On the value of pageviews as proxies for audience interest in news: A Relevance Theory approach

Carlos Chernij

A Thesis

In

The Department

Of

Journalism

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Arts (Journalism Studies) at
Concordia University
Montreal, Quebec, Canada

April 2020

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Carlos Chernij

Entitled: On the value of pageviews as proxies for audience interest in news: A Relevance Theory approach

and submitted in partial fulfillment of the requirements for the degree of

Master of Arts (Journalism Studies)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

Dr. Elyse Amend _____ Chair

Dr. Andrea Hunter _____ Examiner

Dr. Roberto G. de Almeida _____ Examiner

Dr. David Secko _____ Supervisor

Approved by _____

Dr. David Secko – Chair of Department

_____

Dr. André Roy – Dean of Faculty

Date _____

ABSTRACT

On the value of pageviews as proxies for audience interest in news:
A Relevance Theory approach

Carlos Chernij

This thesis discusses the common assumption that the quantity of pageviews a news article gets is proportional to the audience's interest in its topic. It proposes instead that the decision to click on a headline is directly proportional to the interest in the topic it suggests, and inversely proportional to the cognitive effort required to find a personal context in which the expected information is likely to be useful. The theoretical linkage between these three concepts is established by the application of Sperber and Wilson's (1986/1995) Relevance Theory. The validity of the hypothesis is discussed by fitting the observations made by Kormelink and Meijer (2018) regarding audience members' rationales for clicking or not clicking on a news article, and by Boczkowski and Mitchelstein (2013) regarding the gap between editors' and audiences' content choices. It also suggests further applications of Relevance Theory in Journalism Studies, particularly concerning the development of metrics and key performance indicators for news analytics.

Acknowledgments

This thesis would not have been possible without the endless patience of my supervisor, Dr. David Secko. I'd also like to also express my gratitude to Dr. Mike Gasher, who endured long discussions in a time when this thesis was just a hazy idea. While many people supported me throughout this project, it was you both who made it ultimately possible. For that, my deepest appreciation.

Table of contents

List of tables

1. Introduction

The availability of web metrics data on individual news articles raises the question of whether the number of pageviews an item gets correlates with the interest of audiences in this type of content. In some research concerning different aspects of web metric data and its influence on journalistic practices, the existence of this correlation underlies some researcher and journalist assumptions (Anderson, 2011; Boczkowski & Mitchelstein, 2013; Ferrer-Conill & Tandoc Jr., 2018; Lee et al., 2014; MacGregor, 2007; Singer, 2011; Tenenboim & Cohen, 2015; Tewksbury, 2003; Welbers et al., 2016; Zamith, 2017). However, other research points out that audience members do not agree with this assumption when asked about it, saying that sometimes they also click on items that do not necessarily interest them, and other times they do *not* click on content considered interesting or important (Kormelink & Meijer, 2018).

This apparently contradictory behaviour, combined with the observed audience tendency to click more on topics such as entertainment, sports and crime (Anderson, 2011; Boczkowski & Mitchelstein, 2013; Hamilton, 2004; Tenenboim & Cohen, 2015; Tewksbury, 2003), makes it difficult to judge the performance of news articles. This is particularly apparent for public-affairs topics, which are expected from and valued by journalists but usually generate considerably fewer pageviews (Boczkowski & Mitchelstein, 2013, p. 6). Editors and journalists are therefore unsure about how to interpret and react to these metrics (Graves, Kelly & Gluck, 2014, p. 12; Petre, 2015). The situation is made worse by traffic-centric newsroom goals being adopted by either their employers or their competitors, such as reaching a certain number of unique visitors per month (Usher, 2012, p. 1908) or each journalist being made responsible for attracting a certain minimum amount of traffic (Petre, 2015, p. 18).

The correct interpretation of web metrics, while taking into consideration journalism's particularities, is the goal of *editorial analytics* (Cherubini & Nielsen, 2016) — an emerging field of study that faces the same challenge as general web analytics: raw data only shows what people *did*, not what they *intended* to do or whether they got what they wanted (Kaushik, 2010, p. 5, Cherubini & Nielsen, 2016, p. 37). As Petre points out, "a number on its own does not mean anything without a conceptual framework with which to interpret it" (2015, p. 8).

This research project proposes to explore the application of Sperber and Wilson's Relevance Theory (RT) as a suitable framework to discuss to what extent pageviews can be reliable as indicators of audience interest in the content. A linguistics theory based on cognitive psychology and with a detailed mechanism, RT aims to explain the process employed by humans to make sense of messages. As articulated by Sperber and Wilson, RT makes some strong claims about the priorities of our cognition (Sperber & Wilson, 1986/1995). Such priorities could lead to cognitive biases which could explain some of the seemly conflicting clicking behaviour of news audiences (Boczkowski & Mitchelstein, 2013; Kormelink & Meijer, 2018).

Statement of purpose

The present research will build a theoretical model of audiences' clicking behaviour based on Relevance Theory to discuss the value of pageviews as a proxy for interest in content. Specifically, it will argue that the decision to click on a headline[1] depends not only on the

[1] Images are often part of headline blocks, and visual appeal was listed by Kormelink and Meijer as one of the factors in the decision of clicking or not to clicking on a link (2018, p. 678). While some cognitive aspects of RT could be applicable to images, it was not Sperber and Wilson's focus when developing the theory. Consequently, the present research will deal only with the linguistic aspects.

audience member's interest in the subject it suggests, but also on the cognitive effort required to find a personal context in which the expected content is likely to yield useful conclusions.

This model will be applied to observations made by Kormelink and Meijer's study (2018) in which they interviewed 56 people about their considerations when choosing to click or not to click on news articles, and to Boczkowski and Mitchelstein's research (2013) about the differences in content choices made by audience members (based on "most viewed" lists) and news sites editors (based on the story placement on the page). This will provide additional explanations to their observations and new nuances to their conclusions.

Furthermore, it will discuss the use of RT in potential applications in improving the interpretation of audience behaviour in Journalism Studies, and in developing performance indicators for news content that are more meaningful for both journalist and publishers.

2. Literature review

Editors and journalists

Academic surveys observed that while both journalists and audience members tended to agree on what subjects were more important or newsworthy — generally public-affairs topics such as politics, economy, education, health and environment — journalists believed that non-public affairs such as entertainment, sports and crime were more popular (Atkin et al., 1983; Atwood, 1970; Wurff & Schoenbach, 2014). But in the analog era, backed by successful circulation figures and lacking detailed viewership data that challenged their choices, journalists were largely unaffected by this kind of feedback and could follow their personal and professional beliefs when they felt that a particular story was "needed" by the audience even if not necessarily "wanted" (Atkin et al., 1983; Atwood, 1970; Gans, 1979; Lee et al., 2014).

The introduction of web metrics data brought some support for journalists' suspicion that non-public-affairs content generally gets more audience attention than public affairs (Anderson, 2011, p. 559, Boczkowski & Mitchelstein, 2013; Hamilton, 2004, p. 194, Tenenboim & Cohen, 2015, p. 2012; Tewksbury, 2003). It also allowed advertisers and publishers to negotiate ad prices based on the number of times they are actually shown to audience members, making advertising revenue dependent on the viewership of individual items instead of the edition as a whole, and incentivizing publishers to deliver impressions as quickly as possible (Graves, Kelly & Gluck, 2014). For editors and journalists deciding what content will be produced, this creates the challenge of balancing long-established professional news values with a story's potential for pageviews (Boczkowski & Mitchelstein, 2013; MacGregor, 2007; Petre, 2015; Usher, 2012, p.

1910). Pageview data has also been observed to influence the item's placement on the page (Lee et al., 2014), how long it is going to stay there (Bright & Nicholls, 2014) and the likelihood of the story having a follow-up (Welbers et al., 2016).

Furthermore, having access to viewership data of their individual articles, either directly through the publisher's analytics tools or indirectly through "most viewed" lists and reactions in social media, has been observed to provoke behavioural responses in journalists (Petre, 2015). At the *New York Times*, where web metrics were not widely available and not taken into consideration when evaluating a journalist's performance, they were nevertheless a source of anxiety, raising questions about how many pageviews constitute a good result for a given type of article and whether the editors' judgments about newsworthiness were justified (p. 27). Whereas at *Gawker*, where metrics were prominently shown in the newsroom and used to determine journalists' remuneration and job security (p. 18), news values would be continuously adjusted in order to generate as much traffic as possible (p. 21). But regardless of their influence over working conditions, journalists and editors seem to be shifting towards a growing interest in web metrics data in order to understand what interests their audiences (Cherubini & Nielsen, 2016, p. 7; MacGregor, 2007, p. 295).

The recognition that exposure figures fail to capture many important aspects of journalism has prompted attempts at developing better performance metrics even before the internet. Meyer (2009) has explored the application of several social science methods in trying to correlate the effect of factors such accuracy, easiness to read, staffing levels, readers' trust and community affiliation on newspaper business performance. Some statistically significant correlations between higher standards and profitability could be observed, but they were very small and practically negligible due the natural monopoly conditions. Meyer points out that it is very

difficult to show that "quality journalism is the cause of business success rather than its byproduct" (p. 5).

Newer approaches to measure the impact that news content has on society (e.g. politicians citing the reporting, law changes, corporate actions in response to news articles) are being explored (Keller & Abelson, 2015), but challenges remain in agreeing on what should be measured and how to record it. The development of impact metrics has been led by non-profits, philanthropists, and public media in the United States (Cherubini & Nielsen, 2016, p. 39). However, Keller & Abelson observed that pageview numbers were surprisingly important even for non-commercial organizations, because important donors, such as foundations, would ask for them (p. 25).

Building on pageviews and associated web metrics, news organizations have been making efforts to create indicators that balance organizational needs (whether commercial, non-profit or public service) with the particularities of journalism. That is the goal of the developing field of *editorial analytics,* which also aims at generating insights from data that can be useful in short and long-term news coverage planning (Cherubini & Nielsen, 2016, p. 9). While considerable progress has been made in terms of improving the timing of publication and the reach of content produced, important challenges remain in the current inability to confirm that groups assumed to be part of the audience (e.g. public servants in a capital city) are really there, and in identifying to which demographics certain kinds of content, such as videos, appeal to (Cherubini & Nielsen 2016, pp. 12, 26, 37; MacGregor, 2007, p. 292).

Scholars

News organizations keep precise web metrics data private for commercial reasons, but provide some indirect data publicly in the form of lists (most read, most shared, most commented, most sent by email, and so on). This data became the most granular information about audience behaviour generally available to scholars, who began to incorporate it into new research in Journalism Studies (Zamith, 2017).

Boczkowski and Mitchelstein (2013) produced one of the first large-scale studies of the differences between content choices of journalists and audiences based on web metrics, relying on direct data about audience behaviour (instead of depending on their reporting at a later date) and being able to make observations at the individual story level instead of a whole newspaper edition or broadcast (p. 16). Their method consisted of regularly recording the first ten articles placed on the top area of a news website, from left to right, which they considered to be what editors deemed the most newsworthy, and the list of "most viewed" articles, generated from the number of pageviews they got and deemed to be what the audiences found most newsworthy. The articles were then classified as either public affairs or non-public affairs. The differences between these two lists of articles would then be analyzed in order to identify gaps between the editors' choices — considered the supply of news — and the audience's choices — considered the demand for news (p. 14).

The study was conducted between 2007 and 2009 and recorded data from 20 news sites from seven countries, four languages and different political orientations (p. 13). The authors observed that, in general, journalists favoured public-affairs stories and audiences non-public affairs. However, the width of the gap was variable: up to 30 percent points in times of relative

normality, but decreasing or almost disappearing during events such as national crisis or a presidential election. Once relative normality was reestablished, the gap would widen again (pp. 16-17, 83).

The authors also supplemented the quantitative data with a small ethnography involving 12 editors and 25 audience members from studied websites about the reasoning for their actions (p. 15). Editors mentioned the struggles between giving prominence to content they consider important and dealing with web metrics that seem to show what audiences really “want” or find interesting (p. 75). Audience members pointed out that even though they believed that public-affairs stories were probably more important, factors such as lack of familiarity with the subjects would prompt them to scroll down until they reached “easier” articles, generally about entertainment or sports (p. 81). Boczkowski and Mitchelstein believe that their findings show that "there is active avoidance of public-affairs stories rather than passive lack of interest in them” (p. 144). Their main conclusion is that, in order to adapt to these conditions, newsrooms should be more flexible and adjust their news supply according to the different demands observed in different situations. A major obstacle for that, they believe, is journalists' “rigidity of values and beliefs of the occupation and organizations of journalism” (p. 149).

Kormelink and Meijer (2018), however, problematized a central premise of Boczkowski and Mitchelstein’s work — that clicking on an article correlates with interest and demand for its kind of content, and that lack of clicking means no interest. Their approach consisted of a mix of observations, interviews, sensory ethnography and use of think-aloud protocol with 56 audience members, whom they call news users, in order to get more details about their reasoning for clicking and *not* clicking on a news story (p. 671). They found 30 different considerations divided in three groups:

Table 2.1 - Audience considerations for clicking or not clicking on a news story's link (Kormelink and Meijer, 2018, pp. 677-679)

Cognitive considerations

Recency	Whether the user sees the news as timely or current.
Importance	Whether the user sees the news as something they ought to know.
Personal relevance	Whether the topic has a relation to the user's everyday life.
Geographical proximity	Whether the user sees the news as concerning their immediate surroundings.
Cultural proximity	Whether the user recognizes a kinship with the news.
Unexpected	Whether the user sees the news as surprising.
This is logical	The user thinks the news is obvious.
Follow-up	The user wants to know the sequel to a story she has been following.
Already know	The user has already heard the news elsewhere.
Ring a bell	Whether the protagonist or subject matter of the news rings a bell with the user.
More detail on particulars	The user wants to know what exactly is going on.
Join in conversation	The user expects to be able to bring the news up in conversation.
Own opinion	The user wants to see how a topic they have an opinion about is discussed in the news.
Supersaturation	The user thinks the news repeats itself too often.

New perspective	The headline offers a different perspective that sheds new light on the topic.
Participatory perspective	The user wants to witness the news event.
Just an opinion	The user wants facts rather than opinions.
Disjointed news fact	The user wants the whole story, not an isolated update.
Informational completeness	The user has no need to click because the headline says it all.
Associative gap	The user is unable to connect the headline to the topic.

Affective considerations

Disheartenment	The user is saddened by the news.
Feel-good	The light-hearted news makes the user feel good.
Visual appeal	The image evokes the urge to want to see more.
Bemusement	The user feels excitedly puzzled by the headline.
Bullshit	The user instantly dismisses the pettiness of the headline.
Categorical welcome/rejection	The user feels either enthusiasm or aversion towards the beat or the topic of the news.
Gleeful annoyance	The user is delightfully enraged by the news.

Pragmatic considerations

Disruption	Clicking will interrupt a smooth user experience.
Data-heaviness	Clicking will use up too much data.
Does not fit routine	Clicking does not match with the user's schedule.

In the light of these observations, Kormelink and Meijer conclude that “even if one seeks a rough estimate of people's news interests, clicks are a flawed instrument” and that “in terms of news interests, then, the news gap between news makers and news users may not be as wide or unbridgeable as Boczkowski and Mitchelstein (2013) point out” (p. 680).

Furthermore, they noticed that “although most participants had little problem verbalizing their motivations [for clicking or not clicking on an item], news users may not know precisely what they want and why they want it” (p. 671). Petre (2015) also pointed out that while professionals who define and create metrics and the people who use these metrics tend to correlate pageviews with demand, there was no research yet showing evidence that audience members agree that the fact they click on certain articles should be interpreted as them wanting more like it (p. 37).

Audience Experience

Analyzing news production in terms of supply and demand is not new, due to the fact that there are infinitely many events that could be turned into news but newsrooms resources are limited, resulting in a resource allocation problem that is at the essence of economics (Hamilton, 2004). However, because of the natural monopoly conditions of analog media, the demand function was centred not on the audience — who contributed little to revenue through subscriptions or single-copy purchases, or nothing in the case of many radio and TV stations — but on advertisers, who would only pay for audiences believed to have a certain consumer profile (Barnhurst & Nerone, 2001, p. 105; Hamilton, 2004, p. 29; Picard, 2011, p.140). At the same time, the supply side was influenced by the available technology, which limited the number of

pages in a publication, the duration of a program, and the number of publications or analog channels available (Hamilton, 2004; Picard, 2011).

The main challenge in applying the same framework to digital content is that several constraints on its supply, such as medium space, high capital costs of entry in the market and geographical distances that would limit the distribution range, are reduced or virtually inexistent on the internet (Hamilton, 2004, pp. 190,192,199; Picard, 2011, p. 93). This results in a sharp increase in supply from the audiences' point of view, who now have infinitely more content choices than the time to consume them — a phenomenon referred to as *scarcity of attention* (Hamilton, 2004, p. 10; Picard, 2011, p. 4; Simon, 1978).

Two aspects of this shift from a market in which content was scarce to one where attention is limited have been particularly studied. The first is the effect of the breakdown of mass audiences into smaller groups spread out through more media outlets, making them more difficult to reach (*audience fragmentation*) (Picard, 2011, p. 127; Napoli, 2012, p. 81). The second is the industrial and commercial impact of new technical capabilities for choosing the time and frequency of content consumption, as well as the possibility of blocking ads (*audience autonomy*) (Napoli, 2012, p. 84). However, in both cases the focus is mainly on how publishers and advertisers are affected. There has been less emphasis on whether this explosion of content makes it easier or harder for audience members to choose what content to invest their attention in.

The diversity and quantity of content available on the internet magnify the nature of news content as an *experience good* (whose actual value cannot be properly assessed without it being consumed first), which increases the uncertainty of the choice process (Hamilton, 2004, p. 9). Herbert A. Simon (1978), one of the first scholars to use the concept of scarcity of attention,

argued that decisions made “in complex, dynamic circumstances that involve a great deal of uncertainty, and that make severe demands upon [one’s] attention” are deeply affected by the procedures that compose the decision process (p. 14). These procedures, in humans, are defined by the way our cognitive system works and its limitations (p. 8). Explaining the decision maker’s behaviour, according to Simon, would require building theories about the decision process itself (p. 14).

Methodology

The present research will adopt a theory-building approach, as described by Shoemaker et al. (2006). Its main steps, in resumed form, are:

1. Start with a problem, some unexpected results, an anomaly, an observation of something unusual, something you would like to know the effects of, or something you would like to know the causes of.
2. Identify (or formulate) the key concepts involved in the phenomenon of interest.
3. Specify theoretical definitions for all concepts.
4. Specify operational definitions for all concepts.
5. Link some of the concepts to form hypotheses. These hypotheses might involve two, three, or four variables. The hypotheses will often state or imply causal relationships. Specify the form of the linkage—linear, curvilinear, power curve, or other.
6. Specify the theoretical rationale for the hypotheses.
7. Try to put the hypotheses in some kind of organized system.
8. Evaluate the theory using criteria such as testability, falsifiability, scope, explanatory power, predictive power and heuristic value.

This method will be used in order to propose the inclusion of a third variable in the relationship between pageviews and audience interest, which would account for the variations in

clicking behaviour observed by Boczkowski & Mitchelstein (2013) and Kormelink & Meijer (2018).

This third variable will be related to the concept of cognitive effort required to find a personal context in which the expected information is likely to be useful. This will be derived from Sperber and Wilson's Relevance Theory (1986/1995).

3. How relevance guides the selection of news

Relevance Theory (RT) is a linguistics theory based on cognitive psychology. Its initial goal was to provide a detailed mechanism that explains how people select the right context in which sentences should be interpreted in order to reconstruct the intended meaning, particularly concerning implicatures. In order to do so, the theory makes strong claims about how our cognitive system decides what to focus its attention on (Sperber & Wilson, 1986/1995, pp. vii).

This chapter will explore the use of RT as the foundation for a model about how news audiences select what to click or not to click on in an online environment. Its main principles will be presented in the next section, and additional concepts will be mobilized as needed during the development of the model. Emphasis will be more on the cognitive aspects of the theory (i.e. claims about what attracts or retains one's attention), which are the focus of the present research, and less on the meaning-making ones[2].

Relevance

The concept of relevance, either in general or as one of the characteristics of news, is mentioned by different scholars in Social Science and Journalism Studies (some examples can be found in Berger & Luckmann, 1966, p. 45; Kormelink & Meijer, 2018, p. 673; MacGregor, 2007, p. 294; Petre, 2015, p. 11; Schulz, 1982, p. 146; Singer, 2011, p. 624). For instance, in Galtun and Ruge's classic taxonomy of news values, the fourth factor — meaningfulness for the

[2] See chapters 2 and 4 in Sperber & Wilson (1986/1995) for detailed discussions of these aspects.

audience members — would be mainly based on cultural proximity, but with a second dimension "in terms of *relevance*: an event may happen in a culturally distant place but still be loaded with meaning in terms of what it may imply for the reader or listener" (1965, p. 6). However, instead of providing formal definitions, these mentions rely only on the intuitive notion that, in a given context, some pieces of information are more important or consequential than others.

Dan Sperber and Dreide Wilson formulated their definition of relevance as part of their research about how people make sense of utterances — uninterrupted chains of spoken or written language, such as sentences. This is the domain of the linguistics field of *pragmatics*, which has traditionally relied on theories based on the *code model* of comprehension (Sperber & Wilson, 1986/1995, p. 2).

In the code model, a sender[3] encodes a thought using a language (the code) and then transmits it through a medium (such as sound waves or printed words) to a receiver, who will then use their own copy of the code to decode the message in order to reconstruct the sender's meaning (Sperber & Wilson, 2004, p. 607). In this approach, after the message is linguistically decoded, the receiver first defines a context for interpretation, then interprets the message, and finally assess its relevance (Sperber & Wilson, 1986/1995, p. 141).

Sperber and Wilson argue, however, that in many cases this is insufficient to explain how *implicit* meanings are recovered from the message. A particular example is making sense of figurative language such as poetry, humour and irony — when what one *says* is not what he or she *means*. Trying to fit these linguistic effects into the code model results in an inconsistent patchwork of ad hoc rules to explain how the right context for interpretation was transmitted by the sender or inferred by the receiver. In many instances, they rely on additional layers of

[3] In linguistics, the usual terminology is *speaker* and *hearer*. As the present research focuses on written language intended for an unknown audience, these terms will be replaced, respectively, by *sender* and *receiver*.

information that are not part of the message, or in rather complex inferential processes that are unlikely to be performed on a routine and automatic basis (pp. 15-21, 230).

Travellers in a country whose language they don't know, people who interact with small children and owners of dogs or cats can all attest that basic communication does not depend on a formal language — if one clearly signals that he or she wants to communicate something, the other will make an effort to try find what that something could possibly be. This notion was formalized by the linguist and philosopher Paul Grice, implying that communication was not simply a coding-decoding task, but an inferential process in which meaning is reached through reasoning with pieces of evidence — the message, the sender's behaviour and the context. Language, then, would be just one of the tools used by the sender to guide the receiver in the task of recognizing and interpreting his or her intentions (Griece, 1967; Sperber & Wilson, 1986/1995, p. 33).

This approach raises the question of how the receiver is going to decide which one of the multiple possible meanings that can be derived from a message is the one intended. Grice suggested that utterances automatically create expectations which guide the hearer toward the speaker's meaning, and which the speaker is expected to observe (Sperber & Wilson, 1986/1995, 33; 2004, p. 607). He described these expectations in a series of nine maxims:

Maxims of quantity

1. Make your contribution as informative as required.

2. Do not make your contribution more informative than is required.

Maxims of quality

1. Do not say what you believe to be false.

2. Do not say that for which you lack adequate evidence.

Maxim of relation

Be relevant.

Maxims of manner

1. Avoid obscurity of expression.

2. Avoid ambiguity.

3. Be brief.

4. Be orderly.

Real-life examples of communication that contradict one or more of these maxims and yet succeed are so common that, while believing that Grice's main idea is correct, Sperber and Wilson doubt that the maxims correspond to the expectations raised — except for relevance, which was used by Grice as, once again, an intuitive notion and not properly defined (Sperber & Wilson, 1986/1995, pp. 35-38).

Basic Relevance Theory concepts

According to Sperber and Wilson, *relevant* is the information that, when processed in a given context, yields useful conclusions to an individual at a particular time (Sperber & Wilson,

1986/1995, p. 121; 2004, p. 608). These conclusions should not be reachable through reasoning with the new information alone, or the context alone, but only as a result of an interaction between the two (Sperber & Wilson, 1986/1995, p. 107). It can be thought as unknown information that, when combined with a set of already known information, allows for the production of *new* information.

The more important or more numerous the useful conclusions, the higher is the relevance of the information. Likewise, the lower the cognitive effort to process the information, the more relevant it becomes. Overall, relevance is the *ratio* between the importance of the useful conclusions one can derive from a piece of information over the effort to process it (Sperber & Wilson, 1986/1995, pp. 124-125).

Useful conclusions, called *contextual effects*, are the ones that improve one's mental representation of the world. This can be done by adding new beliefs, adjusting the degree of confidence in already existing beliefs, or by abandoning beliefs that do not seem to be valid (Sperber, Cara & Girotto 1995, p. 48; Sperber & Wilson, 1986/1995, p. 117).

Effort refers to the cognitive work related to acquiring and processing information. This includes, for example, the use of perception systems such as sight and hearing to get information from the environment, the use of linguistic processes to decode written and spoken words, the storage and retrieval of information from memory, and the mental computations that use available premises in order to derive new conclusions. All these systems are based on biological systems that required the expenditure of energy to function, so their use is not free (Sperber & Wilson, 1986/1995, pp. 124, 126).

Information relevance depends on the *context* used in its interpretation. A context is a subset of information coming from the environment or from the receiver's working or long-term

memory in which the message can be used to derive useful conclusions. “Classes have been cancelled today because of the snowstorm” has implications if contextualized with “I have class today,” but not with “Ottawa is Canada’s capital.” Contexts can contain pieces of information, called *premises*, of any type — assumptions taken as facts, beliefs, doubts, hopes, wishes, plans, goals, intentions, questions, etc. (Sperber & Wilson, 1986/1995, p. 15; Wilson, 2016, p. 4)

Different people can assemble different contexts depending on their knowledge, experiences and cognitive abilities (Sperber & Wilson, 1986/1995, p. 16). The collection of all the premises that are available to an individual — either stored in the working or long-term memory, derivable from stored premises, or acquirable from the environment through the senses — is called their *cognitive environment* (p. 39).

The individual’s experiences and cognitive abilities affect not only the content of his or her cognitive environment, but also which information is *accessible* off the top of one’s head and which information is archived deeper into memory. This makes some premises more accessible than others, meaning that they require less effort to be recovered. Consequently, some contexts for utterance interpretation will be more accessible than others, and a single utterance can evoke rather different initial contexts for interpretation for different people (pp. 76-78, 138).

Initial contexts can be expanded by fetching more information from memory or the environment, by computing new premises from existing ones, or by actively looking for more information. However, each of these operations increases the effort invested in interpreting the utterance, potentially decreasing its relevance (pp. 140, 142).

One can try different contexts for processing an utterance in order to use the one that yields the most contextual effects, but at some point the process must stop. Besides the energy spent in each attempt, trying to process “Ottawa is Canada’s capital” with every single bit of information

in one's cognitive environment would take an extremely long time for very few useful conclusions (p. 137). The cut-off points, according to Sperber and Wilson, are when either the value of the conclusions is at least worth what was spent to process the information — a situation they call *optimal relevance* — or when processing begins to consume too much energy compared to the usefulness of the conclusions it yields; in this case, processing halts and the brain starts working in the next piece of information claiming one's attention (1986/1995, pp. 46-47, 144; 2004, p. 613).

The main claim of Relevance Theory is that the only expectation raised by an utterance, following Grice's model, is the *expectation of optimal relevance*, and this is enough to guide the receiver towards the intended meaning (1986/1995, pp. 158; 2004, 603). If a sender makes manifest its intention to communicate something, the receiver, in a reflex, assumes that there must be a context in which the message is relevant, otherwise the sender wouldn't bother crafting and sending it. The relevance of a message, therefore, is presumed to be guaranteed; what the receiver does is try to find the context in which the message yields useful conclusions. When the worth of the conclusions is at least equal to the effort, processing stops and the conclusions are assumed to be the meaning that the sender intended. If it takes too long with no encouraging results, the receiver might look for additional information or abandon processing (Sperber & Wilson, 2004, p. 613). This guarantee of relevance is one of the two principles proposed by RT, called the *Principle of Relevance*, later renamed to *Communicative Principle*[4] (1986/1995, p. 158).

[4] This principle is the basis for Sperber and Wilson's alternative to the code model: the *ostensive-inferential model*. In the code model, first the context is determined, then the message is interpreted, and then its relevance is assessed. In the ostensive-inferential model, relevance is taken for granted, and this assumption then prompts the receiver to look for a context in which the interpretation of the message achieves relevance (Sperber & Wilson, 1986/1995, p. 141).

The need to balance the endless flow — throughout one's lifetime — of available information with limited time available to process it motivates RT's other main principle, called the *Cognitive Principle*: "Human cognitive processes, we argue, are geared toward achieving the greatest possible cognitive effect for the smallest possible processing effort. To achieve this, the individual must focus his attention on what seems to him to be the most relevant information available" (1986/1995, p. 1; 2004, p. 610).

When many stimuli compete for one's attention, Sperber and Wilson argue that decisions about what to pay attention to will be made based on expectations of contextual effects — which depend on which premises are most accessible in one's cognitive environment at a particular time — and considerations of effort (Sperber & Wilson, 2004, p. 609; Sperber et al., 1995, p. 49).

When there are no particular expectations of effects, or when expected effects are few, choices would tend to be made entirely based on considerations of effort; the lower the effort to find a context to process a message, the more likely it will attract or keep the receiver's attention (Sperber et al., 1995, p. 49).

Criticism of RT points out that while this approach provides a comprehensive mechanism for context determination, it does not provide detailed explanations about how particular implicatures are retrieved, making the theory too vague and general to be falsified (Burton-Roberts, 2007; Levinson, 1989; Soria & Romero, 2010). For the present research's purposes, however, these meaning-making aspects are secondary. The more important aspects are the cognitive, which have been shown to be empirically testable (Sperber et al., 1995) and provide a rich theoretical literature to explore news audience behaviour.

Audiences' clicking behaviour as a relevance-guided process

The present research will propose a reevaluation of the meaning of pageview data from a Relevance-Theoretical point of view. It will follow the theory-building approach proposed by Shoemaker et al (2004). Its main steps, in resumed form, are:

1. Start with a problem, some unexpected results, an anomaly, an observation of something unusual, something you would like to know the effects of, or something you would like to know the causes of.
2. Identify (or formulate) the key concepts involved in the phenomenon of interest.
3. Specify theoretical definitions for all concepts.
4. Specify operational definitions for all concepts.
5. Link some of the concepts to form hypotheses. These hypotheses might involve two, three, or four variables. The hypotheses will often state or imply causal relationships. Specify the form of the linkage—linear, curvilinear, power curve, or other.
6. Specify the theoretical rationale for the hypotheses.
7. Try to put the hypotheses in some kind of organized system.
8. Evaluate the theory using criteria such as testability, falsifiability, scope, explanatory power, predictive power and heuristic value.

Initial problem and basic concepts

The starting point is the current working hypothesis for the interpretation of pageviews, from the literature reviewed in chapter 2, which will be defined as the following:

$$\text{Pageviews} \propto \text{Audience's interest in the expected content}$$

The following RT-inspired changes will be applied these initial concepts:

Pageviews are the recorded result of a decision process. In the present research, which focuses on the process itself, the concept of pageviews will be replaced by the procedural *Decision to clicking on an article's link*.

The audience's interest in the expected content will be redefined as the *Expectation of contextual effects for one or more particular purposes*.

Finally, a third concept will be included: the *Cognitive effort to find an appropriate context*.

The proposed RT-based relationship of these concepts, to be later detailed in the present research's hypothesis, is the following:

$$\text{Decision to click on an article's link} \propto \frac{\text{Expectation of contextual effects for one or more particular purposes}}{\text{Cognitive effort to find an appropriate context}}$$

These three concepts represent continuous variables. For any given news article, their values will vary from individual to individual. In the present research, these concepts will instead be operationalized as categorical variables, whose values will come from the list of considerations compiled by Kormelink & Meijer (2018) and observations made by Boczkowski & Mitchelstein (2013). This will allow the use of that data as arguments for or against the validity of the proposed relationship, in a discussion that will take place in chapter 4. If the hypothesis is supported, its predictive power could be empirically tested in further research.

Concept 1 - Decision to click on an article's link - Dependent variable

Theoretical definition: The audience member's decision to click on a particular item amongst all available items, meaning that some threshold of expected worth was crossed.

Operational definition: the relative position a category of story gets in the "most viewed" lists collected by Boczkowski & Mitchelstein (2013).

Concept 2 - Expectation of contextual effects for one or more particular purposes - Independent variable

Theoretical definition: The degree to which the audience member expects the linked article's content to yield useful conclusions for one or more particular purposes. It represents the level of interest the person has in the suggested content.

Interest can be broadly defined as "the state of wanting to know or learn about something or someone" (Stevenson & Lindberg, 2010). Different disciplines have developed different theories about why news content interest people and what kinds of uses they might have for it. A few examples will be provided in the next paragraphs.

One approach, based on Economics, proposes that audience members look for information in order to fulfill four types of purposes: consumption (information about things or services one wants to buy), production (information that can be used at work), choosing whom to vote for, and entertainment (a large category which encompasses everything left out by the other three, in which information is consumed for its own sake, without being used as an aid in making another type of decision) (Hamilton, 2004, p. 10).

In Communication Studies, James Carey (1989) proposed that news have an important role not only in transmitting information, but also in forging relationships between members of society and creating a feeling of community. Carey calls this the ritual mode of communication, in which the audience reads the newspaper not to learn about the world — although this happens as part of the process — but as a means to fulfill needs related to concepts such as sharing, participation, association and fellowship (p. 15).

Another example of theory that deals with personal, non-economic uses of news is the Uses and Gratifications framework. Researchers following this approach observed that audience members fulfill many different purposes through media content, such as escaping from routine or problems, getting emotional release, companionship, self-reference, reality exploration and value reinforcement (McQuail, 2010, pp. 423-425).

The present research proposes that all these uses are compatible and valid from a RT point of view. Since one's cognitive environment can contain not only encyclopedic information about concepts but also premises concerning the person's beliefs, doubts, hopes, wishes, plans, goals, intentions, questions, etc., there are many different ways in which information can achieve relevance for an individual. Wilson argues that there seems to be no difference in importance between contextual effects of different categories (2016, p. 5). Therefore, the expected

contextual effects from a news article can vary depending on the audience member's cognitive environment and which of its premises are more accessible at that particular moment.

Operational definition: Certain types of reasons – the ones that specifically mention the intended use of the suggested content – given by audience members to click or not to click on a headline (Boczkowski & Mitchelstein, 2013; Kormelink & Meijer, 2018) and observations about the kinds of articles that are most shared / commented / sent by email (Boczkowski & Mitchelstein, 2013).

Concept 3 - Cognitive effort to find an appropriate context - Independent variable

Theoretical definition: The amount of cognitive work (e.g. fetching information from memory, the environment or other sources) to assemble a context in which the usefulness of the conclusions reached after consuming/processing the content are likely to at least match the time/effort invested in processing it.

Operational definition: Certain types of reasons – the ones that specifically mention perceptions of effort – given by audience members to click or not to click on a headline (Boczkowski & Mitchelstein, 2013; Kormelink & Meijer, 2018).

Research hypothesis

Two initial assumptions will be made:

(i) Headlines are messages that carry a guarantee of optimal relevance

A headline is not a random feature of the environment. It is a message crafted by an editor, in the hopes of attracting the audience's attention to it and to the content that follows it. Therefore, a headline can be classified as an ostensive act in RT terms — that is, an act that explicitly calls the receiver's attention to something — which should then carry an implicit guarantee of optimal relevance.

(ii) - With competing content, decisions about what headlines will get clicked on will be made with respect to expectations of effects and considerations of effort

This follows from (i), where it was established that headlines are signals that can be studied from a RT perspective, and therefore inherit all of the theory's explanations (Shoemaker et al., 2004, p. 52).

Assuming (i) and (ii), the present research proposes the following hypothesis:

The decision to click on a news article link is directly proportional to the audience member's interest in its topic, i.e. the expectation of cognitive effects from the suggested content, and inversely proportional to the cognitive effort it takes to find an appropriate personal context to process the headline.

Theoretical linkage:

1. In a communication process, humans do not assess the relevance of messages; they assume that the message must be relevant if processed in a suitable context. As relevance is taken for

granted, the receiver's task when interpreting a message is to look for a suitable context (Sperber & Wilson, 1986/1995).

2. Human cognition's goal is to maximize relevance — i.e. get the most useful conclusions with the least amount of effort (Sperber & Wilson, 1986/1995)

3. In the absence of specific expectations of effect, or even in the presence of weak expectations of effect, considerations of effort may play the central role in directing attention and the retrieval of background knowledge (Sperber et al., 1995)

In order to assess the validity of this hypothesis, a RT-based procedure for news selection will be proposed in the following section, creating a system to group and classify the observations made by Boczkowski & Mitchelstein (2013) and Kormelink & Meijer (2018). The data will be fitted into this model in chapter 4.

RT-based procedure for news selection

Two additional assumptions will be made:

(iii) A headline is the first segment of a larger text

During the comprehension of a series of utterances, the context retrieved for the interpretation of the first one, as well as its resulting contextual effects, are taken as the *immediately given context* for the interpretation of the next sentence. This is just an initial, tentative context that can be expanded in different directions (Sperber & Wilson, 1986/1995, p. 139).

In assuming that a headline is the first segment of a larger text, it follows that:

— A click on the headline means that the audience member expects more content segments to follow;

— The context assembled for the interpretation of the headline, as well as any contextual effects produced by its interpretation, should be the basis used for developing the expectation of contextual effects for the rest of the article.

(iv) In the absence of specific expectations of contextual effects, the effort to find a context to process the headline will tend to be minimal

Considering (ii) and the fact that all other competing headlines also communicate expectations of optimal relevance, it seems reasonable to assume that an audience member is unlikely to invest additional effort in a headline that did not raise any particular expectations of contextual effects and failed to achieve relevance in the first attempt.

This assumption is based on Sperber's assertion that "in the absence of specific expectations of effect, or even in the presence of weak expectations of effect, considerations of effort may play the central role in directing attention and the retrieval of background knowledge. Since the perception and conceptual representation of information involves a processing cost, the most salient and easily representable information at a given time is, *ceteris paribus*, likely to be the most relevant information at that time" (Sperber et al., 1986/1995, p. 49).

In this case, it will be assumed that either a context is easily accessible to the audience member, or attention is redirected to another headline.

The steps of the proposed relevance-guided news selection procedure are the following:

Table 3.1 - Proposed relevance-guided news selection procedure

1. Reading/decoding a headline will launch the search for a context (C) in which it can achieve relevance.
2. If (C) is not easily found, the tendency is to skip to the next headline. *A pageview is unlikely*.
3. If (C) is found, it is used to interpret the headline and derive contextual effects from it. Possible outcomes:
3.1. No significant contextual effects can be derived from it. *A pageview is unlikely*.
3.2. Contextual effects are derived from it: Possible outcomes:
3.2.a. The headline is relevant on its own but no additional effects are expected from the rest of the article. *A pageview is unlikely*.
3.2.b. The contextual effects include at least one relevant question whose answer is expected to be in the article. ***A pageview is likely***.

Theoretical linkage

1. Reading/decoding a headline will launch the search for a context (C) in which it can achieve relevance.

This comes from a direct application of RT's ostensive-inferential model of communication, based on its Communicative Principle (Sperber & Wilson, 1986/1995, p. 158).

This search will be influenced by the premises retrieved from memory for the linguistic decoding of the headline. If someone utters "Ottawa is Canada's capital," the receivers will automatically retrieve the most accessible premises they have in memory about "Ottawa," "Canada" and "capital" (1986/1995 p. 138; 2004, p. 615). The exact content can greatly vary from one individual to the next, according to their individual knowledge, experiences and cognitive abilities (1986/1995, pp. 16, 38). These premises will remain in the working memory and will be accessible during the interpretation of the next text segment (p. 139).

2. If (C) is not easily found, the tendency is to skip to the next headline. A pageview is unlikely.

This is based on assumption (ii).

3. If (C) is found, it is used to interpret the headline and to derive contextual effects from it.

Possible outcomes:

3.1. No significant contextual effects can be derived from it. A pageview is unlikely.

There are at least three situations in which the conclusions derived from an utterance can be deemed *irrelevant* in the chosen context (1986/1995, pp. 121, 142).

The first is when the conclusion is already present in the working memory, most likely because it was already part of the premises recovered during linguistic decoding in step 1. The conclusion might be useful, but the receiver *already knew it*, so the effort was not worth it.

The second case is when new information can be derived from the utterance, but the receiver *does not see a connection with the current context*. A roadblock in the receiver's usual way home might be generally relevant, but not if it lasts only one day and the receiver is away on vacation for the week.

The third case is when the conclusion *contradicts other information in the context without enough strength to upset it*. If you have a meeting with John tomorrow at 5pm, it is easy to find a context in which the utterance "I think John is out of town this week" could be relevant to you. However, if you saw him walking down the street earlier today and he later sent you an email confirming the meeting, it is highly unlikely that you would conclude that he is not going to show up.

Based on assumptions (ii) and (iii), it is unlikely that an audience member would be willing to invest time in the subsequent segments of the article if the headline is deemed irrelevant.

3.2. Contextual Effects are derived from it:

Possible outcomes:

3.2.a. The headline is relevant on its own but no additional effects are expected from the rest of the article. A pageview is unlikely.

In sentence comprehension, each element can achieve relevance in three different ways. The most straightforward is by being *relevant on their own*, that is, by adding premises that the receiver can use to derive useful conclusions (1986/1995, pp. 142, 207-209). A second way is by *evoking a context which will lower the effort to process the subsequent elements*. A classic example is the opening paragraph of a novel, whose relevance might be low when analyzed in

isolation but might evoke a context in which all the subsequent sentences will become more relevant (pp. 160, 207-209). The third way is by raising a *relevant question* — a question whose *answer* is relevant to the reader and presumed to be delivered in the subsequent segments (pp. 207-208).

If the interpretation of the headline yields one or more useful conclusions but *no relevant questions*, the receiver does not have reasons to expect that more contextual effects will be derived by clicking on the link and investing attention in the subsequent segments of the article.

This is different from deeming the headline irrelevant, as in step 3.1. Here, the headline achieves relevance, but whatever expectations of contextual effects developed by the receiver at that particular moment are fully satisfied by the headline itself. This distinction will become important when analyzing cases when audience members do *not* click on content they consider important or interesting.

b. The effects include at least one relevant question whose answer is expected to be in the article. ***A pageview is likely****.*

Following assumption (iii), that the headline is the first segment of a larger text, it is likely that a headline whose interpretation yields one or more relevant questions in the context assembled by the audience member will be deemed worth of a click, in the expectation that these questions will be answered by the article.

The possible consequences of succeeding or failing in delivering these expectations will be discussed in chapter 5.

4. Evidence of relevance-guided news selection in Kormelink & Meijer (2018) and Boczkowski & Mitchelstein (2013)

In this chapter, the 30 considerations for clicking or not clicking on news articles found by Kormelink and Meijer (2018) and some of the observations made by Boczkowski and Mitchelstein (2013) will be fitted in the RT-based news selection procedure defined in chapter 3 (table 3.1). Additionally, observations from other research in linguistics that applied RT to the study of news headlines will be mobilized in order to substantiate some of the present research's claims. In doing so, the procedure's empirical validity will be gauged through a secondary analysis of the findings in these studies. Its implications for the authors' main findings and conclusions will be discussed in the next chapter.

For Kormelink and Meijer's 30 considerations (see table 2.1), numbers in parentheses will be added after their first mention. They are a simple counter and they do not correspond to the order that the authors present them in their research, as they were rearranged according to the procedure's steps. Numbers in parentheses after names indicate the age of the participant[5].

1. Reading/decoding a headline will launch the search for a context (C) in which it can achieve relevance.

Ifantidou (2009) observed direct evidence of this process in her RT-based study about the effectiveness of headlines (i.e. how much interest or curiosity they stir) from the audience's

[5] The authors called the participants *news users*. For consistency, in the present study the term was adapted to *audience member*.

perspective. The participants — 137 university students — were given 36 headlines, with no accompanying texts, to rate according to their perceived effectiveness in attracting their attention[6]. Additionally, they were asked to provide details about their interpretation of three headlines.

For example, for the headline "Kate's back in therapy," the interpretations given comprehended the following[7]:

a. Kate Moss in drug therapy

b. Kate Moss in psychological treatment

c. Kate Winslet in psychotherapy

d. Kate Winslet is starring in a film called ''therapy''

e. Kate Winslet in diet program

f. Kate X is treating a physical or mental illness

g. Kate is treating her back after an accident

h. drug addicts program

i. Hollywood actress with serious health problem seeing an acupuncturist

j. injured athlete resumes therapy sessions

k. aromatherapy

l. chemotherapy

[6] Some implications of the present research's model for this concept of headline effectiveness are discussed in chapter 5, p. 65.

[7] The headline is from an article from 2006, and the experiment was conducted in Athens in 2008. Had it been a few years later, it could be expected that the disambiguation of "Kate" would yield results mentioning Kate Middleton, whose marriage with Prince William took place in 2011.

This illustrates not only that participants searched for a context to interpret the headline, but also that its contents seem to depend on which premises are more accessible in their cognitive environments. According to Ifantidou:

> Increasingly narrower interpretations [from *l* to *a*] suggest that headline readers using the relevance-theoretic comprehension heuristic narrow the encoded concept THERAPY just so far as is required to satisfy their expectations of relevance and no further. Given their background assumptions and interests (e.g. interest in top-model Kate Moss, or familiarity with film star Kate Winslet feature stories), readers choose narrowings that cost them relatively little effort and provide them with a plausible and relevant interpretation (e.g. anti-drug therapy for Kate Moss, or weight reduction program for Kate Winslet, respectively) (p. 714)

Ifantidou also observed that, in general, participants rated more vague headlines (i.e. more open to multiple interpretations) as more interesting (p. 704). The implications of this will be discussed in chapter 5.

2. If (C) is not easily found, the tendency is to skip to the next headline. A pageview is unlikely.

Boczkowski and Mitchelstein (2013) made observations consistent with the idea that contextualization, and particularly the level of cognitive effort required for it, plays a role in the process of choice between clicking or ignoring a story's link. This is manifest in mentions to the

relative difficulty in accessing background knowledge or in working out the possible consequences of the expected content:

> These various reasons [for hesitating when choosing between public-affairs and non-public-affairs] are tied to the divergent cognitive demands associated with the consumption of these two types of news. A 30-year-old clerical employee who regularly looks at Lanacion.com says "Other people might like that politics stories are [at the top of the homepage] because they understand [them]. But since I don't understand them well, I scroll down [looking for non-public-affairs news]. When I reach a story I like [at the bottom part of the homepage, mostly populated with sports and entertainment stories], I stop and read." María, a 22-year-old college student, says "It's difficult for me to do a deep analysis of anything having to do with the economy and form an opinion because I don't understand much about economics ... [But entertainment stories] are understandable by everybody. They're very basic, and it's not necessary to do any analysis." For Sebastián, "Reading a news story about international politics you have to pay more attention ... and even go to another site to get information and only then you understand the article ... With sports [stories] it is much more simple; I follow them every day and I don't need an introduction to the subject." (p. 81)

Four considerations found by Kormelink and Meijer (2018) relate to the result of the initial contextualization of the headline and whether a context that makes it minimally relevant is found or not. In these considerations, to be discussed in the next paragraphs, audience members imply that either a minimum degree of familiarity with the subject or the protagonist of the headline is established, or they see little point in clicking on the article. This familiarity can be either in

terms of general knowledge, geographical or cultural proximity, or personal affinity with the subject.

Conversely, and consistent with the assumption that the existence of an easily accessible context in the audience member's cognitive environment is a consideration in deciding what to click on, when some kind of familiarity was established, audience members pointed out that it was a reason *to click* on a link. Kormelink and Meijer classify this kind of considerations as *dual*, meaning that they can be responsible for either clicking or not clicking on a news article (p. 673).

The four considerations are:

Rings a bell (1) — Whether the protagonist or subject matter of the news rings a bell with the audience member:

> This concerned famous people but also names or events the participants recognized but could not quite place, as Nina (54) illustrates: 'That Benno L., you've heard something about that before and then [you're] like, gosh, who was that Benno L. again?' Conversely, Eddy (53) asks why he would click if the subject matter does not ring a bell: '"Fight parenting clinic and insurer resolved," well, I wouldn't know what a parenting clinic is, so (laughs) I'm like, why should I read that?' (p. 674)

Geographical proximity (2) — Whether the audience member sees the news as concerning their immediate surroundings:

First, participants tended to click if they saw the headline as concerning news taking place within their immediate surroundings, regardless of absolute distance. Bianca (54) clicked on a headline about a dead body found 20 km away from her hometown: '[City] is so close, I just wanna know. [...] And if it's not so close then it's not interesting'. Yet, Tracy (53) skipped a headline about an accident that happened within a similar distance because she did not experience it as nearby: 'I think it didn't happen in this region but somewhere in the south. No, that doesn't really interest me.' (p. 673)

C*ultural proximity (3)* — Whether the audience member recognizes a kinship with the news:

(...) Golding and Elliot's (1979) 'cultural proximity' depends 'on what is familiar and within the experience of journalists and their audience', but for our participants, more specifically, it refers to whether they recognize a kinship with the subject of the news, again regardless of absolute distance (p. 166). Leonard (24) clicked on sports news concerning compatriots: 'I like cycling, especially if Dutch people are participating. [...] I don't have to know if some Slovak won a round in Poland'. Conversely, Dutch native Andrew (58) did not click on a headline regarding Antilleans in the Netherlands because he does not feel a connection: 'It may be important, but [...] not for me right now. [...] Because I don't do anything with Antilleans. [...] I mean, I don't know one Antillean and I don't know if they're good or bad.' (p. 674)

Categorical welcome/rejection (4) — The audience member feels either enthusiasm or aversion towards the beat or the topic of the news:

> [Categorical rejection] was often the case with sports news, as Ruth (24) illustrates: 'The last [headline] is sports news, sports mean nothing to me.' Anita (21), on the contrary, categorically welcomes news about sports with which she has affinity but rejects others: 'I don't find soccer interesting, so I skip those headlines automatically. But ice skating and tennis, those I do follow.' (p. 678)

The need for an easily accessible context could also explain the fact that audience members deemed *recency* (5) — whether the news is seen as timely or current — as a general requisite for news, but not as a major consideration when deciding whether or not to click on an article's link[8] (p. 673).

Another consideration that was not deemed by the participants as dominant when making decisions about what to click on was the perceived *importance* (6) of the news — whether they see the news as something they *ought* to know — which was generally correlated with the prominence of the headline in the page (p. 673). However, Kormelink and Meijer bring interesting observations:

> Sandra (25) illustrates how the placement of news on a website influences how important she perceives it to be: 'Cabinet: no clear picture of money laundering', I couldn't care less, so wouldn't click on that. [...] If it was REALLY important it would have been big at the top [of the homepage]. Then maybe I would've clicked on it.' Online news presented as important

[8] According to RT, what really matters is whether the information allows you to reach useful conclusions (and whether you can access a context in which this processing could take place). The information's recency is not a pre-requisite for this. However, it might indeed become a requisite for topics that the audience member follows or knows a lot about (i.e. rich contexts are accessible), in order to avoid only repeating information that is already known.

> through prominent placement on the website or news app is experienced as more worthy of knowing; if the same news is placed less prominently, it apparently is not significant enough to deserve a click. (p. 673)

This relationship between headline prominence and story importance was also observed by Boczkowski and Mitchelstein:

> (...) Diego, a 30-year-old human-resources consultant, says "Usually, the articles that interest me are those placed at the beginning of the homepage, because they are considered most important. They are (the stories about) national politics, international politics, eventually a crime story." (2013, p. 80)

> Sebastián, a 25-year-old lawyer who visits news sites every morning, declares "First, I read all the most important headlines, which always are about national politics. Then, I look at those that interest me the most." (In his case, the latter tend to be sports stories.) (p. 79)

This suggests that receivers gauge the intensity of the communicative intention manifested by the editors. More prominence possibly boosts the expectation of contextual effects to some extent, consequently encouraging audience members to put some extra effort in contextualizing the headline.

3. If (C) is found, it is used to interpret the headline and derive contextual effects from it.

Possible outcomes:

3.1. No significant contextual effects can be derived from it. A pageview is unlikely.

In this case, the receiver succeeds in assembling a context for the headline interpretation, but the resulting conclusions are deemed irrelevant in RT terms.

As seen in chapter 3, one reason for that is when the resulting contextual effects already are present in the working memory, having been made accessible during the linguistic decoding of the headline. In other words, the audience member *already knows* it (7). This could also happen when the news was consumed elsewhere and the person sees no reason to expect additional contextual effects from getting more information from another source — "as Karen (50) indicates: '[This] I already just heard, so I'm not going to read that again'" (Kormelink and Meijer, 2018, p. 674).

A similar situation happens in the consideration called *this is logical* (8) — when the audience member does not click because from their perspective the news is (too) obvious:

> Regarding the headline 'Nokia unsure about brand name for the future,' Nanda (21) noted that she already knew Nokia was not doing well: 'Then this seems like a logical continuation. Then I don't have to read it, because I already know why that is.' (p. 674)

In RT terms, this means that the contextual effects expected from the content are easily derivable from the headline combined with the selected context, to the point that investing further attention in it does not seem worth it.

Another possibility is that the context assembled by the audience member does not connect with the content of the headline (or the article's). As Sperber and Wilson argue, the sender might point the receiver in the direction of a initial context for interpretation, but this context might be expanded and enriched in different ways depending on the receivers' cognitive environment and what's going on in their minds (1986/1995, pp. 138, 140). One possibility is what Kormelink and Meijer called an *associative gap* (9), in which the audience members are unable to connect the headline to the article's topic:

> Ella (51) read, 'Pieterburen [location of a famous seal crèche] will possibly move to [island]' and said, 'The headline doesn't tell me much, that's why I don't click it.' However, later in the interview she did click on a headline that explicitly mentioned 'seals' and said she was fascinated by them. Clearly, she had not made the connection between Pieterburen and seals. Based on clicks, it would be tempting to conclude that Ellen was not interested in this article, but based on her comments about how much seals 'intrigue' her, it seems safe to assume that she is. Similarly, Matthew (25) was clear about his interest in clicking the headline 'Warning Kerry about Cold War Ukraine,' claiming he was following all news about the country because he planned to visit its city Chernobyl, 'and of course I'm not gonna go there if there is almost a civil war.' Yet, he did not click on a headline about former Ukrainian president Janoekovitsj because 'I don't know exactly who that is, so I think I would skip that.' While this consideration is similar to 'ring a bell', the focus here is not the topic; instead, it is about

not being able to make a connection between the headline and the user's (pre-existing) interest in the topic. (p. 676)

However, even if the connection between headline and an article topic is well established and a personal context to process it is accessible, some premises resulting from the headline interpretation might decrease the expectation of contextual effects from the article to a point where it is not deemed worthy of further attention. This is one aspect of the consideration called *personal relevance* (10):

> Henry (55), who invests, clicked on a news item about the stock market, but skipped a headline concerning the shares of a specific company: 'I [don't invest] in companies, so the particular company mentioned here I couldn't care less about.' Matthew (25) clicked on a headline about Samsung Galaxy S5: 'because I want to buy a new phone,' but skipped news about rented housing because 'after [I leave my student house] I'm not going to rent, I will buy something immediately.' (p. 673)

Conversely, some premises resulting from the headline interpretation might connect with some salient aspect of the audience member's cognitive environment, consequently increasing the expectation of contextual effects from the article — for instance, the person in the previous paragraph clicking on an article about a cell phone because the context he had assembled included a premise about his current desire to buy a new phone. This was observed by Kormelink and Meijer, who classified personal relevance as a dual consideration responsible for both clicking and not clicking on an article (p. 673).

3.2. Contextual effects are derived from it:

Possible outcomes:

3.2.a. The headline is relevant on its own but no additional effects are expected from the rest of the article. A pageview is unlikely.

If an audience member has a limited (or a very specific) interest in a subject — and consequently a rather restricted context in which it could achieve relevance — all expectations of relevance could be fulfilled from the interpretation of the headline itself, with no reason to believe that investing more attention in the rest of the article would yield enough contextual effects to justify the additional cognitive effort. Kormelink and Meijer observed that:

> A frequent occurrence was that the participants showed interest in the news itself but the headline was informationally complete and consequently, they did not expect to be better informed by clicking. Lauren (26) noted, '"More than 4 million viewers for Olympic finals 1500 meters," that's a fun fact to know, but I know that this is usually all the information you're gonna get, so I don't really have to click it anymore'. This is the opposite of clickbait: Lauren is interested in the topic, but there is no need to click because the headline tells the whole story. Nick (24) similarly illustrates, 'I see it says "Final will be great," so I already know they're in the final so I don't necessarily have to click it.' (p. 676)

They called this consideration *informational completeness* (11), when there is no need to click because "the headline says it all."

The lack of expectation of further cognitive effects can also explain what Kormelink and Meijer label *supersaturation* (12), which happened when participants would not click on stories deemed repetitive:

> Bruce (55) noted about the ongoing crisis in Syria: 'Because every day it's the same, same, same, at some point it becomes less interesting. Even though it's not less terrible'. This is less about 'compassion fatigue' (Moeller, 1999) than about how hearing about it again does not provide new insights. The headline does not invite a click anymore, as Jeff (58) illustrates: 'You actually drown in that kind of news. At some point you're like, it's not going to stop anyway. It's not that it's not important, but it doesn't stop'. As we will elaborate later, not wanting to click on a headline does not mean the user does not want to see it. But for now the headline itself provides a sufficient update about the situation; it is not until 'something completely new' happens that Jeff (58) will click again. (p. 675)

3.2.b. The contextual effects include at least one relevant question whose answer is expected to be in the article. ***A pageview is likely.***

As mentioned in chapter 3, the processing of a headline might yield one or more relevant questions, whose *answers* are relevant to the reader and presumed to be provided in the subsequent segments of the article. A direct equivalent of this concept was identified by Kormelink and Meijer in their consideration *more details on particulars* (13):

> More detail on particulars comes into consideration when the headline raises a question in the participant's mind, causing them to want to know more about the situation, as Jack (56) illustrates: 'Heavy weather in Italy, I see [...] (clicks) What is going on here?' For a similar reason, Karen (50) clicked on a headline about a fishing ban: 'Then I'm like, what do we catch there? [...] What kind of fish is swimming there?' (p. 674)

The questions seem to be triggered by different considerations. Two examples are when the suggested subject of the article has *personal relevance* for the audience member, or when it at least *rings a bell*. And if the person is already somewhat familiar with a particular story, the lower effort required to contextualize a new development — if combined with an expectation of a minimum of new contextual effects in order to avoid *supersaturation* — might motivate a *follow-up* (14):

> Like journalists selecting stories already in the news (Galtung and Ruge, 1965; Harcup and O'Neill, 2001), participants regularly clicked on follow-ups to stories they had read before. Lauren (26) illustrates, 'What catches my eye immediately is the headline [....] "Exam fraud [school] costs 3 million euro." I've followed [that story] before.' (p. 674)

Furthermore, questions about why a particular event takes place or unfolds in a particular way — possibly upsetting premises that were recovered or inferred during context assembly for the interpretation of the headline — may rise when it is *unexpected* (15):

> Whereas for journalists unexpected refers to rare, out-of-the-ordinary developments (Galtung and Ruge, 1965), from a user perspective it is about whether the news fits their idea of what is common. Lilly (26) clicked on a headline about a joint action from a trade union and an employers' organization: 'Seems interesting, I'm curious why [they] are on the same page here, seems a bit illogical'. It is important to stress that what is unexpected to journalists may not be experienced as unexpected by users, and vice versa. For instance, Anita (21) did not click on news about a man lighting himself on fire: 'Yeah, it's bad, but it's, I don't care [...] because uhm, yeah it happens regularly.' (p. 674)

Content relevance for a particular purpose

According to the present research's hypothesis, expectations of contextual effects developed by audience members are with respect to one or more particular purposes. The observations of Kormelink and Meijer (2018) and Boczkowski and Mitchelstein (2013) provide evidence of different uses that audience members have for content, fitting the three different theoretical approaches mentioned in chapter 3 (see pages 26-27).

The use of news as an input for decision planning, as is the case in the purposes of consumption, production and voting mentioned in the Economics approach cited by Hamilton (2004, p. 10), was observed by both groups of researchers:

> Participants often clicked on news that had *personal relevance*, relating to their everyday life, including work. This consideration is dual, meaning that it counts as reason to click when present and as reason not to click when absent. (Kormelink & Meijer, 2018, p. 673)

> "(…) the reader's top choices are marked by a strong predilection for non-public-affairs topics and 'news you can use' (meaning stories with direct implications for everyday life)." (Boczkowski and Mitchelstein, 2013, p. 23)

Examples of "news you can use" include "how to remove one's profile from Facebook, how a new car performs, how to avoid being kicked off a plane, how to install a home network, and how to find a new life after retirement" (p. 44).

The fourth purpose mentioned by Hamilton, entertainment (in which information is consumed for its own sake, without being used as an aid in making another type of decision), has similarities with the Uses and Gratifications approach. In RT terms, when we talk about information that allows for "useful conclusions," this usefulness is more dependent on what's salient in the person's cognitive environment at the time — and which premises were assembled for the interpretation context — than on the seriousness or frivolousness of the content, or whether it has practical implications (Wilson, 2016, p. 5). This seems in line with several considerations from Kormelink and Meijer's list:

Own opinion (16) — The audience members want to see how a topic they have an opinion about is discussed in the news:

> Jenna (27) clicked on the headline 'World Bank freezes aid to Uganda over gay law' because 'I personally have an opinion about it, so I'm curious on what grounds the World Bank does something like that.' However, this consideration was uncommon; like in Donsbach's (1991)

study that relativized the influence of cognitive dissonance on readers' selections, our participants rarely expressed strong opinions about headlines. If they did, disagreement was not a reason not to click. (p. 675)

New perspective (17) — The headline offers a different perspective that sheds new light on the topic:

> This is not about the news event being unexpected but about the headline offering 'the other side' of a topic. Such news inspires because it adds to your knowledge or broadens your horizon and as such enables an *aha-erlebnis* (cf. Costera Meijer, 2013).
>
> Corbin (24) illustrates,
>
> > Here's an article called 'According to these three imams the Koran has nothing against gays.' That's interesting to me [because] you have this image that in the Koran it says that homosexuality is wrong and here it says something completely different, and I'm curious how that is substantiated by those imams.
>
> Rather than the topic of homosexuality and the Islam it is the original angle of the headline that makes Corbin click. (p. 675)

Both the considerations above are connected to the very definition of contextual effects. By comparing their mental representations of a given topic with other people's opinions or with different perspectives, audience members get the chance to learn new things and add new premises, increase or decrease the level of confidence of premises already present, and consider excluding premises that do not seem to be valid.

Also, as some other considerations suggest, it seems to be the case that contexts assembled by audience members might include premises — needs or desires — related to purposes proposed by the Uses and Gratifications approach: escaping from routine or problems, getting emotional release, companionship, self-reference, reality exploration and value reinforcement (McQuail, 2010, pp. 423-425). That includes:

Participatory perspective (18) — The audience member wants to witness the news event:

> Sometimes, participants clicked on a headline because they wanted to see for themselves or 'experience' what happened. We labelled this participatory perspective. An example is Nick (24), who clicked on the headline 'Man makes illegal base jump from moving ski lift' because he 'can't really picture how anyone would do that' and hoped to see it in a video. (p. 675)

Disheartenment (19) — The audience member is saddened by the news:

> Sarah (21) illustrates, 'This one I would read: "Biker killed by car." That's just sad.' However, if participants found the headline too disheartening, they skipped it: 'It's such a heavy text, "Dragging patients is risky." I prefer starting with happy news' (Jeff, 58). (p. 677)

Feel-good (20) — The light-hearted news makes the audience member feel good:

Isabel (30) illustrates, ‘Something about self-cleaning plastic for cars. [...] Yeah, that’s a fun news item. [...] It’s light, [...] just nice to read’. While this corresponds to the news value ‘good news’ (Harcup and O’Neill, 2001), from a user perspective ‘feel good’ is about the impact of the news rather than its genre. (p. 677)

Gleeful annoyance (21) — The audience member is delightfully enraged by the news:

Lilly (26) clicked on the headline ‘President of Uganda will sign antigay law’ because she found it ‘particularly bothersome that again there is a country that does not understand that homosexuality is not something you should draft a law against, so yeah, I’ll read that news and be very irritated by it.’ Isabel (30), similarly, clicked rather than ignored a headline that annoyed her: ‘Bart Veldkamp once again has an opinion. [...] Now he thinks that the Netherlands should share their ice-skating knowledge. [...] It does evoke a bit of irritation, that headline. I’m like, you became a Belgian.’ (p. 678)

A particularly interesting case is the consideration *bemusement* (22) — when the audience member feels excitedly puzzled by the headline. It not only adds support for the existence of non-practical contexts in which content can achieve relevance, but also shows that *relevant questions* can be raised in these contexts:

Eva (19) illustrates, ‘Something provocative like “Anders Breivik: Playstation 2 instead of Playstation 3 is torture,” [...] then I think what is this about? And then I click it and read it’. Such headlines usually concern remarkable or bizarre news, which might partially explain

> why this type of news is so heavily clicked (cf. Tenenboim and Cohen, 2015). The colloquial term for this is clickbait – headlines with a 'what-the-hell' factor that makes the user want to click, as Martin (24) illustrates: 'Actually it never has any news value, but it's usually those headlines that make you think, yeah, I'm curious what it is exactly.' (p. 678)

However, the same reasons that made people click on content based on the bemusement consideration would sometimes be responsible for making audience members *not* click on similar articles, deeming them *bullshit* (23) — when the audience member instantly dismisses of the pettiness of the headline. This in line with the idea that different contexts can be assembled depending on what premises are more accessible at a particular time, therefore changing the way in which a particular piece of content can achieve relevance.

> Leonard (24) explains, 'Now I see "German cat survives 30-meter fall." Then you're like, I don't care. [...] I think it's a bit rubbish actually.' We classified [the consideration 'bullshit'] as affective instead of cognitive because it is a gut reaction dismissing the pettiness of the headline rather than a cognitive deliberation about whether or not the topic is nonsense. However, this consideration was mentioned less often than 'bemusement,' where the silliness of the headline was exactly what does make users click. (p. 678)

One particular consideration that can be connected to multiple purposes is the *visual appeal* (24) of accompanying pictures — when the image evokes the urge to want to see more:

> Danny (25) is not interested in the news itself, but the picture evokes arousal: 'On Nu.nl you often have these dumb items about, I don't know, the New Year's dive. Couldn't care less, but if it happens to have a picture of a lady, I do click on it.' (p. 677)

Relevance Theory does not propose a specific mechanism for studying images. However, it seems plausible that the Communicative Principle (the guarantee of relevance) should be applicable: Like news articles, pictures are not random features of the environment. Both photographers and editors choose certain subjects and framings in attempts to communicate something. Similarly, it seems reasonable to assume that the interpretation of pictures yields premises that will be added to the ones coming from the headline interpretation, therefore affecting the expectations of relevance related to the article and the decision to click or not to click on its link. But as previously mentioned, this discussion falls outside the scope of the present research.

Finally, there is also evidence of uses that align with the ritual model of communication proposed by James Carey (1989). It appears in the consideration *join in conversation* (25) — when the audience member expects to be able to bring the news up in conversation:

Rod (24) explains why he clicked on a headline about the Winter Olympics:

> Because if you start a conversation with people then often you want to talk about things that uh are recent and speak to a lot of people and uh the Winter Olympics I think are a part of that, so uhm to be able to join in the conversation, so to speak.

> Rod's reason for clicking is the social utility function the topic provides: fodder for conversation. Teacher Joe (26) similarly clicked on a headline about the 'largest lunar impact ever recorded' 'because I also talk about that with my students.' (p. 675)

Boczkowski and Mitchelstein also observed similar situations in which the possible value resulting from the use of the content in a social interaction — the value the audience member attributes to *the results of the interaction* — seems to be part of an article's contextual effects. For instance, after examining the websites of CNN, The Washington Post and USA Today in 2008 and 2009, they noticed that:

> (…) [The] comparison of the most clicked, most emailed, and most commented on articles on the three sites reveals that users take advantage of these interactive features in different ways. They tend to click on what they deem interesting — most often non-public-affairs stories in the straight-news format. They prefer to email what they find either bizarre or useful — typically non-public-affairs stories told in feature style. They post comments on what they consider to be controversial— often commentary-style or straight-news-style articles about high-profile public-affairs topics. (2013, p. 114)

Furthermore, journalists also noted that the potential for social interaction might be a factor in audiences' choices:

> When asked why they believe that consumers' news preferences diverge from theirs during ordinary times, journalists often note that people seem to enjoy stories that entertain them,

> help them further their leisure interests, and connect them with others at home and at work. They add that public-affairs news appears to be less suitable for these purposes than non-public-affairs stories. According to Daniel Vittar of Clarín, "People ... want ... things that have to do with their milieu, their taste, [and] their interest in music or entertainment."(...) On the effectiveness of non-public-affairs content in fostering social interactions, an editor at Nación notes that "in general, people are more relaxed talking about soccer or about what happened in [the television series] Lost [than talking about national news]." (p. 77)

Other considerations of effort

As predicted by Sperber and Wilson, considerations of effort — cognitive or otherwise — also seem to affect the expected relevance of content. The remaining considerations from Kormelink and Meijer's list illustrate this, with cases in which even if a context for interpretation is found and expectations of cognitive effects are raised, the audience members decide not to click on the headline for deeming that it won't be worth it at that particular time.

If the content is expected to yield premises whose usefulness or reliability is doubtful, the audience member might conclude that it is not worth investing attention on it. For example, the content might be dismissed as *just an opinion* (26) — when the audience member wants facts rather than opinions.

> Regarding a developing story about the possible resignation of a minister, Tara (20) noted, 'If a decision really has been taken, I'll find it interesting, but [...] nine out of ten times it's blether. [...] If [prime minister] says "[He] is staying," then that's not a truth but just an

opinion.' What keeps Tara from clicking is the lack of validity or decisiveness. (Kormelink & Meijer, 2018, p. 676)

A similar evaluation of whether the effort required will be worth it seems to happen when the audience member sees the headline as a *disjointed news fact* (27) — the audience member wants the whole story, not an isolated update:

> Tara (20) is not interested in clicking on isolated updates about developments she is already aware of: 'I don't need to have that information in between, [...] I want the answer, you know, the conclusion.' From a user perspective, even the conclusion of a story can be a disjointed news fact. Mark (52) did not click on a headline concerning a resolved conflict, because he was not aware of the problem in the first place: 'You have to know what the problem is [and] then you can also know: what is the solution? [...] But yeah, just an isolated little fact, I would never read that'. About a headline regarding the Ukraine, he similarly argued that it concerned a detail too small to warrant a click. If he were to click, he would also want to know the context: 'What is the cause? How did it happen? What happened? Why do they do it? What do they want to achieve?' This suggests that Mark would appreciate a headline like '5 things you should know about the crisis in the Ukraine' that allowed him to get a full picture of the situation within one article. (p. 676)

The final three considerations from Kormelink and Meijer's list are also related to considerations of effort, although not necessarily of cognitive nature. What they do is to add premises of practical order to the resulting context (C) from headline interpretation which seem

to decrease the expected relevance of the linked article — that is, the effort required to consume it at that particular moment seems to be too high when compared with the expected contextual effects.

One case is when consuming the content at a particular time *does not fit routine* (28) — when clicking does not match with the audience member's schedule:

> Josh (62) only has a few minutes to check headlines before he leaves for work, where the radio is playing the whole day. He skipped a headline about a poison gas attack in Syria, explaining, 'That's very important, [...] but I'm sure I'll hear it on the radio'. Similarly, Jenna (27) skipped a headline noting she would only click on it if she 'really took the time to really dive into it'. While interested, clicking right now did not fit her schedule. (p. 679)

Another case is when delays between the click on the headline and the delivery of the content, such as long loading times or the presence of pre-roll ads in videos, might discourage the audience member from consuming it due the *disruption* (29) it causes in the experience:

> Bruce (55) illustrates, 'Then you have to sit through commercials before you can watch something. Well, I won't do that, I don't want to.' (p. 679)

While advertisers might be pleased with audience members being exposed to their ads — and, according to the RT hypothesis, engaging in a spontaneous and automatic attempt to find a context in which these messages could achieve relevance for them — the expectation of additional cognitive work unrelated to the article's content might decrease its expected relevance

to a point that the audience member might rather invest this time in something else. The same goes for any content that takes too long to load.

The idea that this kind of practical consideration can be a premise in the context (C) resulting from the headline interpretation is also consistent with the last consideration in Kormelink and Meijer's list: *data-heaviness* (30) — clicking will use up too much data:

> A related reason not to click mostly associated with videos was that the item was data-heavy. Here platform-specificity also plays a role. Joe (26) does click videos about wrestling news on his computer, but not on his smartphone: 'Videos [...] I'd rather not watch on my phone because, well, data heavy.' Clicking would cost him too much. (p. 679)

While a premise such as "I have a fast internet connection on my smartphone, with a large data package" might not be salient when choosing content, a more consequential "My phone is rather slow, and my data package is usually already over by the middle of the month" might be significant enough to make its way into the resulting context from headline interpretation. This might affect the considerations of effort to a point where the expected cognitive effects won't seem enough to warrant a click at that particular time.

***

Based on the above data, the present research makes the claim that this RT-based procedure for news selection is consistent — i.e. non-contradictory — with what Kormelink and Meijer (2018) and Boczkowski and Mitchelstein (2013) observed empirically. In the next chapter, its theoretical implications will be discussed, as well as ways in which it could be empirically tested.

5. Discussion and conclusions

Some important additional inputs for the present research's conclusions can be obtained by discussing the main findings from Boczkowski and Mitchelstein (2013) and Ifantidou (2009) from the proposed model's point of view. This will be the focus of the following sections.

Relevance Theory and Boczkowski and Mitchelstein's News Gap

Boczkowski and Mitchelstein (2013) found that, in general, journalists favoured public-affairs stories and audiences non-public affairs, resulting in a gap whose width could be up to 30 percent points in times of relative normality, but decreasing or almost disappearing during events such as national crisis or a presidential election. Once relative normality was reestablished, the gap would widen again (pp. 16-17, 83).

If journalists and audience members tend to agree on what subjects are most newsworthy when formally asked about it (Atkin et al., 1983; Atwood, 1970) and assuming that both are following RT's Cognitive Principle and trying to maximize relevance, could they both be sincere in their efforts and yet make rather different content choices? According to the present research's model, the answer is yes, due to differences in their cognitive environments.

Journalists and editors process much more information than what ends up published. During this process, many premises might be discarded for being deemed not relevant to the audience at that particular moment. However, the fact that they were processed might make them more accessible to the journalist at a later time — Sperber and Wilson argue that the more one process a premise, the more accessible it becomes; think, for instance, multiplication tables (1986/1995,

p. 76). Consequently, journalists regularly covering a beat are likely to be able to assess or forecast the consequences of information, particularly those not obvious or immediate, with less effort than their audiences.

Furthermore, journalists and editors perform a task that does not figure amongst the audience's main concerns when reading news: They are supposed to constantly rank the relative relevance of events and their related premises. It could be in terms of whether an event is worth developing into news or not, which aspects of it to include or leave out, and in which order and with how much emphasis it will be presented to the audience.

This extra cognitive effort required to explore different possibilities might lead to better assessments on the consequentiality of news, but it is warranted only by the fact that journalists and editors are doing it because it is *relevant to their work* of content production. The procedure used for choosing competing content for *consumption*, as proposed by the present research, might halt the interpretation process before the audience member reaches the same conclusions that made the journalists and editors deem a news item more important than another.

In other words, journalists and their audiences operate in slightly different cognitive environments. This suggests that even if both groups were perfectly aligned in terms of preferences regarding public or non-public affairs and what purposes each news item would be used for, a certain degree of fluctuation in content choices between them would happen due to differences in cognitive effort required to access certain contexts. Therefore, from a RT point of view, the existent of a gap such as the one described by Boczkowski and Mitchelstein is to be expected.

The dynamic nature of the gap is also consistent with RT's approach. The gap will close at times because as the situation changes, certain premises will become more salient. For instance,

it might take some level of specialized knowledge or a work-related need to contextualize news about a housing market bubble that might go bust in maybe five years time and arrive at conclusions that are useful now. However, if the potential crisis evolves toward a real one, facts such as people seeing their mortgages being more expensive than what their houses are worth or news of people losing their homes in foreclosures might turn the subject into something visibly consequential to far more people.

A similar case could happen in an election. While some members of the audience might be able to find relevant contexts to process news about the candidates and their proposals very early on, others might find it only worth of their attention when it becomes very manifest that big changes might be on the horizon — that is, the closer it gets to election day.

In other words, people's interest in news might switch as its consequentiality becomes apparent to *them* — which will depend on *their* cognitive environment, with all its different needs and variances according to what the person is going through at a particular time.

It's important to notice that this is not the same as the audience member not being able to work out the consequences of the news. It means that there might be other contexts in the person's cognitive environment — which might include premises about more pressing needs, even if seemly frivolous — that are more accessible at that particular time. While journalists' efforts to make more manifest "why this matters" are likely to affect the degree of effort that audiences will be willing to make when contextualizing headlines, some people might nevertheless conclude that they will be better off investing their attention in something else at that particular time.

This view adds new nuances to Boczkowski and Mitchelstein's proposition that their findings show that

> (...) there is active avoidance of public-affairs stories rather than passive lack of interest in them, and that in a 'high-choice' media environment such avoidance is easier than it was in the past. This avoidance is an outcome of the perception among consumers that public-affairs stories make greater cognitive demands than stories on non-public-affairs topics (2013, p. 144).

According to the present research, perceived differences in cognitive effort are indeed in play, but they are more complex than deeming public affairs as higher-effort and non-public affairs as lower-effort. An article about changes to a country's voting system might be low-effort for a political scientist, and a headline about sports might be very high-effort for someone who has no interest in it. Furthermore, audience members might meet all their expectations of cognitive effects from processing the headline alone (step 3.2.a in the procedure from chapter 3) — in which case no pageview would be recorded.

Preference for vague headlines

One of the starting points of Ifantidou's research on the effectiveness of headlines from the audience's point of view was a study from Dor (2003), in which he makes a similar analysis from the editors' perspective. Both used Relevance Theory as the theoretical framework, but arrived at rather different conclusions.

Dor (2003) has argued that Relevance Theory could explain traditional guidelines for headline writing adopted by editors from both broadsheets and tabloids. According to Dor's observations, the headline properties prized by editors are:

[1] be as short as possible

[2] be clear, easy to understand, and unambiguous

[3] be interesting

[4] contain new information

[5] not presuppose information unknown to the readers

[6] include names and concepts with high 'news value' for the readers

[7] not contain names and concepts with low 'news value' for the readers

[8] 'connect' the story to previously known facts and events

[9] 'connect the story' to prior expectations and assumptions

[10] 'frame' the story in an appropriate fashion

Dor proposes that these properties can be synthesized in a single one: Make the headline such that it renders the story *optimally-relevant* for the audience members (p. 696). However, this would mean different things for different types of newspapers. Tabloid headlines would often attain that by suggesting contexts based on "cliches and prejudices, and feelings of fear, passion and hatred," low on information but also very accessible to the targeted audience; hence its appeal (p. 717). For broadsheets, that would mean the best possible combination of high information, low processing effort, and pointing the readers toward the right context for interpreting the article (p. 716).

Ifantidou questions Dor's view by pointing out that audience members seem not to use headlines primarily as a way to get a summary of full content, as commonly though by broadsheet editors, but as independent elements whose main effect is to arouse interest (p. 716).

According to Ifantidou's observations, most of the guidelines included in Dor's list, such as being short, clear an unambiguous, do not seem to have the effect expected by editors (p. 702).

From the present research's perspective, the apparent contradiction in Dor and Ifantidou's observations is explainable by a claim made by Sperber and Wilson: That when multiple contexts for interpretation are available — which is the case in more ambiguous or vague headlines — the one assumed to be intended by the sender is *the most accessible one for the receiver* (2004, p. 614).

For instance, when interpreting the headline "One baby lives because another died; One mother offers another the circle of life (Kirkey, 2020)," the text does not allow for confirming or ruling out different circumstances in which one baby would have to die for another to live.

If I have some premise accessible in my cognitive environment which would result in the headline yielding a relevant question, a relevance-guided interpretation would encourage me to develop expectations of cognitive effects with respect to that premise, as it is the most accessible context to me in which the headline — guaranteed to be relevant in some context — would be relevant to me. According to the present research's model, as a suitable personal context was found, I would be inclined to click on the headline.

In the mentioned example, the headline was linked to an article about a heart transplant between a baby who became brain-dead two days after birth due to sudden infant death syndrome and another with a congenital heart condition. If I assemble a context based on the premise that the article was about transplants — something alluded to in the headline, but in a very vague way — or more specifically about heart transplants, my expectations of cognitive effects are likely to be fulfilled to some degree; how much exactly will depend on the specifics of the content.

However, if I assemble a different context that turns out to be unrelated to the information in the article — maybe there wasn't enough food for two babies, or I have a particular interest in kidney transplants because a child in my family needs one — the content could be deemed less than optimally relevant, or not relevant at all. I would get less than I expected from the attention I invested in the article, or even a feeling that my time was wasted. But a pageview, and my supposed interest in that particular article produced, would have been recorded.

In the same example, a more specific headline[9], such as one following the properties proposed by Dor, could have different results both in cognitive and in web analytics terms. Following the proposed procedure, different things could happen if it specifically mentions, for instance, "heart transplant." It could confirm my expectations and give me more confidence that my time won't be wasted if I click on the link. Or, I could fail to easily assemble any context in which it could be useful to me (step 2 in the procedure from chapter 3), or it could be deemed irrelevant right away (step 3.1). I could also develop limited expectations of cognitive effects that could be fully met with the processing of the headline itself (step 3.2.a). In these cases, I would probably move on to consider the next headline. It would be optimal to me in terms of getting the most from my time's and attention's worth, but a pageview would not be recorded.

In sum, according to the present research's model, the more vague variants of a headline are likely to get more clicks due to audience members tendency to, when multiple interpretations are allowed, assemble contexts that they would *like* to be covered by the article's content. Only after the content was consumed the audience members will be able to evaluate whether their clicks were good decisions and if the article really reflects their interests.

[9] One example slightly more specific can be found in the article's URL: https://nationalpost.com/news/canada/when-your-baby-needs-a-heart-it-means-waiting-for-another-child-to-die

Conclusions

This thesis concludes by proposing a redefinition of the meaning of pageview data according to the present research's model. In addition, it proposes the incorporation of audience loyalty as a meaningful factor when evaluating the performance of news content. Then, an evaluation of the model is conducted using criteria suggested by Shoemaker et al. (2004). Finally, some practical applications of the model in news analytics and journalistic practice are discussed, as well as further applications of Relevance Theory in Journalism Studies. Each of these sections points to future studies.

What pageviews do mean from an RT perspective

Based in the evidence provided in chapter 4, the present research's hypothesis — that the number of pageviews a news article gets depends not only in the audience's interest in it, but also in the cognitive effort required to find a suitable context to process it — is consistent with the audience reasonings observed by Kormelink & Meijer and Boczkowski & Mitchelstein. The applicability of Relevance Theory to the study of news audience's behaviour seems to be supported. Which brings us back to the central question: what do pageviews mean?

From a RT perspective, a click on a news article link does not mean that the content suggested by it was deemed worthwhile. It means that the interpretation of the headline has created an *expectation* of cognitive effects — conclusions useful to the audience member to one or more purposes — with respect to a particular context assembled by the person. This expectation may or may not be fulfilled, depending on the content that follows after the click.

Only after engaging with the rest of the content — that is, *after* the pageview was recorded — the audience member will be able to evaluate its relevance. Therefore, according to the present research's model, pageviews are not reliable indicators of content value, interest or demand when analyzed in isolation.

If the present research's model is correct, simply directing newsrooms to produce more content similar to what gets more pageviews can be counterproductive in the long run. There's the risk of creating a bias for developing content that emphasizes context accessibility at the expense of meaningful cognitive effects. While such low-effort, low-return content might have some value for the audience depending on their purpose for its consumption, its success depends more on the author's skill in manipulating the audiences' cognitive tendencies than producing quality information. Trying to operate in this market segment would put news organizations against many competitors who might not depend on producing content based on facts or in following ethical standards in order to succeed in business terms.

For editors in particular, the present research's model implies that writing headlines that points audience members towards the correct contexts in which the content could be relevant might result in fewer pageviews overall but, counter-intuitively, that *this not necessarily bad*. If the time saved by *not* clicking on content that is less than optimally relevant is invested in other content that turns out to be relevant or very relevant, the overall interaction is likely to be deemed more relevant. This in turn should affect the audience members' loyalty and the site's or app's priority with respect to other competing content sources.

For news analytics, accounting and controlling for these cognitive tendencies becomes even more important as news organizations, like many other companies in different sectors, are increasingly exploring the applications of *machine learning* in order to understand and predict

the behaviour of the people who interact with their services. Machine learning (ML) is a subfield of artificial intelligence (itself a field of Computer Science) concerned with the development of algorithms to go through massive amounts of data in order to find useful patterns. These patterns could then be used to develop computational models to either classify new data or make predictions about it (Kelleher & Tierney, 2018, pp. 97-98). A classic example is a spam filter: By examining large numbers of messages manually labelled as "spam" or "not spam" — called the *training set* — ML algorithms can create models that establish statistical correlations between the labels and the values of certain features, such as the presence of specific words (e.g. "Viagra") or combinations of words (e.g. "double your income"), message length, and so on. These models can then be applied to new, unlabelled messages in order to predict whether they should be flagged as spam (pp. 104-114).

For publishers, particularly those whose revenue comes mostly from online advertising, it might be tempting to label news articles as "good" or "bad" based on a certain threshold of pageviews and then apply machine learning techniques in order to find which features have the most influence in the number of visualizations an article gets. However, if the present research's model is correct, this is likely to lead to conclusions that characteristics such as vague headlines — allowing the audience members to come up with a context that they would like to be related to the article — or suggested subjects that are rather frivolous — but very accessible and guaranteed to be relevant — are part of the features that define the "good" articles. Once again, these conclusions might lead to decisions that could be detrimental to the news organization in the long run.

Loyalty as an indicator of relevance

According to the present research's model and the situations discussed above and in previous chapters, pageviews alone cannot be taken as indicators of demand, interest or assessments of content value. And, to complicate matters further, as proposed in step 3.2.b of the procedure from chapter 3, a *lack* of clicking does not necessarily mean that the content is irrelevant, as the headline alone might fulfill the audience member's expectations of relevance. However, despite the complications that RT brings to the interpretation of web analytics, it also provides clues on how to control for the factors presented in this chapter and in chapter 3. They are based on the Cognitive Principle: "Human cognitive processes, we argue, are geared toward achieving the greatest possible cognitive effect for the smallest possible processing effort. To achieve this, the individual must focus his attention on what seems to him to be the most relevant information available" (1986/1995, p. 1; 2004, p. 610).

If the present research's hypothesis and model are correct, it is very hard to use raw pageview data from any single article to make any inferences about whether audience members wanted it or found it relevant. As discussed in previous sections, pageviews are recorded before the audience member can assess whether the cognitive effects resulting from the content meets, surpasses, or falls short of the expectations. One possible solution, based on the Cognitive Principle, could be taking into consideration the interactions of each audience member with the news website across time.

A brief sketch of how this interaction could be analyzed from an RT point is as follows. If an audience member interacts with more than one article in a visit to a website or app, one of them might be just optimally relevant (the useful conclusions simply offset the invested effort),

another one highly relevant (many useful conclusions, or a few very important ones, for a rather small effort), another is irrelevant (the article's content has nothing to do with what the headline suggested), and so on. It seems reasonable to assume that the overall relevance of this visit (or, in web analytics terms, this *session*) would be the sum of all cognitive effects obtained divided by the time and effort invested.

One might have more than a single news app available on their phones, or more than a news website that could be consulted. Even if the audience member uses many different ones per day, people will have their favourites, which might be the first one they check in the morning, the ones that they allow to send push notifications, and so on. The present research proposes that, based on RT's Cognitive Principle — that cognition's long-term goal is to maximize relevance — it is reasonable to assume that this ranking of news sources' priority is connected to the cumulative results of the "relevance score" of multiple visits to the site or app. If a news organization constantly delivers high relevance, it is likely to rise in the ranking. If the results are inconsistent, it might fall in priority; and if it wastes one's times with irrelevant content too many times in a row, it might be abandoned altogether.

Therefore, a possible method to enrich pageview data is *to consider whether it came from a loyal user of the site or app or from "drive-by" traffic*, and analyze these two data groups separately. If the present research's model and hypothesis are correct, data coming from loyal audience members should carry more information about what content is succeeding at creating value for the audience.

While this hypothesis requires confirmation, its empirical testability is very straightforward. Most analytics platforms used by news organizations allow for filtering data by criteria such as loyal or returning users. Organizations with more advanced technological platforms are able to

record each visit and action taken by each device or registered user, including more accurate estimates of time spent with the content or how far down the text was scrolled. Similarly, the application of machine learning techniques to pageview data should benefit from first grouping content by preferred by loyal audience members, non-loyal, and both, and *then* proceeding to analysis of which features and subjects seem to be the most important ones to each group. All this could be a starting point to further research regarding what subjects the news organization seems to be succeeded at creating value for the audience, and to what purposes.

New scholarly research could focus on verifying whether there are indeed differences in preferences between loyal and non-loyal audience members and if their reasonings support or refute the present research's model.

Model evaluation

A final step in the method for theory development from Shoemaker et al. is to try to evaluate the resulting theory or model, with several criteria being proposed (2004, pp. 171-176). They will now be applied to the present research's model in an effort to identify its potential to contributing to further research, as well as points that require further testing or development.

Testability

In order to be testable, a theory needs to be stated in terms of concepts or variables that can be measured (p. 171). The proposed model's concepts — the decision to click on an article's link, the expectation of contextual effects for particular purposes and the cognitive effort to find

an appropriate context— are not directly measurable. However, their theoretical definitions and the relationship between them can be used to derive more specific hypotheses based on measurable data.

For instance, a new research similar to Kormelink and Meijer's (2018) could also ask the participants to rate from zero to ten their level of expectations regarding the content suggested by the headline. Likewise, they could be asked to rate from zero to ten the difficulty in finding a way (a context) in which such content could be useful or significant to them. Qualitative aspects could be measured by asking what the participants think the linked article is about (in order to observe the influence of headline vagueness or specificity in context-finding) to what purposes they believe the content would be of use for them. Furthermore, additional questions could be developed in order to verify how well the steps of the proposed RT-based procedure for news selection correspond to what participants actually do.

Web analytics experiments to observe the differences in clicking behaviour between loyal and non-loyal audience members are reasonably simple from a technical point of view. Further analysis would require the definition of procedures to create new indices. For instance, one could attempt to create a score of vagueness or specificity based on the number of occurrences (or the absence) of verbs, adverbs and adjectives in the headline.

Falsifiability

In the experimental settings discussed above, there are different ways in which the results could disprove the present research's hypothesis. For instance, the most-clicked headlines could be rated by the participants as low in expected value and high in effort to find a context in which

they could be significant. Another possibility is an absence of statistically significant differences in the reported levels of expectations and effort of contextualization between the headlines that get clicked and the ones that do not.

Parsimony

Shoemaker et al. define parsimony as attempting to reach an ideal balance between theory complexity and its explanatory or predictive power (p. 172). In the present research's case, efforts were made particularly in two fronts. The first was in limiting the quantity of concepts to be imported from Relevance Theory to the most pertinent ones to this study's goal. The second was in keeping the number of steps of the relevance-guided procedure for news selection as low as possible while still able to explain all Kormelink and Meijer's (2018) observations.

Explanatory Power

The proposed model provides a mechanism, inherited from Relevance Theory, that explains a wide range of audience reactions when selecting content. It also reconciles many different observations from other researchers that were difficult to be simultaneously explained by a single theory. And, as previous discussed, it does so in a way that is empirically testable. To the author's knowledge, the present research is one of the first in Journalism Studies that explores pageview data in an explanatory way instead of a mainly descriptive one.

Predictive Power

While the research allows for many predictions based on the relationships established in the hypothesis and the steps of the RT-based news selection procedure, its actual predictive power remains to be tested.

Scope

Shoemaker et al. point out that "the more phenomena that a theory helps us understand, the better the theory," and that "in the social sciences, most theories deal with a limited range of behaviours and therefore are low in power" (p. 173). While the present research is indeed limited in scope to the use of Relevance Theory to the interpretation of pageview data, RT might be a useful theoretical tool for other problems in Journalism Studies. This will be discussed in the next section.

Cumulative Nature of Science

> Theory is not static but is changing and growing. Research is cumulative, with later studies building on earlier studies. Through this process, theory is continuously refined as we test hypotheses with appropriate evidence. New studies probe the unanswered questions left by earlier studies. In this way, theories move to a closer approximation to the truth (pp. 173-174).

One contribution of the present research to the cumulative knowledge in Journalism Studies is showing that observations from rather different studies — such as Boczkowski & Mitchelstein (2013), Kormelink & Meijer (2018), Ifanditou (2009) and Dor (2003) — are consistent from the proposed theoretical model. This could add independent support to those studies and also provide new ideas about how to further improve them. Furthermore, it introduces a new theoretical tool to news audience studies based on pageview data, which might be useful in the design of new studies.

Degree of Formal Development

> Theories range greatly in their degrees of formal development. We can visualize a continuum of degree of formal development of theories. At the lower end, we would have "areas of research" in which concepts are being developed, hypotheses are being formulated, and data are being gathered, but there is not an effort to be exhaustive about the parts of a theory or to arrange propositions in a logical system. At the higher end, we would have theories made up of systems of propositions, with some logically deduced from others. These more formal theories would also include most or all of the elements of theory discussed in this book—concepts, theoretical definitions, operational definitions, hypotheses, theoretical linkages, operational linkages, limits, and assumptions (p. 174).

Based on Shoemaker et al.'s classification, the present research has a fairly high level of formal development. Improvements could be made in terms of operational definitions of variables, and some possibilities were discussed in the section regarding testability.

Heuristic Value

> A theory is valuable when it helps us generate ideas for research and when it leads to other theoretical ideas. The more new hypotheses that can be generated from a theory, the better the theory (p. 176)

Many hypotheses for news and web analytics experiments could be derived from the present research's model. Some could be simple, such as repeating previous experiments splitting the datasets based on the audience member's loyalty. An example of a more elaborated one could be predicting that some new users of a news app are likely to have an initial period of high engagement, which then suddenly drops and ends with the user abandoning the app — that is, the initial high activity was the audience member *looking for* relevance, and not necessarily *finding* it. Further analysis of what content users that become loyal consumed versus what the abandoning ones consumed could lead to insights about where the news organization is succeeding at creating value for its audience and where it is lacking.

Practical applications in news analytics and journalistic practice

The practical uses of the present research for journalists would not be direct, but through the incorporation of its concepts into newsrooms KPIs — *key performance indicators*; in analytics, the way that one measures success with respect to a certain goal (Kaushik, 2010, p. 37). The present research's contribution lies in providing both publishers and analytics teams with a rationale to reexamine the prominence of pageview data as a KPI not based on a normative basis,

but on a mechanism that points out possible factors to be considered and generates hypotheses that are empirically testable. One example of such an approach was provided by *The New York Times* (2017), when outlining their goals for 2020:

> (…) We are not trying to maximize clicks and sell low-margin advertising against them. We are not trying to win a pageviews arms race. We believe that the more sound business strategy for The Times is to provide journalism so strong that several million people around the world are willing to pay for it (p. 3).

> (…) The newsroom needs a clearer understanding that pageviews, while a meaningful yardstick, do not equal success. To repeat, The Times is a subscription-first business; it is not trying to maximize pageviews. The most successful and valuable stories are often not those that receive the largest number of pageviews, despite widespread newsroom assumptions. A story that receives 100,000 or 200,000 pageviews and makes readers feel as if they're getting reporting and insight that they can't find anywhere else is more valuable to The Times than a fun piece that goes viral and yet woos few if any new subscribers. The data and audience insights group, under Laura Evans, is in the latter stages of creating a more sophisticated metric than pageviews, one that tries to measure an article's value to attracting and retaining subscribers. This metric seems a promising alternative to pageviews (pp. 24-25).

The development of the aforementioned metric is a good example of a practical application that could benefit from the present research's model. While the concept that pageviews from loyal audience members are more important than non-loyal ones is already incorporated — by

stating that their success is measured in terms of how many subscribers they have — the present research, and RT at large, could provide an explanation about why audience members behave the way they do. And, more importantly, it would suggest other variables that could be manipulated in order to generate news analytics experiments aiming at determining which content features are responsible for converting casual readers into subscribers.

It's noteworthy that this approach has allowed the *Times* to, starting in 2012, get more of its revenue from its subscribers than from its advertisers (p. 4). While the difference was marginal at first, it has been growing ever since, and in 2019 the company projected the ratio to reach 70% from subscriptions and 30% from advertising by 2025 (Doctor, 2019). While most news organizations have far more limited reach (and potential for subscriptions) than the *Times*, focusing on content favoured by a more loyal audience, even if the organization does not charge a subscription, is likely to allow for higher advertising CPM rates through better targeting, based on better knowledge about the audience's profile.

Finally, the *Times* example is also important as an illustration that, even if the organization's business model, editorial values and guidelines are not based on maximizing traffic, simply telling journalists that they are not being judged by the number of pageviews their content gets is not enough. Another indicator must be provided, and hence the importance of theory that might guide its development process.

Further possibilities for audience segmentation and content value estimation

An important hypothesis derivable from this work and going deeper into Relevance Theory is that the content consumed by audience members — that is, what content they spend time with

and go well beyond the headline — might give indications of what kinds of contexts are accessible to them. If confirmed, this could allow for a way to *measure not only the audience quantity, but also its quality*. Each pageview could have a different weight according to a value — commercial or institutional — attributed to the audience segment from which the audience member comes from, and a score could be calculated based on each pageview multiplied by this weight. This way, a news item with fewer pageviews but appealing to audience segments considered to be important to the news organization could be comparable with another one with much more pageviews, but from segments deemed to be of less value; examples could include drive-by traffic, or consumers of content that can be easily produced without the use of journalistic techniques, such as humour and opinion.

One example of research that has shown results that encourage testing this view is a large-scale survey about what audiences in the Netherlands expect from journalists and what they should typically cover, conducted by Wurff and Schoenbach (2014). They grouped various expectations into two groups. The first was called Civic Demands, related to what scholars and journalists consider to be important democratic functions of the press and generally aligned with public affairs topics as defined by Boczkowski and Mitchelstein (2013). The second was named Citizen Demands, covering the "complaints and wishes of the citizens" in their private lives, aligning with non-public affairs (pp. 441-443). One of their observations was that

> The findings show that better educated respondents have stronger Civic expectations, whereas lower educated ones show stronger Citizen expectations (Table 4). We also note that older people favour both types more strongly, in particular Civic expectations of news media. Finally, we note that, after the other sociodemographics are controlled for,

> respondents from a higher social class and males tend to express marginally weaker Citizen Demands too (p. 443).

This type of enrichment of pageview data would allow for the development of performance indicators for news content that could have a significant impact in news organizations. Editors would have tools for planning and evaluating news production in which stories of high societal importance but with a smaller audience could be much easier to justify in business terms, based on the value attached to its members' profile. And once such profiles are determined, publishers could be able to use them to maximize revenue from all their content — a person who reads more complex articles and is likely to belong to a certain educational level and have a higher potential as a consumer conserves this potential when reading simpler or more frivolous content.

This kind of audience segmentation and ad targeting could be helpful in mitigating the effects of the dynamic nature of the gap between audience's and journalists' choices observed by Boczkowski and Mitchelstein (2013). The authors propose that, since journalists' preferences seem to be constant (favouring public affairs) and the audiences' are variable, the best way to deal with it would be by producing more non-public-affairs content in times of normality and switching back to public affairs when important events take place (p. 149). The approach suggested in the above paragraphs could be an alternative to that. Ad targeting would be more focused on reaching the best audience for the ad instead of the largest, therefore minimizing the pressure on newsrooms to constantly redeploy its resources based on what's popular at a given moment.

Further applications of Relevance Theory in Journalism Studies

Relevance Theory, the present research's author believes, could provide new insights for different questions in Journalism Studies. For instance, the mechanisms proposed by RT to explain the behaviour of senders and receivers support the idea that journalists have some influence on how people will invest their attention, with audience members responding to signals sent by editors. However, actual engagement and cognitive effects might be rather limited when the subject doesn't evoke a context for the audience member. These nuances could be interesting points to be discussed in the context of agenda-setting theory.

The most interesting questions that RT could possibly help with are the ones about how the journalistic techniques of content processing create value for its audiences, and what subjects or events could benefit the most from this processing. One possible start is assuming that journalists maximize relevance for their audiences by, first, selecting the events most likely to be consequential for them; second, by selecting the event's aspects most likely to be consequential and, finally, by using this information to produce content in a way that minimizes the cognitive effort required to process it. This is consistent with traditional, professional-values guided journalists practice. What RT could add is a mechanism that could explain how this would work on a cognitive level and allow for new hypotheses about the factors that influence this process.

The present research only skims the surface of Relevance Theory, particularly concerning the details of its proposed mechanisms. The author believes that RT could be very useful in developing explanatory models and theories in Journalism Studies for the above-mentioned questions and others, and hopes that this study is an example that such undertaking would be worthwhile.

***

REFERENCES

Allern, S. (2002). Journalistic and Commercial News Values: News Organizations as Patrons of an Institution and Market Actors. *Nordicom Review, 1-2/2002*, 137-152.

Anderson, CW (2011). Between creative and quantified audiences: Web metrics and changing patterns of newswork in local US newsrooms. *Journalism, 12*(5), 550-566.

Atkin, C.K., Burgoon, J. K., & Burgoon, M. (1983). How Journalists Perceive the Reading Audience. *Newspaper Research Journal, 4*(2), 51-63.

Atwood, L. E. (1970). How Newsmen and Readers Perceive Each Others' Story Preferences. *Journalism & Mass Communication Quarterly, 47*, 296-302.

Barnhurst, K. G., & Nerone, J. (2001). *The Form of News*. New York: The Guilford Press.

Boczkowski, P. J., & Mitchelstein, E. (2013). *The News Gap: When the Information Preferences of the Media and the Public Diverge*. Cambridge: The MIT Press.

Bright, J., & Nicholls, T. (2014). The Life and Death of Political News: Measuring the Impact of the Audience Agenda Using Online Data. *Social Science Computer Review, 32*(2),170-181.

Burton-Roberts, Noel (ed.) (2007). *Pragmatics*. Basingstoke: Palgrave.

Cherubini, F., & Nielsen, R. K. (2016). *Editorial analytics - how news media are developing and using audience data and metrics*. Digital News Reports 2016 (series of reports). Oxford: Reuters Institute for the Study of Journalism.

Doctor, K. 2019. Newsonomics: CEO Mark Thompson on offering more and more New York Times (and charging more for it). *Niemam Lab*. Retrieved March 1, 2020 from https://www.niemanlab.org/2019/11/newsonomics-ceo-mark-thompson-on-offering-more-and-more-new-york-times-and-charging-more-for-it/

Ferrer-Conill, R., & Tandoc Jr., E. C. (2018). The Audience-Oriented Editor. *Digital Journalism, 6*(4), 436-453.

Gans, H.J. (1979). *Deciding What's News: A Study of CBS Evening News, NBC Nightly News, Newsweek, and Time*. Northwestern University Press, Evanston.

Graves, L., Kelly, J., & Gluck, M. (2014). *Confusion Online: Faulty Metrics and the Future of Digital Journalism*. New York: Tow Center for Digital Journalism - Columbia University.

Grice, H. P. (1967). *Logic and conversation*. William James Lectures, reprinted in Grice 1989: 1-143.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge: Harvard University Press.

Hamilton, J.T. (2004). *All the News that's Fit to Sell: How the Market Transforms Information into News*. Woodstock: Princeton University Press.

Interest [Def. 1] (2010). In Stevenson, A., & Lindberg, C. (Eds.), *New Oxford American Dictionary*. Oxford: Oxford University Press. Retrieved August 14, 2019, from https://www-oxfordreference-com.lib-ezproxy.concordia.ca/view/10.1093/acref/9780195392883.001.0001/m_en_us1258561.

Kaushik, A. (2010). *Web Analytics 2.0 : The Art of Online Accountability & Science of Customer Centricity.* Indianapolis: Willey Publishing.

Kelleher, J. D, & Tierney, B. (2018). *Data Science*. The MIT Essential Knowledge Series. Cambridge: The MIT Press.

Keller, M., & Abelson, B. (2015). *Newslynx: A Tool For Newsroom Impact Measure*. New York: Tow Center for Digital Journalism - Columbia University.

Kormelink, T. G., & Meijer, I. C. (2018). What clicks actually mean: Exploring digital news user practices. *Journalism, 19*(5), 668–683.

Kirkey, S. (2020, February 14). One baby lives because another died; One mother offers another the circle of life. *National Post*. Retrieved March 1, 2020, from https://nationalpost.com/news/canada/when-your-baby-needs-a-heart-it-means-waiting-for-another-child-to-die

Lee, A. M., Lewis, S. C., & Powers, M. (2014). Audience Clicks and News Placement: A Study of Time- Lagged Influence in Online Journalism. *Communication Research, 41*(4), 505-530.

Levinson, S. C. (1989). A Review of Relevance. *Journal of Linguistics*, *25*(2), 455-472.

MacGregor, P. (2007). Tracking the online audience: Metric data start a subtle revolution. *Journalism Studies, 8*(2), 280-298.

Matthews, P. H. (2014) ”Relevance Theory." The Concise Oxford Dictionary of Linguistics. : Oxford University Press, January 01, 2014. Oxford Reference. Retrieved March 1, 2019, from http://0-www.oxfordreference.com.mercury.concordia.ca/view/10.1093/acref/9780199675128.001.0001/acref-9780199675128-e-2862>

Meyer, P. (2009). *The Vanishing Newspaper: Saving Journalism in the Information Age*. Columbia: University of Missouri Press.

Napoli, P. M. (2012). Audience Evolution and the Future of Audience Research. *The International Journal on Media Management*, *14*, 79–97.

Petre, C. (2015). *The Traffic Factories: Metrics at Chartbeat, Gawker and The New York Times*. New York: Tow Center for Digital Journalism - Columbia University.

Schulz, W. F. (1982). News structure and people's awareness of political events. *Gazette*, 30, 139–153.

Shoemaker, P. J., Tankard, J. W., & Lasorsa, D. L. (2004). *How to build social science theories* : SAGE Publications, Inc.

Simon, H. A. (1978). Rationality as Process and as Product of Thought. *American Economic Association, 68*(2), 1-16.

Soria, B., & Romero, E. (2010). *Explicit Communication: Robyn Carston's Pragmatics*. Basingstoke: Palgrave.

Sperber, D., & Wilson, D. (1986/1995). *Relevance: Communication and Cognition*. Oxford: Blackwell.

Sperber, D., & Wilson, D. (2004). Relevance Theory. In L. Horn & G. Ward (eds), *Blackwell's Handbook of Pragmatics* (pp. 607-632). Oxford: Blackwell.

Singer, J. B. (2011). Community Service: Editor pride and user preference on local newspaper websites. *Journalism Practice, 5*(6), 623-642.

Tenenboim, O., & Cohen A. A. (2015). What prompts users to click and comment: A longitudinal study of online news. *Journalism, 16*(2), 198-217.

Tewksbury, D. (2003). What Do Americans Really Want to Know? Tracking the Behavior of News Readers on the Internet. *Journal of Communication, December 2003*, 694-710.

The New York Times. (2017). *Journalism that Stands Apart: The Report of the 2020 Group*. Retrieved March 1, 2020, from https://www.nytimes.com/projects/2020-report/

Usher, N. (2012). Going Web-First at The Christian Science Monitor: A Three-Part Study of Change. *International Journal of Communication, 6*, 1898–1917.

Welbers, K., van Atteveldt, W., Kleinnijenhuis, J., Ruigrok, N., & Schaper, J. (2016). News selection criteria in the digital age: Professional norms versus online audience metrics. *Journalism, 17*(8), 1037-1053.

Wurff, R., & Schoenbach, K. (2014). Civic and Citizen Demands of News Media and Journalists: What Does the Audience Expect from Good Journalism? *Journalism & Mass Communication Quarterly, 91*(3), 433–451.

Zamith, R. (2017). A Computational Approach for Examining the Comparability of 'Most-Viewed Lists' on Online News Sites. *Journalism & Mass Communication Quarterly, 95*(1), 122–141.