ZESZYTY NAUKOWE POLITECHNIKI POZNAŃSKIEJ
Nr 93 Organizacja i Zarządzanie 2026

Arkadiusz MORAWSKI[1], Rafał MICHALSKI[2]

THE IMPACT OF ANTAGONIZATION ON THE PERCEPTION OF ADVERTISING MATERIALS IN FOOTBALL RIVALRIES: A FIELD EXPERIMENT IN GOOGLE ADS

DOI: 10.21008/j.0239-9415.2026.093.03

Football rivalries create a highly emotional context in which antagonistic or comparative advertising is expected to increase attention and engagement. Consequently, such strategies are considered vital for brands seeking to stand out in saturated markets by appealing to deep-seated consumer identities. Despite the ongoing debate regarding the risks of these advertisements, they remain widely used in rivalry-driven contexts under the assumption that they deliver superior results compared to traditional forms. Our study examines whether creative antagonistic advertising increase click-through rates in a real-life setting. We conducted a 2 × 2 × 2 full factorial field experiment using Google Ads display campaigns targeting a 30-km radius around Kraków (Poland). Eight different static banner variants manipulated emotional tone (antagonistic vs. neutral), use of color (club-related shirt colors vs. grayscale), and format size (200 × 200 px vs. 250 × 250 px). Click-through rate (CTR) was selected as the primary dependent variable. The campaign generated a total of 9,550 impressions and 283 clicks (overall CTR = 2.96%) across all variants. The results were formally analyzed using a series of nonparametric tests. In general, the analysis revealed that factor-level contrasts were non-significant, with the exception of one pairwise comparison. Contrary to some previous studies, we found no evidence of a systematic CTR advantage for antagonizing messages or club colors in this football rivalry setting. These findings highlight the need for more in-depth local testing and caution against assuming that performance necessarily improves as a result of antagonization in sports advertising.

Keywords: sports advertising, antagonization, comparative ads, football rivalry, field experiments, Google Ads, CTR

[1] ORCID: 0009-0005-4671-6973.
[2] Wrocław University of Science and Technology, Department of Psychology and Ergonomics. ORCID: 0000-0002-0807-1925.

1. INTRODUCTION

Advertisement based on rivalry targets competition between groups (e.g., politicians or sports clubs), to increase the level of emotional arousal, sharpen brand differentiation and influence behavioral responses of the target group. Antagonistic advertising is a subtype of comparative advertising that clearly attacks the rival. It may boost engagement but also result in a series of risks (e.g., reduced brand trust, ethical concerns, among others). This kind of advertisement has attracted growing attention as many online ad platforms facilitate high-tempo testing and make the effects of those experiments measurable.

Football rivalries create a unique emotional and attentional context for advertising. Fans identify strongly with their club and often react more intensely to messages related to opponents – making their behavior an interesting field of research (Mills et al., 2018; Sanford, Scott, 2016). To target such audiences, advertisers may be interested in using antagonizing or comparative messages to attract attention and convince football fans to buy their products or services (Nichols et al., 2016; Kwak, Pradhan, 2021).

While interest in negative and comparative appeals is growing, there is still a lack of field evidence or data-based tests in the sports industry. Most studies rely on lab experiments, self-reported attitudes, or general consumer samples, rather than real click behavior in a particular rivalry context (Johnson, 2023). Moreover, most advertising-based science articles focus on rivalries between politicians (Pinkleton, 1997; Wattenberg, Brians, 1999). In parallel, online display advertising is a challenging environment for field experiments, because exposure is shaped by auctions, platform optimization, and competition, with the true effects often being small (Johnson, 2023).

This paper presents a field experiment conducted in Google Ads in the context of a football rivalry in Kraków. We manipulate the level of antagonizing emotions, the use of club-related colors, and banner size in an eight-variant factorial design. The main dependent variable is CTR, observed under real campaign conditions. The study provides ecologically grounded evidence on when antagonizing messages do – or do not – lead to higher CTRs in the sports industry.

The aim of this paper is to empirically assess how antagonizing creative messages and rivalry-linked visual cues (club colors) lead to higher click-through rates compared to neutral creatives in a real Google Ads display campaign conducted in a local football rivalry market. A secondary aim is to evaluate the role of banner size and potential interaction effects between these three factors.

2. LITERATURE REVIEW

For a long time, comparative and negative advertising have been studied as an interesting way to gain attention, create differentiation, and stimulate behavioral responses in customers. Comparative advertising evaluates how a brand is perceived

against a competitor, while negative or "attack" variants add attacking cues directed at the target (Starcevic, 2006; Casais, Proença, 2015). Despite that interest, it is hard to make a general conclusion about this type of advertising. Some studies report behavioral advantages, such as increased purchase rates, while others find equivocal results or outcomes dependent on context and highlight risks such as unfairness perceptions, credibility loss, or unintended competitor salience (Demirdjian, 2015; Soscia et al., 2010).

Empirical findings on negative and comparative appeals are mixed. In an early experiment with 165 participants and three levels of negative ad intensity, Pinkleton (1997) showed that negative comparative political advertising consistently lowered evaluations of the targeted candidate without harming evaluations of the sponsoring candidate, despite the ads themselves receiving low credibility ratings. The effect on the sponsor was non-linear: moderately negative messages increased favorability toward the sponsor, while strongly negative messages reduced it, suggesting a curvilinear relationship between message intensity and sponsor backlash. Soscia et al. (2010), in an experiment with 280 participants, found that comparative advertising's effect depends on consumers' perceived differentiation between the compared brands and on their involvement with the product category; under some conditions, comparative ads paradoxically increased perceived similarity between the advertised brand and its competitor, rather than producing the intended differentiation. Muehling et al. (2018) showed across three empirical studies that individual differences in "tolerance of negativity" moderate consumer responses to comparative attack advertising, with low-tolerance consumers reacting unfavorably even to mild comparative cues. Taken together, these findings indicate that the effectiveness of comparative or antagonizing messages depends jointly on message intensity, audience characteristics, and the perceived alignment between advertised and referenced brands, rather than translating into uniformly positive effects.

Comparative ads can backfire by increasing perceived similarity instead of differentiation (Soscia et al., 2010). The structure of the comparison may also influence the results: direct or indirect formats can interact with the comparison strategy (market-leader vs. multi-brand) and affect ad perception (Kalro et al., 2017). Other studies stress informational mechanisms – comparative messages may work better when they provide clear, meaningful contrasts (e.g., via message informativeness) rather than relying on tone alone (Hsu, 2020).

Individual differences also play a crucial role: tolerance of negativity depends on the target group and shapes responses to comparative attack advertising, especially when the ad is an object of deeper analysis, implying that negative executions can polarize audiences (Muehling et al., 2018). In digital marketing, reactance and perceived intrusiveness may further reduce effectiveness (Chen et al., 2019; Bambauer-Sachse, Heinzle, 2018), while repeated exposure to the same ad can trigger processes that diminish engagement (Alwreikat, Rjoub, 2020). These constraints are particularly important for display advertising (ads where an image plays dominant role), where

attention is scarce (due to daily ad saturation) and creative effects may be dampened by clutter and low-involvement browsing (Filiopoulou et al., 2019; Le, Vo, 2017).

Other experiments show that online display environments often yield small, hard-to-isolate effects that are hard to interpret. Field experiments suffer from limited control over the setting because outcomes are shaped by users, algorithms, and market competition (Johnson, 2023). Ad settings are often limited or can be affected by dozens of uncontrolled variables. Studies in other online advertising domains report near-null effects (Broockman, Green, 2014), while evidence from mobile banners suggests that effects can vary across outcomes: ad networks, locations, devices, and channels (Osinga et al., 2019). In addition, targeting complexity and data quality can constrain performance gains; many segments would require implausibly large CTR lifts to outperform broad targeting (Ahmadi et al., 2024). The increasing influence of AI incorporated into ad platforms should also be considered. The amount of control over campaigns is being reduced, while the majority of important decisions are increasingly delegated to AI systems. Together, this motivates cautious interpretation of small CTR differences and supports real-platform field experiments to validate creative assumptions.

Quantitatively, the magnitude of effects reported in display-advertising field experiments is consistently small. Broockman and Green (2014) conducted two randomized field experiments with Facebook ads in U.S. legislative campaigns; their first experiment randomized 32,029 voters and completed 2,984 telephone interviews, yet found no statistically detectable increase in candidate name recognition or favorability, with only a 5.3-percentage-point increase in ad recall in the second experiment and no recall effect in the first. Osinga et al. (2019) delivered over 3.5 million mobile banner ads across the Netherlands and tracked twenty months of sales data; using a difference-in-differences design with matching, they observed an increase of approximately 2% in offline sales but no statistically significant effect on online sales. Johnson (2023), summarizing the methodological state of the field, characterizes display advertising as a "hostile medium" for experimentation, one in which effect sizes are inherently tiny and where exposure is shaped jointly by advertisers, users, algorithms, and competitive auction dynamics. This set of findings establishes a clear empirical baseline: even very large field experiments often produce small or null behavioral effects. This has direct implications for the interpretation of click-based outcomes in smaller-scale studies such as the present one.

In sports marketing, rivalry and fan identity provide a theoretical basis for anticipating strong reactions to opponent-related cues. Rivalry contexts intensify in-group/out-group distinctions and can shape behavioral choices toward out-groups (Mills et al., 2018) as well as driving willingness-to-pay patterns that reflect rivalry intensity (Sanford, Scott, 2016). Team identification can condition responses to sponsor advertising, and rival cues may suppress supportive responses in certain campaign contexts (Kwak, Pradhan, 2021; Nichols et al., 2016). Fan involvement is also heterogeneous, suggesting that casual and highly identified fans may respond differently (Clowes, Tapp, 2003; Lu et al., 2012). In Kraków, rivalry between Wisła

Kraków and Cracovia has been documented through local markers and fan attitudes, supporting the ecological relevance of rivalry cues in this market (Krauthausen et al., 2019; Farbaniec, Kozioł, 2019).

Empirical research in sports marketing has documented several mechanisms by which rivalry-related cues shape consumer responses to advertising. Kwak and Pradhan (2021), in three experiments examining sponsor advertisements that featured a team after a pivotal loss, showed that high-identifier fans responded more favorably than low-identifier fans to advertising acknowledging the loss, while low-identifier fans showed less favorable evaluations of the same brand, indicating asymmetric processing of rival-related communications across fan segments. Nichols et al. (2016), in three experiments on cause-related sports marketing, demonstrated that the presence of a rival team in league-wide campaign imagery can reduce fans' intentions to support the cause, with perceived sincerity acting as a mediating mechanism. These findings suggest that the same rivalry cue may produce engagement, indifference, or backlash depending on fan typology and on the specific advertising context, complicating the simple expectation that rivalry-related stimuli always enhance attention and engagement.

This combined evidence highlights a clear research gap. Comparative or negative advertising research rarely tests antagonizing executions in real, paid digital environments using platform-native behavioral metrics. Similarly, sports marketing research documents rivalry and identification effects but seldom evaluates antagonistic advertising strategies via controlled ad-platform experiments. Field evidence therefore remains limited on whether antagonizing messages and rivalry-linked visual cues (such as club colors) reliably improve CTRs in club rivalry contexts under real display delivery dynamics. The present study addresses this gap through a factorial field experiment in the Kraków rivalry market, using CTRs as the primary behavioral outcome.

3. RESEARCH METHODOLOGY

3.1. Research questions and hypotheses

As outlined in the Introduction, prior work on negative and comparative appeals suggests that such messages can possibly increase engagement and behavioral responses compared to more neutral communication (Demirdjian, 2015; Yousef et al., 2021). At the same time, the evidence has been generated mostly in non-sports contexts – such as political campaigns (Pinkleton, 1997; Wattenberg, Brians, 1999) and health promotion (Yousef et al., 2023) – and under controlled laboratory or survey-based conditions.

To the best of our knowledge, no prior study has directly examined how antagonizing versus neutral messages perform on real-time bidding advertising platforms in the context of football club rivalries. Adjacent research, however, provides relevant evidence:

comparative and antagonistic appeals have been tested in political and laboratory contexts (Pinkleton, 1997; Soscia et al., 2010; Muehling et al., 2018), display-advertising field experiments have documented small and often null behavioral effects in non-sports settings (Broockman, Green, 2014; Osinga et al., 2019; Johnson, 2023), and rivalry-related cues have been shown to shape consumer responses to sponsorship and brand messages in surveys and laboratory studies (Kwak, Pradhan, 2021; Nichols et al., 2016).

The design manipulates three factors – emotional tone (antagonistic vs. neutral), use of color (club-related shirt colors vs. grayscale) and format size (200 × 200 px vs. 250 × 250 px) – and observes differences in campaign CTRs. Because real-world online advertising effects are often small and difficult to isolate under auction- and algorithm-driven delivery, field-based validation is particularly important in this setting (Johnson, 2023; Broockman, Green, 2014).

Building on the literature reviewed above and the identified research gap, we formulated four research questions and three corresponding hypotheses guiding the empirical analysis:

RQ1: How does the emotional tone of a creative (antagonistic vs. neutral) influence the click-through rate in a real football rivalry advertising campaign?

RQ2: How does the use of club-related shirt colors (vs. grayscale) influence the click-through rate?

RQ3: How does banner format size (200 × 200 px vs. 250 × 250 px) influence the click-through rate?

RQ4: Do banner size, use of color, and emotional tone interact with each other in shaping click-through rate, beyond their main effects?

Based on prior empirical and theoretical work, the first three research questions translate into directional hypotheses, while RQ4 is examined exploratorily:

H1: Antagonistic ads will achieve a higher CTR than neutral ads.

H2: Using club colors in advertising creatives will increase CTR compared to grayscale colors.

H3: The larger ad format (250 × 250 px) will achieve a higher CTR than the smaller format (200 × 200 px).

3.2. Setting and platform

The experiment was conducted in Google Ads, using standard display campaigns. The ads promoted Football Team, a Polish free-to-play browser-based football management game with MMO elements, in which users develop both a club and individual players over time.

The campaign was targeted to users located within a 30-kilometre radius of the center of Kraków, a city characterized by an unusually dense and historically rooted football rivalry scene, including multiple professional clubs and the long-standing "Święta Wojna" between Wisła Kraków and Cracovia (Krauthausen et al., 2019; Farbaniec, Kozioł, 2019).

The test campaign ran for one week, from 3 September 2025 to 10 September 2025. Ads were eligible to be shown on all device types, with no explicit device restrictions; given the square banner formats, a high share of impressions was observed on mobile phones. There were no explicit age or gender restrictions (18-65+), and no detailed audience targeting beyond the geographic radius; audience segments were only observed ex post in the reporting interface.

To minimize the impact of automated optimization implemented by the advertising platform, the campaign relied on standard display delivery, disabled optional goal-based optimizations where possible, and avoided additional automated elements such as responsive creatives or dynamic formats (Johnson, 2023).

3.3. Experimental design

The study followed a 2 × 2 × 2 full factorial design. Three experimental factors were manipulated: banner size, use of color and emotional tone. Banner size had two levels: a smaller 200 × 200 px format (S200) and a larger 250 × 250 px format (S250), corresponding to two commonly used square display formats in the Google Display Network.

The use-of-color factor distinguished creatives rendered in club-related shirt colors (color condition) from their grayscale counterparts (gray condition). The emotional tone factor contrasted antagonistic creatives, which explicitly leveraged the rivalry context, with neutral creatives that did not contain antagonizing content.

Crossing these three binary factors resulted in eight experimental variants (V1-V8). Each variant represented a unique combination of size (200 × 200 px vs. 250 × 250 px), use of color (color vs. gray) and emotional tone (antagonistic vs. neutral). All eight variants were placed within the same Google Ads campaign, shared identical targeting and bidding settings, and were eligible for delivery throughout the full campaign period. Ad rotation and other delivery settings were configured to reduce systematic favoring of creatives, thereby allowing performance differences to emerge primarily from user responses.

3.4. Materials

The stimuli consisted of static display banners prepared in two sizes: 200 × 200 px and 250 × 250 px. For each size, four versions were created: two in club-related colors and two in grayscale, combined with antagonistic versus neutral messaging. In the color condition, players' shirts and key visual elements used club-related colors associated with the Kraków rivalry context, whereas in the grayscale condition these visual cues were removed and replaced with neutral tones.

Antagonistic creatives contained provocative, rivalry-oriented slogans that implicitly or explicitly referenced the opposition between local clubs, aiming to activate existing antagonisms among viewers. Neutral creatives used more generic

football-related messaging that invited users to try the game without referring to any specific club rivalry.

Antagonistic (PL original): “KTO RZĄDZI W KRAKOWIE?”, “GRAJ ZA DARMO”.

Antagonistic (EN translation): “WHO RULES IN KRAKÓW?”, “PLAY FOR FREE”.

Neutral (PL original): “ZOSTAŃ LEGENDĄ MIASTA!”, “GRAJ ZA DARMO”

Neutral (EN translation): “BECOME A CITY LEGEND!”, “PLAY FOR FREE”.

Apart from the manipulated elements (color and text), the layout, imagery style and branding elements of the banners were held constant across conditions. Each unique combination of size, use of color and emotional tone formed one experimental variant (V1-V8), as summarized in table 1.

Table 1. Experimental variants (V1-V8) in a 2 × 2 × 2 factorial design (Size × Disp × Emo)

Variant	Size (px)	Display – Emotional tone	Design
V1	200 × 200	Color – Antagonistic	
V2	200 × 200	Color – Neutral	
V3	200 × 200	Gray – Antagonistic	

Variant	Size (px)	Display – Emotional tone	Design
V4	200 × 200	Gray – Neutral	
V5	250 × 250	Color – Antagonistic	
V6	250 × 250	Color – Neutral	
V7	250 × 250	Gray – Antagonistic	
V8	250 × 250	Gray – Neutral	

3.5. Measures

The primary dependent variable was CTR. For each variant, CTR was calculated as the number of recorded clicks divided by the number of impressions, expressed as a percentage. The key input variables for the analysis were therefore, for each variant V1-V8, the total number of impressions and the total number of clicks. These data were obtained from standard Google Ads campaign reports exported in CSV format at the end of the campaign period.

In addition to CTR, the reporting interface also provided information on devices, age and gender splits and automatically inferred audience segments. These variables were used descriptively (for example, when characterizing the sample composition) but did not form the primary focus of the inferential analyses. The main inferential tests were based on click versus non-click counts per variant, as described in the Results section.

3.6. Procedure

Throughout the campaign, delivery and basic performance indicators were monitored to ensure that all creatives received non-trivial numbers of impressions and that no obvious technical issues occurred. At the end of the one-week period, aggregated data on impressions and clicks for each variant were exported from Google Ads as CSV files. These aggregated counts formed the basis for the descriptive statistics (impressions, clicks, CTR per variant) and the inferential tests (chi-square tests and pairwise contrasts) reported in the Results section.

4. RESULTS

Across the eight variants (V1-V8), the campaign reached a total of 9,550 impressions and generated 283 clicks, resulting in an overall CTR of 2.96%. At the individual variant level, CTR values were narrowly distributed, with no single creative clearly dominating the others. The maximum spread between the best- and worst-performing variants within any size group stayed below one percentage point, indicating a relatively limited spread of results.

Table 2 summarizes impressions, clicks and CTR for the 200 × 200 px formats (S200). In this size group, the highest CTR was obtained by V3 (Gray – Antagonistic, 3.00%), whereas V1 (Color – Antagonistic, 2.28%) achieved the lowest CTR, with a difference of 0.72 percentage points between the two. V2 (Color – Neutral) reached a CTR of 2.98%, and V4 (Gray – Neutral) reached 2.75%, placing them between these extremes.

Table 2. Impressions, clicks and CTR for variants V1-V4 (200 × 200 px format)

Variant	Impressions	Clicks	CTR (%)
V1 – Color – Antagonistic	1,273	29	2.28%
V2 – Color – Neutral	1,575	47	2.98%
V3 – Gray – Antagonistic	1,202	36	3.00%
V4 – Gray – Neutral	981	27	2.75%

Table 3 presents the corresponding outcomes for the 250 × 250 px formats (S250). Here, the top-performing variant was V6 (Color – Neutral, 3.69%), followed by V5 (Color – Antagonistic, 3.21%) and V8 (Gray – Neutral, 3.15%), whereas V7 (Gray – Antagonistic) produced the lowest CTR at 2.77%. The spread between the best and worst performing creatives in this size group amounted to 0.92 percentage points. Taken together, the descriptive results suggest that while some variants perform slightly better than others, the overall variation in CTR is limited.

Table 3. Impressions, clicks and CTR for variants V5-V8 (250 × 250 px format)

Variant	Impressions	Clicks	CTR (%)
V5 – Color – Antagonistic	936	30	3.21%
V6 – Color – Neutral	1,084	40	3.69%
V7 – Gray – Antagonistic	1,229	34	2.77%
V8 – Gray – Neutral	1,270	40	3.15%

When comparing the two sizes, the larger 250 × 250 px format achieved a higher average CTR of approximately 3.21%, compared to 2.75% for the smaller 200 × 200 px format. While this pattern is directionally consistent with the expectation that larger formats attract more engagement, the magnitude of the difference is modest. Within each size group, variations between antagonistic & neutral and color & grayscale creatives were minimal, showing no systematic advantage for any specific combination of emotional tone and use of color.

To formally test whether CTR differed across variants, we conducted a chi-square test of independence on click versus non-click frequencies aggregated by variant. The omnibus test across all eight creatives produced a chi-square statistic of $\chi^2 = 4.74$ with 7 degrees of freedom and a p-value of 0.692, indicating that the null hypothesis of equal click probabilities across variants cannot be rejected at conventional significance levels (0.05). This result is consistent with the relatively small descriptive differences in CTR reported above and provides no evidence for substantial differences in CTR at the global level (fig. 1).

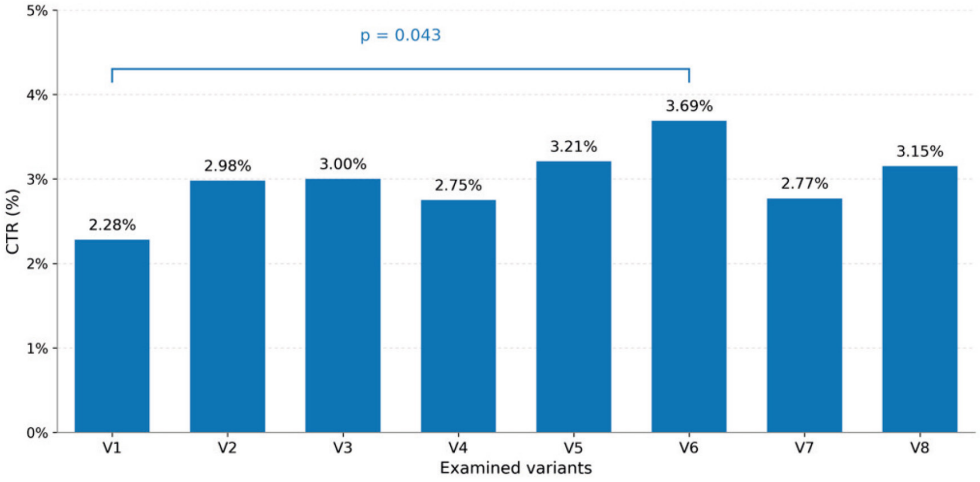

Fig. 1. Observed CTR values by Variant (V1-V8) the bracket indicates the only statistically significant pairwise difference (V1 vs. V6, p = 0.043)

A matrix of pairwise contrasts presented in table 4 indicated that only one comparison reached conventional significance: V1 (Color – Antagonistic, 200 × 200) versus V6 (Color – Neutral, 250 × 250 px), with p = 0.043. Even in this case, however, the absolute difference in CTR remained small in practical terms. Overall, the pattern of results points to modest performance differences between variants rather than to clear winners and losers.

Table 4. Matrix of pairwise contrasts (chi-squares and p-values)

	V1	**V2**	**V3**	**V4**	**V5**	**V6**	**V7**	**V8**
V1	×	0.245	0.265	0.473	0.182	0.043	0.436	0.176
V2	1.35	×	0.987	0.734	0.756	0.315	0.733	0.799
V3	1.24	<0.001	×	0.736	0.78	0.355	0.736	0.824
V4	0.514	0.116	0.114	×	0.56	0.23	0.984	0.582
V5	1.78	0.096	0.078	0.34	×	0.552	0.551	0.941
V6	4.11	1.01	0.857	1.44	0.353	×	0.208	0.471
V7	0.608	0.117	0.113	<0.001	0.356	1.59	×	0.572
V8	1.83	0.065	0.049	0.303	0.005	0.52	0.319	×

P-values are reported above the diagonal, and chi-square (χ^2) statistics are reported below the diagonal. The highlighted cell indicates the only pairwise comparison reaching $p < 0.05$ (V1 vs V6; p = 0.043).

Additionally, we examined selected factor-level contrasts by collapsing variants along specific dimensions. When comparing formats (tab. 5), the smaller 200 × 200 px banners yielded 139 clicks out of 5,031 impressions (CTR ≈ 2.75%), whereas the larger 250 × 250 px banners yielded 144 clicks out of 4,519 impressions (CTR ≈ 3.21%). A 2 × 2 chi-square test for size indicated no statistically significant difference ($\chi^2 = 1.49$, $p = 0.22$).

Table 5. Aggregated 2 × 2 contingency table for Size (S200 × S250px) $\chi^2 = 1.49$, $p = 0.223$

Level	Clicks	Non-Clicks	Total
S200	139	4,892	5,031
S250	144	4,375	4,519
Total	**283**	**9,267**	**9,550**

Similarly, for emotional tone (tab. 6), antagonistic creatives produced 129 clicks out of 4,640 impressions (CTR = 2.78%), while neutral creatives produced 154 clicks out of 4,910 impressions (CTR = 3.14%), with the corresponding chi-square test again non-significant ($\chi^2 = 1.05$, $p = 0.30$).

Table 6. Aggregated 2 × 2 contingency table for emotional tone (Antagonistic vs. Neutral) with $\chi^2 = 1.05$, $p = 0.305$

Level	Clicks	Non-Clicks	Total
Antagonistic	129	4,511	4,640
Neutral	154	4,756	4,910
Total	**283**	**9,267**	**9,550**

Finally, for use of color (tab. 7), color versions achieved 146 clicks out of 4,868 impressions (CTR = 3.00%), and gray versions achieved 137 clicks out of 4,682 impressions (CTR = 2.93%), with a chi-square statistic of 0.04 and p = 0.83.

Table 7. Aggregated 2 × 2 contingency table for Disp (Color vs. Gray) with $\chi^2 = 0.044$, $p = 0.833$

Level	Clicks	Non-Clicks	Total
Color	146	4,722	4,868
Gray	137	4,545	4,682
Total	**283**	**9,267**	**9,550**

In line with the omnibus test, these factor-level analyses confirm that observed differences are small and not statistically significant.

As an exploratory step, we inspected whether the observed patterns were consistent across devices and over time (for example, comparing early versus late days of the campaign). Descriptive breakdowns by device, age and basic audience segments did not reveal any systematic divergence that would qualitatively change the conclusions drawn from the main analysis. Given the relatively short campaign duration and the limited sample size, these robustness checks are reported as auxiliary evidence only and should be interpreted with caution rather than as a basis for strong claims.

5. DISCUSSION

Overall, the experiment did not reveal clear and consistent advantages of antagonizing messages or club colors in terms of CTR. Differences between variants were small in absolute terms – below one percentage point within each size group – even though one pairwise comparison (between a color–antagonistic 200 × 200 px variant and a color–neutral 250 × 250 px variant) reached conventional significance levels. Banner size also revealed only modest differences in average CTR, rather than strong, systematic format effects. Taken together, the results suggest that creative changes of the type tested here do not generate large performance shifts in this setting, which is consistent with the general observation that measurable effects in display advertising field contexts are often small and difficult to isolate (Johnson, 2023; Broockman, Green, 2014).

Addressing the research questions, the results allow us to revisit each of the four research questions in light of the observed CTR patterns.

Regarding RQ1 (and H1), antagonistic creatives did not produce a systematically higher CTR than neutral creatives. The aggregated factor-level analysis (tab. 6) showed that antagonistic banners achieved a CTR of 2.78% compared to 3.14% for neutral banners ($\chi^2 = 1.05$, $p = 0.305$) – directionally opposite to the hypothesized effect, although the difference was not statistically significant. H1 is therefore not supported. Even in the only pairwise comparison reaching conventional significance (V1 vs. V6, $p = 0.043$), the higher CTR was observed for the neutral, not the antagonistic, creative – further reinforcing the conclusion that antagonistic advertising does not automatically translate into stronger click behavior in this setting.

Regarding RQ2 (and H2), the use of club-related shirt colors did not yield a meaningful CTR advantage over grayscale visuals. Color creatives achieved 3.00% CTR versus 2.93% for grayscale ones ($\chi^2 = 0.04$, $p = 0.833$), an essentially null result. H2 is not supported. This suggests that in a saturated rivalry market like Kraków, club-related visual cues may not function as a strong attention-driving signal in display banners, possibly because audiences are already heavily exposed to such cues across multiple media.

Regarding RQ3 (and H3), the larger 250 × 250 px format yielded a slightly higher average CTR (3.21%) than the smaller 200 × 200 px format (2.75%), consistent with the hypothesized direction. However, the difference did not reach statistical significance ($\chi^2 = 1.49$, $p = 0.223$), so H3 is only weakly and descriptively supported. The directional consistency suggests a possible advantage for larger formats, but the magnitude is too small to be considered conclusive evidence under the present sample size.

The practical significance of the size manipulation deserves additional reflection. The shift from 200 × 200 px to 250 × 250 px corresponds to a 25% increase in side length but a 56% increase in surface area (from 40,000 to 62,500 pixels). The perceived prominence of these formats, however, does not depend on pixel dimensions alone: the same banner occupies a different proportion of the screen, depending on the device's resolution, physical screen size, and pixel density. On a high-resolution desktop display, the banner occupies a small fraction of the available screen area, whereas on a smartphone the same banner may cover a substantial part of the visible page. Because mobile placements accounted for a considerable share of impressions in this campaign, the effect of banner size on user attention is likely to differ between mobile and desktop contexts. This heterogeneity in how size is actually experienced by users may partly explain why the descriptive CTR advantage of the larger format did not reach statistical significance, and it points to device-level analysis as a useful direction for future research.

Regarding RQ4, the exploratory analysis of interactions between size, color, and emotional tone did not reveal systematic patterns. The omnibus chi-square test across all eight variants was non-significant ($\chi^2 = 4.74$, $df = 7$, $p = 0.692$), and only one of 28 pairwise contrasts reached conventional significance (V1 vs. V6). Given the number of comparisons, this isolated result should be interpreted with caution. Substantively, however, the V1 vs. V6 contrast aligns with the observation that the worst-performing combination (color, antagonistic, small) and the best-performing combination (color, neutral, large) sit at opposite corners of the design space; this suggests that if any interaction matters, it may be the cumulative effect of multiple unfavorable choices rather than a single dominant factor.

These findings nuance prior work on negative and comparative appeals, which has been conducted primarily in controlled laboratory or survey-based designs (e.g., Kalro et al., 2017; Soscia et al., 2010) and has often reported significant creative effects. In a real football rivalry context, the audience may already be highly saturated with club-related and rivalry-related stimuli, and antagonizing messages may not stand out as strongly as theoretical expectations would suggest. Moreover, delivery on an automated display platform and the intrinsic variability of online attention may further dampen the observable impact of relatively subtle creative manipulations (Johnson, 2023). The results therefore highlight the importance of context, audience composition and platform dynamics when evaluating antagonistic advertising strategies.

6. CONCLUSION

This study reported a field experiment on antagonistic advertising in sports advertising that used a real Google Ads display campaign. Using a 2 × 2 × 2 factorial design, we manipulated banner size, use of color and emotional tone across eight creatives shown in the Kraków football rivalry context and compared their click-through rates. Overall, the results did not reveal clear and robust advantages of antagonizing messages or club colors. Differences in CTR between variants were small in absolute terms and the omnibus chi-square test across all creatives was non-significant, with only a single pairwise comparison reaching conventional significance. Larger 250 × 250 px banners achieved slightly higher CTR on average than 200 × 200 px banners, but the effect size was modest.

These findings suggest that theoretical expectations regarding the effectiveness of antagonizing messages in rivalry markets should be treated with caution. Context, audience composition and platform dynamics appear to constrain the impact of relatively subtle creative manipulations. Beyond its specific numerical results, the study illustrates how carefully designed field experiments on real ad platforms can inform both academic discussions and practical campaign decisions in sports advertising.

Theoretical Implications

From a theoretical perspective, the study suggests that social identity and negativity bias do not automatically translate into higher click behavior for antagonizing ads. Antagonistic advertising may interact with factors such as perceived legitimacy of the message, current fan sentiment, campaign clutter and users' short-term goals when browsing. For some segments, antagonistic messages may be engaging; for others, they may be ignored or even perceived as inappropriate. This is consistent with evidence that responses to negative messaging depend on tolerance of negativity and the conditions of processing (Muehling et al., 2018). In sports contexts, identification and rivalry cues can also affect sponsor-message processing in heterogeneous ways (Kwak, Pradhan, 2021; Nichols et al., 2016). Future research should examine these potential moderators more systematically, combining field experiments with richer measures of attitudes, emotions and fan typologies.

Practical Implications

The findings of this study translate into several practical implications for advertisers operating in sports rivalry contexts. Firstly, advertisers should test antagonizing messages locally rather than assuming they will automatically increase CTR, because creative concepts that seem intuitively powerful may produce only small performance differences in real campaigns, especially in display environments where effects are often modest (Johnson, 2023). Secondly, larger banner formats

may provide some incremental value, but their contribution should be validated against cost, available inventory, and creative constraints, rather than treated as a guaranteed improvement.

Finally, when using rival-related messaging, practitioners need to balance potential attention gains with brand risk and possible backlash among neutral or mixed audiences, particularly in highly polarized rivalry markets (Nichols et al., 2016; Kwak, Pradhan, 2021). For this reason, field experimental tests on real platforms, with clear experimental control and consistent targeting, should become a standard part of campaign planning in such contexts, allowing teams to verify creative assumptions with actual behavioral data (Johnson, 2023).

Limitations

Several limitations of this study should be acknowledged when interpreting the results.

Firstly, the experiment was conducted in a single city and rivalry context (the Kraków area) and for a single product category (an online football management game), so the findings may not generalize to other leagues, markets, or types of campaigns.

Secondly, the campaign ran for one week and generated a moderate number of impressions and clicks; while sufficient for basic comparisons, the study was not powered to detect very small effects or to support detailed subgroup analyses by device, age, or audience segment. This is commonly noted for field experiments in online advertising (Johnson, 2023).

Thirdly, CTR was the primary outcome measure, and we did not collect data on brand attitudes, recall, perceived entertainment value, or emotional responses, nor did we focus on downstream behavioral metrics such as registrations or purchases, so the conclusions pertain to click behavior only.

Fourthly, because the study relied on a live Google Ads environment, allocation of impressions remained partly influenced by the platform's delivery mechanisms; although we attempted to limit automated optimization and kept targeting constant across variants, residual platform effects cannot be fully ruled out (Johnson, 2023).

Finally, the creative manipulations were constrained to specific combinations of size, use of color, and antagonistic versus neutral messaging; consequently, the results may not extend to more extreme forms of antagonistic advertising, different visual styles, or other advertising formats such as video or social media placements.

Directions for Future Research

Building on the present findings, several avenues for future research appear particularly promising. Future studies should replicate this design in different cities, leagues, and rivalry contexts, since antagonizing messages may interact with local

fan cultures, rivalry intensity, and club histories; multi-site comparisons would therefore help assess the generalizability of the results (Mills et al., 2018; Sanford, Scott, 2016).

Subsequent experiments should also go beyond CTR and include richer outcome measures, such as attitudes, recall, perceived entertainment value, emotional responses, and downstream behavior (e.g., registrations, repeat visits, purchases). For example, Broockman and Green (2014) demonstrated the value of measuring name recognition and candidate evaluations alongside exposure data, which allowed them to distinguish between attention and persuasion.

Segmentation by fan typology deserves special attention because dedicated supporters, casual followers, and neutral users may react very differently to antagonizing messages, and aggregated CTR can mask these patterns; combining field experiments with survey-based profiling or panel data could help uncover heterogeneous effects across segments (Kwak, Pradhan, 2021; Nichols et al., 2016).

Further work should examine other platforms and formats, including video, social media placements, and in-app environments, where social dynamics and engagement norms differ from display advertising; cross-platform experiments could clarify whether the limited effects observed here are specific to display banners or more general (Johnson, 2023).

Methodologically, future research could integrate field experiments with more advanced modelling approaches and longitudinal designs, tracking how responses to antagonizing creatives evolve over longer periods and across multiple campaign waves (Johnson, 2023).

LITERATURE

Ahmadi, I., Abou Nabout, N., Skiera, B., Maleki, E., Fladenhofer, J. (2024). Overwhelming targeting options: Selecting audience segments for online advertising. *International Journal of Research in Marketing*, 41(1), doi: 10.1016/j.ijresmar.2023.08.004.

Alwreikat, A.A.M., Rjoub, H. (2020). Impact of mobile advertising wearout on consumer irritation, perceived intrusiveness, engagement and loyalty: A partial least squares structural equation modelling analysis. *South African Journal of Business Management*, 51(1), doi:10.4102/SAJBM.V51I1.2046.

Bambauer-Sachse, S., Heinzle, P. (2018). Comparative advertising for goods versus services: Effects of different types of product attributes through consumer reactance and activation on consumer response. *Journal of Retailing and Consumer Services*, 44, doi:10.1016/j.jretconser.2018.05.007.

Broockman, D.E., Green, D.P. (2014). Do online advertisements increase political candidates' name recognition or favorability? Evidence from randomized field experiments. *Political Behavior*, 36(2), doi:10.1007/s11109-013-9239-z.

Casais, B., Proença, J.F. (2015). A model to classify television social advertisements according to the use of positive or negative appeals. In: *Advances in Advertising Research*, vol. V, doi:10.1007/978-3-658-08132-4_2.

Chen, Q., Feng, Y., Liu, L., Tian, X. (2019). Understanding consumers' reactance of online personalized advertising: A new scheme of rational choice from a perspective of negative effects. *International Journal of Information Management*, 44, doi:10.1016/j.ijinfomgt.2018.09.001.

Clowes, J., Tapp, A. (2003). Looking through the hourglass of fan segmentation: Research findings and marketing implications for live spectator sports. *International Journal of Sports Marketing and Sponsorship*, 5(1), doi:10.1108/ijsms-05-01-2003-b004.

Demirdjian, Z.S. (2015). Attitudinal and behavioral response to comparative advertising: An experimental field analysis. In: *Developments in Marketing Science: Proceedings of the Academy of Marketing Science*, doi:10.1007/978-3-319-11101-8_61.

Farbaniec, S., Kozioł, L. (2019). Wizerunek marki klubu przeciwnika i postawa wobec jego sponsora wśród kibiców Wisły Kraków i MKS Cracovia. In: *Sportowe zarządzanie: młodzi o sporcie 2019*.

Filiopoulou, D., Rigou, M., Faliagka, E. (2019). Display ads effectiveness: An eye tracking investigation. In: *Advances in E-Business Research Series*, doi:10.4018/978-1-5225-7262-6.ch012.

Hsu, T. (2020). Investigating the effectiveness of comparative advertising: The roles of structural alignability and message informativeness. *Journal of Marketing Development and Competitiveness*, 14(5), 20-27, doi:10.33423/jmdc.v14i5.3982.

Johnson, G.A. (2023). Inferno: A guide to field experiments in online display advertising. *Journal of Economics and Management Strategy*, 32(3), doi:10.1111/jems.12513.

Kalro, A.D., Sivakumaran, B., Marathe, R.R. (2017). The ad format – strategy effect on comparative advertising effectiveness. *European Journal of Marketing*, 51(7-8), doi:10.1108/EJM-11-2015-0764.

Krauthausen, P., Leitner, M., Ristea, A., Curtis, A. (2019). Collecting and analyzing soccer-related graffiti with the spatial video technology and GIS: A case study in Krakow, Poland. *Urban Development Issues*, 62(1), doi:10.2478/udi-2019-0010.

Kwak, D.H., Pradhan, S. (2021). "If you ain't first, you're last!" Understanding identity threat, team identification, and advertisement messages when your favorite team loses. *Journal of Sport Management*, 35(2), doi:10.1123/JSM.2019-0445.

Le, T.D., Vo, H. (2017). Consumer attitude towards website advertising formats: A comparative study of banner, pop-up & in-line display advertisements. *International Journal of Internet Marketing and Advertising*, 11(3), doi:10.1504/IJIMA.2017.085654.

Lu, Z., Kahle, L.R., Lee, S.M., Lee, S.-Y. (2012). Football fans' contrasting motivations: China, S. Korea, and the USA. *Asia Pacific Journal of Innovation and Entrepreneurship*, 6(1).

Mills, B.M., Tainsky, S., Green, B.C., Leopkey, B. (2018). The ultimatum game in the college football rivalry context. *Journal of Sport Management*, 32(1), doi:10.1123/jsm.2017-0094.

Muehling, D.D., Vijayalakshmi, A., Laczniak, R.N. (2018). The effects of tolerance of negativity on consumers' responses to comparative attack advertising. *Journal of Marketing Communications*, 24(7), doi:10.1080/13527266.2016.1216462.

Nichols, B.S., Cobbs, J., Raska, D. (2016). Rival team influence on team identification and cause-related sports marketing. In: *Developments in Marketing Science: Proceedings of the Academy of Marketing Science*, doi:10.1007/978-3-319-11815-4_22.

Osinga, E.C., Zevenbergen, M., van Zuijlen, M.W.G. (2019). Do mobile banner ads increase sales? Yes, in the offline channel. *International Journal of Research in Marketing*, 36(3), doi:10.1016/j.ijresmar.2019.02.001.

Pinkleton, B.E. (1997). The effects of negative comparative political advertising on candidate evaluations and advertising evaluations: An exploration. *Journal of Advertising*, 26(1), 19-29, doi:10.1080/00913367.1997.10673515.

Sanford, K., Scott, F. (2016). Assessing the intensity of sports rivalries using data from secondary market transactions. *Journal of Sports Economics*, 17(2), doi:10.1177/1527002514527112.

Soscia, I., Girolamo, S., Busacca, B. (2010). The effect of comparative advertising on consumer perceptions: Similarity or differentiation? *Journal of Business and Psychology*, 25(1). doi:10.1007/s10869-009-9130-4.

Starcevic, S. (2006). Possibilities for application of comparative advertising and its effects. *Economic Annals*, 51(171), doi:10.2298/eka0671090s.

Wattenberg, M.P., Brians, C.L. (1999). Negative campaign advertising: Demobilizer or mobilizer? *American Political Science Review*, *93*(4), 891-899, doi:10.2307/2586117.

Yousef, M., Dietrich, T., Rundle-Thiele, S. (2021). Social advertising effectiveness in driving action: A study of positive, negative and coactive appeals on social media. *International Journal of Environmental Research and Public Health*, 18(11), 5954, doi:10.3390/ijerph18115954.

Yousef, M., Rundle-Thiele, S., Dietrich, T. (2023). Advertising appeals effectiveness: A systematic literature review. *Health Promotion International*, 38(4), doi:10.1093/heapro/daab204.

WPŁYW ANTAGONIZACJI NA POSTRZEGANIE MATERIAŁÓW REKLAMOWYCH W RYWALIZACJACH PIŁKARSKICH: EKSPERYMENT POLOWY W GOOGLE ADS

Streszczenie

Rywalizacje piłkarskie tworzą silnie emocjonalny kontekst, w którym reklamy antagonizujące lub porównawcze mają zwiększać uwagę i zaangażowanie. Strategie te uznaje się za kluczowe dla marek chcących wyróżnić się na nasyconych rynkach poprzez odwołanie do głęboko zakorzenionej tożsamości konsumentów. Mimo debaty nad ryzykiem takie reklamy są powszechnie stosowane w przekonaniu o ich wyższej skuteczności względem form tradycyjnych. Nasze badanie sprawdza, czy kreatywne reklamy antagonizujące realnie zwiększają liczbę kliknięć w rzeczywistych warunkach. Przeprowadziliśmy pełnoczynnikowy eksperyment polowy 2 × 2 × 2 w sieci reklamowej Google Ads, skierowany na obszar w promieniu 30 km od Krakowa. Osiem wariantów banerów różniło się tonem emocjonalnym (antagonizujący vs. neutralny), kolorystyką (barwy klubowe vs. skala szarości) oraz formatem (200 × 200 px vs. 250 × 250 px). Główną zmienną zależną był współczynnik klikalności (CTR). Kampania wygenerowała 9550 wyświetleń i 283 kliknięcia (całkowity CTR = 2,96%). Analiza statystyczna testami nieparametrycznymi wykazała, że różnice między czynnikami były nieistotne, z wyjątkiem jednego porównania na poziomach

czynników. W przeciwieństwie do wielu wcześniejszych badań nie udało się udowodnić systematycznej przewagi komunikatów antagonizujących ani barw klubowych. Wyniki te wskazują na potrzebę lokalnych testów i sugerują ostrożność w zakładaniu, że antagonizacja w reklamie sportowej zawsze poprawia wyniki.

Słowa kluczowe: reklama, antagonizacja, reklamy porównawcze, rywalizacja piłkarska, Google Ads, CTR (współczynnik klikalności), eksperymenty w polu