and analyses

A3.1. Participants perceptions of the most disruptive ad

Participants could also name the most disruptive ad of their website experience. In the Article experience (Figure A.1), the majority of participants perceived the in-content ad as being most disruptive, even when they did not experience an in-content ad (Figure 12, third panel). It is possible that participants misunderstood the question and reported what they generally perceived to be the most disruptive ads or they incorrectly assumed that an in-content ad was present in their web experience.

In the Game experience (Figure A.2), if the ad refresh occurred in the top banner ad, the top banner ad placement was considered more disruptive than if the top banner ad was static. The same pattern was evident for the right rail ad: when the ad refresh occurred in the right rail ad, it was considered more disruptive than when it was the static ad.

Figure A.1: Proportion of responses to the question: "Please choose the ad that was most disruptive to your web experience" for the Article experience

With 95% confidence interval

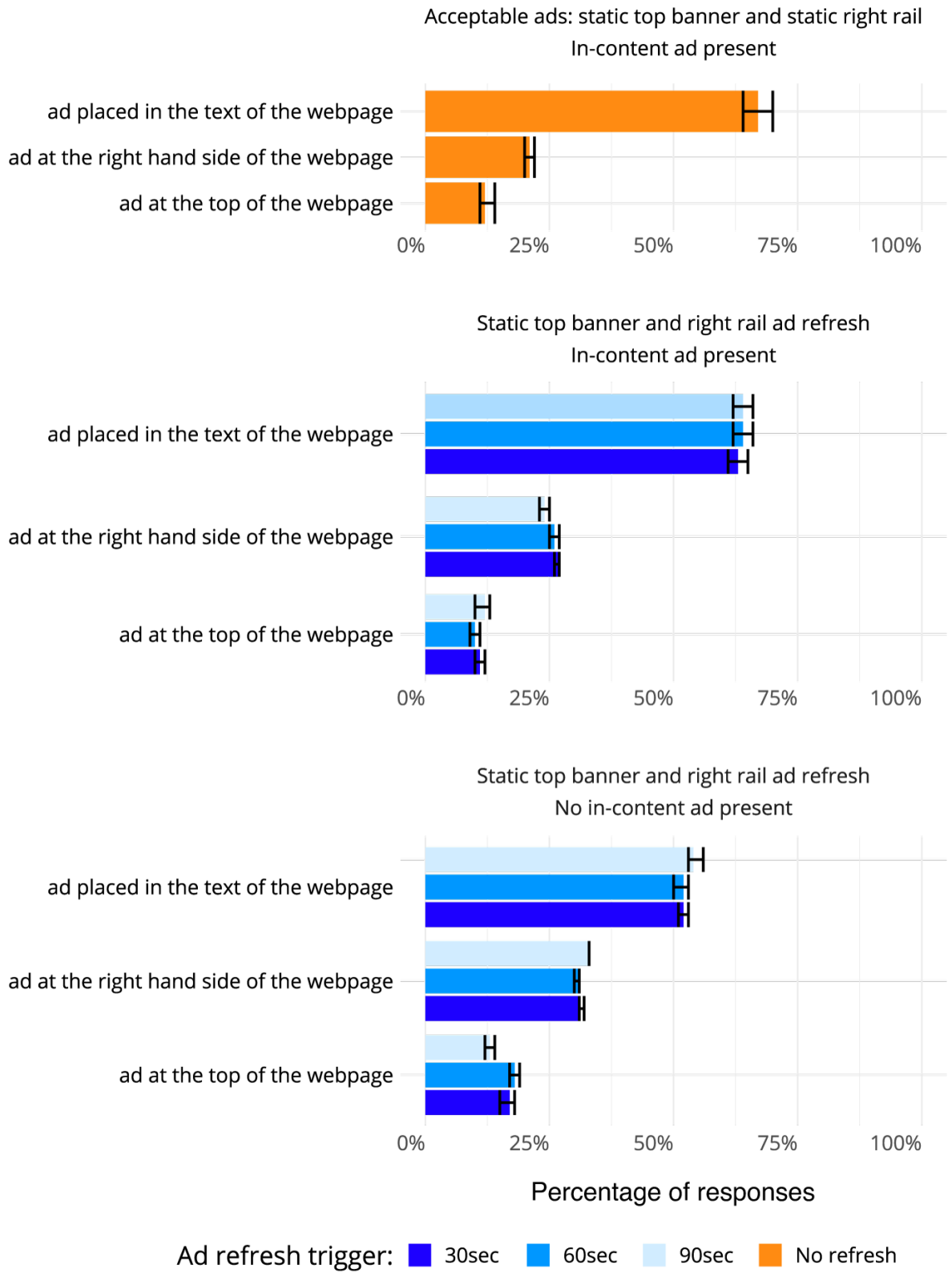
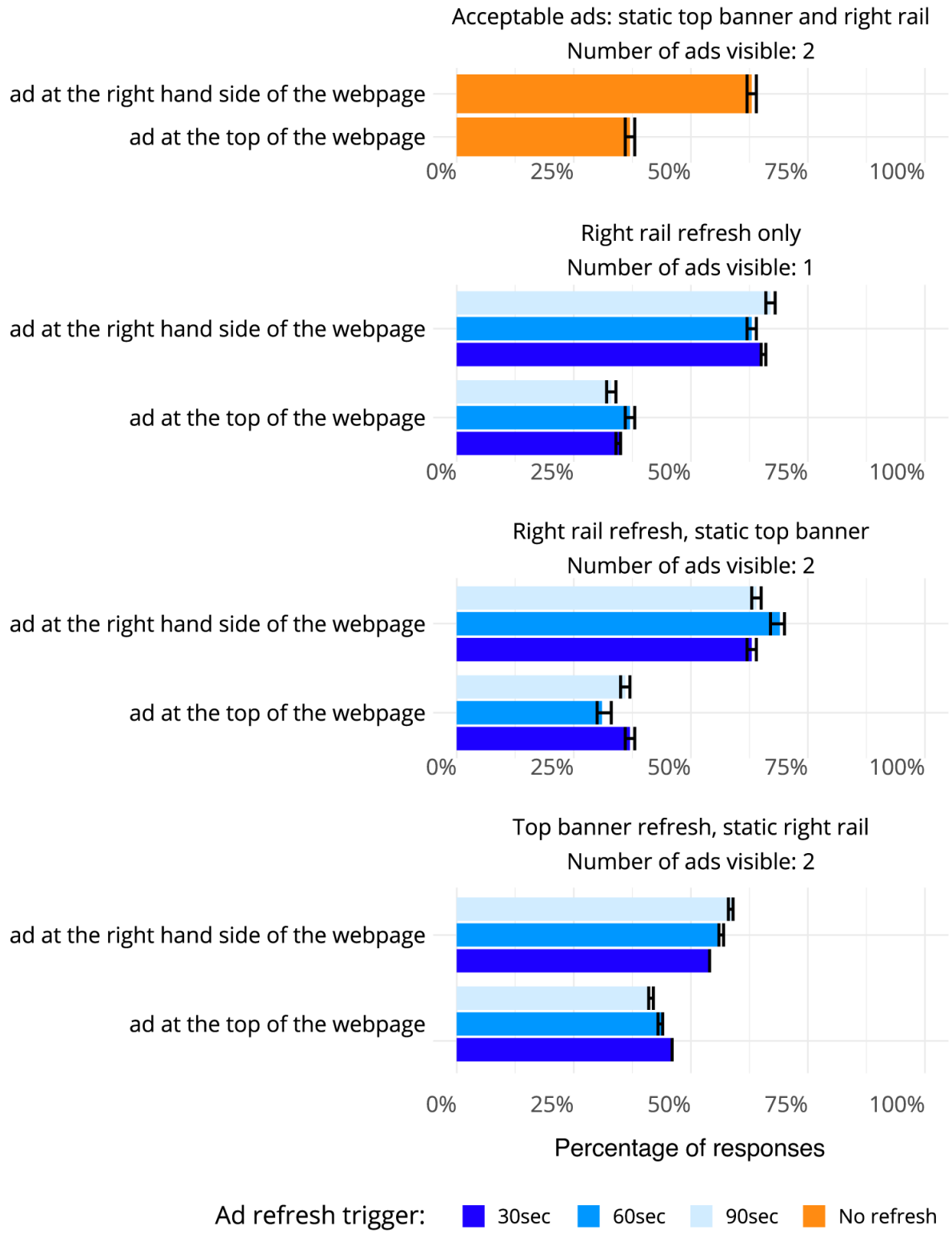


Figure A.2: Proportion of responses to the question: "Please choose the ad that was most disruptive to your web experience" for the Game experience

With 95% confidence interval



A3.2 Perceptions of online advertising

Participants answered a set of 11 questions assessing their perceptions of online advertising on 5-point scales ranging from "completely disagree" to "completely agree" (e.g., "Generally, I consider internet advertising to be a good thing"). Details for each item can be found in Appendix 2. An average across the responses to the 11 items were calculated representing participant's perception towards online advertising with lower values representing more positive attitudes. Table A2 shows that in all three countries, the younger generation have more positive attitudes towards online advertising. Participants in the US show the most positive attitudes. There are no meaningful differences between genders or education levels. For income, in Germany and France there is a tendency that participants with lower income tend to have more negative views on online advertising.

Table A.2: Median general attitude towards online advertisement by country and generation

With 95% confidence interval in brackets

	Generation	
Country	**18-44 years**	**45 years or older**
US	3 *[3.0, 3.1]*	3.3 *[3.3, 3.4]*
DE	3.1 *[3.1, 3.2]*	3.5 *[3.4, 3.6]*
FR	3.2 *[3.2, 3.3]*	3.7 *[3.6, 3.7]*

Note: Scores on a 5-point scale with lower scores represent more positive attitudes towards online advertising.

A3.3. Engagement with the web experience

To ensure participants would see the ad refresh effect, they were required to stay on the webpage for a minimum of 2.5 minutes up to a maximum of 5 minutes. The average time spent on the Article (3.69 minutes; 95%-CI: [3.66, 3.72]) and the Game (3.90

minutes; 95%-CI: [3.87, 3.93]) exceeded the minimum time by more than a minute, suggesting that many participants were engaged with the content. Importantly, there were no meaningful differences in how long participants spent on the webpage across the different ad experiences. Further, for the Article experience, it appears that the majority of participants (n = 3,231) interacted with the webpage as expected: they slowly scrolled down the website, stopped at certain heights to interact with the content before moving on to explore the website further. In addition, the majority of participants indicated that they were slightly or very satisfied with the Article (76.2%) or Game (69.6%) experience.

We also checked how well participants comprehended the webpage content. In total, 65.8% of participants in the Article experience and 80.3% in the Game experience[4] answered 2 out of the 3 comprehension items correctly and there were no meaningful differences across the different ad experiences. In summary, these results suggest that participants were sufficiently engaged with the web experiences.

A3.4. Engagement with the ad experience

To evaluate engagement with the ads in the web experiences, participants were asked to recall where they had seen ads and to indicate which ads they recognized appearing on the webpage. Interestingly, 29.9% of the participants in the Article experience and 45.7% in the Game experience reported that they had not seen any ads on the webpage.

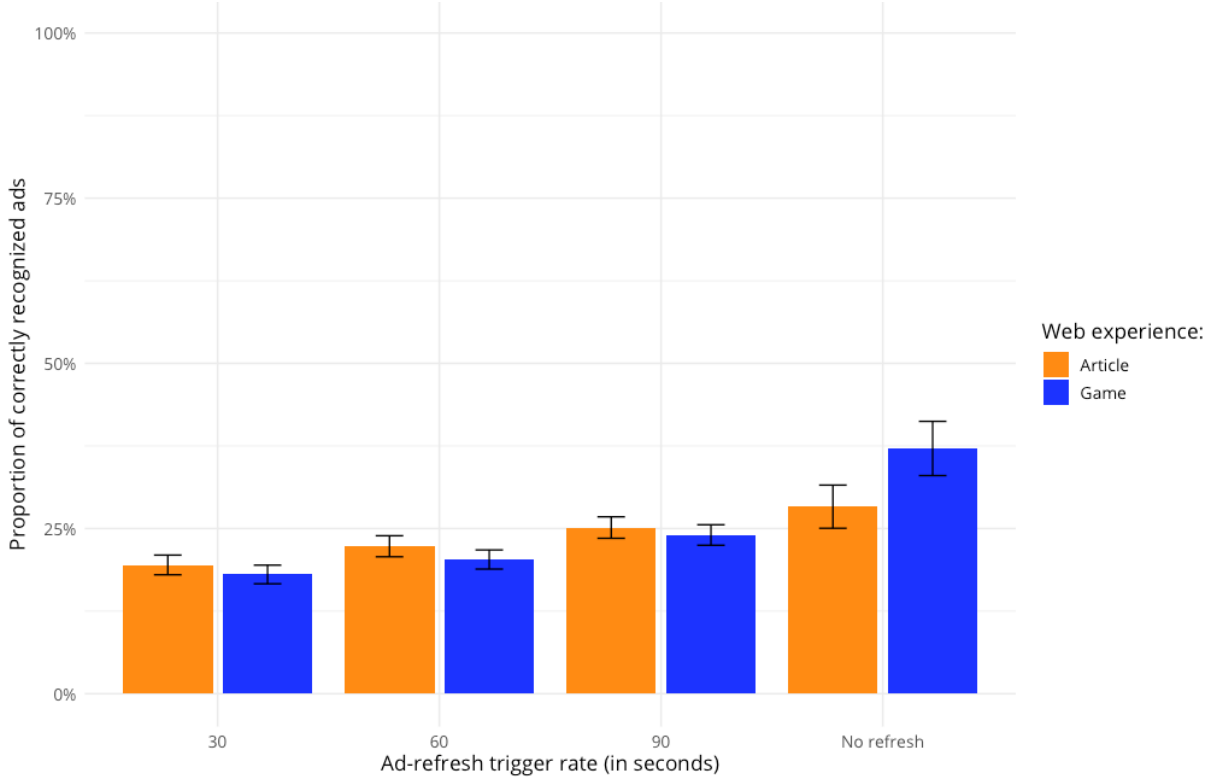
For the Article experience, only 46% of participants remembered seeing the top banner ad, consistent with an effect known as (top) banner blindness. The right rail sticky ad was remembered better, with 65% of participants remembering the ad. If the in-content ad was present, only 41% recalled seeing the ad. For the Game experience, when the top banner ad was present, 60% of participants remembered seeing the ad, similar to

[4] The 15% discrepancy between the Article and Game experience could be explained by the greater difficulty answering questions about the article content relative to a game experience where immersion in the game play enabled participants to understand the mechanics of the game in order to answer the comprehension questions.

memory of the right rail ad (64%). For each experience, the ad refresh time-based trigger rate did not improve recall of the ad locations. However, younger participants were more likely than older participants to remember any of the ads.

Figure A.3: Average share of correctly recognized ads based on three paired comparison tasks where participants were shown an ad from their web experience and a decoy ad and asked to indicate which of the two ads appeared on the webpage.

With 95% confidence interval



Finally, to examine how well participants recognized the ads that they were shown, participants were presented with three pairs of ads with each pair containing an ad that was actually presented to them or a decoy ad (an ad that never appeared in their web experience). As shown in Figure A.3, recognition of ads was greatest in the Acceptable Ads experience where both ads were static and did not refresh. Further, recognition was lower for shorter ad refresh time-based trigger intervals. This result is consistent with industry market patterns where CPMs decrease for higher ad refresh time-based trigger rates as advertisers account for the fact that users may not see the ad for as long and therefore may be less likely to remember the ad's content.