OPEN-RATE CONTROLLED EXPERIMENT IN E-MAIL MARKETING CAMPAIGNS

KONTROLIRANI EKSPERIMENT STOPE OTVARANJA U MARKETINŠKIM KAMPANJAMA PUTEM E-POŠTE

MARKET TRŽIŠTE

Antun Biloš[a], Davorin Turkalj[b], Ivan Kelić[c]

[a] Faculty of Economics, University of Osijek, Gajev trg 7, 31000, Osijek, CROATIA, abilos@efos.hr

[b] Faculty of Economics, University of Osijek, Gajev trg 7, 31000, Osijek, CROATIA, davorin@efos.hr

[c] Faculty of Economics, University of Osijek, Gajev trg 7, 31000, Osijek, CROATIA, ikelic@efos.hr

Abstract

Purpose – The main purpose of this paper is to test the controlled experiment (A/B split) methodology in B2C oriented e-mail marketing campaigns.

Design/Methodology/Approach – E-mail marketing techniques have been a substantial part of e-marketing methodology since the early Internet days of the mid-1990s. From the very beginning of Internet utilization for business purposes, e-mail was one of the most widely used communication techniques in B2B and B2C markets alike. Due to high volumes of spamming and progression of online communication clutter, some practitioners began to question the usability of e-mail as a marketing communication channel, while others embarked on working on improving the message itself. Efforts were invested into improving message quality, as well as into better understanding user expectations. One of the most commonly used techniques to test specific e-mail message elements is the controlled experiment.

Findings and implications – This paper explores several types of controlled experiments in a specific Croatian B2C market. Tests were run to determine subscriber behavior towards several newsletter components, including sending time, sending day, sender's name, and subject line. Open and click rates for tested campaigns, and several other metrics were investigated using MailChimp software. An N – 1 two-proportion test

Sažetak

Svrha – Glavna svrha ovoga rada jest testiranje metodologije kontroliranoga eksperimenta (A/B split) na marketinškim kampanjama putem e-pošte na B2C tržištu.

Metodološki pristup – Marketinške tehnike putem e-pošte činile su značajnu cjelinu metodologije e-marketinga od ranih internetskih dana sredine devedesetih godina prošloga stoljeća. Od samih početaka primjene interneta u poslovne svrhe e-pošta bila je jedna od najčešće korištenih komunikacijskih tehnika na B2B, ali i na B2C tržištima. Uslijed značajne količine neželjene pošte i jačanja internetskog komunikacijskog zagušenja, neki su počeli sumnjati u njezinu uporabnu vrijednost kao marketinškog komunikacijskog kanala, dok su drugi počeli raditi na usavršavanju same komunikacijske poruke. Uložen je napor da bi se usavršila kvaliteta poruke i bolje razumijevanje onoga što očekuju korisnici. Jedna od najčešće korištenih tehnika testiranja pojedinačnih elemenata poruke e-pošte jest kontrolirani eksperiment.

Rezultati i implikacije – Ovaj rad proučava nekoliko kontroliranih eksperimenata na specifičnom B2C tržištu u Republici Hrvatskoj. U radu je istraženo korisničko ponašanje na bazi nekoliko elemenata newslettera poput vremena slanja, dana slanja, imena pošiljatelja i teme poruke. Uz pomoć specijaliziranog softvera MailChimp testirana je razina otvaranja poruke, razina klikova na poveznice unutar poruke i nekoliko dodatnih pokazatelja. Korišten je N – 1 dvostruki proporcijski test koristeći

using an adjusted Wald confidence interval around the difference in the proportions was used for comparing the open-rate measure in the controlled experiments between subjects.

Limitations – Controlled experiments (A/B split tests) showed a lot of potential as a way of measuring behavior and preferences of subscribers, although several apparent limitations (the data-set scope, comparability issues) indicated a clear need for standardization on a managerial and scientific level.

Originality – This paper provides an up-to-date e-mail marketing effectiveness literature review, describes and tests the methodology and metrics for e-mail campaigns measurement, and suggests several important guidelines for further research.

Keywords – controlled experiment, A/B split test, E-mail marketing campaigns, newsletter

prilagođeni Waldov interval pouzdanosti oko razlike u proporcijama za usporedbu stope otvaranja u kontroliranom eksperimentu između grupa.

Ograničenja – Kontrolirani eksperimenti (A/B split testovi) pokazali su značajan potencijal za mjerenje ponašanja i preferencija pretplatnika, iako nekoliko evidentnih ograničenja (opseg seta podataka, poteškoće kod usporedbe) označavaju jasnu potrebu za standardizacijom na upravljačkoj i znanstvenoj razini.

Doprinos – Rad pruža ažuran pregled literature efektivnosti marketinga putem e-pošte, opisuje i testira metodologiju i metriku mjerenja kampanja putem e-pošte i upućuje na nekoliko značajnih smjernica za daljnja istraživanja.

Ključne riječi – kontrolirani eksperiment, A/B split test, marketinške kampanje putem e-pošte, newsletter

1. INTRODUCTION

The constant development of new media helped marketers to evolve from traditional to digital marketing. Nowadays, it helps the evolution from digital to digitally interactive marketing, which facilitates relationships between marketers and customers or users while also taking privacy issues into consideration.

E-mail is an asynchronous and one-to-one medium (Huang, Lin & Lin, 2009, p. 160) and the basic tool of Internet-based digital communication. The first electronic application for message exchange among computers appeared over 40 years ago (Partridge, 2008, p. 3-6), but the popularity of e-mail as a communication tool has not declined since then. Furthermore, e-mail remains the ubiquitous form of business communication, and is still among the most widely used internet services globally (Radicati, 2014).

E-mail is a vital form of communication within a marketing channel where businesses can communicate their value propositions downstream to target audiences, and for customers to communicate their needs upstream to businesses (Dapko & Artis, 2014, p. 254-255). Its efficiency is based on several important advantages (Ružić, Biloš & Turkalj, 2014, p. 189-191): e-mail is cost effective, highly measurable, and suitable for personalization and accurate user segmentation. The use of e-mail as a preferred method of digital communication is expected to grow, given the increased use of technology by young adults (millennials) (Dapko & Artis, 2014, p. 254-255). In addition, it will continue to play an informative and influential role on recipients' behavior (De Bruyn & Lilien, 2008, p. 151-152).

E-mail marketing clearly offers great opportunities for different businesses. Various marketing activities supported by e-mail communication allow companies to directly communicate with their target groups without time or location barriers. However, e-mail marketing techniques should be guided by high ethical principles and used in compliance with permission-based marketing rules. Furthermore, the digital environment of contemporary business processes allows marketers to measure, test, and evaluate assumptions using accurate analytical models.

The main research problem addressed in this paper is the accurate measurement of efficiency within e-mail marketing communication through controlled experiments on different e-mail message elements. E-mail message efficiency in terms of marketing communication goals can be measured and analyzed in order to refine and improve the communication process. Changes in different e-mail message elements can directly influence the reach and message reception of the target audience. However, there are a lot of uncertainties and partial arbitrary judgments and suggestions on which e-mail message element contributes, and to what extent, to e-mail communication improvement. Furthermore, there is a clear lack of standardization in e-mail marketing metrics. This paper explores the controlled experiment methodology and tests several metrics for the efficiency measurement of e-mail campaigns using a specific retail B2C market.

2. LITERATURE REVIEW

2.1. Permission-based e-mail marketing

E-mail is a highly valuable marketing tool for conveying short and simple messages that guide a recipient toward some type of desirable behavior. E-mail marketing can be used to achieve a wide range of specific business goals. These goals make up a component of internet marketing strategy and usually include the following (Pantea & Pop, 2010, p. 737-738; DiGuido, 2003):

o   Increased revenues through promotions and up- or cross-sell efforts;

o   Increased traffic to a web-site, brick-and-mortar retail location, or call center;

o   Improved brand awareness and preference;

o   Conversions (actions);

o   Customer loyalty programs;

o   Deeper customer preference and profile information through surveys, promotions, and sweepstakes;

o   Relationship building and management.

Godin (1999) proposed the term "permission marketing" and advised marketers to seek customers' permission before sending them any type of promotional messages. Permission marketing creates a platform for two-way interaction and engagement as a solution to the communicational challenge faced by traditional marketing (Kumar, Zhang & Luo, 2014). Consequently, permission-based e-mail marketing represents an e-mail marketing approach where messages are sent only to those users who have directly asked to receive them.

Marketers can use permission-based marketing to harness a number of its advantages. Harris and Dennis (2008, p. 222) suggested several important advantages of e-mail permission-based promotional communication:

o   Acquiring e-mail addresses with the permission of the user;

o   Targeting a specific, appropriate audience for an e-mail marketing campaign;

o   Developing personalized communication with tailored content;

o   Executing and administering campaigns with relative ease;

o   Imitating a response to customers' replies;

o   Easy flow as information is passed along to others;

o   Constant maintenance of e-mail lists.

Strauss and Frost (2009, p. 307) argue that micro-segmentation of e-mail marketing recipient lists leads to the sending of e-mail messages to relevant consumers. These messages are usually personalized, and the communicated content is tailored. Permission-based e-mail marketing offers marketers an opportunity to strengthen brand loyalty (Merisavo & Raulas, 2004). Clearly,

e-mail marketing should be used to enhance the consumer experience and not alienate the consumer from the company. In addition, Yildiz (2007, p. 5). suggests that, by filling out a subscription form, the prospective customer places "confidence in his applicant and commits himself for the future".

E-mail marketing can be used as a powerful marketing tool, but it can also be extremely challenging. Groves (2009) even called the e-mail inbox a hostile environment. However, a reputation of being consumer-oriented is more important than having large subscriber lists: message quality is more important than the quantity of e-mail contacts (Strauss & Frost, 2009, p. 307).

A study by Kent and Brandal (2003, p. 500-502) showed that many permission-based e-mails are not read by their recipients. Furthermore, these e-mails are not found to be interesting, illustrating that there is a lot of room for improvement. Marketers should clearly get to know their customers' preferences better and develop a mutual relationship where customers are encouraged to respond or engage. Whether an e-mail message is considered noteworthy or not depends on marketers' ability to adhere to the fundamentals of authentic relationship building with respective customers (Groves, 2009, p. 1-2).

Research by Micheaux (2011, p. 45-46) revealed that perceived pressure from a commercial e-mail sender is an individual phenomenon. Under conditions of low relevance, the unsubscribing effect is moderated by the execution of e-mail advertising, as marketers can control their own advertising but not overall e-mail advertising volume.

Due to general clutter increase within communication channels and attention-grabbing struggles of all sorts, customers (or users) clearly have less time to check their e-mails. With nearly half (48%) of e-mail massages being read on mobile devices, the virtual inbox is physically getting smaller (Jordan, 2013). In addition, that

amount will increase as smart watches and similar modern communication gadgets continue to promote mobile communication.

Permission-based e-mail marketing will most probably continue to play a powerful role in the consumer purchase process. E-mail supports the newer platforms, acting as a portal for important updates, billing notifications, and password resets. Several authors (Aufreiter, Boudet & Weng, 2014; Marnell, 2015) have argued that consumers open their inbox only when they are in the mindset to engage, so segmentation and personalization will continue to be marketers' key e-mail tools in the future. Marnell (2015) concludes that today's inbox has evolved and that so should e-mail marketing strategies.

2.2. E-mail marketing metrics

The digital environment allows marketers to accurately measure users' actions. This is not the case in many situations with traditional media; hence, it is one of the biggest and most important competitive advantages of conducting marketing activities digitally (Ružić et al., 2014). Measuring results achieved in the digital environment can and should directly explain the success rate of any e-marketing activity. Therefore, accurately measuring achieved results of any given e-mail marketing campaign can help companies understand and improve the marketing activities they conduct in order to ultimately reach their business goals. This approach allows marketers to increase the effectiveness of marketing efforts and measure the return on investment (ROI) of marketing expenditure. It is impossible to determine the success rate of marketing efforts if its effects are unknown or its metrics unspecified. Once a commercial e-mail message has been received, individuals decide whether or not to open it. Actually opening the message acts as the trigger factor in the process (José-Cabezudo & Camarero-Izquierdo, 2012, p. 97-98).

Several authors (Cole, Nordfelt, Ring & Fair, 2005; Groves, 2009; Fabian, Bender & Weimann, 2015) agree that e-mail tracking within the activities of

e-mail marketing is a crucial part of the process. According to Fabian et al. (2015), e-mail tracking uses personalized links and pictures for gathering information on user behavior: where, when, on what kind of device, and how often an e-mail message has been read. In other words, e-mail tracking enables marketers to remotely observe whether an e-mail has been opened, the time when a user reads an e-mail, the application in which the user opens it, and identify the links on which the user clicks. This information is very useful for a business in order to understand customer behavior in more depth (Cole et al., 2005, p. 316) and thus very useful for marketing purposes. Gathering this type of information is not only critical for determining the engagement rate but also for learning and refining the e-mail marketing process.

In addition, when an e-mail message is sent, the responsible E-mail Service Provider (ESP) automatically adds a special code that enables the tracking of certain recipient responses. Groves (2009, p. 172) suggests that it is possible to track several metrics at the ESP level:

o Which e-mail messages bounced and why they bounced;

o Which e-mail messages received spam complaints;

o Who opted out of receiving future e-mail messages (unsubscribed);

o Who enabled the images to display in their e-mail messages;

o Who clicked the links in an e-mail message;

o Who forwarded an e-mail to someone else.

A study by Fabian et al. (2015) showed that both tracking links and tracking pixels are widely used in commercial practice. Furthermore, almost 98% of all e-mails analyzed in the study contained at least one e-mail tracking method. These e-mail tracking techniques could potentially create massive privacy concerns on the consumer side, in particular with privacy-sensitive users. However, the understanding of users'

privacy issues is one of the cornerstones of permission-based e-mail marketing, and businesses simply need to approach privacy issues with a lot of attention.

Ellis-Chadwick and Doherty (2012, p. 843-848) conclude that the subject line and the sender of any given e-mail message are directly responsible for the opening of e-mail. The subject line of an e-mail message must grab the initial attention of the recipient and motivate the desired user behavior; if such is the case, this behavior is the act of opening and reading an e-mail. If that does not happen, there is no opportunity for sustained attention, and the message will most probably be deleted and never seen again, unlike print media messages, which can be returned to later.

Although e-mail engagement rates started out very high, they have declined and stagnated over time (Direct Marketing Association, 2005; eMarketer, 2013). Finding different ways to increase these engagement rates is crucial for e-mail marketers. An effective e-mail testing methodology is a useful tool to achieve this. Identifying potential strengths and weaknesses of the content (i.e. the e-mail's creativity) and the target group at full scale can help marketers improve the engagement rates for their campaigns (Bonfrer & Drèze, 2009, p. 251-252).

The marketers' decision on KPIs (i.e. key performance indicators) represents a major part of the metrics used to determine the efficiency of any marketing activity on a company level. KPIs may be heavily influenced by the marketing goals that a business tries to achieve, but it is essential to derive KPIs for a longer period of time. The most important KPIs of e-mail marketing success level (Stokes, 2011, p. 173; Rita & Rita, 2003) are as follows:

o   Delivery rate (number or percentage of delivered e-mails);

o   Bounce rate (number of percentage of undelivered e-mails);

o   Open rate (% of opened e-mails);

o   Click-through rate or click rate (% of clicks on links within an e-mail message);

o   Number of emails forwarded;

o   E-mail message replies;

o   ROI (return on investment);

o   Number of social shares;

o   Database growth;

o   Conversion rate (website activity generated by the e-mail).

Small and large companies alike tend to use one of the popular software solutions for conducting their e-mail marketing activities. These software applications offer quite simple solutions for managing e-mail campaigns and subscriber lists. The implementation process is straightforward and allows marketers to focus on their marketing goals and not on technical issues. There are a number of different solutions, of which Campaigner, MailChimp, Get Response, iContact are among the most popular (Rashid, 2015). These e-mail marketing solutions vary according to their capabilities and price range.

2.3. Studies of e-mail marketing effectiveness

Based on a review of the recent academic and professional literature, it is apparent that the number of e-mail marketing effectiveness papers is somewhat limited. In addition, published research papers on e-mail marketing and metrics related to e-mail marketing effectiveness tend to be heterogeneous in terms of scope, achieved goals, and the methodology used. An overview of the most important available studies concerning e-mail marketing effectiveness is provided below.

Several papers (Sigurdsson, Menon, Sigurdarson, Kristjansson & Foxall, 2013; Kumar et al, 2014; Končar, Vukmirović & Petrović Katai, 2009; Mailchimp, 2015a; Chittenden & Rettie, 2003; Nanji, 2015) have focused primarily on determining the most important e-mail message elements influencing the success rates of e-mail

campaigns. These studies pinpoint several factors that have a positive impact on open rates, click rates, conversions, and other metrics.

Sigurdsson and others (2013, p. 299-303) conducted an e-mail marketing experiment based on the behavioral perspective model to test the influence on conversion rate. Specialized e-mail messages were sent to two segments from the same subscriber database of registered consumers interested in children's books. The process consisted of subscribers receiving the e-mail message, opening it, clicking on a link and buying the target books. The results showed that the informational stimuli in the e-mail message were more successful in motivating consumers to open the e-mails (open rate metric), whereas the utilitarian stimuli were beneficial in increasing buying behavior (conversion rate metric).

Kumar and others (2014, p. 416-417) studied the effect of marketing intensity and customer characteristics on e-mail marketing efficiency. The study showed that customers under high marketing intensity are less likely to become subscribers. Furthermore, after customers have subscribed, high exposure to commercial e-mail messages (high e-mail message frequency) can make them withdraw more quickly. Valid formulation of e-mail marketing messages positively contributes to activities of achieving marketing goals (Končar et al., 2009). In addition, Kumar et al. (2014) found that higher e-mail open rates lead to higher spending levels, suggesting that businesses should focus on delivering marketing messages that are relevant to their e-mail subscribers.

MailChimp (2015b) published an extensive study based on a robust sample size. The research piece, based on hundreds of millions of emails delivered by their e-mail delivery system, calculated the average unique open rates, average unique click rates, average unique soft bounces, average unique hard bounces, and average unique abuse complaint rates by the particular industry. MailChimp tracked campaigns with at least 1000 subscribers and ranged from small startups to Fortune 500 companies, thus

creating a comprehensive review of each industry. The study showed significant differences among industries especially in open and click rates. The industry that showed the best performing open rates was the Hobbies category (29.42%), followed by Arts and Artist (27.93%), and Photo and Video (27.06%). According to this research, the worst performing industry is, interestingly enough, the Daily Deals/E-coupons category (13.89%) (Appendix 1) but several minor discrepancies regarding the reported data can be found in other reports (Silverpop, 2014). However, significant differences among industries are apparent.

Recent studies (Foreman, 2014; MailChimp, 2015a) have also shown that, even though campaign success is highly linked to the industry type and the subscriber characteristics, there are several best practice rules that generally apply (based on the open-rate metric), namely:

o Campaigns are more successful on weekdays than during the weekend.

o Late mornings (10 to 12) are better than late afternoons or evenings.

o Simple and straightforward subject lines work best.

o The subscriber should be able to recognize the sender's name.

A research study based on 30 e-mail marketing campaigns following qualitative research among industry experts (Chittenden & Rettie, 2003, p. 205-215) identified that subject line, e-mail message length, incentive, and number of images can be associated with increased response rate. For several campaigns it was possible to link demographic and lifestyle data to response rate (it was suggested that users who have bought online had higher response rates). A more recent study (Nanji, 2015) was based on data gathered from a survey of 303 marketing, sales, and business professionals on a global scale (68% B2B-focused, 32% B2C-focused). According to that report, creating a meaningful call-to-action is cited as the best way to increase

on conversions and more efficient split tests (a conversion is a predefined desired action completed by the user). In a common controlled experiment, the interface version which generates the most conversions is considered best (Speicher et al., 2014, p. 545-546).

Control/Treatment tests have been used for testing marketing efficiency in various fields for decades (Nielsen, 2005). This has been a highly used method in direct mail, where companies often split their mailing lists and send out different versions of a mailing to different recipients. Control/Treatment testing became popular in website optimization, where it was used to show different web page versions to different visitors and measure the results of that exposure (Kohavi et al., 2009; Nielsen, 2005).

In the case of e-mail marketing metrics, a controlled experiment (A/B split test) refers to testing elements of an e-mail message on two segments of recipients. This approach allows marketers to send two different versions of the e-mail message to two segments of the recipient list. Those two tested segments can be smaller samples of the total recipient list or the two can accrue to the total number of the list (in which case, every segment is 50% of the list). If the first approach is used, marketers can select the version that gets the most opens or clicks to send to the rest of the subscriber list, while the second approach allows the maximization of the sample size.

While only one element of the e-mail message is tested, almost every e-mail message element can be tested in a controlled experiment (A/B split test) (Shivdasani, 2014):

o   Subject line – wording variation and use of special characters;

o   From name – sender variation (a person or an organization);

o   Delivery date/time – time related variation;

o   Content – images, layouts, and messages variations.

By changing aspects of the campaign between the two different groups, it is possible to determine which variation the recipients responded to and to what extent. The controlled experiment (A/B testing) in e-mail marketing has several major benefits (Eisenberg & Eisenberg, 2006, p. 226-228; Nielsen, 2005; Peterson, 2004, p. 76-78):

o   It measures the user behavior under real-world conditions.

o   It can measure very small performance differences.

o   It can resolve trade-offs between conflicting guidelines or qualitative usability findings by determining which one carries the most weight under the circumstances.

o   It is cost-effective and relatively easy to implement.

However, several authors suggest a number of limitations of the controlled experiment in e-mail marketing to be considered in addition to the advantages (Nielsen, 2005; Kohavi et al., 2009, p. 157-158):

o   Applicability for testing with one significant key performance indicator;

o   Lack of behavioral insights;

o   Implications of measuring short-term effects;

o   Provision of data on the tested element only.

Accurately measuring the outcomes of controlled experiments (A/B split tests) and the attribution of measured variation is only the first part of efficiency calculation. The crucial moment is determining the statistical significance of the measured data and deriving fact-driven conclusions.

Sauro and Lewis (2012) suggest that comparing the two outcomes of dichotomous variables for two independent groups is one of the most frequently computed procedures in applied statistics. However, there is little agreement on the best statistical test for this situation. In addition, the heterogeneity of research goals and the perspectives of the papers described in the

MARKET TRŽIŠTE

literature review imply the heterogeneity of the methodology used. The latest research (Sauro & Lewis, 2012, p. 75-81) suggests that a slight adjustment to the standard chi-square test and an equivalent adjustment to the two-proportion test generate the best results for almost all sample sizes. Similar to the two-sample t-test, where the difference between the means was compared to the t-distribution, the N − 1 chi-square test is equivalent to an N − 1 two-proportion test. Instead of using the chi-square distribution to generate the p-values, Sauro & Lewis (2012, p. 75-81) suggest using the normal (z) distribution. The same authors (Sauro & Lewis, 2012, p. 96) finally conclude that, in order to compare a binary outcome, a measure such as a task completion rate or conversion rate – as used in a between-subjects controlled experiment (A/B test) – the N − 1 two-proportion test using an adjusted Wald confidence interval around the difference in the proportions is recommended (Agresti & Caffo, 2000, p. 280-288).

4. CONTROLLED EXPERIMENT AND RESULTS

An exploratory analysis of a controlled experiment (A/B split test) was conducted in a specific Croatian B2C market using a company in retail industry. Based on the literature review and prior experience, the main research goal was to test four elements of e-mail marketing messages and determine if the variations used had statistically significant differences.

The company used in this research has been in operation since 2011, and e-mail marketing is an important part of its marketing activities on all levels. The company reported using e-mail marketing activities to accomplish various business goals, namely communicating with target groups; promoting products, services, and/or events; providing customers with feedback; and testing subscriber preferences.

Data used in this study was gathered from 45 e-mail campaigns sent during a four-year period (11/2011 to 05/2015) using the MailChimp e-mail marketing software. The number of subscribers varied over time due to continuous newsletter promotion and growing popularity on the one hand, and subscriber list cleanups on the other. Popularity and promotion of the newsletter led to a rise in the number of subscribers over time. However, the company reported periodical list cleanups based on subscriber behavior. Even though subscriber list cleanups mean directly losing a part of the potential target group as subscribers are removed from the list, this approach is generally considered to be a long-term maintenance tactic and should provide benefits to senders and subscribers alike.

Several conclusions can be formed based on the aggregated data analysis from 45 e-mail campaigns. It is clear that there are significant differences between open rates and click rates among different campaigns, ranging between 25.0% and 36.0% within the open-rate metric and between 9.5% and 20.3% within the click-rate metric. The average open rate is 30.0% (x=0.3, sd=0.03) and the average click rate is 13.1% (x=0.13, sd=0.03). Interestingly enough, there is only a weak correlation between the open and click rates (0.44). The industry open rate is estimated at around 17.97% (MailChimp, 2015b) even though Silerpop (2014) reported a slightly higher industry average for the retail industry, between 18 and 20%. Compared to the industry average, this data set (campaign open rates) shows much better performance.

Furthermore, a clear influence of seasonality can be observed. The best performing campaigns based on the open-rate metric are those sent in the late fall and early winter period (November to January). This seasonality is highly influenced by industry characteristics and specific subscriber interest in that time period. Minor exceptions to this seasonality pattern could be linked to other factors, namely the content of the sent message (however, this influence was not tested). The open-rate and click-rate metrics for sent campaigns in the observed time period are presented in Graph 1.

GRAPH 1: Campaign open-rate and click-rate metrics
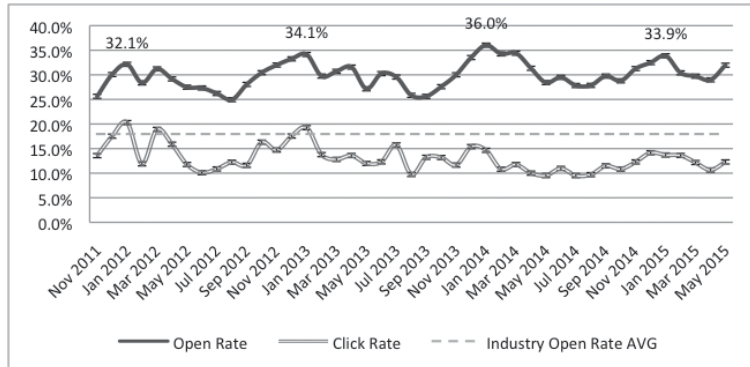
Source: Authors' research

The main focus of this research paper was on 10 e-mail controlled experiments (A/B split tests). These tests were run among the observed 45 campaigns in the time period running from 2012 to 2014. The experiments were run to determine the subscriber behavior towards several newsletter components, namely:

o   Sending time;

o   Sending day;

o   Sender's name;

o   Subject line.

Every experiment (A/B split test) was focused on a single tested component in order to determine the significance of the varied item (as argued earlier). The evaluation criterion (OEC) used in every controlled experiment was the open-rate metric. However, 6 experiments targeted list segments to test the open-rate metric and then used the better performing group for the rest of the subscriber list. The better performing group is the group with the higher open-rate metric and it can be analyzed in two ways: as the difference between two group open rates (the difference between score A and score B) or as the percentage difference between the better performing group score and the other one. The remaining 4 tests were run on the complete subscriber list (total). The final scores for the 10 controlled experiments (A/B split tests) are submitted below (Table 1).

TABLE 1: Controlled experiment results

Nr	Testing item	Group A	Group B	Measure (OEC)	Score A	Score B	z	p-value
1	Sending time	Thu 10:30	Thu 14:00	Open rate (total)	0,316	0,308	0,336	0,737
2	Subject line	Generic	Specific	Open rate (total)	0,283	0,331	0,865	0,387
3	Subject line	Generic	Specific	Open rate (segment)	0,300	0,213	2,035	**0,042**
4	Sender	Full name	Short name	Open rate (segment)	0,277	0,259	0,339	0,734
5	Subject line	Generic	Specific	Open rate (segment)	0,358	0,264	1,832	**0,067**
6	Sender	Generic	Special characters	Open rate (segment)	0,361	0,319	1,011	0,312
7	Sender	Generic	Special characters	Open rate (segment)	0,250	0,280	0,636	0,525
8	Subject line	Generic	Special characters	Open rate (segment)	0,318	0,331	0,265	0,791
9	Sending time	Thu 10:00	Thu 17:00	Open rate (total)	0,280	0,288	0,416	0,677
10	Sending day	Tue 10:00	Wed 10:00	Open rate (total)	0,296	0,294	0,142	0,887

Source: Authors' research

N – 1 two-proportion test (using the normal (z) distribution) was used to test the statistical significance of the measured data (with an adjusted Wald confidence interval around the difference in the proportions).

A controlled experiment (A/B split test) of sending-time variation was conducted twice (test 1 and 9). Tests were run on the same day, with group A sent at mid-morning (at 10:00, i.e. 10 a.m.) and group B in the afternoon (a 14:00 or 17:00, i.e. 2 p.m. or 5 p.m.). Neither test revealed significant differences in open rates: $a_1=0.316$, $b_1=0.308$ and $a_9=0.280$, $b_9=0.288$ ($z=0.336$, $p=0.737$; $z=0.416$, $p=0.677$, respectively). The last test (test 10) was focused on the sending day, where group A was sent on Tuesday and group B on Wednesday (at the same times of day). The goal of this test was to test different weekdays. The test showed no difference in open rates ($a_{10}=0.296$, $b_{10}=0.294$) and no statistical significant difference ($z=0.142$, $p=0.887$). However, it is crucial to point out that all campaigns were sent on weekdays (as opposed to weekends), based on prior research.

Subject-line variation was tested 4 times (tests 2, 3, 5, and 8). The first 3 tests were focused on the difference between a generic subject line (A) and a newsletter-specific subject line (B). Even though the first test (test 2) showed no statistical difference between the groups ($z=0.865$, $p=0.387$), additional tests (test 3 and 5) showed that best-performing subject lines are in fact generic ones. The difference in open rate in test 3 ($a_3=0.300$, $b_3=0.213$) was statistically significant ($z=2,035$, $p=0,042$). The probability that the two variations (A and B) have the same open rate is around 4%. There is about a 96% probability the open rates are different and not a result of mere chance. The 90% confidence interval around the difference in open rates (critical value of 1.64) is 0.086 ± 0.069 (ranging between 0.017 and 0.156). Consequently, if variation A was used on all users, we could expect an open rate of between 1.7% and 15.56% higher than for variation B. Similar conclusions can be drawn from test 5. The difference in open rate in test 5

($a_5=0.358$, $b_5=0.264$) was statistically significant ($z=1.832$, $p=0.067$). The probability that the two variations (A and B) have the same open rate is around 7%, so there is about a 93% probability that the open rates are different and not a result of mere chance. The 90% confidence interval around the observed difference of 0.093 ranges between 0.010 and 0.176. If variation A was used on all users, we could expect an open rate of between 1% and 17.61% higher than for variation B. The final test concerning the subject-line variation compared a generic subject line with a special-characters subject line. Special characters include visual stimuli with both alphanumeric characters and/or small images (icons). Test 8 showed no statistically significant difference between groups A and B ($z=0.265$, $p=0.791$).

A controlled experiment (A/B split test) with sender-name variation was conducted three times (tests 4, 6, and 7). The first test compared full names to short versions of the senders' names. The test showed no statistically significant difference between groups A and B ($z=0.339$, $p=0.734$). In test 6 and test 7, special characters were used with sender-name variation. Interestingly enough, even though the tests showed opposite open-rate results ($a_6=0.361$, $b_6=0.319$ and $a_7=0.250$, $b_7=0.280$), those were not statistically significant ($z=1.011$, $p=0.312$ and $z=0.636$, $p=0.525$, respectively); therefore, there is not enough evidence to reject the assumption that there is no difference between them.

Finally, out of 10 controlled experiments (A/B split tests), only two found a statistically significant difference between the tested variants (test 3 and test 5). According to these experiments, the best-performing subject lines are in fact generic ones. The rest of the experiment showed no statistically significant difference among the tested variants (including sending time/day, sender's name, and the use of special characters). As other researchers have reported different results and different conclusions accordingly, this field clearly needs additional research on a larger scale and additional experiment runs.

4.1. Research limitations

Although this controlled experiment (A/B split test) study provides some new insights into the techniques and metrics used in e-mail marketing campaigns development, it clearly suffers from a number of limitations. First and foremost, it is based on the data-set of a single company in a specific market, which limits the scope of derived conclusions. It could be argued that the length of the time period and the number of analyzed campaigns could have caused the study's limitations, but there is not enough evidence to support this theory.

Furthermore, the lack of metric standards across markets affects the comparability of similar research pieces. One of the biggest problems on the global scale is that studies of e-mail marketing campaigns are limited and heterogeneous, which will hopefully change in the future.

Further research into e-mail marketing success metrics is clearly needed. Potential guidelines for a new research study should involve a larger sample of companies across different industries. The use of a single methodology with strict control of the reported measurement scales is absolutely crucial. Controlled experiments (A/B split tests) have shown a lot of potential as a way of measuring behavior and preferences of subscribers. However, several test results seem to be inconclusive and, therefore, there is a clear need for standardization on a practical and scientific level.

5. CONCLUSIONS

E-mail marketing techniques have been a substantial part of e-marketing methodology since the early Internet days in the mid-1990s. From the earliest applications of the Internet application for business purposes, e-mail was one of the most widely used communication techniques in B2B and B2C markets alike. At the same time, businesses started using e-mail to create specialized newsletters and build specific online target audiences while improving metrics for measuring the rate of success of every e-mail campaign sent. As the spamming grew in volumes and online communication clutter also progressed, some practitioners started questioning the usability of e-mail as a marketing communication channel, while others embarked on improving the message itself. Considerable efforts were invested not only into improve the quality of the message but also into better understanding user expectations.

The Internet and the digital environment provide an unprecedented opportunity to evaluate assumptions using controlled experiments. The controlled experiment is often referred to as an A/B test (especially in professional studies), which is a colloquial name for a testing technique under which different elements are compared on two identical segments of the same group. This paper explores several types of controlled experiments in a specific Croatian B2C retail market. The tests were run in order to determine subscriber behavior towards several e-mail message components, including sending time, sending day, sender's name, and subject line. Open and click rates for tested campaigns were investigated using the MailChimp e-mail marketing specialized software. The N – 1 two-proportion test, using an adjusted Wald confidence interval around the difference in the proportions, was used for comparing the open rate measure in controlled experiment (A/B test) for between-subjects. A/B split tests showed a lot of potential as a way of measuring behavior and preferences of subscribers.

Antun Biloš, Davorin Turkalj, Ivan Kelić

References

1. Agresti, A., & Caffo, B. (2000). Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures. *The American Statistician*, *54*(4), 280-288.
2. Aufreiter, N., Boudet, J., & Weng, V. (2014). *Why marketers should keep sending you e-mails*. Available at: http://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/why-marketers-should-keep-sending-you-emails (15/5/2015).
3. Bonfrer, A., & Drèze, X. (2009). Real-Time Evaluation of E-mail Campaign Performance. *Marketing Science*, *28*(2), 251-263.
4. Brkić, N., & Unkić, F. (2009). Recipients Attitudes Towards Electronic Mail As Instrument Of Direct Marketing. *Ekonomska revija*, *7*(13), 25-37.
5. Chittenden, L., & Rettie, R. (2003). An evaluation of e-mail marketing and factors affecting response. *Journal of Targeting, Measurement and Analysis for Marketing, 11*(3), 203-217.
6. Cole, E., Nordfelt, M., Ring, S., & Fair, T. (2005). *Cyber Spying Tracking Your Family's (Sometimes) Secret Online Lives*. Rockland: Syngress Publishing.
7. Dapko, J. L., & Artis, A. B. (2014). Less is More: An Exploratory Analysis of Optimal Visual Appeal and Linguistic Style Combinations in a Salesperson's Initial-Contact E-mail to Millennial Buyers Within Marketing Channels. *Journal of Marketing Channels, 21*(4), 254-267.
8. De Bruyn, A., & Lilien, G. L. (2008). A multi-stage model of word-of-mouth influence through viral marketing. *International Journal of Research in Marketing, 25*(3), 151-163.
9. DiGuido, A. (2003). *Goals, Benchmarks, and E-Mail Marketing Success.* Available at: http://www.clickz.com/clickz/column/2138301/goals-benchmarks-email-marketing-success (17/5/2016)
10. Direct Marketing Association (2005). *The DMA 2005 Response rate report*. Direct Marketing Association, New York.
11. Eisenberg, B., & Eisenberg, J. (2006). *Call to Action: Secret Formulas to Improve Online Results*. Thomas Nelson Inc.
12. Ellis-Chadwick, F., & Doherty, N. F. (2012). Web advertising: the role of email marketing. *Journal of Business Research*, *65*(6), 843-848.
13. eMarketer (2013) *Email Marketing Benchmarks: Key Data, Trends and Metrics*. Available at: http://www.emarketer.com/public_media/docs/eMarketer_Email_Marketing_Benchmarks_Key_Data_Trends_Metrics.pdf (22/2/2016)
14. Fabian, B., Bender, B., & Weimann, L. (2015). E-Mail Tracking in Online Marketing – Methods, Detection, and Usage. *12. Internationale Tagung Wirtschaftsinformatik (Wirtschaftsinformatik 2015)*, At Osnabrück, Germany. Available at: https://www.researchgate.net/publication/267694659_E-Mail_Tracking_in_Online_Marketing_Methods_Detection_and_Usage (25/5/2015)
15. Foreman, J. (2014). *Insights from MailChimp's Send Time Optimization System.* Available at: http://blog.mailchimp.com/insights-from-mailchimps-send-time-optimization-system/ (4/6/2015)
16. Godin, S. (1999). *Permission Marketing: Turning Strangers Into Friends And Friends Into Customers*. New York, NY: Simon & Schuster.
17. Groves, E. (2009). *The Constant Contact Guide to E-mail Marketing*. Hoboken: John Wiley & Sons.
18. Haq, Z. (2009). E-mail advertising: A study of consumer attitude toward e-mail advertising among Indian users. *Journal of Retail & Leisure Property*, *8*(3), 207-223.
19. Harris, L., & Dennis, C. (2008). *Marketing the e-Business*. 2nd ed. New York, NY: Routledge.
20. Huang, C. C., Lin, T. C., & Lin, K. J. (2009). Factors affecting pass-along email intentions (PAEIs): Integrating the social capital and social cognition theories. *Electronic Commerce Research and Applications, 8*(3), 160-169.

21. Jordan, J. (2013). *48% of Emails are Opened on Mobile; Gmail Opens Down 20% Since Tabs*. Available at: https://litmus.com/blog/48-of-emails-are-opened-on-mobile-gmail-opens-down-20-since-tabs (1/6/2015)

22. José-Cabezudo, R. S., & Camarero-Izquierdo, C. (2012). Determinants of opening-forwarding e-mail messages. *Journal of Advertising, 41*(2), 97-112.

23. Kent, R., & Brandal, H. (2003). Improving email response in a permission marketing context. *International Journal of Market Research*, *45*(4), 489-503.

24. Kohavi, R., Longbotham, R., Sommerfield, D., & Henne, R. M. (2009). Controlled experiments on the web: survey and practical guide. *Data mining and knowledge discovery, 18*(1), 140-181.

25. Končar, J., Vukmirović, G., & Petrović Katai, Z. (2009). Direct E-Mail Marketing Development In The Republic Of Serbia. *4th Aspects and Visions of Applied Economics and Informatics*, March 26-27, Debrecen, Hungary.

26. Kumar, V., Zhang, X. A., & Luo, A. (2014). Modeling Customer Opt-In and Opt-Out in a Permission-Based Marketing Context. *Journal of Marketing Research*, *51*, 403-419.

27. MailChimp (2015a). *Subject Line Comparison: The Best and Worst Open Rates on MailChimp*. Available at: http://mailchimp.com/resources/research/email-marketing-subject-line-comparison/ (10/6/2015)

28. MailChimp (2015b). *Email Marketing Benchmarks: Average Unique Open, Click, Bounce, and Abuse Complaint Rates*. Available at: http://mailchimp.com/resources/research/email-marketing-benchmarks/ (4/6/2015)

29. Marnell, A. (2015). *Why Your Email Marketing Must Adapt to Recent Changes*. Available at: http://www.marketingprofs.com/opinions/2015/27639/why-your-email-marketing-must-adapt-to-recent-changes (1/6/2015)

30. Merisavo, M., & Raulas, M. (2004). The impact of e-mail marketing on brand loyalty. *Journal of Product & Brand Management*, *13*(7), 498-505.

31. Micheaux, A. L. (2011). Managing E-mail advertising frequency from the consumer perspective. *Journal of Advertising, 40*(4), 45-66.

32. Nanji, A. (2015). *The Best Tactics for Improving Email CTRs*. Available at: http://www.marketingprofs.com/charts/2015/27811/the-best-tactics-for-improving-email-ctrs (2/6/2015)

33. Nielsen, J. (2005). Putting A/B Testing in Its Place. Normal Nielsen Group. Available at: https://www.nngroup.com/articles/putting-ab-testing-in-its-place/ (17/5/2015)

34. Pantea, C., & Pop, A. N. (2010). Email Marketing Campaigns: The Easiest Path From Organizations to Consumers – An Exploratory Assessment. *Annals of Faculty of Economics*, *1*(1), 737-742.

35. Partridge, C. (2008). The Technical Development of Internet Email. IEEE Annals of the History of Computing, *IEEE Computer Society*, *30*(2), 3-29.

36. Peterson, E. T. (2004). Web analytics demystified: a marketer's guide to understanding how your web site affects your business. Ingram.

37. Radicati, S. (2014). Email Statistics Report 2014-2018. The Radicati Group, INC. Available at: http://www.radicati.com/wp/wp-content/uploads/2014/01/Email-Statistics-Report-2014-2018-Executive-Summary.pdf (22/2/2016)

38. Rashid, F. Y. (2015). The Best Email Marketing Software of 2015. Available at: http://www.pcmag.com/article2/0,2817,2453354,00.asp (22/11/2015)

39. Rita, P., & Rita, C. (2003). *E-Mail Marketing Campaigns: An Application to the Tourism Industry*. Available at: https://www.researchgate.net/publication/267829733_E-Mail_Marketing_Campaigns_An_Application_to_the_Tourism_Industry (20/5/2015)

40. Roy, R. K. (2001). *Design of experiments using the Taguchi approach: 16 steps to product and process improvement*. Boston, MA: John Wiley & Sons.

41. Ružić, D., Biloš, A., & Turkalj, D. (2014). *E-marketing*. 3rd ed. Osijek: Ekonomski fakultet u Osijeku.

42. Sauro, J., & Lewis, J. R. (2012). *Quantifying the user experience: Practical statistics for user research*. Boston, MA: Elsevier.

43. Shivdasani, N. (2014). *A/B Testing in MailChimp: 7 Years of Successful Experiments*. Available at: http://blog.mailchimp.com/ab-testing-in-mailchimp-7-years-of-successful-experiments/ (4/6/2015)

44. Sigurdsson, V., Menon, R. G. V., Sigurdarson, J. P., Kristjansson, J. S., & Foxall, G. R. (2013). A Test of the Behavioral Perspective Model in the Context of an E-Mail Marketing Experiment. *The Psychological Record*, *63*, 295-308.

45. Silverpop (2014). *2014 Email Marketing Metrics Benchmark Study*. Available at: http://www.silverpop.com/marketing-resources/white-papers/all/2014/email-metrics-benchmark-study-2014/ (7/6/2015)

46. Speicher, M., Both, A., & Gaedke, M. (2014). WaPPU: Usability-Based A/B Testing. In: *Web Engineering* (pp. 545-549). London: Springer International Publishing.

47. Stokes, R. (2011). *eMarketing: The essential guide to digital marketing*. 4th ed. Quirk (Pty) Ltd.

48. Strauss, J., & Frost, R. (2009). *E-marketing*. Boston, MA: Pearson Prentice Hall.

49. Yildiz, H. (2007). Permission marketing and commitment: proposal for a theoretical framework applied to e-mail opt-in. *Recherche et Applications en Marketing*, *22*(3), 5-26.

Appendix 1

Industry	Open	Click	Soft Bounce	Hard Bounce	Abuse	Unsub
Agriculture and Food Services	25.77%	3.49%	0.68%	0.51%	0.03%	0.28%
Architecture and Construction	25.19%	3.16%	1.66%	1.18%	0.04%	0.37%
Arts and Artists	27.93%	2.99%	0.78%	0.56%	0.03%	0.27%
Beauty and Personal Care	19.50%	2.27%	0.52%	0.50%	0.05%	0.33%
Business and Finance	21.59%	2.93%	0.81%	0.65%	0.03%	0.24%
Computers and Electronics	21.68%	2.62%	1.13%	0.79%	0.03%	0.30%
Construction	22.28%	2.09%	1.73%	1.34%	0.06%	0.44%
Consulting	19.95%	2.58%	1.06%	0.80%	0.03%	0.29%
Creative Services/Agency	23.26%	3.01%	1.22%	0.95%	0.04%	0.36%
Daily Deals/E-Coupons	13.89%	1.96%	0.14%	0.09%	0.01%	0.09%
eCommerce	16.89%	2.62%	0.43%	0.27%	0.03%	0.21%
Education and Training	22.70%	3.01%	0.64%	0.55%	0.03%	0.20%
Entertainment and Events	21.53%	2.45%	0.58%	0.47%	0.03%	0.27%
Gambling	17.66%	2.70%	0.50%	0.57%	0.04%	0.17%
Games	22.14%	3.54%	0.56%	0.60%	0.04%	0.23%
Government	26.88%	3.71%	0.59%	0.48%	0.02%	0.13%
Health and Fitness	23.36%	3.11%	0.53%	0.52%	0.04%	0.37%
Hobbies	29.42%	5.66%	0.39%	0.31%	0.03%	0.23%
Home and Garden	25.77%	4.17%	0.73%	0.51%	0.05%	0.38%
Insurance	19.82%	2.16%	0.77%	0.81%	0.04%	0.22%
Legal	22.50%	3.04%	0.80%	0.66%	0.03%	0.21%
Manufacturing	23.41%	2.67%	1.61%	1.09%	0.04%	0.37%
Marketing and Advertising	18.58%	2.19%	0.88%	0.70%	0.03%	0.28%
Media and Publishing	22.76%	4.75%	0.34%	0.22%	0.01%	0.12%
Medical, Dental, and Healthcare	23.08%	2.73%	0.83%	0.83%	0.05%	0.29%
Mobile	22.09%	2.70%	0.80%	0.74%	0.04%	0.40%
Music and Musicians	22.99%	2.89%	0.70%	0.52%	0.04%	0.30%
Non-Profit	25.66%	2.98%	0.56%	0.47%	0.03%	0.19%
Other	23.39%	3.11%	0.90%	0.69%	0.04%	0.28%
Pharmaceuticals	19.29%	2.71%	0.77%	0.68%	0.03%	0.21%
Photo and Video	27.06%	4.24%	0.81%	0.68%	0.04%	0.39%
Politics	22.83%	2.28%	0.51%	0.48%	0.04%	0.22%
Professional Services	21.25%	2.75%	1.04%	0.81%	0.03%	0.31%
Public Relations	20.52%	1.86%	0.90%	0.69%	0.02%	0.24%
Real Estate	22.08%	2.22%	0.78%	0.65%	0.05%	0.33%
Recruitment and Staffing	20.88%	2.53%	0.68%	0.69%	0.04%	0.32%
Religion	26.59%	3.28%	0.22%	0.20%	0.02%	0.12%
Restaurant	23.95%	1.58%	0.36%	0.30%	0.03%	0.30%
Restaurant and Venue	22.71%	1.47%	0.67%	0.57%	0.04%	0.39%
Retail	22.13%	2.85%	0.45%	0.37%	0.03%	0.28%
Social Networks and Online Communities	22.24%	3.75%	0.43%	0.36%	0.03%	0.23%
Software and Web App	22.47%	2.69%	1.12%	0.89%	0.03%	0.39%
Sports	26.15%	3.60%	0.59%	0.51%	0.03%	0.27%
Telecommunications	20.54%	2.27%	1.20%	0.96%	0.03%	0.26%
Travel and Transportation	20.66%	2.46%	0.76%	0.53%	0.03%	0.24%
Vitamin Supplements	17.12%	2.11%	0.40%	0.33%	0.05%	0.25%

Source: MailChimp, 2015a