Summer 6-15-2016

READ THIS PAPER! A FIELD EXPERIMENT ON THE ROLE OF A CALL-TO-ACTION IN PAID SEARCH

Darius Schlangenotto
University of Paderborn, darius.schlangenotto@upb.de

Dennis Kundisch
University of Paderborn, dennis.kundisch@upb.de

READ THIS PAPER! A FIELD EXPERIMENT ON THE ROLE OF A CALL-TO-ACTION IN PAID SEARCH

Research

Schlangenotto, Darius, University of Paderborn, Paderborn, Germany, darius.schlangenotto@upb.de

Kundisch, Dennis, University of Paderborn, Paderborn, Germany, dennis.kundisch@upb.de

Abstract

Current research on ad copy design in the context of paid search highlights in particular the integration of an explicit call-to-action in the last part of the ad. Research on this topic is in its infancy and, so far, has only investigated click-through-rates rather than conversions. This paper considers conversion rates in ad copy design by examining the impact of a call-to-action on user behavior, conducted in a field experiment. Contrary to prevalent advice in the industry, our research shows that a call-to-action does not necessarily enhance paid search performance. Using logistic regression we identify a phrasing scheme which minimizes costs (clicks) while simultaneously maximizing profits (conversions) for advertisers. The diametric user behavior we observe provides first indicative evidence of a self-selection mechanism at play when paid search users respond to differently phrased ad copies. Future research in the field of paid search with regard to ad copy design should be cognizant of our findings.

Keywords: paid search, ad copy design, call-to-action, field experiment, logistic regression

1 Introduction

This paper presents our study of the effect of different ad copy designs in the context of search advertising, or paid search. Today, paid search – the mechanism of placing ads in response to user search queries on search engine result pages (SERP) – is one of the main sources of internet advertising revenue (GroupM, 2014). In 2014 paid search accounted for 48% (US$82 bn) of global internet advertising revenue and is expected to grow by 10% annually over the next four years (eMarketer, 2015).

As an emerging technology, paid search opened up numerous new avenues for research (Dhar and Ghose, 2010). Its key features, such as the pricing mechanism and the phrasing of ads in response to users' search intents, have generated numerous studies, especially in the fields of Information Systems and Marketing (Rutz and Bucklin, 2013). Most of the research in this area has focused on the generalized second-price auction (Varian, 2007) applied to assign ad slots on a SERP and its effect on bidding behavior (Edelman and Ostrovsky, 2007), or on success determinants of paid search campaigns, e.g. keyword characteristics, position of the advertisement, and the quality of the landing page (Ghose and Yang, 2009). One key aspect of every paid search campaign that has received far less scholarly attention, so far, has been ad copy design itself. (Rutz and Trusov, 2011) note that while ad copy design is widely discussed in practice, research on the effects of its design remains in its infancy. Typically, the design of paid search ads incorporates three textual elements: a title, a URL, and a description. As depicted in figure 1 the appearance of an ad changes, depending on its positioning on a SERP. In addition, the ad cannot be enhanced with graphics or other visual elements. Search engines place even further restrictions on design options by enforcing writing standards and limiting the number of characters.

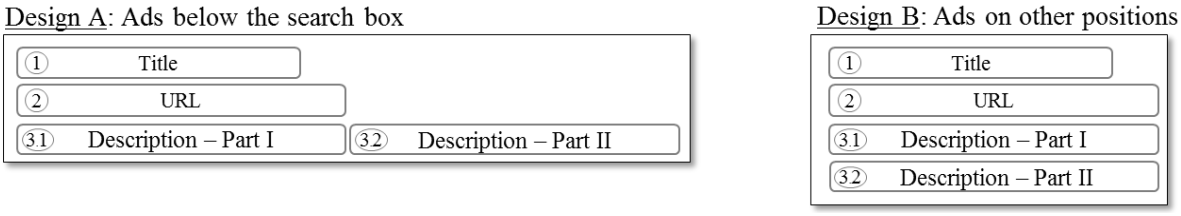


Figure 1. Ad Copy Designs on a SERP

From an advertiser's perspective, a well-crafted ad is one success-determinant amongst others (Animesh et al., 2011) to reach the target audience and convince potential buyers to click the ad. The effectiveness of ads is measured in terms of click-through-rates (CTR) and conversion-rates (CVR). CTR is defined as the percentage of users who actually click on the ad out of the total number of user who were exposed to it (impressions). In some cases, marketers aim to engage users in a specific action (conversion) after clicking the ad, such as buying a product, or enter a lottery. CVR is defined as the number of users who engage in a desired action out of the total number of people who clicked the ad. To improve ad effectiveness a general advice which is often presented as a de-facto standard is to include a call-to-action in the ad (Google, 2015; Bing, 2015). A recent study by (Atkinson et al., 2014) suggests that while a call-to-action only enhances CTR when placed in the last part of the description (see figure 1, 3.2) it decreases CTR if it is placed in the first part of the description (see figure 1, 3.1). Up to now, research studies focusing on the benefits of incorporating a call-to-action in an ad copy have limited their analyses to differences in click-behavior. (Rutz and Trusov, 2011) argue that a conversion decision is driven by the search intent of the user and the offer itself. In their analysis the ad copy design is not considered as an influential factor in regard to CVR. This might be problematic due to the findings by (Haans et al., 2013) who conclude that different ad copy description texts (see figure 1, 3.1 & 3.2) entail different CVRs. We argue that the use of different wordings, and positioning those differently in the ad, might influence ad perception not only in terms of CTR but also the associated user behavior in terms of CVR. A call-to-action might particularly impact CVR due to its conversion-oriented phrasing. Therefore we would like to augment call-to-action research by investigating the following research question: *What is the impact of using a call-to-action in an ad copy in the context of paid search on user behavior?*

In order to shed light on the question of how a call-to-action affects ad-perception as well as user behavior we teamed up with a well-known mid-sized bricks and mortar business-to-consumer (b2c) furniture retailer operating in Germany. As suggested by (Sudhir, 2016) we made use of the benefits of experimental design and conducted a field experiment resulting in 67,711 observations. Using logistic regression while controlling for potential confounding factors we reason that a call-to-action does not necessarily impact ad perception and user behavior in general. Comparing the position effect of a call-to-action reveals that users show diametric behavior in terms of their ad perception and conversion behavior. When a call-to-action is placed in the last part of the ad it is 7% more likely that a user clicks the ad than if the call-to-action is placed in the first part of the description. On a conversion level this behavior is inverted. Users who clicked on an ad with a call-to-action in the last part are 15% less likely to engage in the desired action after all. Our work has several implications for theory and practice. First, we find evidence that ad copy phrasing schemes can affect ad perception as well as conversion behavior. Second, we report an ad copy phrasing which has the capability to reduce advertising costs by simultaneously maximizing conversions. Third, the observed diametric user behavior opens up avenues for further research to determine the root causes of this behavior.

The paper is structured as follows: Section 2 provides an overview of the relevant streams of literature. Section 3 presents the hypotheses. Section 4 introduces the research setup and section 5 analyzes the results. Finally, conclusions and implications are set out in section 6.

2 Related Literature

Ad copy design in the form of offline ads in newspapers or on billboards, for example, is a well-studied field of research. However, researchers only recently started to investigate the success-factors of ad copy designs for paid search as one of the leading forms of online advertising. Three lines of research are relevant to our study: The positioning of the ad on the SERP, the relationship between keywords and ad copy design, and the structure and wording of the ad itself.

Ad positioning in the context of ad copy design

On a SERP a multitude of ads might be presented to the user. The number of shown ads depends on the quantity of advertisers who bought a keyword that matches the search query. Google Search, for example, presents up to three ads in the most prominent slots directly below the search query. In addition, up to eight other ads might be placed on less prominent slots on the side or at the end of the SERP. Many scholars (e.g., Jansen et al., 2013) point out that CTR depends on the visual placement of the ad and increases for more prominent slots. Yet, those position effects seem to be weaker for smaller firms and more specific search queries (Narayanan and Kalyanam, 2015). A field experiment conducted by (Animesh et al., 2011) studies the relationship between ad copy designs and different ad positions. The authors conclude that crafting an ad with a unique selling proposition is not sufficient to affect click behavior, whereas an ad copy that differentiates the firm from others is moderated by its visual placement on a SERP. The researchers argue that in respect of the ad position, different types of customers are likely to click on the ad, and that, therefore, the ad position parameter lends itself to segmenting customers into groups. The idea of customer segmentation based on ad positions is in line with findings from (Rutz and Trusov, 2011) who suggest that the advertised offer should be changed in respect of its position on the SERP. Many scholars (e.g., Ghose and Yang, 2009) argue that the ad position impacts CTR as well as CVR. (Rutz et al., 2012) present empirical evidence that CVR decreases for lower positions. These findings stand in direct contrast with the results of (Agarwal et al., 2011) who also conclude that CVR depends on rank but increases for lower ad ranks. Conceptually speaking, the current body of knowledge suggests that it may be necessary to differentiate between the intentions of users who reach the website via a top-positioned ad in comparison to those who were acquired via an ad which was placed less prominently.

Keywords in the context of ad copy design

In the context of paid search, advertisers have to define keywords for which they want to be listed on the SERP. Whenever a user enters a search query the search engine will link this query to contextually matching keywords, and display ads of marketers who bought those keywords. (Rutz and Bucklin, 2011) group the search-phase keywords by distinguishing between generic terms (e.g., "buy furniture") and branded terms (e.g., "buy furniture of company X"). Finding evidence for spillover effects from generic keyword requests to branded keyword queries, they conclude that this indicates a systematic difference between search engine user characteristics in terms of the wording of their query. For example, if a user incorporates the brand name of the company in their search it is obvious that they are already familiar with the brand and might have formed specific associations towards the brand and/or its products. This user heterogeneity in relation to keyword characteristics is supported by (Nottorf and Funk, 2013; Lu and Zhao, 2014) who empirically show that users act differently in terms of CTR and CVR relative to the keyword characteristics entered. (Jansen et al., 2011) particularly focus on the interrelation between keywords and the ad copy. Their correlational analysis reveals that brand keywords in combination with a branded ad copy design is the major driver of sales revenue. Conceptually speaking the current body of knowledge suggests that keyword characteristics might be a pivotal determinant of user behavior.

Ad copy design

In a first exploratory study on ad copy design (Turnbull and Bright, 2008) state that ad copies should not incorporate questions in the title (see figure 1). Their analysis performed on different paid search campaigns reveals that using a question reduces CTR significantly compared with a statement-based phrasing. (Yoo, 2011) performed an experiment in order to shed light on the question of how messages should be framed in the context of paid search in general in order to enhance clicks. The study differentiates messages in terms of a customer's low or high level of involvement as the degree of interest of a customer in a product category or brand. They find that a positive framing of ad proposals increases CTR in the context of low-involvement products or brands by 25% in comparison to a negative framing. For high-involvement products or brands a negative framing is more effective compared to a positive one and increases CTR by 20%. In a field experiment in cooperation with a b2c retailer in the Netherlands (Haans et al., 2013) investigate the influence of different description texts (see figure 1, 3.1 & 3.2) in an ad copy. Amongst other aspects, the scholars investigate the effectiveness of alternative descriptions (e.g., causal evidence, expert advice, statistical evidence) in terms of conversions and report that for the tested description types the CTR is not suited as an approximation of CVR. An effect description (causal evidence), for instance, is associated with the lowest CTR but leads to the highest CVR. (Rutz and Trusov, 2011) apply a Bayesian framework to implement a two-stage consumer model. The model is used to analyze paid search ads placed by the mobile ringtone industry. The study distinguishes between different ad components and analyzes their effect on CTR, accounting for consumer heterogeneity. The authors conclude that a low ad density is favorable. A reduction of just one word from the title increases CTR by 2%, whilst reducing the description by one word increases CTR by 4%. Furthermore this study concludes that a call-to-action increases CTR by 33% as opposed to ads where a call-to-action is not included. Based on these findings the authors advise marketers to incorporate attention-grabbing content in their ads. (Atkinson et al., 2014) refine the findings from (Rutz and Trusov, 2011) by performing an odds ratio analysis and conclude that a call-to-action increases clicks on a paid search ad only when placed in the last part of the description (see figure 1, 3.2). Their analysis is based on secondary data of a b2c automotive retailer chain based in Australia. (Atkinson et al., 2014) measure ad effectiveness in terms of CTR without differentiating between distinct types of keywords.

Our work adds to the current body of knowledge on ad copy design by analyzing the impact of a call-to-action – in various positions – on user behavior. Previous studies on ad positioning and on keyword characteristics have stressed the need to account for two important factors, namely keyword heterogeneity and the visual placement of ads on a SERP. These are, therefore, the factors we account for in our study.

3 Hypotheses Development

(Atkinson et al., 2014) link their finding of a call-to-action enhancing clicks in the last part of the description while it decreases clicks in the first part to the tendency of people following an Attention-Interest-Desire-Action (AIDA) cognition scheme. The AIDA mechanism dates back to 1898 where St. Elmo Lewis proposed a hierarchical framework on how people react to ads (Strong, 1925). Today, this model is often used to assess the attractiveness of ad designs (Hofacker and Murphy, 1998; Rutz and Bucklin, 2013). The AIDA cognition scheme can also be transferred to the detailed structure and phrasing of ad copies in paid search. (Rutz and Bucklin, 2013) note that the main purpose of a headline is to draw attention and generate interest, while the main function of the body of the ad is to create desire for the product. Following the AIDA concept and in line with the findings of (Atkinson et al., 2014) we hypothesize that people who are exposed to an ad are most likely attracted by a phrasing which matches their cognitive process. Therefore, it should be important to them that the phrase which motivates the desired action is shown as the last part of the ad. The strongest claim to motivate an action is a direct call-to-action. But an action could also be motivated using a neutral statement which highlights the benefits of the action. Research indicates that a call-to-action should positively impact CTR (e.g., Rutz and Trusov, 2011; Atkinson et al., 2014). Hence, we have the first hypothesis:

> **H1**: A direct call-to-action in the last part of the ad increases CTR significantly in comparison to a neutrally phrased statement which only focusses on the benefits of the action.

Furthermore, are the users who were captured by a direct call-to-action more likely to actually engage in the desired action after clicking on the ad? These users are after all directed towards the desired action. Up to now no research has been conducted that assesses user-behavior on web pages involving a call-to-action phrase. It is known, though, that CTR and CVR might differ substantially depending on the textual elements used in the description (Haans et al., 2013). We argue that a more goal oriented wording with a call-to-action might positively impact CVR and formulate our second hypothesis:

> **H2**: Users who were engaged by a direct call-to-action in the last part of the ad show an increased CVR compared to users who were engaged by a neutrally phrased statement.

When search engine users are unable to process information in accordance with the AIDA scheme, i.e., the AIDA scheme is mixed up by using the call-to-action in the middle of the ad, they should be less likely to click on the ad. Therefore we propose a negative impact of a mixed up scheme and formulate our third hypothesis:

> **H3**: Phrasing an ad which incorporates a call-to-action in the last part of the ad increases CTR significantly in comparison to an ad in which the call-to-action is placed in the middle of the ad copy.

Furthermore, are users who were attracted by the mixed up advertising scheme different from those who follow the established AIDA principle? We argue that users who were exposed to a mixed up AIDA scheme might not even be aware of the call-to-action due to its less prominent position in the ad copy. These users should show a decreased CVR as opposed to those where the call-to-action is placed more prominently in the last part of the description. Hence, we can formulate our fourth hypothesis:

> **H4**: Users who were engaged by a call-to-action in the last part of the ad show an increased CVR as opposed to users who were engaged by the same textual elements, but where the call-to-action was placed in the middle of the ad copy.

4 Research Setup

To test the four hypotheses we teamed up with a mid-sized bricks'n'mortar b2c furniture retailer with several stores in Germany. The company is a well-known local brand and has requested to remain anonymous. Up to two million customers visit their stores each year. Over a time span of 64 days users who searched furniture-related keywords using Google were randomly exposed to different ad copies. The specific conversion goal of our campaign was to turn paid search website visitors into lottery participants. To test for the call-to-action effect as regards to wording (hypothesis H1 and H2) and positioning within the ad (hypothesis H3 and H4), three ad variants were crafted. Table 2 depicts the variously changed elements of the ad copy design, used in relation to our hypotheses. The neutrally phrased statement in the last part of the ad (Neutral+) is compared to a call-to-action wording placed in the last part of the ad (CTA+). The position effect of a call-to-action statement is evaluated comparing CTA+ to a similar worded ad which incorporates the call-to-action in the middle position (CTA-). As explained above, three lines of research are relevant to our study: ad copy design, the positioning of the ad on the SERP and the relationship between keywords and ad copy design. All those aspects were incorporated in our experimental design in order to distinguish between differential effects which might trigger different user behaviors.

		Used to test Hypothesis				
Ad Copy Design: Description – Part I & II		H1	H2	H3	H4	**Referred to as**
31 Grand opening lottery	32 Win prices up to 4000€	X	X			Neutral+
31 Grand opening lottery	32 Participate now! Win 4000€	X	X	X	X	CTA+
31 Participate now! Win 4000€	32 Grand opening lottery			X	X	CTA-

Table 2. *V*ariously changed elements of the ad copy design

Ad copy design

To prevent potentially confounding effects driven by improper ad contents, current findings on effective ad copy design were incorporated in all ad variants. According to (Yoo, 2011) we positively framed the general ad copy text. Following (Atkinson et al., 2014) the ad description contained the offer and a phrase to motivate the action. In line with (Jerath et al., 2011) the ad copy contained the retailer's brand name.

Ad positioning in the context of ad copy design

As argued by many scholars (e.g., Animesh et al., 2011) the ad position is considered to be one of the main success determinants of paid search campaigns. Due its importance ad position could be considered as a control variable. However, controlling for ad position effects might not be sufficient in our research environment in which the appearance of an ad changes in respect of its placement. This dynamic change could systematically bias the results. To prevent potential confounding effects driven either by consumer heterogeneity in dependence of ad position (e.g., Agarwal et al., 2011) or by dynamic changes in ad appearance (see, figure 1), our study focuses on the most prominent ad positions (slots 1-3) on a SERP which is illustrated in Design A (see, figure 1).

Keywords in the context of ad copy design

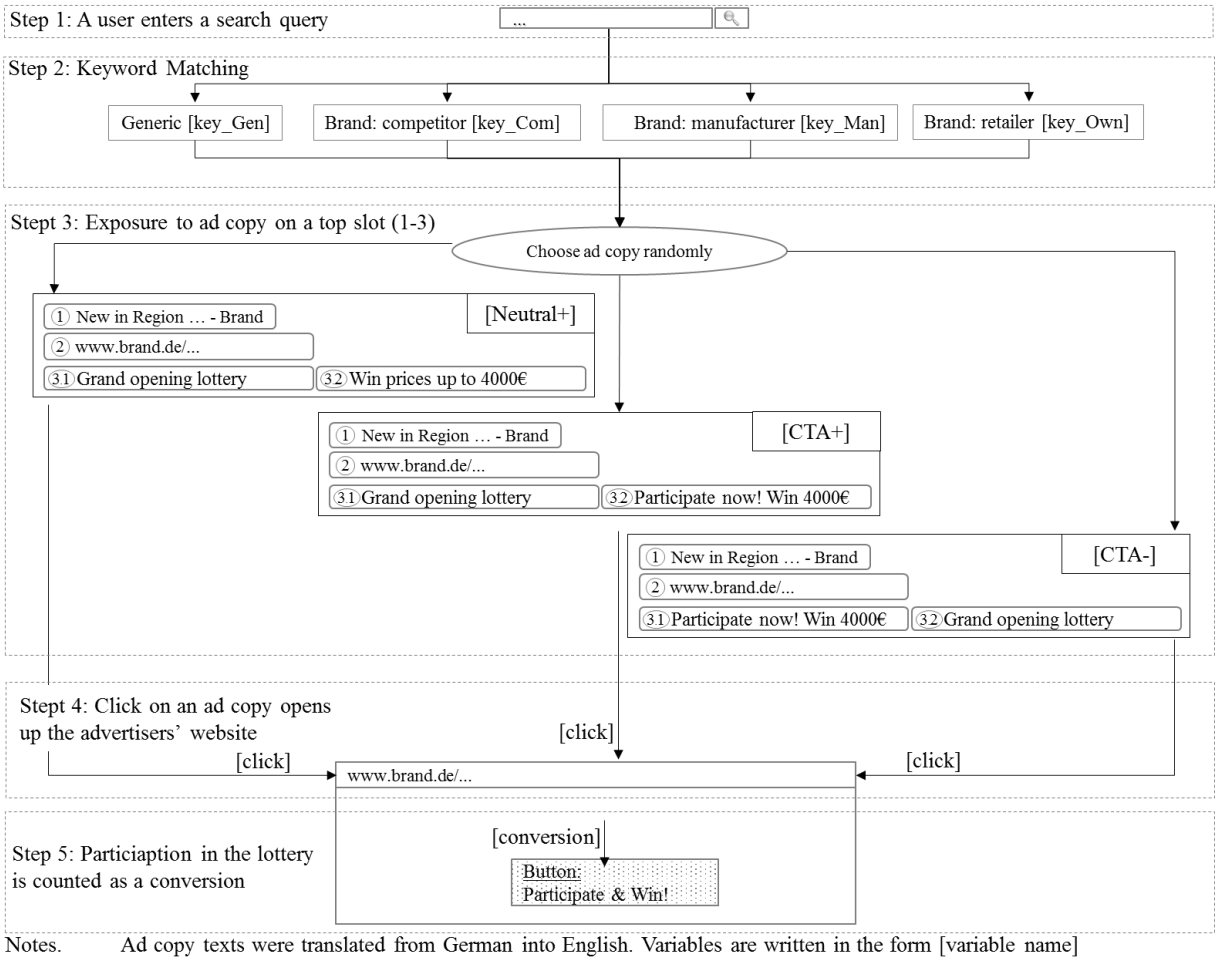
As highlighted by a number of scholars (e.g., Ghose and Yang, 2009; Rutz and Bucklin, 2011), search engine users show different behaviors in terms of CTR and CVR, dependent on their query. To account for potential keyword effects all keywords used in the experiment are grouped into four distinct clusters. All Clusters were designed using segmentations used by other scholars. In their work (Rutz and Bucklin, 2011) differentiate keywords by its brand information and compare generic to branded terms. (Ghose and Yang, 2009) focus on branded queries and differentiate keywords further into manufacturer and retailer brands. In addition, recent studies focus on the impact of firms who place their own ads in respondence to keywords of competitor brands on market outcome (e.g., Desai et al., 2014; Sayedi et al.,

2014). In line with those researchers we differentiate keywords into generic and three brand specific groups to account for user heterogeneity driven by keyword characteristics. Table 1 lists the distinct groups and the number of keywords assigned to each group.

Type	Example	# distinct keywords	Referred to as
Generic	"buy furniture"	436	key_Gen
Brand: competitor	competitor name	10	key_Com
Brand: manufacturer	"buy furniture of Rolf Benz"	39	key_Man
Brand: retailer	retailer name	2	key_Own

Table 1. Keyword characteristics

Focusing on top positioned ads we aim to analyze the wording and position effect of a call-to-action on user behavior in terms of clicks and conversions using three ad variants, while controlling for keyword heterogeneity. Hence, we have our overall research setup which is depicted in figure 2. During the execution of the experiment all aspects which could be influenced by the researcher and thus possibly biasing the results, were held constant. The keywords (see figure 2, Step 2) were not altered, the three different ad copy designs (see figure 2, Step 3&4) and the website with all its components remained unchanged (see figure 2, Step 5). Due to randomization, an individual's likelihood of assignment to a condition is equal for all three conditions – i.e., the likelihood of being exposed to Neutral+ is 0.33, the likelihood of being exposed to CTA+ is 0.33, and the likelihood of being exposed to CTA- is also 0.33 when searching for furniture related topics using Google.



Notes. Ad copy texts were translated from German into English. Variables are written in the form [variable name]

Figure 2. Research Setup

5 Empirical Analysis

Throughout the experiment the three ad copies were presented nearly 70,000 times in a top position directly below the search box. Each design was randomly delivered to search engine users via Google's automatic experimentation mechanism applying an equal split. Having applied the Wilcoxon rank sum test confirmed that, when we analyzed group differences in terms of impressions per day, the equal split was indeed performed appropriately. Impression differences between Neutral+ and CTA+ are non-significant with p=0.96 and also non-significant for the groups of CTA- and CTA+ with p=0.99. Two percent of all observations were assigned to the keyword group of manufacturers' specific keywords. Those observations had to be excluded from the dataset due to an insufficient number of clicks (69) and an inadequate number of conversions (2). Apparently users searching for manufacturer brands were not interested in a lottery of a b2c-retailer. Excluding manufacturers' specific keywords the resulting dataset contains 67,711 observations. During the experiment more than 8,000 users clicked the ads which led to a CTR of 12.2%. Of all those users more than 1,000 participated in the lottery and were therefore counted as conversions which lead to a CVR of 12.5%. In line with other researchers who investigated ad copy design (e.g., Atkinson et al., 2014; Haans et al., 2013) the observed performance metrics indicate a well-performing ad campaign overall. The observed CTRs and CVRs can be seen as an indication that the advertised lottery is well suited to catch the attention of potential customers. Table 3 displays descriptive statistics grouped in accordance to ad copy and keyword characteristics.

Ad Copy	Impressions	Clicks	Conversions	CTR	CVR
Neutral+	**22,367**	**2,849**	**345**	**12.7%**	**12.1%**
key_Gen	17,743	1,230	118	6.9%	9.6%
key_Com	1,343	86	10	6.4%	11.6%
key_Own	3,281	1,533	217	46.7%	14.2%
CTA+	**22,685**	**2,805**	**328**	**12.4%**	**11.7%**
key_Gen	17,929	1,253	126	7.0%	10.1%
key_Com	1,455	87	14	6.0%	16.1%
key_Own	3,301	1,465	188	44.4%	12.8%
CTA-	**22,659**	**2,627**	**362**	**11.6%**	**13.8%**
key_Gen	18,076	1,188	145	6.6%	12.2%
key_Com	1,400	78	14	5.6%	17.9%
key_Own	3,183	1,361	203	42.8%	14.9%
Sum	67,711	8,281	1,035	12.2%	12.5%

Table 3. Paid Search Performance Metrics

The descriptive statistic highlights that throughout the three different ad copies retailer specific keywords (key_Own) are an influential keyword group which accounts for 53% of all clicks and is associated with CTRs of above 40% for each ad copy. In regard to conversions retailer specific keywords do not deviate substantially from other keyword groups and are associated with an average CVR of approximately 14%. These high CTRs, coupled with average CVRs, might be seen as an indication that users searching for the retailer use these ads as a navigational shortcut to the website in general. When excluding retailer specific keywords, CTR roughly halves to 6.8% as opposed to CVR which only decreases by one-tenth, to 10.9%. This observation is in line with a recent study by (Blake et al., 2015) who report that users tend to click on ads as navigational shortcuts.

In regard to ad copies a neutral statement in the last part of the ad (Neutral+) is associated with the highest CTR of 12.7% and a CVR of 12.1%. On the descriptive level the performance metrics of a call-to-action in the last part of the ad (CTA+) do not differ substantially from Neutral+ and are associated

with a CTR of 12.4% and a CVR of 11.7%. However, metrics for a call-to-action in the midst position (CTA-) differ more pronounced. CTA- is associated with the lowest CTR (11.6%) but yields to the highest CVR (13.8%) overall. Taking keyword characteristics into account the observed behavior persists. For each keyword group CTA- is associated with the lowest number of clicks and yields to the highest number of conversions. To analyze those findings further and make predictive claims regression analysis is used.

5.1 Model

In our research setup the dependent variables of interest are clicks and conversions which are by its nature dichotomous distributed. When exposed to an ad a user has two distinct options, either clicking the ad (click = 1) or refusing to click (click = 0). When a user clicked the ad he will be exposed to the webpage which offers the opportunity to participate in the lottery (i.e., the action). Again, the user has two distinct options, either participating in the lottery (conversion = 1) or refusing to take part (conversion = 0). Accordingly, to test for the effect of ad copy design while accounting for possible keyword effects on those binary outcome variables we use multiple binary logistic regression. By doing so it is possible to distinguish between effects that are driven by the ad copy and effects which are caused by keyword characteristics. Thus, we consider the following model in latent variable form (Wooldridge, 2015) as our main model:

$$Y^* = \beta_0 + \beta_1 AdCopy_i + \beta_2 Keyword_j + \varepsilon, \qquad (1)$$
$$Y = 1[Y^* > 0].$$

where Y equals one when a user clicked the ad (or converted, when conversion is used as dependent variable). $AdCopy_i$ is coded as the independent ad copy design i (Neutral+, CTA+, CTA-), $Keyword_j$ incorporates the keyword characteristics j (key_Gen, key_Com, key_Own), and ε is the random error term.

5.2 Main Results

Our first set of hypotheses test for the wording effect of an ad copy in accordance to AIDA using a neutral statement (Neutral+) in comparison to a call-to-action (CTA+) placed at the end of the ad. Column (1) and (4) of table 4 present the estimates using either clicks or conversions as dependent variables. The coefficients show that for clicks as well as conversions user behavior is primarily driven by retailer specific keywords. The ad copy effect is associated with coefficients close to zero which are not significant. Accordingly, hypotheses H1 and H2 are rejected. Apparently, a call-to-action in the last part of the ad does not necessarily impact user behavior in comparison to a neutrally phrased statement.

Our second set of hypotheses is used to test for the position effect of a call-to-action. Using a fixed set of textual elements the position effect is isolated and analyzed comparing an ad for which the call-to-action is placed in the first part of the description (CTA-) to an ad where the call-to-action is placed in the last part (CTA+). Column (1) of table 5 lists the coefficients using clicks as dependent variable. Consistent with Hypothesis H3 and in line with the findings of (Atkinson et al., 2014) a call-to-action in the last part of the ad affects clicks significantly positive. In addition, retailer and competitor keywords do also positively impact click behavior. Column (4) of table 5 lists the coefficients using conversions as dependent variable. As stated in hypothesis H4 a positive influence of CTA+ on conversions is expected. However, the CTA+ coefficient is significantly negative with a value of -0.165. Accordingly, an AIDA phrasing scheme proved only beneficial in terms of click behavior. On a conversion level the mixed up scheme of CTA- outperforms it while controlling for possible keyword effects. Those findings are in line with (Haans et al., 2013) experimental results who also conclude that in their experiment on ad copy design one of the highest CTRs is associated with the lowest CVR. Hence, at least in our case the commonly used performance metric of CTR cannot be viewed as a suited proxy to determine paid search success and predict user behavior in regard to different ad copy designs.

	Dependent variable:					
	click			conversion		
Model (Column)	Main (1)	Interaction (2)	Reduced (3)	Main (4)	Interaction (5)	Reduced (6)
CTA+	**-0.035**	**0.010**	**0.005**	**-0.015**	**0.119**	**0.117**
	(0.031)	**(0.042)**	**(0.040)**	**(0.083)**	**(0.137)**	**(0.132)**
key_Comp	-0.124	-0.082	-0.125	0.277	0.290	0.277
	(0.081)	(0.115)	(0.081)	(0.243)	(0.351)	(0.243)
key_Own	2.415***	2.469***		0.394***	0.505***	
	(0.032)	(0.046)		(0.087)	(0.124)	
CTA+ * key_Comp		-0.084			-0.024	
		(0.162)			(0.486)	
CTA+ * key_Own		-0.106			-0.221	
		(0.065)			(0.174)	
Constant	-2.577***	-2.600***	-2.597***	-2.249***	-2.318***	-2.317***
	(0.026)	(0.030)	(0.029)	(0.080)	(0.100)	(0.098)
Observations	45,052	45,052	38,470	5,649	5,649	2,652
Log Likelihood	-14,185	-14,184	-9,650	-2,024	-2,023	-840
Akaike Inf. Crit.	28,379	28,380	19,306	4,056	4,058	1,687

Notes. Ad Copy Neutral^{+} and keyword category key_Gen are used as baselines. 6 click observations could not be matched to impressions. *p<0.1; **p<0.05; ***p<0.01.

Table 4. Logistic Regression Results for Hypotheses: H1 & H2

	Dependent variable:					
	click			conversion		
Model (Column)	Main (1)	Interaction (2)	Reduced (3)	Main (4)	Interaction (5)	Reduced (6)
CTA+	**0.067****	**0.067**	**0.067***	**-0.165****	**-0.172**	**-0.184**
	(0.032)	**(0.042)**	**(0.041)**	**(0.082)**	**(0.131)**	**(0.125)**
key_Comp	-0.170**	-0.174	-0.170**	0.344	0.418	0.345
	(0.083)	(0.120)	(0.083)	(0.230)	(0.317)	(0.230)
key_Own	2.363***	2.363***		0.274***	0.264**	
	(0.033)	(0.047)		(0.085)	(0.119)	
CTA+ * key_Comp		0.008			-0.152	
		(0.166)			(0.462)	
CTA+ * key_Own		-0.001			0.020	
		(0.065)			(0.171)	
Constant	-2.656***	-2.656***	-2.656***	-2.030***	-2.027***	-2.022***
	(0.026)	(0.030)	(0.030)	(0.076)	(0.091)	(0.089)
Observations	45,344	45,344	38,860	5,427	5,427	2,603
Log Likelihood	-13,986	-13,986	-9,546	-2,028	-2,028	-901
Akaike Inf. Crit.	27,980	27,984	19,099	4,065	4,069	1,808

Notes. Ad Copy CTA- and keyword category key_Gen are used as baselines. 5 click observations could not be matched to impressions. *p<0.1; **p<0.05; ***p<0.01.

Table 5. Logistic Regression Results for Hypotheses: H3 & H4

5.3 Economic Significance of Main Results

Because we estimate a logistic regression model, coefficients cannot be interpreted as the direct impact on the change in the output variable for a one unit increase in the respective predictor variable, while all other predictors remain constant (Hosmer et al., 2013). To interpret our results economically we need to calculate the odds-ratio, which is equal to the exponent of the coefficient of the respective variable.

The first set of hypothesis is rejected and results are not significant. Therefore, H1 and H2 is not analyzed any further. In regard to hypothesis H3, column (1) of table 5 shows that the coefficient associated with CTA+ is 0.067. Thus, the odds ratio for this variable is equal to exp (0.067) =1.069 (Confidence Intervals: 2.5% =1.005; 97.5% = 1.137). In others words, the odds that a user clicks on an ad with a call-to-action in the last part increases by 7%. When interpreting the economic impact of the positioning in regard to conversions, column (4) of table 5 is considered. The coefficient for CTA+ is significantly negative and translates into an odds ratio of 0.848 (Confidence Intervals: 2.5% =0.721; 97.5% = 0.996). Odds below 1 indicate a negative impact. In other words, when a user is exposed to an ad with a call-to-action in the last part it is about 15% less likely that he will take part in the advertised lottery. Summing

up, at least in our case, when incorporating a call-to-action, the midst position is highly favorable. This design choice reduces overall campaign costs for advertisers by 7% (every click is billed by the search engine) while increasing profits (conversions translate into monetary values for the firm) by 15%.

5.4 Robustness Checks

To enhance credibility of our findings we report several robustness checks. First, we report two alternative logit models which are used to ensure the credibility of our main model. Second, we report further robustness checks to enhance credibility of the field experiment itself.

Interaction Model

As noted by (Jansen et al., 2011) their might exist interaction effects between the ad copy presented to the search engine user and the keyword characteristics used. To control for those potential confounds the main model is extended into an interaction model which accounts for possible second-order effects resulting from a keyword category in combination with the ad copy design used. Columns (2) and (5) in table 4 highlight the results of the interaction model in regard to hypotheses H1 and H2. The robustness check shows that the CTA+ coefficients even changed from negative to positive. Thus, enhancing credibility that there is no wording effect at all. In regard to hypotheses H3 and H4, columns (2) and (5) of table 5 are considered. The results show that reported effects remain qualitatively unhanged while interaction effects are not significant. Using conversions as the dependent variable the CTA+ coefficient even increases from -0.165 to -0.172 indicating that the reported CTA+ effect might be underestimated in our main model. However, the interaction model was not considered as our main model due its lack in ability to significantly ($p>0.05$) enhance general model fit.

Reduced Model

The reported dataset is strongly influenced by retailer specific keywords which account for 53% of all clicks and 59% of all conversions. To assure that our findings are not driven by those keywords an additional logit model is reported. This reduced model is identical to our main model but excludes all observations associated with retailer specific keywords (key_Own). Columns (3) and (6) in table 4 and 5 list the model coefficients. Analyzing the reduced model, we report that in all cases the observed effects persist. Once again, the model coefficients displayed in table 5 highlight that the observed diametric user behavior is not driven by retailer specific keywords. On a conversion level, excluding retailer specific keywords, the observed effect is even stronger and decreases the likelihood that users who were exposed to a CTA+ design participate in the advertised lottery by -17%.

Further Checks

One might argue that even when focusing on top positioned ad slots (Slot 1-3) biases might be induced due to a systematic placement of one ad copy in a more favorable position (e.g., Slot 1) in comparison to another one which is placed in a less favorable slot (e.g., Slot 3). Our research environment provides information on the average position of the ad copies. Throughout the experiment the ad position of all ad copies did not differ systematically. On average each ad was placed on an ad slot of 1.5. Wilcoxon signed rank tests yield to no significant group differences in regard to the average ad position on a daily basis. In addition, result pages for all 487 keywords were monitored throughout the experiment to account for possible confounds driven by visual changes on SERPs. No visual changes could be recognized. It might also be claimed that even when all users were exposed to the same webpage which offered the lottery and stayed unchanged throughout the experiment, the observed effects could be driven by special website attributes. To mitigate this argument a second webpage design was used in a post-test in which the lottery was offered less clearly and in an unobtrusive way further down the page. Our results remained qualitatively unchanged. In the post-test no wording effect is found and CTA- is still favorable in terms of CVR enhancing conversions by 12%.

6 Discussion

The reported field experiment was conducted to investigate the impact of using a call-to-action in paid search on user behavior. The empirical analysis provides evidence that the common industrial advice of incorporating a call-to-action is not necessarily beneficial. When comparing user behavior in terms of clicks and conversions no significant difference in regard to the wording effect can be found. A neutral statement in the last part of the ad seems to be as good as a call-to-action in the same position. Using a call-to-action in the last part of the ad might even lead to an inefficient capital allocation. Analyzing the position effect of a call-to-action using logistic regression the prediction shows that a call-to-action in the last part of the ad increases the likelihood of a user clicking the ad by 7%. On the webpage of the advertiser those users tend to act differently as others and the likelihood that they perform the desired action (conversion) decreases by 15%. Since search engines will bill every click on the ad and conversions can be seen as a monetary equivalent for firms, a call-to-action in the last part results in an inefficient capital allocation. Those findings are robust and persist even when controlling for possible second-order effects or using a reduced model.

Our results have various practical implications. First, our experiment suggests that a call-to-action is not necessarily needed to enhance clicks and conversions. Second, when using a call-to-action in an ad copy the general advice of incorporating a call-to-action is too vague. As a marketer when the campaign goal is to enhance website traffic in general an ad in accordance to the AIDA phrasing scheme proved beneficial. Instead, when the campaign goal is to maximize the number of conversions a mixed up phrasing with a call-to-action in the first part of the description outperforms all tested ads. At least in our case, a less pronounced call-to-action in the middle of the ad copy maximizes conversions by simultaneously reducing advertising costs to a minimum. Our findings also have theoretical implications. First, our study adds to the current body of literature on ad copy design in the context of paid search. To our knowledge we are the first to investigate user behavior in response to call-to-actions in regard to clicks as well as conversions. Second, our findings show a diametric behavior of users in terms of clicks and conversions dependent on the call-to-action placement within the ad. This suggests that a less salient call-to-action in the middle of the ad copy is more likely to be clicked by users who are interested in the advertised conversion goal. In our research environment using a call-to-action in the first part of the description is considered unusual. Analyzing 803 competitor ads which were placed along with our ads during the experiment revealed that only in 5% of all cases a call-to-action is placed in the first part of the ad. Maybe users do not expect such a hidden call-to-action and it is therefore more likely that only those of them click the ad who take the time to read the ad copy carefully. In turn those users are more likely to be truly interested in the advertised action and might show an increased conversion likelihood. From our dataset, we are not able to draw conclusions upon the root causes of the observed user behavior. Instead our results open up new avenues for further research.

The executed field experiment has several limitations. First, the external validity of the experiment is inherently low due to its specific context (furniture retail), its specific conversion goal (lottery participation) and a limited local area in Germany where the experiment was carried out. Second, our claims are restricted to the most prominent positioned ads directly below the search box on a SERP. Third, a possible confound might be a specific competitive advertising situation on Google during experiment execution which we cannot control for. One possible avenue of future research is to test our findings in other research environments. Beyond that, further research could assess the deviating behavior in depth to shed light on the question which patterns affect the observed user behavior. Previous scholars concluded that keywords characteristic (e.g., Nottorf and Funk, 2013) as well as the ad position (e.g., Agarwal et al., 2011) can be used to distinguish user by its intents and the associated behavior. Our findings might be seen as a first indication that the ad copy design could also be suited as a self-selection mechanism for search engine users. Finding designs which are suited as a self-selection mechanism, such as ours, could enhance profits (conversions) by simultaneously reducing costs (clicks) of paid search for marketers.

References

Agarwal, A.; K. Hosanagar; M. Smith (2011). "Location, location, location: An analysis of profitability of position of online advertising markets." *Journal of Marketing Research* 48 (6), 1057–1073.

Animesh, A; S. Viswanathan; R. Agarwal (2011). "Competing "creatively" in sponsored search markets. The effect of rank, differentiation strategy, and competition on performance." *Information Systems Research* 22 (1), pp. 153–169.

Atkinson, G.; C. Driesener; D. Corkindale (2014). "Search Engine Advertisement Design Effects on Click-Through Rates." *Journal of Interactive Advertising* 14 (1), 24–30.

Bing (2015). *Writing an effective ad.* URL: http://advertise.bingads.microsoft.com/en-us /cl/240 /training/writing-an-effective-ad (visited on 11/01/2015).

Blake, T; C. Nosko; S. Tadelis (2015). "Consumer Heterogeneity and Paid Search Effectiveness: A Large-Scale Field Experiment." *Econometrica* 83 (1), 155–174.

Desai, P.; W. Shin; R. Staelin (2014). "The Company That You Keep. When to Buy a Competitor's Keyword." *Marketing Science* 33 (4), 485–508.

Dhar, V. and A. Ghose (2010). "Sponsored search and market efficiency." *Information Systems Research* 21 (4), 760–772.

Edelman, B. and M. Ostrovsky (2007). "Strategic bidder behavior in sponsored search auctions." *Decision Support Systems* 43 (1), 192–198.

eMarketer (2015). *Advertising Forecast.* URL: http://www.emarketer.com/Article/Google-Will-Take-55-of-Search-Ad-Dollars-Globally-2015/1012294 (visited on 11/03/2015).

Ghose, A. and S. Yang (2009). "An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets". *Management Science* 55 (10), 1605–1622.

Google (2015). *What makes up a text ad.* URL: https://support.google.com/adwords/answer/1704389 (visited on 11/03/2015).

GroupM (2014). *Global advertising spending from 2010 to 2015.* URL: http://www.groupm.com/news/groupm-news/groupm-revises-2014-global-ad-spend-forecast-downward-39-growth (visited on 11/03/2015).

Haans, H.; N. Raassens; R. van Hout (2013). "Search engine advertisements: The impact of advertising statements on click-through and conversion rates." *Marketing Letters* 24 (2), 151–163.

Hofacker, C. and J. Murphy (1998). "World Wide Web banner advertisement copy testing." *European Journal of Marketing* 32 (7/8), 703–712.

Hosmer, D.; S. Lemeshow; R. Sturdivant (2013). *Applied logistic regression.* Third edition. Wiley.

Jansen, B.; K. Sobel; M. Zhang (2011). "The Brand Effect of Key Phrases and Advertisements in Sponsored Search." *International Journal of Electronic Commerce* 16 (1), 77–106.

Jansen, B.; Z. Liu; Z. Simon (2013). "The effect of ad rank on the performance of keyword advertising campaigns." *Journal of the American Society for Information Science & Technology* 64 (10), 2115–2132.

Jerath, K.; L. Ma; Y. Park; K. Srinivasan (2011). "A "Position Paradox" in Sponsored Search Auctions." *Marketing Science* 30 (4), 612–627.

Lu, X. and X. Zhao (2014). "Differential Effects of Keyword Selection in Search Engine Advertising on Direct and Indirect Sales." *Journal of Management Information Systems* 30 (4), 299–326.

Narayanan, S. and K. Kalyanam (2015). "Position Effects in Search Advertising and their Moderators: A Regression Discontinuity Approach." *Marketing Science* 34 (3), 388–407.

Nottorf, F. and B. Funk (2013). "A cross-industry analysis of the spillover effect in paid search advertising." *Electronic Markets* 23 (3), 205–216.

Rutz, O.; R. Bucklin; G. Sonnier (2012). "A Latent Instrumental Variables Approach to Modeling Keyword Conversion in Paid Search Advertising." *Journal of Marketing Research* 49 (3), 306–319.

Rutz, O. and R. Bucklin (2011). "From Generic to Branded: A Model of Spillover in Paid Search Advertising." *Journal of Marketing Research* 48(1), 87–102.

Rutz, O. and R. Bucklin (2013). "Paid Search Advertising." *Advanced Database Marketing: Innovative Methodologies and Applications for Managing Customer Relationships*, 229–245.

Rutz, O. and M. Trusov (2011). "Zooming In on Paid Search Ads - A Consumer-Level Model Calibrated on Aggregated Data." *Marketing Science* 30 (5), 789–800.

Sayedi, A.; K. Jerath; K. Srinivasan (2014). "Competitive Poaching in Sponsored Search Advertising and Its Strategic Impact on Traditional Advertising." *Marketing Science* 33 (4), 586–608.

Sudhir, K. (2016): "Editorial—The Exploration-Exploitation Tradeoff and Efficiency in Knowledge Production." *Marketing Science* 35 (1), 1–9.

Strong, E. (1925). The Psychology of Selling and Advertising. McGraw-Hill.

Turnbull, D.; L. Bright (2008). "Advertising academia with sponsored search: an exploratory study examining the effectiveness of Google AdWords at the local and global level." *International Journal of Electronic Business* 6 (2), 149–171.

Varian, H. (2007). "Position auctions." *International Journal of Industrial Organization* 25 (6), 1163–1178.

Wooldridge, J. (2015). Introductory Econometrics. A Modern Approach. Cengage Learning.

Yoo, C. (2011). "Interplay of message framing, keyword insertion and levels of product involvement in click-through of keyword search ads." *International Journal of Advertising* 30 (3), 399–424.