Are our students really interested in Science? Or does Google Trends show a social desirability bias in Brazilian public opinion surveys?

Renato P. dos Santos

ABSTRACT

Discovering public understanding of science allows both to formulate public policies and to understand the acceptance of technological innovations, as to improve ways of Science popularizing & teaching and to understand the reasons that lead young people to choose or not scientific careers. People do like Science and Technology but are happy enough not to know very much about it, even at the risk of a huge price tag for the ordinary citizen. General public opinion surveys on interest in Science in Brazil intended measuring public interest in predefined topics previously selected by experts using "forced-choice" questions. However, this methodology is subject to a "socially desirability bias" that may lead respondents to inform a preference for more "socially desirable" responses to certain sensitive issues. Conversely, there is evidence that the same respondents feel at ease in the privacy of their Internet searches. In this paper, Google Trends was employed as a non-survey-based methodology to verify if those results on public opinion surveys on interest in Science and Technology in Brazil have been overestimated. Combined methodologies of search engine data with other forms of inquiry may be more suited "to the analysis of our fast-moving technological and socio-political context." If significant portions of the population are reluctant to disclose views which could be construed as socially unacceptable, such polls may overestimate the actual levels of support for public policies in Science & Technology, with immediate impact on Science Education.

Keywords: Science education. Interest in Science. Google Trends. Public opinion surveys. Socially desirability bias.

Nossos estudantes estão mesmo interessados em Ciência? Ou o Google Trends indica um viés de desejabilidade social nas pesquisas brasileiras de opinião?

RESUMO

Descobrir o entendimento público da ciência permite tanto a formulação de políticas públicas quanto compreender a aceitação de inovações tecnológicas, melhorar as formas de popularização e ensino da Ciência e compreender os motivos que levam os jovens a escolher ou não carreiras científicas. As pessoas gostam de Ciência e Tecnologia, mas estão satisfeitas em não saber muito a

Renato P. dos Santos é Doutor em Física. Atualmente, é Professor Adjunto do Programa de Pós-Graduação em Ensino de Ciências e Matemática – ULBRA. Universidade Luterana do Brasil/PPGECIM. Endereço para correspondência: ULBRA/PPGECIM, Av. Farroupilha, 8001, prédio 14, sala 338, 92450-900, Canoas, RS.
E-mail: renatopsantos@ulbra.edu.br
Recebido para publicação em 16/08/2016. Aceito, após revisão, em 1/09/2016.

respeito dela, mesmo com o risco de um preço enorme para o cidadão comum. Pesquisas abrangentes de opinião pública sobre o interesse pela ciência no Brasil tiveram o propósito de medir o interesse público em temas previamente definidos e selecionados por especialistas, usando perguntas "de escolha forçada". No entanto, esta metodologia está sujeita a um "viés de desejabilidade social" que pode levar os entrevistados a informar uma preferência para respostas "mais socialmente desejáveis" com relação a certas questões sensíveis. Por outro lado, há evidências de que os mesmos entrevistados se sentem à vontade na privacidade de suas pesquisas na Internet. Neste artigo, o Google Trends foi empregado como uma metodologia não baseada em inquéritos para verificar se esses resultados em pesquisas de opinião pública sobre o interesse em Ciência e Tecnologia no Brasil foram superestimados. Metodologias combinadas de dados de motores de busca na Internet com outras formas de investigação podem ser mais adequadas "para a análise do nosso contexto tecnológico e sociopolítico em rápida mudança". Se porções significativas da população estão relutantes em divulgar pontos de vista que possam ser interpretadas como socialmente inaceitáveis, tais pesquisas podem superestimar os verdadeiros níveis de suporte para políticas públicas em Ciência e Tecnologia, com impacto imediato na educação científica.

Palavras-chave:

INTRODUCTION

Discovering public understanding of science and technology is important not only in the political sphere as in the academic. It allows both to formulate public policies and to understand the acceptance of technological innovations, as to improve ways of Science popularizing and teaching and to understand the reasons that lead young people to choose or not scientific careers (BRASIL, 2015).

It is common sense that Science and Technology have significant and increasing presences in our daily life. Unfortunately, however, Science and Technology may be less graspable today than it was before. Unlike the typewriter in which the connections between the keys and the characters that strike the paper were visible and understandable, the computer is a "black box" whose inner logic and workings remain a mystery to most of us even after inspection of its innards (CAIAZZA, 2005).

Quoting Clarke's Third Law ("Any sufficiently advanced technology is indistinguishable from magic.") (1973, p.21, footnote), Szerszynski argues that contemporary technology, being so mysterious, is seen as 'magical' by the layperson (2005). As a consequence, as Mowbray, former editor of *Popular Science* once said, people do like Science and Technology but are happy enough not to know very much about it (MOWBRAY, 2004). As a matter of fact, "you don't need to know whether a particle is subatomic to charge your cell phone," [*sic*] as Mowbray said (2004).

Science, however, is an enterprise that is run not only by scientists but also by members of the society, who reflect on and shape social perspectives and values (BOERSEMA, 1998). Ignorance about important scientific and ethical issues of global warming and reproductive cloning may carry a hefty price tag for the ordinary citizen as debates seem more shaped by political and economic interests than reason (MOWBRAY, 2004). Therefore, even if people are "happy enough" with this ignorance, they may become much less 'happy' as the consequences of it increasingly affect their lifestyle.

There is a long history of interest being linked with achievement motivation and with educational achievement (DEWEY, 1913). Students' attitudes towards scientific issues are generally regarded as a psychological trait supporting and maintaining the learning process and, consequently, taken as central themes in science education across the world (OLSEN; LIE, 2011). Understanding how interest in science is related to participation in science has practical implications for science educators concerned with students' involvement in both curricular and extracurricular science activities (AINLEY; AINLEY, 2011).

Scientific literacy is one of the three core competencies included in the Programme for International Student Assessment (PISA) study and was the focus of its 2006 edition (OLSEN; LIE, 2011). Students with very low scores on this scale were indicating little interest in learning science (AINLEY; AINLEY, 2011). Nevertheless, we would like to emphasise that, according to Valsiner, interest is not in the object of study, nor in the mind of the student, "but it emerges as a result of processes that link the two in irreversible time" (1992, p.33). Fortunately, then, interest is dynamic and those who lack interest in Science can have their interest triggered and developed (RENNINGER; HIDI, 2015).

However, even if Brazil presented the third largest developments in the overall performance on the PISA exam until 2009, the results regarding scientific literacy and attitudes towards science still situate Brazil at a very distant position from that ambitioned by society (COUTO, 2012).

On the other hand, It is a historical fact that "Brazil doesn't have the tradition of citizen's participation in debates and controversies that involve science" [*sic*] (SIMON et al., 2014). According to SIMON et al. (2014), this could explain why, contrary to the relevance that the public perception of science has assumed in Europe and United States (EUROBAROMETER, 2014; NATIONAL SCIENCE BOARD, 2014), there is a very limited number of general surveys on public's interest in Science in Brazil (ALVES; TOLMASQUIM, 1987; BRASIL, 2010; BRASIL, 2007; BRASIL, 2015).

Our concern here is that those Brazilian surveys were carried out through face-to-face, household, personal interviews using a structured questionnaire, which included asking the person how much they are interested in predefined topics previously selected by experts, an idea of easily measuring that Renninger & Hidi (2015) consider a misconception. To use an example provided by King, on asking someone "did you exercise yesterday?" the researcher has to assume that the respondent has an answer, is willing to tell him, intends to give a genuine response, and that the response happens to be accurate. Possible, but not likely (KING, 2016).

Alternatively, there is empirical evidence that the same person might feel at ease in the privacy and naturalness of her Google searching (CONTI; SOBIESK, 2007) and explore freely the web for news, websites, discussion boards, and other sources of information related to her genuine interests (SCHEITLE, 2011), which may be hardly elicited by other means of data collection (STEPHENS-DAVIDOWITZ, 2014).

Our hypothesis is that interest in Science & Technology may have been perceived as a socially sensitive issue by part of the Brazilian public and, therefore, the results of public

opinion interest in S&T in Brazil obtained through those general surveys may have suffered from a "social desirability bias" and have been overestimated. That was our motivation for using Google Trends[1] (GT) as a non-survey-based data-mining methodology to test it. We are definitely not proposing to completely abandon face-to-face surveys, nor that Internet search data may be more accurate concerning interest in Science & Technology, but that combined methodologies of search engine data with other forms of inquiry may be useful for social science research.

FACE-TO-FACE INTERVIEWS VS. INTERNET SEARCHES

The usual face-to-face interview builds itself as a setting that inevitably constitutes a form of *social* interaction between the interviewer and the respondent. Consequently, some of the interviewees, for fear of being "censured" or socially shunned for their choices, may just bow to the social pressures of the interview and give answers aimed at pleasing the interviewer rather than painting themselves in an unfavourable light (BERINSKY, 1999; CEREZO; HURTADO, 2009, p.100). This tendency is known as a "social desirability bias" (BERINSKY, 1999; EDWARDS, 1953). Tourangeau (2007) found that some 30%–70% of those who test positive for cocaine or opiates by urinalyses deny having used drugs recently. Even it this may be an extreme example, as it involves not only very private but also legal issues, it nevertheless adds to the argument that public opinion obtained through questionnaires, interviews, or self-reports may be a pale reflection of genuine collective public sentiment on particular socially sensitive issues (BERINSKY, 1999; GUO; ZHANG; ZHAI, 2010). As Stephens-Davidowitz (2014) puts it, "How can we know how much racial animus costs a black presidential candidate if many people lie to surveys?"

Edwards (1953) found evidence that the probability of a subject endorsing an item increases linearly with her judged desirability of that item. It is worth emphasizing, however, that social desirability is usually conceptualized a societal construct (KULAS; STACHOWSKI, 2012). As a matter of fact, Edwards (1953) considers that the desirability of a behaviour is defined within and by a group: if a pattern of behaviour is prevalent among members of a group, it will be considered desirable; if it is uncommon, it will be regarded as undesirable. Now, if certain behaviours that are not common among members of the group or culture to which the subject taking the inventory belongs, notwithstanding, are perceived as having a high social desirability scale value by the group or culture to which the interviewer belongs, according to Edwards (1953), this subject may then try, consciously or unconsciously, to give a good impression of himself, by twisting his answers in such a way as to make himself out as having more of the behaviours perceived as socially desirable and fewer of those as socially undesirable.

[1] http://www.google.com/trends/

According to Trevisan (2014), several academic studies see advantages of employing combinations of search engine data accessed via GT as a non-survey-based methodology for social science research in comparison with traditional forms of inquiry "that are unsuited to the analysis of our fast-moving technological and socio-political context." Firstly, the behaviour of individual search for terms online occurs in a natural environment rather than in an experimental laboratory or a setting for purposeful survey or questionnaire. Secondly, GT provides an understandable, transparent and vivid display of covert human curiosity. Trevisan (2014) argued that these outcomes stem from the fact that those relied on what Internet users actually did with search engines, as opposed to discussing what they said they do in questionnaires, therefore mitigating research bias and the incidence of incomplete or false responses. Mellon (2013) concluded that certain search terms could adequately capture public opinion trends in the contexts of US, UK, and Spain.

The arguable assumption here is that if people are interested in a particular issue, they will likely google the web for resources, news, websites, discussion boards, and other types of information related to it (SCHEITLE, 2011). An argument for this is given by Ainley & Ainley (2011) when they affirm that students with individual interest in science will have acquired a reasonable body of scientific knowledge and understanding, enjoy participating in science activities, will generate 'curiosity questions' and extend their knowledge through the process of seeking answers to those questions. However, even if using search queries to infer users' interests by which "the topic in which the user was interested can only be imputed by the researcher," Rose and Levinson (2004) found evidence that the search term alone is sufficient to classify the presumed intent of the query.

Admittedly, the aim of doing a survey is to discover one's view, whereas the purpose of a search is usually to find some information (MELLON, 2013). Paradoxically, researchers may only be able to unlock the potential of the search data when its validity can be tested with enough data generated with conventional survey tools, being subject to draw false conclusions otherwise (MELLON, 2013).

On the other hand, GT has its well-known limitations and biases. Even if the Internet and search engines are increasingly used to find information, people might prefer going online (or, more frequently, using Google) to entertain themselves. According to Waller (2011), there is empirical evidence that, on average, "only about half of search queries are carried out to fill an actual knowledge gap" between what a user knows and what she needs to know. Besides, Internet search engine is not only an interface to information or a shortcut to websites; it is equally a site of leisure, which amounts to about one in six of all searches (WALLER, 2011).

Thousands of scientific works have been made using GT as an understandable, transparent, and vivid display of covert human curiosity (GUO; ZHANG; ZHAI, 2010), in various areas of knowledge. For recent representative examples, consider Public Health (CAVAZOS-REHG et al., 2015; NSOESIE; BROWNSTEIN, 2015), Economics (HEIBERGER, 2015; SCOTT; VARIAN, 2014), Education (YIN et al., 2013; ZHANG et al., 2015), and Politics (BANTIMAROUDIS, 2015; SINCLAIR; WRAY, 2015), among others. More in line with the scope of the present study, Guo, Zhang, & Zhai (2010)

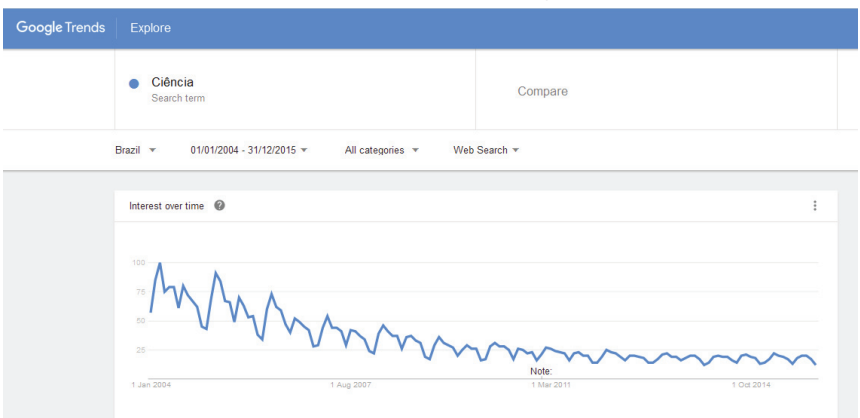
have used GT for the study of human curiosity, understood as a desire to acquire new information and knowledge, and its measurement, while Segev & Baram-Tsabari (2009a, b) have used GT to explore the public interest in Science.

MATERIAL AND METHODS

Google search engine stores about one hundred billion Web searches monthly, all identified by time and place of origin, to be later used by its highly profitable advertising programs, such as *Google AdWords, DoubleClick*, *Google Analytics*, and *Google AdSense*, from which comes 90% of Google Inc. revenue (GOOGLE INC., 2015). Fortunately, this stored information has also been used by various public analytical tools released the last few years, such as *Google Trends* and *Google Correlate* (available at http://www.google.com/trends/correlate).

GT was officially launched in May 2006, built on the idea behind *Google Zeitgeist* (closed in 2007)(MAYER, 2006), and since 2012 also includes resources from what was previously known as *Google Insights for Search* (CLAIBORNE, 2008; MATIAS, 2012) and *StateStats*[2]. GT allows users to sort through several years of Google search queries from around the world or from a particular country or State. It then provides a graphical plotting showing the popularity of particular search terms over time; see an example of the search by the word 'Science' in Figure 1. It also provides a list of 'related topics' (Figure 2), that is, popular search terms that are similar to the one entered, as well as a list of 'related queries', i.e., terms that are most frequently searched together with the one introduced in the same search session, within the chosen category, country, or region (GOOGLE INC., [S.d.]).

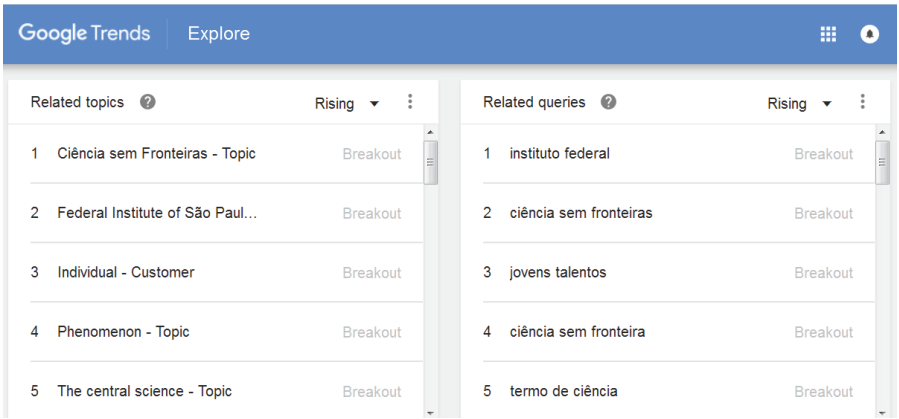
FIGURE 1 – Search example with Google Trends.



Data Source: Google Trends (www.google.com/trends).

[2] Initially launched as an experimental project from the extinct Google Labs in 2008, it tracked the popularity of Google searches per state and then correlated the results to a variety of metrics. Its address (http://statestats.appspot.com/) now redirects to Google Correlate.

FIGURE 2 – Searches related to the above one.



Data Source: Google Trends (www.google.com/trends).

The datasets generated by these searches can be downloaded as comma-separated-value (.csv) spreadsheet-friendly files for further processing.

It must be noticed, however, that GT does not provide absolute search volume raw data, but only normalized search data to the total number of searches done on Google Search Engine (GOOGLE INC., [S.d.]) on the geography and time range it represents, the resulting numbers then scaled to a range of 0 to 100 (GOOGLE INC., [S.d.]). As an example, if, at most, 10% of searches for the given region and time period were for "pizza," GT considers this 100 (GOOGLE INC., [S.d.]).

After using GT, search data was downloaded and analyzed with the resources of the R statistical data analysis language. The interested reader in the data files, the details of this analysis, and/or the R code used can access them from a GitHub repository[3].

Search data from GT was compared with those from the national surveys, which measured public interest in topics previously selected by experts, in an attempt to test our hypothesis that general studies on public's interest in Science & Technology in Brazil may have suffered from a "social desirability bias" that led to overestimated results. However, as GT time span starts on 2004, the 1987 survey (ALVES; TOLMASQUIM, 1987) was discarded. The 2006 (BRASIL, 2007) (field period from November 25 through December 9, 2006), 2010 (BRASIL, 2010) (field period from June 23 through July 6, 2010), and 2015 surveys (BRASIL, 2015) (field period from December 22, 2014 through March 16, 2015) were analysed.

[3] http://github.com/RenatoPdosSantos/GT-and-Brazilian-public-opinion-surveys.git

RESULTS AND DISCUSSION

All 2006, 2010, and 2015 surveys posed the same "forced-choice" question "Please tell me if you have high interest, little interest or no interest in each of these subjects." Those topics, translated into English, are listed in the first column of Table 1. The values shown in the remaining columns of this table represent the average measures of interest to this question for each survey (0 = no interest, 3 = very interested). As one sees from the table, there were only small variations in positioning from one survey to another. These results were interpreted as indicating that Brazilians have a high specific interest in Science & Technology, Environment, and Medicine (BRASIL, 2015).

TABLE 1 – Measures of public interest to selected topics, according to Brazilian surveys (0 = no interest, 3 = very interested.).

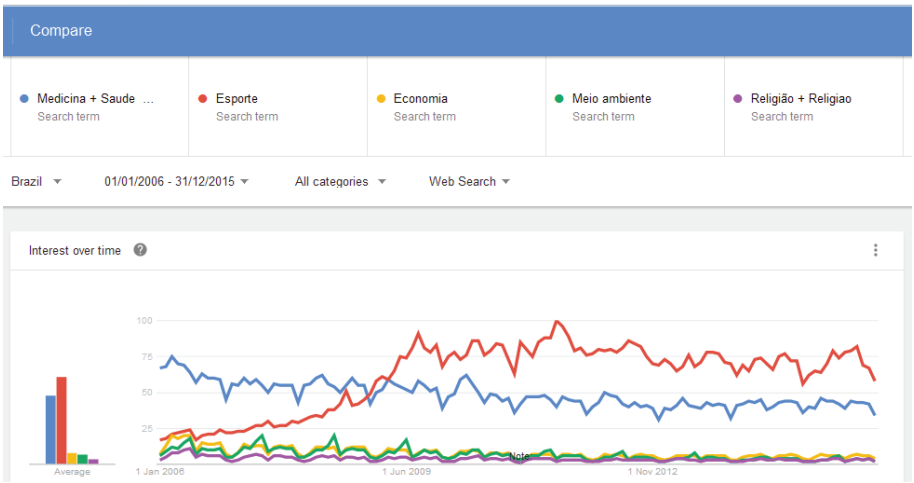
Themes	2006	2010	2015
Medicine & Health	2.52	2.17	2.07
Environment	2.49	2.23	2.07
Religion	2.47	2.08	2.08
Economy	2.38	1.94	1.85
Sports	2.25	1.81	1.65
Science & Technology	2.18	1.80	1.75
Art & Culture	2.15	1.71	1.64
Fashion	1.96	1.38	1.17
Politics	1.84	1.01	0.99

Source: Translated into English from (BRASIL, 2007), (BRASIL, 2010), and (BRASIL, 2015), respectively.

GT was then used to check how much those terms in Table 1 are searched. As GT allows for at most five keywords each search (GOOGLE INC., [S.d.]), only the first five topics from Table 1, namely 'Medicina+Saúde' (Medicine & Health), 'Meio Ambiente' (Environment), 'Religião' (Religion), 'Economia' (Economy), and 'Esporte' (Sports) were included in the first search. As Google differentiates misspellings as well as spelling variations in the search terms (GOOGLE INC., [S.d.]), these were included in this search[4] as for example, 'Medicina+Saúde+Saude' (Figure 3).

[4] https://www.google.com.br/trends/explore?date=2006-01-01%202015-12-31&geo=BR&q=Esporte,Medicina%20%2B%20Saude%20%2B%20Sa%C3%BAde,Economia,Meio%20ambiente,Religi%C3%A3o%20%2B%20Religiao

FIGURE 3 – Search by the terms corresponding to 'Medicine & Health', 'Sports', 'Economy', 'Environment', and 'Religion'.



Data Source: Google Trends (www.google.com/trends).

Since GT provides only relative search volume data (GOOGLE INC., [S.d.]), we included again the 'Medicina+Saúde+Saude' (Medicine & Health) group term in the second search to act as a standard for the normalization of the remaining terms and to provide a means to compare both searches. Therefore, the second search[5] included the group terms 'Medicina+Saúde+Saude' (Medicine & Health), 'Ciência+Ciencia+Tecnologia' (Science & Technology), 'Arte+Cultura' (Art & Culture), 'Moda' (Fashion), and 'Política+Politica' (Politics).

These datasets were averaged for each of the 2006, 2010, and 2015 years, and the results are presented in Table 2.

For the reasons explained above, these figures are relative and cannot be directly compared with the ones in Table 1. However, it makes clear that, among those terms in Table 1, the most searched ones in Brazil on 2006. 2010, and 2015 were 'Medicina +Saúde+Saude' (Medicine & Health), 'Arte+Cultura' (Art & Culture), 'Esporte' (Sports), 'Ciência+Ciencia+Tecnologia' (Science & Technology), and 'Moda' (Fashion), which are not exactly those on the top positions of Table 1.

[5] https://www.google.com.br/trends/explore?date=2006-01-01%202015-12-31&geo=BR&q=Medicina%20%2B%20Saude%20%2B%20Sa%C3%BAde,Arte%2BCultura,Moda,Ci%C3%AAncia%2BCiencia%2BTecnologia,Pol%C3%ADtica%2BPolitica

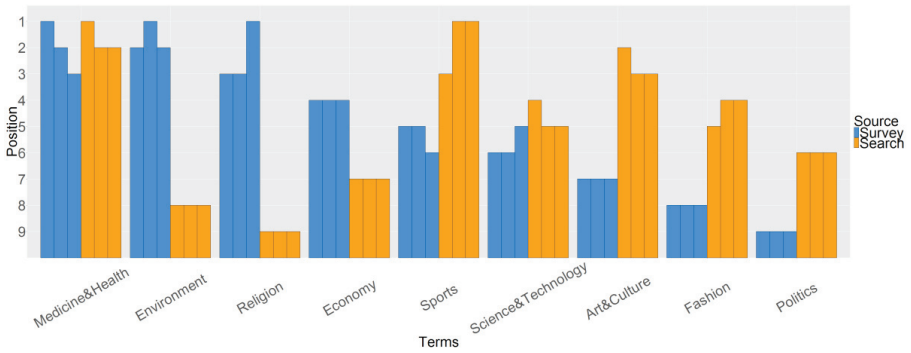
TABLE 2 – Search volumes of the topics from Table 1, according to Google Trends.

Search terms	2006	2010	2015
Medicine & Health	60.8	47.8	39.2
Art & Culture	52.3	32.3	23.3
Sports	20.2	78.6	67.0
Science & Technology	18.4	12.3	8.2
Fashion	15.2	27.6	15.8
Politics	13.6	7.9	5.8
Economy	13.4	7.1	5.3
Environment	9.7	6.4	3.2
Religion	6.2	3.4	2.9

Data Source: Google Trends (www.google.com/trends).

To make comparisons easier, the graph of Figure 4 combines data from Table 1 and Table 2 to show the relative positions of the topics, both from the opinion surveys and Google searches.

FIGURE 4 – Comparison of the relative positions between opinion surveys and Google searches.



Source: Research.

As one sees from Figure 4, apart from Medicine & Health and Science & Technology, the topics in which the public declared the highest interest are visibly not the most searched ones, and vice-versa.

The terms in Table 1 and Table 2 are, however, very general and are not likely to be searched. GT also provides data on the most searched terms on each region and up to five different time periods. Table 3 exhibits the 15 most searched terms in Brazil[6] on

[6] http://www.google.com/trends/explore#geo=BR&date=1%2F2006%2012m%2C%201%2F2010%2012m%2C%201%2F2015%2012m&cmpt=date&tz=Etc%2FGMT%2B3

2006, 2010, and 2015. Figures are again relative: the most searched term for the given region and time period is normalised to be 100 (GOOGLE INC., [S.d.]).

TABLE 3 – Top 15 searches on 2006, 2010, and 2015 in Brazil, according to GT.

2006		2010		2015	
Queries	**Volumes**	**Queries**	**Volumes**	**Queries**	**Volumes**
orkut	100	jogos	100	facebook	100
brasil	75	orkut	70	Youtube	35
fotos	75	youtube	50	Google	35
jogos	60	globo	45	hotmail	30
download	60	hotmail	35	Globo	20
musicas	40	musicas	35	Jogos	15
letras	35	uol	30	tradutor	15
videos	30	msn	30	videos	15
musica	30	tradutor	30	filmes	15
uol	30	google	30	Uol	15
receita	30	jogo	25	frases	10
concurso	30	yahoo	25	Face	10
mensagens	30	baixaki	20	gmail	10
msn	30	caixa	20	Caixa	10
terra	25	terra	20	Olx	10

Data Source: Google Trends (www.google.com/trends).

Table 3 shows that the most searched terms were even more related to social networks, search engines, games ('jogos'), and videos, an entirely different result from the one exhibited in Table 1. This seems compatible with Segev & Ahituv's (2010) transnational study that demonstrated that most popular search queries in Brazil are about entertainment.

Most of the search terms above may have come from 'navigational queries,' that is, queries entered by the user in order to reach a particular site itself that the user may have visited before or assumes to exist (WALLER, 2011), e.g., Facebook, as distinct from obtaining information about that social network. Others may come from 'transactional queries,' a subset of navigational searches in which the user intends to go to a particular website to undertake some web-mediated activity, such as shopping, internet banking, or downloading files (WALLER, 2011), e.g., games, videos, or songs.

Considering the popular culture and entertaining websites or file types that were the target of these navigational or transactional searches, however, a distinct category 'leisure search,' proposed by Waller (2011), could be a better classification for, at least, some of them. Nevertheless, as the frequency of search queries follows a Zipf distribution (WALLER, 2011), one can rightly suppose that there will also be some 'information queries' with search terms related to the subjects of Table 1 in the following 'long tail.'

It is pretty clear that search terms are not the same as categories. GT categorises search terms into 24 different topical categories but, unfortunately, does not provide aggregated search volume data for categories. These categories are: Arts & Entertainment, Autos & Vehicles, Beauty & Fitness, Books & Literature, Business & Industrial, Computers & Electronics, Finance, Food & Drink, Games, Health, Hobbies & Leisure, Internet & Telecom, Jobs & Education, Law & Government, News, Online Communities, People & Society, Pets & Animals, Property, Reference, Science, Shopping, Sports, and Travel, which divide themselves into sub-categories, which are divided again into third-level subcategories, and so on. Naturally, this categorization tree reflects the observations made on May 2016 and may change at any moment in the future.

Then, to get a sense of the most searched categories, for each term in each of the three 50 top searches lists, a search query was submitted to GT, which returned the main and most frequent classifications, allowing its categorization. It is, of course, impossible to be completely sure of what kind of information each user intended to acquire from a search term; however, this classification suggested by GT was based on the majority of search results and carefully controlled by the researcher. Therefore, it is assumed that this assignment of the category to the relevant term has a reasonable degree of confidence. Table 4 exhibits the aggregated results of this classification.

TABLE 4 – Top search categories on 2006, 2010, and 2015.

Category	2006	2010	2015
Arts & Entertainment	410	250	145
Internet & Telecom	150	185	120
Online Communities	100	90	110
Law & Government	95	25	5
Finance	90	45	20
Games	85	150	25
People & Society	65	25	10
Reference	55	40	30
News	15	35	10

Data Source: Google Trends (www.google.com/trends).

It is noticeable that those 'Science & Technology', 'Environment', and 'Medicine & Health' categories that led to the favourable conclusions on Brazilians' interest in Science (BRASIL, 2015) do not even appear in Table 4. However, the high position of the Internet & Telecom category could still provide a base to suppose such interest. To resolve this matter, we drilled down the categories tree to the lowest level, most specific subcategories to those top search terms. The aggregated results are presented in Table 5.

TABLE 5 – Top searches subcategories from 2006 to 2015.

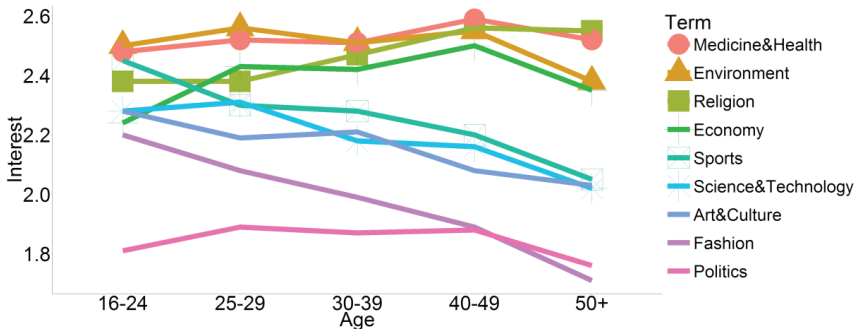
Subcategory	2006	2010	2015
Music Streams & Downloads	150	35	10
Web Portals	100	105	50
Social Networks	100	90	110
Banking	90	45	20
Romance	90	25	10
Games	70	125	15
Text & Instant Messaging	60	75	55
Search Engines	45	55	45
Dictionaries & Encyclopaedias	35	40	20
Vehicle Licensing & Registration	25	15	5
Phone Service Providers	20	35	15
ISPs	10	10	5

Data Source: Google Trends (www.google.com/trends).

Table 5 confirms that most popular search queries in Brazil are about entertainment, even more vehemently contrasting with the results from the one in Table 1. This resonates with Ainley & Ainley findings, in which students with high level of interest express it with sayings such as "I enjoy what I do in class" and "I find that learning is a lot of fun" (AINLEY; AINLEY, 2011).

Susceptibility to social influence decreases with age, and the social desirability bias was observed to vary systematically from youngsters to elders (PARK; LESSIG, 1977). Therefore, one could reasonably expect any discrepancies between queries and searches results coming from this bias to attenuate with interviewee's age, that is, the older the subject, the more genuine the declared interest. Contrary to the other surveys, 2006 one provides a data breakdown of the interviewee's interests by age. The results are shown in Figure 5.

FIGURE 5 – Interviewee's interests vs. age.

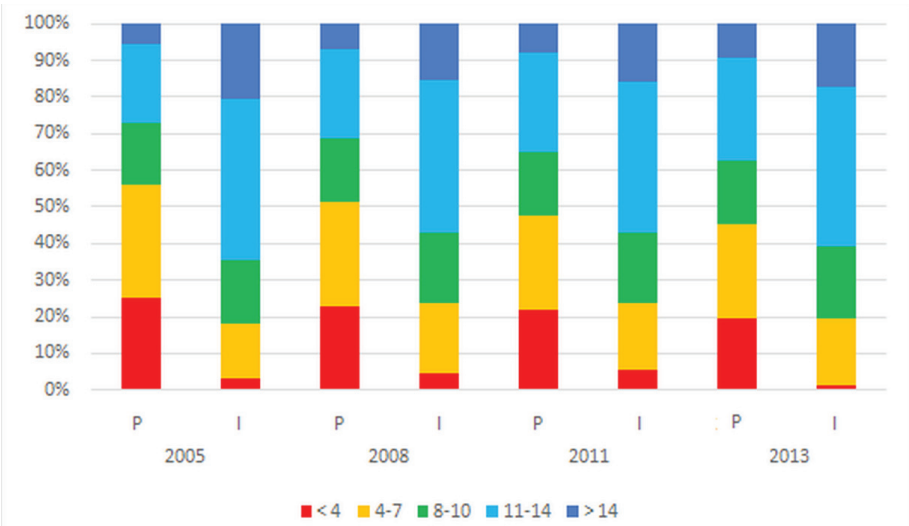


Source: (BRASIL, 2007)

As one clearly sees from Figure 5, apart from 'Medicine & Health,' 'Environment,' 'Religion,' 'Economy,' and 'Politics,' the older the subject, the lesser the declared interest in the topics. Particularly noticeable are the decreases in the reported interest in Fashion and in 'Science & Technology.' It would be surely comforting to assume from this graph that our Science students were even more interested in Science that the general public if they were not also, even more, subject to social desirability bias, making this conclusion even less reliable.

From this data, one could think at first of a digital divide in information uses coming from a socio-demographic bias. According to Segev & Ahituv (2010), there is evidence that higher education and income of users is positively associated with searches for jobs, health, education, news, and other economic and politically-related information, and negatively related to searching the web for entertainment, music, games, sports, and leisure activities. As a result, one could attribute this divergent results to the fact that the Internet user is more likely higher educated and, therefore, unlikely to be representative of the population that was surveyed. Indeed, according to Brazilian surveys on Internet access, (Figure 6), while about 60% of the Internet surfers have had 11 or more years of schooling, less than 40% of Brazil's population have had so many years.

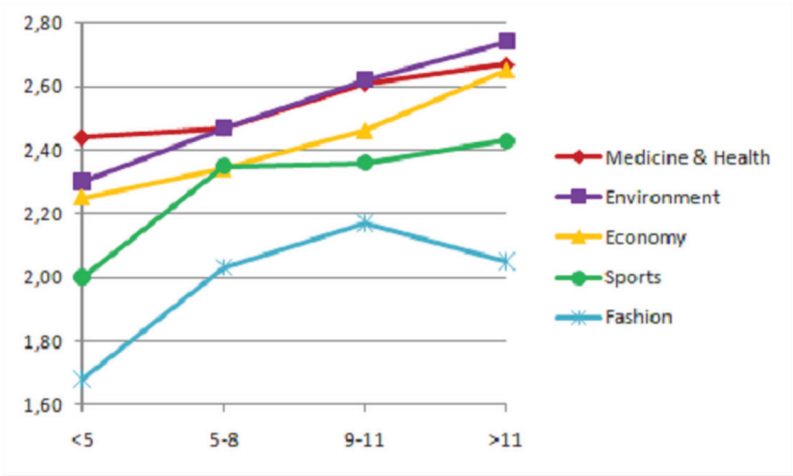
FIGURE 6 – Percentages of the population (P) and Internet users (I) during each reference period, according to the groups of years of schooling.



Source: IBGE – Coordenação de Trabalho e Rendimento, 2007, 2009, 2013, 2015.

To check the soundness of that argument, the survey's own data was used. The 2006 survey also one provides a data breakdown of the interviewee's interests by the number of years of schooling. The results are shown in Figure 7.

FIGURE 7 – Interviewee's interests vs. years of schooling.



Source: Brasil, 2007.

Figure 7 indeed shows that the more schooled the interviewee, the higher interested she expresses herself in scientific themes in 2006. We have no reason to suppose that this trend was different in 2010 or 2015. Therefore, being even more schooled than the general public (Figure 6), the Internet surfers would be expected to be even more interested in scientific and socially relevant themes than in the ones considered of little social prestige (Figure 7). This, however, goes against what the contrasting results of Table 2, Table 3, Table 4, and Table 5 show. Consequently, being more schooled than the general public does not seem to be the cause of the mentioned discrepancy.

Summing up, we interpret this striking divergence between results obtained by express their interests through a face-to-face, 'forced-choice' inventory in topics previously selected by experts, in comparison with the ones from spontaneous, private Internet search behaviours, as an indication of the presence of a high "social desirability bias" (EDWARDS, 1953) induced in part of the interviewees, which led these subjects to hesitate to appear ignorant and inform a preference for more "socially desirable" responses, in accordance with our initial hypothesis.

Albeit those 'forced-choice' questions have been used for decades, as Payne warns, "the fact that something 'works' does not mean that it works correctly" (1951, p.3–4). As mentioned before, the questions included in 2006, 2019, and 2015 surveys are still the same proposed for the 1987 one, while it is beyond question that Brazilian society has suffered many changes during these last almost thirty years. On the other hand, Kreuter, Presser, and Tourangeau (2008) showed that administration of questionnaires through interactive voice recognition (IVR) increased the level of reporting of sensitive information and reporting accuracy relative to conventional computer-assisted telephone interviewing (CATI) and interviewer-administered questionnaires, while Web administration produced

even better results. Eliminating evasive answer bias is certainly inherent to any research procedure, and the researchers should pursue minimizing it.

Opinion polls are undoubtedly useful for analysis and policy (ALBORNOZ; ULLASTRES; ULI, 2009, p.20), and we are definitely not proposing to completely abandon face-to-face surveys, nor that Internet search data may be more accurate concerning interest in Science & Technology. Instead, our point here is that web tools such as GT may provide 'hints' to questionnaires design, and that combined methodologies of search engine data with other forms of inquiry may be useful for social science research and more suited "to the analysis of our fast-moving technological and socio-political context" (TREVISAN, 2013, 2014).

CONCLUSION

As discussed here, even if those opinion polls are useful for analysis and policy, and have been used for decades, they are subject to a 'social desirability' bias (EDWARDS, 1953) that may interfere with the results, and the use of GT may provide new insights into the public understanding of science issues by accessing public's genuine interests as freely expressed in the privacy of its Web searches.

If significant portions of the population are reluctant to disclose views that could be construed as socially unacceptable, such polls may underestimate the real levels of opposition to or overestimate the ones of support for public policies in Science & Technology, with immediate impact on Science Education (BERINSKY, 1999).

ACKNOWLEDGEMENTS

The author would like to thank Dr. Elad Segev from the Hebrew University for many enlightening discussions and suggestions.

REFERENCES

AINLEY, Mary; AINLEY, John. A Cultural Perspective on the Structure of Student Interest in Science. *International Journal of Science Education*, v.33, n.1, p.51-71, Jan. 2011.

ALBORNOZ, Mario; ULLASTRES, Álvaro Marchesi; ULI, Lourdes Arana (Org.). *Cultura científica en Iberoamérica. Encuesta en grandes núcleos urbanos*. . Madrid: FECYT, OEI, RICYT, 2009.

ALVES, Isidoro da Silva; TOLMASQUIM, Alfredo. *O que o brasileiro pensa da ciência e da tecnologia?: A imagem da Ciência e da Tecnologia junto à população urbana brasileira*. Rio de Janeiro e Brasília: MAST e Ministério da Ciência e Tecnologia/CNPq, jan. 1987.

BANTIMAROUDIS, Philemon. A mediated assessment of Samuel Huntington's "Clash of Civilizations": The cultural framing hypothesis. *International Journal of Media & Cultural Politics*, v.11, n.1, p.73-85, 1 Mar. 2015.

BARAM-TSABARI, Ayelet; SEGEV, Elad. Exploring new web-based tools to identify public interest in science. *Public Understanding of Science*, v.20, n.1, p.130-143, 9 Oct. 2009a.
BARAM-TSABARI, Ayelet; SEGEV, Elad. Just Google it! Exploring New Web-based Tools for Identifying Public Interest in Science and Pseudoscience. 2009b, Raanana: The Open University of Israel, 2009. p.20-28.
BERINSKY, Adam J. The Two Faces of Public Opinion. *American Journal of Political Science*, v.43, n.4, p.1209-1230, Oct. 1999.
BOERSEMA, David. The Use of Real and Imaginary Cases in Communicating the Nature of Science: A Course Outline. In: MCCOMAS, WILLIAM F. (Org.). *The Nature of Science in Science Education*. Contemporary Trends and Issues in Science Education. Dordrecht: Kluwer, 1998. p.255-266.
BRASIL. *Percepção Pública da C&T no Brasil 2015:* Ciência e tecnologia no olhar dos brasileiros. Brasília: Ministério da Ciência Tecnologia e Inovação, 2015.
BRASIL. *Percepção pública da ciência e tecnologia no Brasil*. Brasília: Ministério da Ciência, Tecnologia e Inovação/Museu da Vida, 2007.
BRASIL. *Percepção pública da Ciência e Tecnologia no Brasil:* resultados da enquete de 2010. Brasília: Ministério da Ciência e Tecnologia, 2010.
CAIAZZA, John C. Athens, Jerusalem, and the Arrival of Techno-Secularism. *Zygon*, v.40, n.1, p.9-21, Mar. 2005.
CAVAZOS-REHG, Patty, et al. Monitoring marijuana use and risk perceptions with Google Trends data. *Drug and Alcohol Dependence*, v.146, p.e242-e243, 1 Jan. 2015.
CEREZO, José Antonio López; HURTADO, Montaña Cámara. Apropiación social de la ciencia y participación ciudadana. In: ALBORNOZ, MARIO; ULLASTRES, ÁLVARO MARCHESI; ULI, LOURDES ARANA (Org.). *Cultura científica en Iberoamérica. Encuesta en grandes núcleos urbanos*. Madrid: FECYT, OEI, RICYT, 2009. p.81-103.
CLAIBORNE, Trevor. *Announcing Google Insights for Search [Blog post]*. Disponível em: <http://adwords.blogspot.com.br/2008/08/announcing-google-insights-for-search.html>. Acesso em: 13 jun. 2015.
CLARKE, Arthur C. *Profiles of the future:* An inquiry into the limits of the possible. (Revised edition). New York: Harper & Row, 1973.
CONTI, Gregory; SOBIESK, Edward. An Honest Man Has Nothing to Fear: User Perceptions on Web-based Information Disclosure. 2007, New York: ACM Press, 2007. p.112.
COUTO, Sérgio (Org.). *Relatório Nacional PISA 2012:* resultados brasileiros. São Paulo: Fundação Santillana, 2012.
DEWEY, John. *Interest and effort in education*. Boston, MA: Houghton Miffler, 1913.
EDWARDS, Allen Louis. The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology*, v.37, n.2, p.90-93, Apr. 1953.
EUROBAROMETER. *Public perceptions of science, research and innovation*, nº Special Eurobarometer 419. Brussels: European Commission, Public Opinion Analysis Sector, Oct. 2014.
GOOGLE INC. *Find related searches*. Disponível em: <https://support.google.com/trends/answer/4355000?hl=en-GB>. Acesso em: 2 maio 2016a.

GOOGLE INC. *Google Inc. Announces Second Quarter 2015 Results*. Mountain View, CA: [s.n.], Jul. 2015.
GOOGLE INC. *How Trends data is adjusted*. Disponível em: <https://support.google.com/trends/answer/4365533?hl=en-GB>. Acesso em: 11 maio 2016b.
GOOGLE INC. *Search tips for Trends*. Disponível em: <https://support.google.com/trends/answer/4359582?hl=en-GB>. Acesso em: 1 maio 2016c.
GOOGLE INC. *Trends graphs and forecasts*. Disponível em: <https://support.google.com/trends/answer/4355164?hl=en-GB>. Acesso em: 29 abr. 2016d.
GUO, Shesen; ZHANG, Ganzhou; ZHAI, Run. A potential way of enquiry into human curiosity. *British Journal of Educational Technology*, v.41, n.3, p.E48-E52, May 2010.
HEIBERGER, Raphael H. Collective Attention and Stock Prices: Evidence from Google Trends Data on Standard and Poor's 100. *PLoS ONE*, v.10, n.8, p.e0135311, 10 Aug. 2015.
KING, Gary. Preface: Big Data Is Not About The Data! In: MICHAEL ALVAREZ, R. (Org.). *Computational Social Science: Discovery and Prediction*. [S.l.]: CUP – Cambridge University Press, 2016.
KREUTER, F.; PRESSER, S.; TOURANGEAU, Roger. Social Desirability Bias in CATI, IVR, and Web Surveys: The Effects of Mode and Question Sensitivity. *Public Opinion Quarterly*, v.72, n.5, p.847-865, Dec. 2008.
KULAS, John T.; STACHOWSKI, Alicia A. Social Desirability in Personality Assessment: A Variable Item Contamination Perspective. *The International Journal of Educational and Psychological Assessment*, v.11, n.1, p.23-42, Aug. 2012.
MATIAS, Yossi. *Insights into what the world is searching for – the new Google Trends [Blog post]*. Disponível em: <http://insidesearch.blogspot.co.il/2012/09/insights-into-what-world-is-searching.html>. Acesso em: 8 maio 2013.
MAYER, Marissa. *Yes, we are still all about search [Blog post]*. Disponível em: <http://googleblog.blogspot.com.br/2006/05/yes-we-are-still-all-about-search.html>. Acesso em: 18 abr. 2014.
MELLON, Jonathan. Where and When Can We Use Google Trends to Measure Issue Salience? *PS: Political Science & Politics*, v.46, n.2, p.280-290, 28 Apr. 2013.
MOWBRAY, Scott. Ignorance: The Cost Goes Up [Editorial]. *Popular Science*, v.264, n.1, p.6, Jan. 2004.
NATIONAL SCIENCE BOARD. Science and Technology: Public Attitudes and Understanding. *Science and Engineering Indicators 2014*. Arlington, VA: National Center for Science and Engineering Statistics, 2014.
NSOESIE, Elaine O.; BROWNSTEIN, John S. Computational Approaches to Influenza Surveillance: Beyond Timeliness. *Cell Host & Microbe*, v.17, n.3, p.275-278, 11 Mar. 2015.
OLSEN, Rolf Vegar; LIE, Svein. Profiles of Students' Interest in Science Issues around the World: Analysis of data from PISA 2006. *International Journal of Science Education*, v.33, n.1, p.97-120, Jan. 2011.
PARK, C. Whan; LESSIG, V. Parker. Students and Housewives: Differences in Susceptibility to Reference Group Influence. *Journal of Consumer Research*, v.4, n.2, p.102, Sep. 1977.
PAYNE, Stanley Le Baron. *The Art of Asking Questions*. [S.l.]: Princeton University Press, 1951. (Studies in Public Opinion).

RENNINGER, K. Ann; HIDI, Suzanne E. *The Power of Interest for Motivation and Engagement*. London: Routledge, 2015.
ROSE, Daniel E.; LEVINSON, Danny. Understanding user goals in web search. 2004, New York, New York, USA: ACM Press, 2004. p.13-19.
SCHEITLE, Christopher P. Google's Insights for Search: A Note Evaluating the Use of Search Engine Data in Social Research. *Social Science Quarterly*, v.92, n.1, p.285-295, Mar. 2011.
SCOTT, Steven L.; VARIAN, Hal R. Predicting the present with Bayesian structural time series. *International Journal of Mathematical Modelling and Numerical Optimisation*, v.5, n.1/2, p.4, 2014.
SEGEV, Elad; AHITUV, Niv. Popular Searches in Google and Yahoo!: A "Digital Divide" in Information Uses? *The Information Society: An International Journal*, v.26, n.1, p.17-37, Jan. 2010.
SIMON, Fernanda de Oliveira et al. Public Perception of Science: Mapping the concepts of Brazilian undergraduate students of the State of Sao Paulo through Structural Equation Modeling. *American Journal of Theoretical and Applied Statistics*, v.3, n.6-1, p.1-18, 2014.
SINCLAIR, Betsy; WRAY, Michael. Googling the Top Two: Information Search in California's Top Two Primary. *California Journal of Politics and Policy*, v.7, n.1, p.6, 5 Feb. 2015.
STEPHENS-DAVIDOWITZ, Seth I. The cost of racial animus on a black candidate: Evidence using Google search data. *Journal of Public Economics*, v.118, p.26-40, Oct. 2014.
SZERSZYNSKI, Bronislaw. Rethinking the Secular: Science, Technology, and Religion Today. *Zygon*, v.40, n.4, p.813-822, Dec. 2005.
TOURANGEAU, Roger; YAN, Ting. Sensitive questions in surveys. *Psychological Bulletin*, v.133, n.5, p.859-883, Sep. 2007.
TREVISAN, Filippo. *Search Engines and Social Science: A Revolution in the Making*. *Google Forum UK*. Swindon, UK: Economic and Social Research Council, 30 Apr. 2013.
TREVISAN, Filippo. Search engines: From social science objects to academic inquiry tools. *First Monday*, v.19, n.11, p.3, 3 Nov. 2014.
VALSINER, Jaan. Interest: A metatheoretical perspective. In: RENNINGER, K. ANN; HIDI, SUZANNE E. (Org.). *The Role of Interest in Learning and Development*. Hillsdale, NJ: Lawrence Erlbaum, 1992. p.27-41.
WALLER, Vivienne. Not just information: Who searches for what on the search engine Google? *Journal of the American Society for Information Science and Technology*, v.62, n.4, p.761-775, Apr. 2011.
YIN, Chengjiu, et al. Learning by Searching: A Learning Environment that Provides Searching and Analysis Facilities for Supporting Trend Analysis Activities. *Journal of Educational Technology & Society*, v.16, n.3, p.286-300, Jul. 2013.
ZHANG, Meilan, et al. Tracking the Rise of Web Information Needs for Mobile Education and an Emerging Trend of Digital Divide. *Computers in the Schools: Interdisciplinary Journal of Practice, Theory, and Applied Research*, v.32, n.2, p.83-104, 2015.