t. XXVI, nr 4/2019 DOI 10.33226/1231-7853.2019.4.2

mgr Arkadiusz Zaremba
doktorant na Wydziale Zarządznia Uniwersytetu Warszawskiego
ORCID: 0000-0002-6174-798X
e-mail: arkadiusz.zaremba@gmail.com

Conversion attribution in the online environment — identification of crucial decision path stages. Theory and case study

Atrybucja konwersji w środowisku internetowym — identyfikacja kluczowych etapów decyzyjnych. Teoria oraz studium przypadku

This paper presents a review of marketing theories on the topic of consumer behaviour from the perspective of the consumer and the marketer and its adaptation to the online advertising environment. As previous researches have shown, there is a strong need to verify these theories in practice using real consumer data, and not surveys conducted among students — as most researchers do. In order to show the complexity of online advertising and measurement tools, 565 real online consumer journey paths were analysed using several most popular conversion attribution models. The results confirm that classical decision making processes are still suitable to current consumer behaviour but there exist many difficulties in indicating channels responsible for particular decision making process stages and technology limitations require some further research.

Keywords
conversion attribution, customer journey, online conversion path, online decision making process, impact of online media

Artykuł prezentuje przegląd teorii w zakresie zachowań konsumenckich, wraz z ich adaptacją do warunków środowiska reklamy internetowej, jednocześnie z perspektywy konsumenta oraz marketera. Poprzednie badania naukowe w tym obszarze wskazywały na silną potrzebę weryfikacji teorii zachowania konsumentów w obszarze internetowym poprzez wykorzystanie obserwacji realnych zachowań konsumentów. Aby pokazać złożoność narzędzi reklamy internetowej i sposobów pomiaru, dokonano analizy 565 ścieżek decyzyjnych realnych konsumentów z wykorzystaniem kilku popularnych modeli atrybucji. Wyniki potwierdzają, że klasyczne modele decyzyjne są wciąż zbieżne z obecnymi zachowaniami konsumenckimi w środowisku internetowym, jednakże należy pamiętać o wielu trudnościach związanych z analizą wpływu poszczególnych kanałów reklamowych na podejmowanie decyzji przez konsumenta oraz ograniczeniach technologicznych, co wymaga dalszych badań.

Słowa kluczowe
atrybucja konwersji, ścieżka decyzyjna klienta, ścieżka konwersji online, proces podejmowania decyzji online, wpływ mediów internetowych

JEL: M310

Introduction

According to a leading media agency Zenith, this year will see a transition in time spent on media by people — in 2019, the average customer will spend 170 minutes browsing the Internet which will be equal to time spent watching TV. Next year, the dominant medium will be the Internet (https://www.recode.net/2018/6/8/17441288/internet-time-spent-tv-zenith-data-media, 27.12.2018). Marketers have predicted this trend earlier — in 2017, global ad spending on the Internet has exceeded TV ad spending (https://www.recode.net/2018/3/26/17163852/online-internet-advertisers-outspend-tv-ads-advertisers-social-video-mobile-40-billion-2018, 28.12.2018).

Marketers want to spend money wisely — 46% of them want to broaden their knowledge of media-mix modelling (https://www.emarketer.com/content/why-marketers-see-gaps-in-their-attribution, 02.201.2019). Scientific literature does not help — Pomirleanu, Schibrovsky, Peltier, and Nill, (2013) found that 27% of internet-related marketing articles in top marketing journals were

 t. XXVI, nr 4/2019 DOI 10.33226/1231-7853.2019.4.2

related to consumer behavior, but only 7% regarded the internet search or the decision making path. This figures means that the number of researches of the consumer journey is low and that there is a strong need for additional researches. What is more, Darley, Blankson and Luethge (2010) discovered that only 31% of researches was based on experiment, with the rest based on surveys and only 19% using real shoppers' data.

Merely a few of the above-mentioned articles concerning researches conducted by scientist were related to the phenomenon of conversion attribution. Shao and Li (2011) define conversion attribution as a process and a problem which consists in interpreting the influence of advertisements in regards to the user's decision process. Jayawardane, Halgamuge, and Kayande (2015) explain this topic as a process of "assigning credit to one or more advertising channels for influencing a desirable action."

Recent articles referring to the topic of conversion attribution taken up by Danaher and Dagger (2013), Li and Kannan (2014), Zantedeshi, Feit, and Bradlow (2017), Yadagari, Saini, and Sinha (2015) attempt to explain this subject on the basis of many econometrics models which are not widespread amongst marketers — the majority of marketers make their decisions on the foundation of simplified attribution models involving last-click, first-click, linear, position-based, and time-decay approaches (https://www.emarketer.com/content/five-charts-the-state-of-attribution, 03.01.2019). None of the mentioned articles were focused around analysing differences in the results of particular attribution models widely used by marketers.

The first objective of this article is to verify if the existing widespread models of decision-making and the influence of advertising established in the era of offline media still work in the online environment. From the perspective of the brand, it is crucial to know which online channels are responsible for which part of the decision making process — this article shall also examine this problem since this knowledge would prove very useful for marketers to maximise their return from promotional activities — better media budget allocation (for example spending more money on the channels which begin the decision making process) gives the opportunity of higher profits.

The market comprises several conversion attribution models which provide information about marketing channels' contribution to sales and the research shown in this article is focused on differences in results of the most popular attribution model (on the example of an online platforms selling mutual funds). This research also attempts to recognise the usability of each model to specific needs, like general assessment of the channel, the power of starting or closing the path, keeping the consumer in the decision making process.

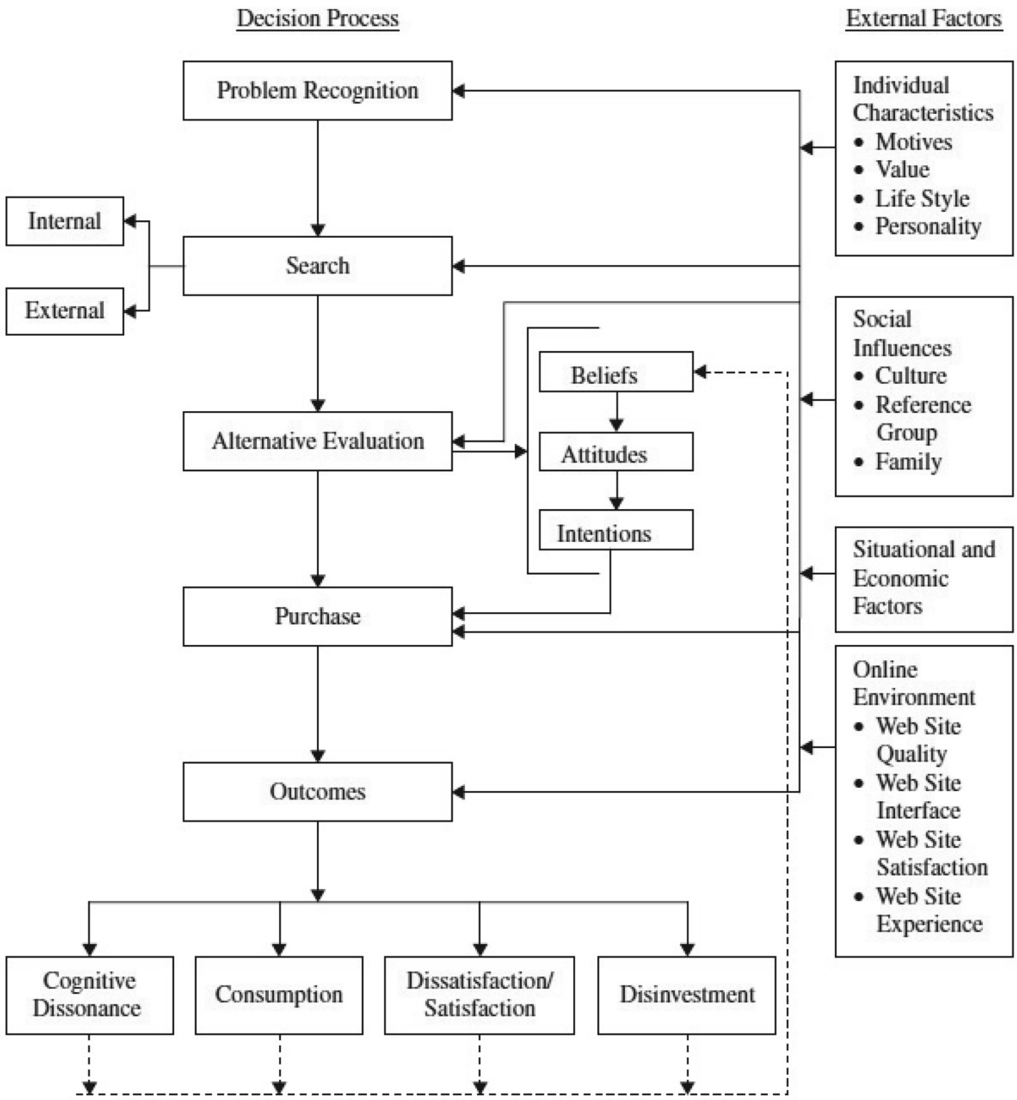
Theoretical background on consumer behaviour

Interactions between the consumer and producer (or marketer) are broadly referred to as consumer behavior (Solomon, 2013, pp. 31–32). Scientist have been trying to quantify the consumer decision journey since at least 1960s (Lemon & Verhoef, 2016), but consumer behaviours are not constant — Peter and Olson (1990, pp. 6–8) state that the process of making purchase decision evolves in time in parallel to economic and social changes. The most broadly spread consumer decision making model is the Engel-Kollat-Blackwell (EKB) model which focuses on five core stages: problem recognition, search, alternative evaluation purchase, choice, outcomes. Darkley, Blankson and Luethge (2010) have adapted the original EKB model (Fig. 1) to the online environment adding several external factors originating from the online environment to exhibit that online consumer's behaviour is a complex phenomenon.

Another common structural decision making model is the one proposed by Howard and Sheth; it attempts to mirror the whole purchase making process. The authors assume that the decision making process is influenced by many external and internal factors, such as beliefs, information, product characteristics which could be easily divided into four groups of variables: inputs (significative stimuli and symbolic stimuli which refer to quality, price, service, availability, distinctiveness and social stimuli referring to family, reference groups and social class which could not be affected by the marketer's activities), perceptual and learning constructs which deal with the psychological variables involved in contemplating a decision, outputs (attention, brand comprehension, attituded, intention) and external variables including personality traits, religion, time pressure. The most important characteristic of this model is the fact, that inputs and outputs are countable (Howard & Sheth, 1969).

An additional analysis of other decision making models, like the above established by O'Shaughnessy, which demonstrates general goals, needs, belief and criteria of the consumer or the model by F.M. Nicosia which distinguishes three major phases (attitude, motivation, behaviour related to shopping) (Rudnicki, 2012, pp. 19–22) shows that general decision making models are divided into three parts: pre-purchase, purchase, post-purchase. The process model for the customer

Figure 1. A modified model of online consumer behaviour and decision making



Source: Darley, Blankson, & Luethge, 2010.

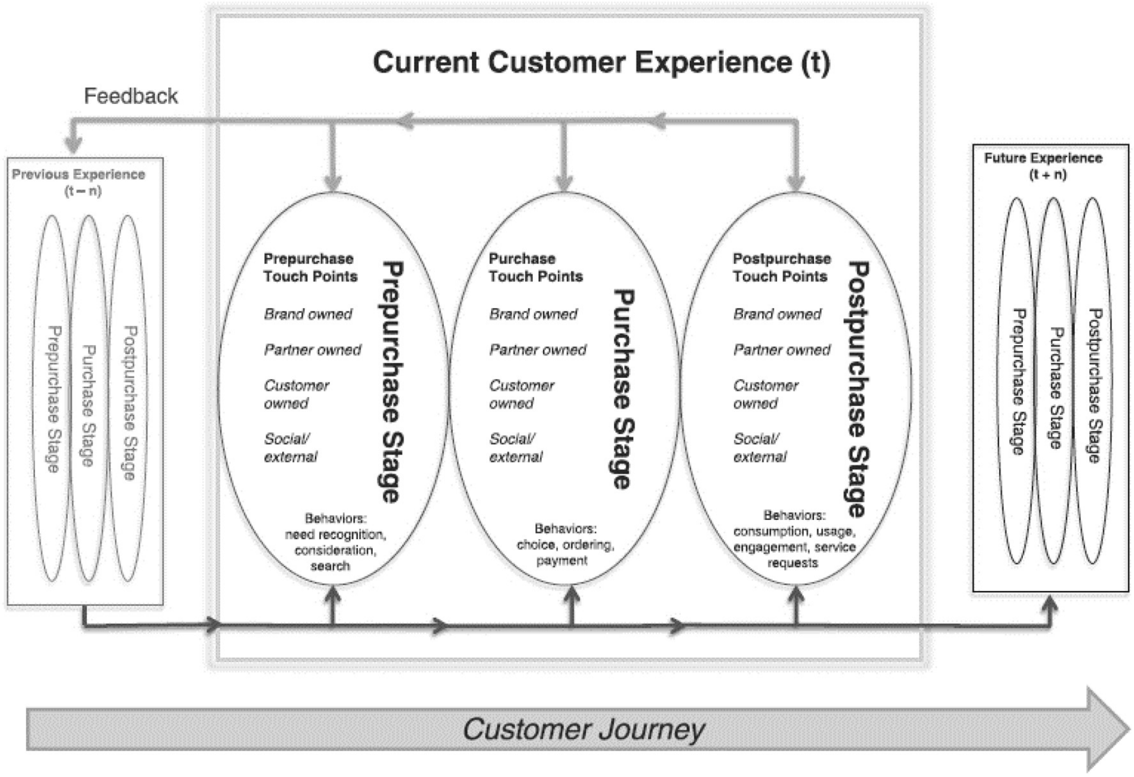
journey and experience described in Figure 2 shows the pre-purchase stage encompasses all aspects before a purchase transaction — interaction with the brand, category, environment, need recognition, search, consideration. The second phase — purchase — includes choice, ordering, payment. The post-purchase phase is related to usage, consumption, engagement, service request (Lemon & Verhoef, 2016).

The marketer's main role is to influence the consumer decision path in order to sell a product by using the marketing-mix tools. The first well-known model of advertising influence was AIDA, whose name stems from four stages of advertising effects: attention, interest, desire, action. The classic AIDA model includes some other modern variants: starting from AICA (attract, interest, convince, action), through AICDA (attract, interest, convince, desire, action) and AIDAS (attract, interest, desire, action, satisfaction) to a more expanded AIDCAS (attract, interest, desire, convince, action, satisfaction) (Gędek, 2013, pp. 475–479). At the same time when the EKB model was invented, Lavidge and Steiner established a new hierarchy of effects model (Lavidge-Steiner model). This model comprise three stages of ad influence: providing facts, changing consumer attitude and finally convincing to purchase (Kall, 2002, p. 22).

Currently, marketers spending money on online media very frequently base on the new consumer behaviour model developed by Lecinski called Zero Moment of Truth. This model extends the traditional approach to the decision making process which is divided into the stimulus phase, purchase (called the first moment of truth) and experience (called the second moment of truth). This extension

t. XXVI, nr 4/2019 DOI 10.33226/1231-7853.2019.4.2

Figure 2. Process model for customer journey and experience



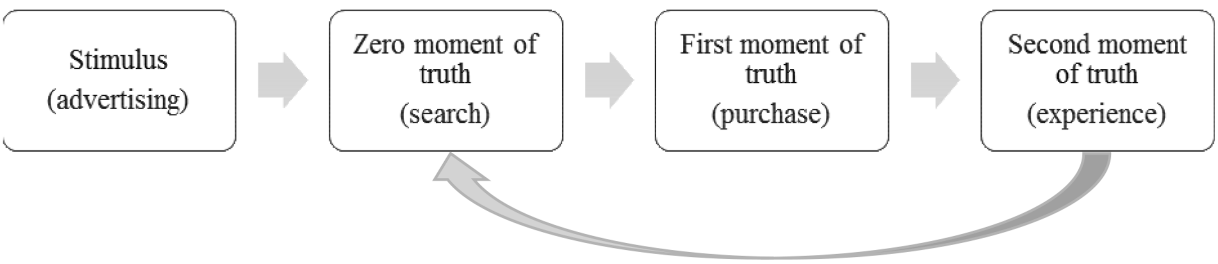
S o u r c e: Lemon & Verhoef, 2016.

Table 1. The Lavidge-Steiner hierarchy of effects model

Area of advertising influence	Stages of customer attitude
Cognitive phase (advertising provides information and facts)	Awareness Knowledge
Affective phase (advertising changes attitude and emotions)	Linking Preference
Conative phase (advertising stimulates or drive desires)	Conviction Purchase

S o u r c e: Kall, 2002, p. 22.

Figure 3. The zero moment of truth approach



S o u r c e: author's own elaboration based on Lecinski, 2011, pp. 15–17.

is done by adding a new stage between stimulus and purchase called the zero moment of truth which refers to the online search (social media, online reviews, WOMM) carried out by customers before buying any product. The search process (the zero moment of truth stage) lasts from a few hours to a few months depending on the industry. Information researched at that stage are created by users in the second moment of truth stage — customers share their experience with other users in the online world (Lecinski, 2011, pp. 15–17).

The online advertising environment, metrics, attribution models and measurement limitations

The review of several decision making models and models of influencing consumers via advertising shows that marketers attempt to affect consumers in particular phases of their decision making process with the use of advertising. In the ecosystem of offline media, marketers mostly attempt to affect the first phase which, depending on the model, is referred to as the stimulus or cognitive phase, or problem recognition, etc. TV, radio, press campaigns are evaluated and measured by indexes like reach, frequency, opportunity to hear/see, intensity (reach x frequency) which mainly focus on the number of potential interactions with the customer. An additional metric is provided by the affinity index which portrays how suitable a particular medium is received by the target group (Gędek, 2013, p. 465).

The Internet allows marketers to measure consumer activities much deeper than traditional media does. Online tracking tools allow to obtain a significantly better perspective — one can measure not only the number of potential interactions of ads with the customer but also the amount of interactions and sales effects.

Online promotion tools are divided into three areas differing by possibility of controlling the scope of information and promotion time as well as the necessity of spending money: owned media, paid media, earned media. Owned media are channels totally manageable by marketers and requiring no money to communicate with customers (except the costs of content production) involving websites, social media channels (Facebook, Instagram, etc.), newsletters, mobile applications available in Google Play and AppStore. Paid media is an area of advertising where marketers pay for promotion, activities are planned and controlled by the marketers. The least controlled and non-paid area are channels which include content and activities generated by the media and online users — some of them are professionals (reporters, influencers) and some are amateurs (opinions, reviews, social media mentions, etc) (Srinivasan, Rutz, & Pauwels, 2016).

In practice, there are many connections between these three areas, especially in search engine optimisation activities which are oriented at improving the accessibility of owned media (for example, a website) in search engine listings but also in the area of earned media in a situation when positive, valuable content was produced by external partners as a review of a product and the marketer wants to promote this material. The same situation occurs in social media — when the marketer publishes a post on Facebook, it is owned media; when he promotes this post by paying for it, this is earned media and when any user shares this post, it is owned media as well. It is very hard as a researcher to recognise if the post was promoted or not (there is no such information provided, and only the marketer it is aware of this fact), therefore it is very difficult to evaluate the social buzz created by the post (the more it is promoted, the more engagement is generated). Influencers also publish content about brands or products and it is not always marked as advertising. Hence, recognising a particular type (owned, paid, earned media) of brand activity in the online environment is a difficult task and requires data and know-how from the brand for thorough assessment.

Marketers have developed many metrics to measure consumer activities and effects of online advertising. As far as the online world is concerned, marketers obtain information on how many users were reached by advertising, how many clicks were generated with a particular web promotion tool (like banners, e-mails) or how many consumer actions turned into conversion (purchase, registration, etc.) in real time. Each type of tool requires separate metrics, but in general, it measures the conversion path from reaching the customer (impression) through interaction (clicks) to sale (conversion) (Bath, Bevans, & Sengupta, 2002). Technology allows to record all users' activities in the right order with the timestamp — these data could be used to imitate and model the customer journey and his decision process from start to end (Anderl, Becker, Wangenheim, & Schumann, 2014).

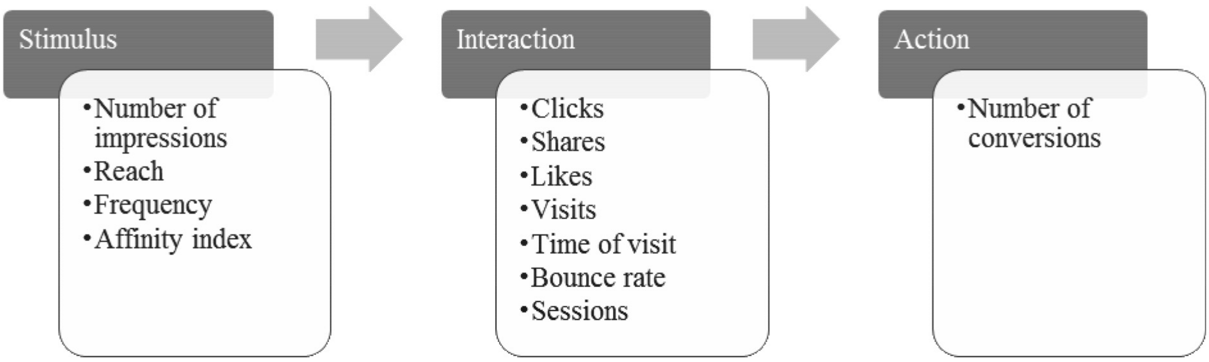
In theory, it is possible to measure almost every activity of the user on the Internet, but in practice, it is very limited due to the type of tracking tools used on the market. According to a survey made by W3Techs, Google Analytics (GA) is the most popular tool used by marketers to track users' activities and advertising effects in the online environment — 55.7% of the websites covering 85.8% of all online traffic use that tool. The second popular tool has only 7.7% market share (https://w3techs.com/technologies/overview/traffic_

Table 2. A general review of digital advertising tools

Area of online promotion	Tool
Owned media	• Website • Blog • Social media channels • Mobile Apps • E-mail marketing/SMS (own database) • Search Engine Optimisation (SEO)
Paid media	• Search Engine Marketing — Pay Per Click (SEM — PPC) • Social Media Ads • Boosted posts • Display • Paid reporters and bloggers • E-mail marketing (external databases) • Affiliate marketing • Video
Earned media	• Social Media — mentions, likes, shares, comments, retweets, etc. • Online Reviews • Word-of-Mouth promotion • Reporters and bloggers writing about business • Search Engine Optimisation

S o u r c e: author's own elaboration based on Srinivasan, Rutz, & Pauwels, 2016, p. 440–453; https://www.titangrowth.com/what-is-earned-owned-paid-media-the-difference-explained, 02.01.2019; Kaznowski, 2007, pp. 118–169.

Figure 4. The stage of customer journey and basic metrics



S o u r c e: author's own elaboration.

analysis/all, 03.01.2019). Google Analytics is an advanced analytic tool with massive options but also many limitations. GA, like other online tracking tools, bases on web cookies (literature name them also cookie files). A web cookie is a file saving data sent by the visited website (or tools implemented on that website, like GA) to the user's computer — every activity of the user on the website could be tracked and described in cookie files — time of entrance, duration of visit, download of attachments, displayed banner, etc. Every user is given a few cookies responsible for the user's identification, passwords, particular activities.

To measure correctly all activities, the GA tracking code (or other tracking codes) should be implemented on all websites containing brand content (paid and non-paid) which is in practice impossible — if the marketers do not pay for publication, the publisher does not want to implement marketers tracking tool. So, the first problem in measuring the customer journey is the fact that this journey cannot not be measured across all media, but only among owned and paid media, with the excluded of earned media.

The second inaccuracy is user recognition. Every internet browser receives its own cookies, so if anyone visits the same page using two browsers, his computer saves two cookies responsible for user identification. Tracking tools interpret those cookies as separate users, which is why cross-device analysis is so difficult. There is a solution for some marketers, namely a login option on their website used to recognise the user as the same person on many devices called UserID, but it works only when the user logs into his account, so it does not work if he is just visiting the website.

GA does not allow to track impressions — only clicks and further activities on the customer journey path can be tracked (see Figure 4). It is still useful but disregards the effect of many tools building brand awareness — when a video ad is displayed to the user on YouTube and he then visits the brand's website directly via the Google search engine, GA is only aware of the visit, without any trace of the video, while the video in fact generated the visit (the same problem occurs when the user reads an article on an external website and then visits the brand's page) (https://support.google.com/analytics/ with subcatalogues, 27.12.2018).

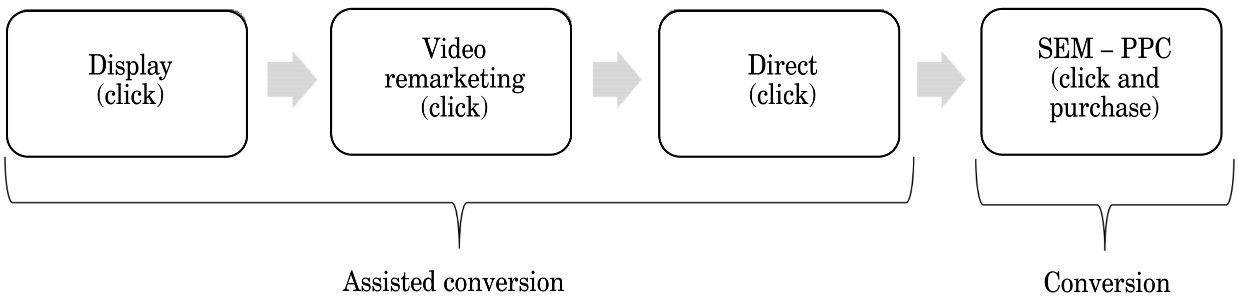
From the marketer's perspective, mapping the customer journey is just the beginning, the crucial issue is to effectively allocate the media budget and create a proper media-mix — according to Kantar research, 46% of marketers perceive marketing mix modelling as a gap (https://www.emarketer.com/content/why-marketers-see-gaps-in-their-attribution, 03.01.2019). Currently, marketers use many conversion attribution models: 43.2% of them use the first-click model, 35.6% use a position-based model, 24.5% base on the last-click model, a little bit less, 24.2% use the time decay model and 23.2% base on the linear model (https://www.emarketer.com/content/five-charts-the-state-of-attribution, 03.01.2019).

The explanation of the conversion attribution concept is based on the example shown in Figure 5. A user saw an ad which attracted his attention, he therefore clicked on the banner and read information about the product contained on the website. Because the user did not convert, he was shown the video by the marketer using remarketing tools which also made the user interested enough to click and read about the product once more. In the next step, the user visited the brand's website directly without making a purchase. Finally, the user purchased a product after clicking on a sponsored link in the Google search engine. The whole journey consists of four touchpoints. The key question is to how to assign credit for the purchase of particular touchpoints on the path.

It is certainly PPC that has led to conversion — this approach is used in the last-click model which only credits the last touchpoint regardless of the customer journey's length and the potential influence of other channels. An opposite approach is shown in the first-click model which recognizes the most important channel as the one which put the user on the path directing to purchase and only credits the first touchpoint. The assumption of the linear model is that every touchpoint is equal, so display, video, direct and PPC get a 25% share in success in the described example. Touchpoints taking part in success but not directly leading to conversion are called assisted conversions. Time-decay and position-based models base on the theory that the closer to conversion, the more important the touchpoint. The first mode focuses on the time to conversion, the second one on the order of touchpoints on the conversion path. All these types of attribution models are available in Google Analytics and the assumptions of credit decay in multi-touch models are shown in Table 3 (time-decay, position-based).

According to the theories of decision making described in this article, the last-click model is not linked with any decision making model or hierarchy of effects model as it ignores the pre-purchase phases. A part of the pre-purchase phase is included in the first-click model, however this approach

Figure 5. An example of the user's conversion path/customer journey



S o u r c e: author's own elaboration.

t. XXVI, nr 4/2019 DOI 10.33226/1231-7853.2019.4.2

focuses only on one touchpoint and ignores further steps of the user. Multi-touchpoint models including linear, position-based and time-decay are generally in line with the marketing theory as they include all touchpoints; however, the subjective approach of crediting particular touchpoints by Google Analytics, which is the most common online traffic analytics tool, is not relevant for all types of activities.

On August 30, 2018, a new online platform selling mutual funds in Poland was launched — in the period of August 30, 2018 to December 31, 2018, the platform acquired 565 customers who bought at least one product. Customers were acquired through a media campaign divided into two major blocks: Public Relations activities (PR) including buzz generated in the blogosphere, media portals and regular push activities including

Table 3. Popular attribution models and their general principles in theory and as per the assumptions of Google Analytics

Attribution model	Attribution rules	Google Analytics rules
Last-click	100% credit to last touchpoint	
First-click	100% credit to touchpoint opening the customer journey	
Linear	Touchpoints equal — percentage of credits goes to all touchpoint in the same proportions	
Time-decay	Credit to all touchpoints, touchpoint closer in time to conversion gets more credits	Credit to all touchpoints based on expotential decay; this model has a default *half-life* of 7 days, meaning that a touchpoint occurring 7 days prior to conversion will receive 1/2 the credit of a touchpoint that occurs on the day of conversion. Similarly, a touchpoint occurring 14 days prior will receive 1/4 of credit of a day-of-conversion touchpoint
Position-based	Credit to all touchpoints, touchpoint closer to conversion gets more credits	40% credit is assigned to each first and last touchpoint, the remaining 20% is distributed equally to the middle interactions

S o u r c e: https://support.google.com/analytics/answer/1662518?hl=pl, 04.01.2019.

This fact has led many scientists to create individual attribution models based on many advanced statistical methods such as: bagged logistic regression model, Bayesian linear regression, Hidden Markov Model (Abhishek, Fader, & Hosanagar, 2012), however, these sophisticated models are not widely used due to their complexity and required time of preparation. Marketers are changing their media-mix on a daily basis, which is why they need daily attribution reporting provided by online tools like Google Analytics.

Comparison of results for the most popular attribution models

Marketers allocate their budgets on the basis of the chosen attribution models. The goal of this white paper is to check the size of differences in results of crediting particular touchpoints in different attribution approaches on the example of a financial product.

a pay-per-click campaign in the Google search engine, display, Social Media ads. The research is focused on conversion paths of the first transactions.

The campaign generated at least 6.4 million impressions (information about the number impressions stems from tools like Facebook Ads, Google Ads, Google Bid Manager) which generated 28,104 clicks by 25,418 users (or unique cookies, to be precise, as described in the section about GA limitations) — the number of clicks and users was counted in GA.

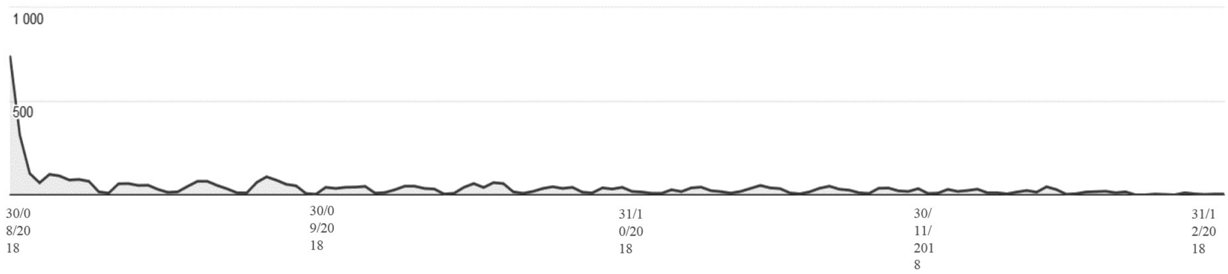
PR activities were ran very broadly and articles about the new product and new platform were published on several biggest websites in Poland. The number of the articles' impressions in unknown, but all of them were published on the same day, August 30, 2018, which explains direct traffic on that date — the brand was completely unknown and it was almost impossible to generate such a big volume of website visits without any promotion.

Table 4. Media activities of the researched brand

Area	Medium	GA shortcut	# Impressions	# Clicks	# Users
Owned Media	WWW	Direct	N / A	N / A	2 914
	WWW — SEO	Organic search	N / A	N / A	4 655
Paid Media	SEM — PPC	Paid search	167 892	15 678	12 313
	Display	Display	3 567 987	10 679	6 774
	Social Media ads	Social	2 678 998	982	458
Earned Media	Reporters and bloggers	Referral	N / A	765	658
TOTAL					25 418

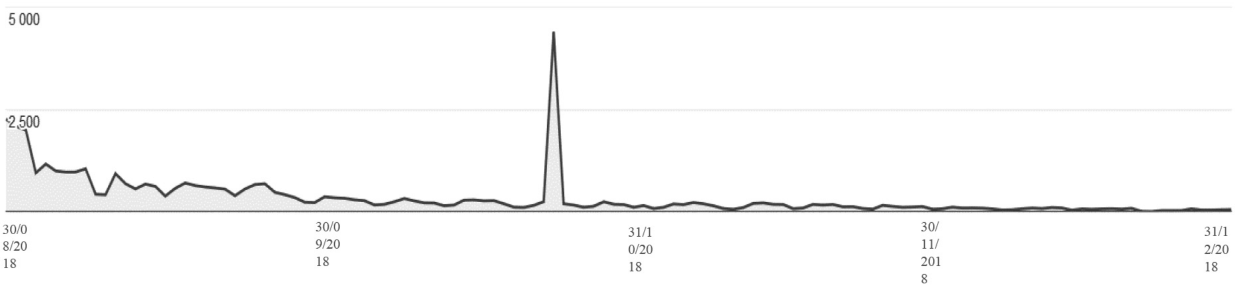
Source: author's own elaboration on the basis of conducted research.

Figure 6. Direct traffic at the time of the campaign (number of sessions, daily)



Source: author's own elaboration on the basis of conducted research.

Figure 7. Whole traffic at the time of the campaign



Source: author's own elaboration on the basis of conducted research.

An analysis of all purchase paths shows that 42.2% of users converted in the same day (less than 24 hours) when they got an information about the product and 37.52% of them needed from 12 to 90 days to make a decision. Only 26.02% of buyers purchased the product after one visit on the brand's website, 45.65% needed from 2 up to 10 interactions. A total of 565 customers purchased the product in 287 different ways — this is the number of unique conversion paths. The most popular conversion path appeared 68 times — the top 10 conversion paths are responsible for 37.7% of all conversions. The total number of touchpoints in all conversion paths is 3,431 (6.07 per path).

In the decision making theory, it is very important to stimulate the customer to start considering the purchase and to finalise sales. Table 8 shows the amount of conversions started by a particular

t. XXVI, nr 4/2019 DOI 10.33226/1231-7853.2019.4.2

Table 5. Conversions time lag (all paths)

Time lag in days	Number of conversions	%
0	238	42.12
1	23	4.07
2	24	4.25
3	11	1.95
4	11	1.95
5	10	1.77
6	16	2.83
7	5	0.88
8–11	15	2.65
12–30	89	15.75
31–60	57	10.09
61–90	66	11.68
TOTAL	565	100.00

Table 6. Paths length (all paths)

Path length (number of touchpoints)	Number of conversions	%
1	147	26.02
2	78	13.81
3	59	10.44
4	29	5.13
5	17	3.01
6	31	5.49
7	13	2.30
8	9	1.59
9	14	2.48
10	8	1.42
11	15	2.65
12+	145	25.66
TOTAL	565	100.00

S o u r c e: author's own elaboration on the basis of conducted research.

Table 7. Twenty most common paths

Path	Length	Number of occurrences	%
direct	1	68	12.04
display	1	44	7.79
direct > direct	2	21	3.72
organic search	1	21	3.72
direct > direct > direct	3	15	2.65
paid search > organic search	2	10	1.77
direct > referral	2	9	1.59
display > direct	2	9	1.59
direct > direct > direct > direct	4	8	1.42
display > direct > direct	3	8	1.42
referral	1	7	1.24
organic search > direct	2	7	1.24
organic search > direct > direct > direct > direct > direct	6	6	1.06
paid search > organic search > referral	3	5	0.88
direct > direct > direct > direct > direct > direct > direct	7	5	0.88
organic search > referral	2	5	0.88
paid search	1	4	0.71
paid search > organic search > direct	3	4	0.71
paid search > organic search > direct > direct > direct > direct	6	3	0.53
direct > direct > referral	3	3	0.53

S o u r c e: author's own elaboration on the basis of conducted research.

channel and finalised. The table also involves data regarding assisted conversions (all conversions on paths longer than 2 touchpoints except the last touchpoint). In the first-click and last-click, direct traffic generated most conversions. Organic search, display, paid search and social media generated more first-click conversions than last-click which means that these sources start the customer journey more often than they finish it. Marketers may conclude that they should allocate a bigger budget to these activities to create more paths which could be finished on other channels. A channel named 'other' is a small set of sources containing test activities and, due to a small number of conversion analyses generated in this channel, any associated conclusions have been omitted in this article.

The analysis of the ratio of assisted conversion to last-click conversion proves to be very useful as it shows how many additional conversions generated a particular source — paid search generates almost 14 assisted conversions per one last-click conversion which means that investing in the media budget in this channel brings direct and non--direct purchases.

Table 8. A comparison of first-click and last-click attribution models (all paths)

Channel	First-click (F-C)	Assisted conversions (AC)	Last-click (L-C)	L-C toF-C dev.	AC / L-C ratio	L-C / F-C ratio
Direct	273	339	392	43.59%	0.86	1.44
Referral	17	122	56	229.41%	2.18	3.29
Organic search	74	167	51	–31.08%	3.27	0.69
Display	98	74	51	–47.96%	1.45	0.52
Paid search	90	122	9	–90.00%	13.56	0.10
Social Media	11	14	1	–90.91%	14.00	0.09
Other	2	14	5	150.00%	2.80	2.50

S o u r c e: author's own elaboration on the basis of conducted research.

The results of multi-touch models are shown in Table 9 — the numbers are not integers due to the adopted methodology described in Table 3. The linear model treats all touchpoints in the same way, so deviations of position-based and time-decay models are related to the linear model. In the case of the position-based model, the highest absolute deviation is observed on social media, but due to a small volume of transactions, this result should be omitted. The second highest deviation is represented by paid search which is higher by 28% in relation to the results in the linear model. Having knowledge about methodology, it may be said that paid search more often occurs at the start or at the end of the path than in the middle.

The time-decay model also brings different results than the linear model but differences are generally lower than in the comparison of position-based versus linear. The highest essential deviation occurs in paid search but the difference is only 12%. The paid search channel is also represented by the highest deviation in the comparison of position--based and time-decay.

But what is the general average difference in results of the position-based and time-decay models to the linear model? To answer this question, the modulo of deviation waged by the share of conversion generated by each source in linear model was counted for the presented models:

$$AD = \sum_{k=1}^{n} \binom{n}{k} \left| \frac{C_a}{C_b} \right| \times S_b ,$$

AD — average deviation of channels between model a and b,
n — number of marketing channels,
C_a — number of conversions generated in channel n in model a,
C_b — number of conversions generated in channel n in model b,
S_b — share of channel n in total number of conversions in model b.

Table 9. Comparison of linear, position-based, time-decay attribution models (all paths)

Channel	Linear (L)	Position-based (P-B)	P-B to L dev.	Time-decay (T-D)	T-D to L dev.	P-B to T-D dev.
Direct	357.22	339.43	–4.98%	369.16	3.34%	–8.05%
Referral	35.02	36.12	3.14%	35.37	1.00%	2.12%
Organic search	63.60	63.42	–0.28%	59.33	–6.71%	6.89%
Display	65.23	71.67	9.87%	63.66	–2.41%	12.58%
Paid search	35.44	45.45	28.24%	30.86	–12.92%	47.28%
Social Media	3.14	5.11	62.74%	2.36	–24.84%	116.53%
Other	5.33	3.79	–28.89%	4.26	–20.08%	–11.03%

S o u r c e: author's own elaboration on the basis of conducted research.

t. XXVI, nr 4/2019 DOI 10.33226/1231-7853.2019.4.2

The position-based model brings in general 6.91% average deviation per channel in comparison to the linear model, and time-decay brings a 4.35% difference. These values are much lower than in the general difference between last-click to first-click which is equal to a significant 57%.

Table 10. The average deviation between pairs of attribution models (all paths)

	F-C	L-C	L	P-B	T-D
F-C	–	56.99%	37.37%	30.92%	41.34%
L-C	56.99%	–	19.74%	26.07%	15.65%
L	37.37%	19.74%	–	6.91%	4.35%
P-B	30.92%	26.07%	6.91%	–	10.69%
T-D	41.34%	15.65%	4.35%	10.69%	–

S o u r c e: author's own elaboration on the basis of conducted research.

Paths whose length is equal to one assign credit for conversion to this individual touchpoint 1 in every approach: last-click, first-click, linear, position-based and time-decay. To remove this effect, only paths including two and more touchpoints should be considered. The results of the analysis are shown in Table 11 — in comparison to the results from Table 8, this shows an increase of the last-click to first-click ratio in the direct channel and a decrease in other channels — last-click to first-click ratio presents information about the impact of a particular channel on the beginning or closing conversion path — the higher value of this metric, the bigger the impact on closing the path, and the lower the ratio, the lower the impact.

Still, the number of first-click conversions in a situation of the company's launch is striking — brand awareness is close to zero (and in general, direct traffic is generated by customers who know the brand/website). The answer is found in data presented in Figure 6 — at the start of the campaign, many PR activities (reporters, bloggers writing about the platform) from earned media are ran and with great probability generate some direct traffic created as an effect of the content read (it means that in reality, the conversion path was started before the first direct visit to the brand's website). As additional analysis was carried out based on this assumption — the results are presented in Table 12.

Table 11. A comparison of first-click and last-click attribution models (path length > 1)

Channel	First-click (F-C)	Assisted conversions (AC)	Last-click (L-C)	AC / L-C ratio	L-C / F-C ratio
Direct	205	339	324	1.05	1.58
Referral	10	122	49	2.49	4.90
Organic search	53	167	30	5.57	0.57
Display	54	74	7	10.57	0.13
Paid search	86	122	5	24.40	0.06
Social Media	10	14	0	–	0.00
Other	0	14	3	4.67	–

S o u r c e: author's own elaboration on the basis of conducted research.

Having the results of five attribution models, it is very difficult to conclude which one fits the brand best. Differences in results between one-touch (last-click and first-click) and multitouch models (linear, position-based, time-decay) are fundamental. However, analysing conversion attribution from the perspective of a few models is helpful to assess the position of a particular channel on the path — some channels occur more often on the start of the path, some in the middle, and some at the end.

Due to the linear nature of the decision making process, the results presented in Table 8 should be analysed more broadly — assuming that each channel is responsible mostly for one function (opening, closing and maintaining the conversion path), direct traffic responsible for 273 first-click and 392 last-click conversions stands in contradiction to this hypothesis.

Table 12 shows (mostly with high precision) which channels are responsible for which function — direct is clearly a channel which closes and maintains the path, referral is mainly responsible for closing, organic search more often begins, rather than closes, and display, paid search and social media are mainly channels which begin the path.

On the other hand, logical assumption refers to the conversion path time lag — if the user purchased the product in a time shorter than 24

Table 12. A comparison of first-click and last-click attribution models (path length > 1, first touchpoint non-direct)

Channel	First-click (F-C)	Assisted conversions (AC)	Last-click (L-C)	AC / L-C ratio	L-C / F-C ratio
Direct	0	134	144	0.93	–
Referral	10	47	34	1.38	3.40
Organic search	53	137	25	5.48	0.47
Display	54	61	5	12.20	0.09
Paid search	86	98	4	24.50	0.05
Social Media	10	10	0	–	0.00
Other	0	0	1	0.00	–

S o u r c e: author's own elaboration on the basis of conducted research.

Table 13. Length of paths (time lag < 1)

Path length (number of touchpoints)	Number of conversions	%
1	147	61.76
2	58	24.37
3	24	10.08
4	4	1.68
5	2	0.84
6	3	1.26
7+	0	0.00
TOTAL	379	100.00

S o u r c e: author's own elaboration on the basis of conducted research.

four touchpoints and the average length of the conversion path of that users is almost four times shorter than the average path length of all paths researched (1.59 to 6.07).

Paths included in Table 13 except the one started by direct traffic are analysed from the perspective of the first-click and last-click models; the results are shown in Table 14.

Results presented in Table 14 are almost similar to the results from Table 12 (path length >1, first touchpoint non-direct), and channels clearly closing and clearly beginning the path have been separated.

Table 14. A comparison of the first-click and last-click attribution models (time lag > 0, first touchpoint non-direct)

Channel	First-click (F-C)	Assisted conversions (A-C)	Last-click (L-C)	AC / L-C ratio	L-C / F-C ratio
Direct	0	124	118	1.05	–
Referral	8	42	18	2.33	2.25
Organic search	39	111	10	11.10	0.26
Display	35	41	3	13.67	0.09
Paid search	63	74	4	18.50	0.06
Social Media	8	8	0	–	0.00
Other	0	0	0	–	–

S o u r c e: author's own elaboration on the basis of conducted research.

hours, there is a possibility that other activities not tracked by GA (like earned media) had a strong impact on the customer's final decision. An analysis of 238 conversions generated in less than 24 hours of the first click interaction with the brand shows that 97% of paths are shorter than

Conclusions, implications, limitations and future research

Conversion attribution does not give clear answers — a review of the results of particular

t. XXVI, nr 4/2019 DOI 10.33226/1231-7853.2019.4.2

models sometimes brings more questions than answers — results are varied, especially single-touch models (last-click, first-click) in comparison to multi-touch (linear, position-based, time-decay). The research results show that it is almost impossible to state which attribution model is the best — each is fit for different purposes and a different type of analysis. The first-click model is recommended when looking for best sources which impact the beginning of the conversion path. but with reservation that purification of data from the external environment is required, especially in the direct traffic channel which opens many paths through the analysis of all paths, which is, however, in contradiction to decision making and brand building theories and finally, requires additional assumptions dependent on the earned media activities — and there is no single rule; rules must be prepared individually.

The last-click model mostly provides information about path-closing channels, but still the analysis should be done on paths including more than one touchpoint (in paths whose lengths are equal one, it is obvious that the first source closes the transaction, hence this data is worthless in the meaning of multi-touchpoint paths). Multi-touchpoint models are useful in obtaining a general view on the usability of each channel.

It is very important to remember about technology limitations during analysis of the results — users are not always "unique" persons — if the person generates two transactions using two different devices (mobile, desktop), it is counted as two different users. This problem, called the cross-device attribution, is well-known but still, there is no tool guaranteeing 100% certainty of real user/person identification and any researches attempting to solve this technological problem are needed.

As the research results show, there is a problem of the analysis of paths whose tracking tools cannot cover the activity of users who purchase a product after reading content in the sphere called earned media are not included, therefore it very hard to assess the influence of earned media. There are some ways of estimating this traffic taking into account increases or decreases in direct traffic, but in the world of thousands of a brand's mentions in earned media, it is impossible to evaluate the impact of separate earned media channels. In the future, the results of similar research may include some polls (for example, using the CAWI method) which may explain the impact of particular earned media sources.

What is more, the customer journey path could start in an area related to the product category but not to the product itself (for example, the user is looking for a financial product and starts searching from his current bank site and then searches for other solutions — this situation is described in the EKB model), but currently it is practically impossible, to use the existing online tracking tools without special permission from the user (for example. an addon could be installed in the web browser and register all activities of consumer).

The results of this research shows that online consumer behaviour is in line with the review models of decision making as well as models of advertising influence — the behaviour of online users can be measured, and with the use of conversion attribution, researchers are able to classify particular channels as the ones taking part in particular phases of the decision making process.

References

Abhishek, V., Fader, P., & Hosangar, K. (2012). *Media Exposure through the Funnel: A Model of Multi-Stage Attribution*. http://dx.doi.org/10.2139/ssrn.2158421

Anderl, E., Becker, I., Wangenheim, F. V., & Schumann, J. H. (2014). *Mapping the Customer Journey: A Graph-based Framework for Online Attribution*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2343077

Bath, S., Bevans, M., & Sengupta, S. (2002). Measuring Users' Web Activity to Evaluate and Enhance Advertising Effectiveness. *Journal of Advertising*, *31*, 97–106. https://doi.org/10.1080/00913367.2002.10673679

Danaher, P. J. & Dagger, T. S. (2013). Comparing the Relative Effectiveness of Advertising Channels: A Case Study of Multimedia Blitz Campaign. *Journal of Marketing Research*, *50*(4), 517–534. https://doi.org/10.1509/jmr.12.0241

Darley, W. K., Blankson, C., & Luethge, D. J. (2010). Toward an Integrated Framework for Online Consumer Behavior and Decision Making Process: A Review. *Psychology and Marketing*, *27*(2), 94–116. https://doi.org/10.1002/mar.20322

Gędek, M. (2013). *Reklama. Zarys problematyki*. Lublin: Wyd. KUL.

Howard, J. A. & Sheth, J. (1969). The Theory of Buyer Behaviour. *Journal of the American Statistical*, (January), 467–473.

Jayawardane, C. H. W., Halgamuge, S. K., & Kayande, U. (2015). *Attributing Conversion Credit in an Online Environment: An Analysis and Classification. Proceedings from the 3rd International Symposium on Computational and Business Intelligence (ISCBI)*. Retrived from https://www.researchgate.net/publication/304294477_Attributing_Conversion_Credit_in_an_Online_Environment_An_Analysis_and_Classification

Kall, J. (2002). *Reklama*. Warszawa: PWE.

Kaznowski, D. (2007). *Marketing w Internecie*. Warszawa: Difin.

Lecinski, J. (2011). *Winning the Zero Moment of Truth*. Google.

Lemon, K. N. & Verhoef, P. C. (2016). Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing: AMA/MSI Special Issue*, *80*, 69–96. https://doi.org/10.1509/jm.15.0420

Li, H. A. & Kannan P. K. (2014). Attributing Conversions in a Multichannel Online Marketing Environment: An Empirical Model and a Field Experiment. *Journal of Marketing Research*, *51*(1), 40–56. https://doi.org/10.1509/jmr.13.0050

Peter, J. P. & Olson, J. C. (1990). *Consumer Behavior and Marketing Strategy*. Boston: Irwin.

Pomirlenau, N., Schibrovsky, J. A., Peltier, J., & Nill, A. (2013). A Review of Internet Marketing Research over the Past 20 Years and Future Research Direction. *Journal of Research in Interactive Marketing*, *7*, 166–181. https://doi.org/10.1108/JRIM-01-2013-0006

Rudnicki, L. (2012). *Zachowania konsumentów na rynku*. Warszawa: PWE.

Shao, X. & Li, L. (2011). *Data-driven Multi-touch Attribution Model. Proceedings from the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Diego. https://doi.org/10.1145/2020408.2020453

Solomon M. R. (2013). *Consumer Behavior. Buying, Having and Being*. Harlow: Pearson.

Srinivasan, S., Rutz, O. J., & Pauwels, K. (2016). Paths to and off Purchase: Quantifying the Impact of Traditional Marketing and Online Consumer Activity. *Journal of the Academic Marketing Science*, *44*, 440–453. https://doi.org/10.1007/s11747-015-0431-z

Yadagiri, M. M., Saini, S. K., & Sinha, R. (2015). A Non-parametric Approach to the Multichannel Attribution Problem. *Web Information Systems Engineering, 9418*, 338–352.

Zantedeschi, D, Feit, E. M., & Bradlow, E. (2017). Measuring Multi-channel Advertising Response. *Management Science*, *63*(8), 2706–2728. https://doi.org/10.1287/mnsc.2016.2451