A socio-technological model of search information divide in US cities

Yong Jin Park
Howard University, Washington, District of Columbia, USA

Received 22 July 2020
Revised 30 October 2020
7 December 2020
Accepted 10 December 2020

Abstract

Purpose – The purpose of the current study is to theorize and apply a socio-technological model – the powerful influence of social determinants in conditioning the effects of information attention on social outcomes. Fundamentally, this study is motivated by the idea that the social determinants of information flow can be used as a predictive tool to inform public socio-policy decisions.

Design/methodology/approach – This study draws upon digital disparity literature and uses publicly available Google search queries in exploring online information attention and its relationships to the HIV/AIDS diffusion in US cities. This study's secondary data collected from extant sources is used to draw attention to a holistic urban ecology under which online search attention represents the variation of information access at the aggregate level.

Findings – The main finding shows that online information attention, as indicated by search trend, is far from being a simple predictor, but operates in complex interactions with existing social environments. A bivariate correlation between AIDS information search and AIDS diffusion rate was found to be significant. However, predictive multivariate models displayed robust effects of social contextual variables, such as income level and racial composition of cities, in moderating the effect of online search information flow.

Practical implications – The importance of these insights is discussed for reducing socio-health disparities at the macro-social level, and policymakers and health administrators are recommended to incubate supportive online infrastructure as an effective preventive measure at the time of a crisis.

Originality/value – The unique contribution of this study is the premise that looks at the aggregate-ecological contour of cities within which the potential benefits of information occur, instead of examining the isolated function of mediated information *per se*. In this vein, online information search, *in lieu* of the exposure to mass media message that is often measured via self-reported items, is a particularly unique and fruitful area of future inquiry that this study promotes.

Keywords Digital inequality, Social determinants, Socio-health disparity, Access, Infodemiology

Paper type Research paper

1. Introduction

Would the use of technology or access to information explain or even predict unequal diffusions of a disease and the spread of a pandemic in metropolitan cities? Castells (2005) offered a socio-technological assessment as follows:

> Technology does not determine society: it is society. Society shapes technology according to the needs, values, and interests of people who use technology. Furthermore, information technologies are particularly sensitive to the effects of social uses on technology itself (2005, p. 3).

This is a powerful statement that helps understand the significance of social use of information technology, i.e., how it might mitigate or exacerbate disparities in urban environments. We can reason, further, that this insight is applied to investigating the potential role of information, access, and its significance in explaining health disparities manifested in social and economic divisions.

This article proposes to test this premise to understand the spread of HIV and AIDS in US cities, with a focus on the predictability of information attention defined in this study as an aggregate level of online attention devoted to searching HIV/AIDS-specific information.

The author expresses his gratitude to anonymous reviewers throughout various versions of this work.

Aslib Journal of Information Management
Vol. 73 No. 2, 2021
pp. 144-159

2050-3806
DOI 10.1108/AJIM-07-2020-0225

Specific questions that this study asks are: (1) the predictive function of information attention in explaining urban health disparities, as indicated by HIV/AIDS rates, and (2) the extent of unequal effects of information. "Search trend data" was used as an operational metric of attention that is indicative of health-related information access embedded in multi-faceted contexts of socioeconomic disparities (van Lent *et al.*, 2017). Doing so, this study attempts to contribute to the field of digital inequality and related policy concerns, as advanced studies (Chen *et al.*, 2014; DiMaggio *et al.*, 2001; Kelley *et al.*, 2016; Rice and Katz, 2000; Rogers, 2003; Scheerder *et al.*, 2017) provided solid grounds to study macro-level social effects of information. We move the discussion of digital disparity into the AIDS epidemic (Whitehead, 1997) in stratified contexts of urban health (Kreiger *et al.*, 2005), asking the extent of differential effects of information in aggregate socio-health outcomes.

Fundamentally, this study is motivated by the idea that the determinants of aggregate information attention can be used as a predictive tool to inform public policy and health administrative decisions (Eysenbach, 2009). In this regard, this study's secondary data collected from extant sources is to draw attention to a holistic urban ecology under which online search attention represents the variation of information access at the aggregate level. Online search queries are the best available measure yet to detect (1) the dynamics of passive yet deliberate information attention *in situ* (Bollier, 2010; cf. Zhu and Bao, 2018) and (2) associated benefits of acquiring information concerning a specific health phenomenon. Importantly, it lends us a capacity for studying sensitive personal behaviors without social desirability bias that individual-level self-reported data entail. Finally, the addition of contextual predictors in predictive multivariate models is to assess the impact of online search attention as it can eradicate or exacerbate disparities embedded in race/ethnicity, income, education, and other aggregate-level socioeconomic indicators (Park *et al.*, 2018; Percheski and Hargittai, 2011).

2. Theorizing socio-technological assessment

2.1 Literature review

Socioeconomic (SES) disparities are dynamic forces that are embedded in social structure as it exerts constant influence on human communicative process. For instance, the history of the telephone illustrates how its use evolved upon socialization practices common to all members of a specific stratum, such as living in urban-working conditions, or being ethnic minorities (Ogburn, 1922). The use of information technology when seen this way is a product of collective-aggregate process, not as an autonomous consumption independent of the environment surrounding an individual. From this premise, we can build understandings of social aspects of information, with stratified environments likely to produce differential outcomes of information access among different populations (Chen *et al.*, 2014; DiMaggio *et al.*, 2001).

Communication scholars (Kelly *et al.*, 2010; Niederdepee, 2008) have been particularly keen on how socioeconomic inequalities, such as education and ethnicities, may influence the effectiveness of communicative messages in preventing the spread of disease at the individual level. Earlier on, renowned researchers (Ettema and Kline, 1977; DiMaggio *et al.*, 2001) pointed out various potential factors that can contribute to widening societal gaps. These may include the differences between low and high SES groups in their characteristics of media exposure, as those with higher educational backgrounds may seek out more informational resources and can be better positioned to acquire knowledge. On one hand, there can be a potential-motivational gap in accessing information because its utility may be deemed as untrustworthy by those from ethnic minority backgrounds.

We can extend the above logic to understand various levels of social dynamics within communities. A study by Valente and Pumpuag (2007), for instance, pointed out that

differentiated flows of health information derive from the availability (or lack) of community resources like "opinion leadership" – local community leaders who can help develop social norms and apply information-awareness into the right community contexts. Related, Rice *et al.* (2012) ran a long-term field experiment and found that in changing public awareness, health-campaign intervention was most effective when there were established opinion leaders in the community. Their finding is significant because it suggests the likely social mechanism through which diverse communities internalize health information resulting in differentiated benefits and outcomes. Likewise, Rogers (2003) documented how influences from social environments can be distinctive as more and more people interact with each other until they reach a critical mass within their community and fully accept a particular health-related norm. Collectively, these insights, as applied at aggregate urban contexts, can set off a solid strand of empirical studies concerning health disparities. What these studies rightly highlighted is the powerful influence of social determinants in conditioning the collective-efficacy of health-related information. Following this line of studies, this study posits that benefit of information (online information search attention-awareness) is moderated through structural contexts of local communities (at the city-level) with unequal diffusions and consequences (outcomes of the HIV/AIDS epidemic).

2.2 Mediated information access and attention

AIDS research had increasingly focused on the patterns of social disparity in HIV-AIDS infection. Socioeconomic disparities in health information access deserve serious attention at the aggregate level, considering the importance of information and public awareness in detecting and preventing the spread of a disease. Earlier studies noted that unmet health attention in concentrated urban environments complicates or even exacerbates existing social disparities. A study by Duh (1991), for instance, convincingly argued that the lower-SES inner-city populations, notably African Americans, remained susceptible to AIDS because their limited financial resources are linked to the lack of health care and low nutritional intake, further creating the condition in which health campaign messages are less likely to be effective among those in inner cities.

Similarly, advanced studies during the 1990s and the 2000s (Freimuth, 1992; Melkote *et al.*, 2000) focused on how to improve the effectiveness in delivering health message and construct persuasive narratives that appeal to target populations. These concerns were also addressed in the studies that measured the effectiveness of AIDS prevention programs disseminated through media channels. Notably, a study by Meekers (2000) tested the effect of a radio program on individual efficacy in rural Africa. Importantly, their study found that perceived significance of oral communication traditions undermined effect of mass media campaign, signaling the need for culturally and socially sensitive health intervention.

Although the past studies varied in their specific focuses, they add up to a largely consistent portrait. First, the access to information is a necessary ingredient in HIV prevention, especially because there is a lack of a medical cure. Second, mass media can play a key role in stimulating or creating public awareness regarding HIV/AIDS infection. Third, health message is not a panacea, of which the effectiveness hinges upon conditional variables of SES and the cultural frames of specific communities. These lessons have been a critical building block for future studies. Yet in identifying gaps in the field, it is also important to note that a considerable number of prior endeavors in AIDS research focused on individual-level behavioral change alone. Subsequently, narrative strategies and persuasion using specifically-tailored media messages had been brought to the forefront of intervention effort (Rice *et al.*, 2012). Analytically, seldom tested in AIDS studies is the function of information and public attention in actual health outcomes, i.e., whether information in the use of technologies mitigates or exacerbates existing health disparity (see DiMaggio *et al.*, 2001;

Park and Shin, 2020; Park *et al.*, 2018). It is fair to say – at least in the field of AIDS research – that little evidence has been documented of structural features of communication as we know very little about the dynamics through which the access to new technologies and SES conditions facilitate or impede the effects of information in health outcomes.

Recently, infodemiology research using online data from Google or other social media sources (e.g., Twitter) opened up a possibility that digital traces can be used to predict or explain the occurrence of health incidents. Various studies investigated such cases as: Ebola detection based on tweet surges (van Lent *et al.*, 2017); tracking the spread of measles with Google queries (Mavragani and Ochoa, 2018); and US state-level cancer incident and Google search. The study by Mavragani and Ochoa (2018) is noteworthy, as they found a positive correlation between public interest, indicated by specific online terms such as "anti-vaccine", and actual health behavior, indicated by the aggregate data (percentages) of immunization.

2.3 US Metropolitan cities: information, Internet and socioeconomic conditions

Our premise is to look at the aggregate-ecological contour of cities within which potential benefits of information occur, instead of examining the isolated function of mediated information *per se*. In this vein, online information search *in lieu* of the exposure to mass media message is a particularly fruitful area of inquiry. As early as the year 2013, Pew surveys (Fox and Duggan, 2013) indicated that more than 61% of US adults relied on online sources when they felt sick or looked for health-related information. Similarly, empirical studies found that health information seeking on Internet was positively related to healthy lifestyle (Laurent and Vicker, 2009; Ramirez *et al.*, 2013), as variances of information access may predict or have an explanatory power of unequal health outcomes. Here it is testable whether these insights gained from individual-level data are also applicable in understanding health disparities at the aggregate city-level.

3. Research questions

Epidemiological studies (Whitehead, 1997) consistently documented the societal gaps deriving from inadequate health attention and resources in concentrated urban environments. Digital divide literature (Chen *et al.*, 2014) also pointed out that access to new technologies, notably, Internet, tends to widen the existing disparities between information "haves" and "have-nots." These differences can be exacerbated, at least in part, by the fact that the population in high SES urban settings may have greater exposure to health-related knowledge and associated information resources (Kelley *et al.*, 2016; Percheski and Hargittai, 2011; Scheerder *et al.*, 2017), with different health consequences at the aggregate level.

Overall, HIV/AIDS may be more than an individual medical problem, but suggestive of SES and technological indicators of the cities related to unequal information diffusion (Ettema and Kline, 1977; DiMaggio *et al.*, 2001; Park, 2018a, b; Percheski and Hargittai, 2011). That is, effects of information do not occur in a contextual vacuum but in the manifold processes through which the use of information technology is being shaped in interaction with existing urban environments. Accordingly, this study asks the extent to which benefits of information can affect particular health outcomes, as conditioned by existing social conditions. Thus, to assess the benefits of information on health outcomes, we ask:

RQ1. What is the potential role that information might play in reducing or exacerbating the disparity of HIV/AIDS diffusion in US cities? Do levels of health-related information attention on Internet predict the disparity in health outcomes at the aggregate level?

RQ2. To what the extent is the impact of health-related information attention on HIV/AIDS diffusion embedded in socioeconomic divisions? Do differences in access to Internet and SES in urban settings moderate the effect of information?

4. Methods

Analytically, one of the main goals in this study is to explicitly test large-scale extant data in a predictive model. Unobtrusive observation of naturalistic consumption may be best suited to test sensitive information attention on such topics as AIDS and HIV. Online search query of Google Trends can be particularly useful for our purpose because they provide us with a direct measure of the volume of public attention in different geographical settings. Search trend data also complement the low external validity of self-reported behavioral measures and allows us to extend individual-level consumption to the city-level aggregate estimates. At a minimum, online information search involves a certain level of awareness – in this case, HIV/AIDS. From a measurement standpoint, what we capture in search query trend is the access to information that is sought in a highly voluntary fashion. Internet has already become a critical source for those who seek health information as well as clinicians, health practitioners, and government officials (Rice and Katz, 2000). Therefore, online queries and their search trend data potentially indicate, not only the degree of information access, but also associated health resources accessible to each city, or even more than just a symptom of the spread of a disease.

Of course, the use of the large-scale extant data has inherent limitations (King *et al.*, 1994). Critics (boyd and Crawford, 2012; Hargittai, 2015) rightly argued that computational analysis of large-scale datasets is often disguised as "more scientific." The core of this criticism concerns a-theoretical quantification of large numbers, which is wide-open to interpretations, as much of the analysis in most of so-called "big data" studies remains data-driven and descriptive. Popular media fueled the debate and tended to treat "big data" as a panacea for predicting any trends. For instance, Google Flu Trend (GFT) report had been publicized for accurately predicting the outbreak of a disease, when in fact correlations are easily bound to occur with any big numbers. Scholars have been well-aware of these limitations, however, and newly-published studies had begun to address much of methodological criticisms through rigorous statistical test (Neuman *et al.*, 2014) and explicit theoretical underpinnings (Scharkow and Vogelgesang, 2011). Following these, we strive to go beyond correlational trend identification and advance analytically through predictive multivariate models. More importantly, this study is guided by a theoretical framework, with analytic inference techniques learned from survey and experiment research as a fundamental building block of this study.

4.1 Measure of independent variable: information search attention

Given the immense universe of information online, the search engine is possibly the only and the cheapest gatekeeper that organizes information sensibly for human intelligence. Logged search queries of Google Trends are used as this study's baseline measure. Certainly, not every online search occurs through Google and other search engine providers like Yahoo and Microsoft also collect and retain search queries. Yet notably, Google makes over 137 billion search queries publicly accessible per year. The tiny market shares of other search engines enable Google to be in a unique position in digital spheres. For example, for the year 2014 alone, Google was responsible for as close as 70% of all information searches conducted in the US, producing over 13.1 billion search results a month. Unless we have empirical data that demonstrate any significant difference in the types of information searches conducted between Google and other search engines, there

is no serious threat to the validity of this study's measure (Scharkow and Vogelgesang, 2011).

This study used search query measures of the two key terms, (1) AIDS (M = 19.50, SD = 12.45; Min = 0, Max = 76.92) and (2) HIV (M = 39.81, SD = 24.28; Min = 0, Max = 83.25), which indicate information attention in related health symptoms (AIDS: acquired immunodeficiency syndrome and HIV: human immunodeficiency virus). It is known that not everyone with HIV infection develops the AIDS. Accordingly, search queries in two dimensions are important to detect differential levels of information access in each city. Specified data intervals were from January 1, 2010 to December 31, 2010, and this period was carefully chosen to match the year 2010-retrieved contextual data at the city-level. The use of generic search terms has the following bases. First, information science literature (Gordon and Pathak, 1999; Schwartz, 1994) indicates that in information seeking, broad terms attract a more common frame of reference among people. Second, narrowly restricted search strings do not generate enough search volume and may be too sensitive to distill certain issue dimensions that otherwise would stand out (Schwartz, 1994). Instead of highly specific medical terminologies, common and simple vocabularies readily accessible to the public will be immune to individual variations and better capture aggregate city-level estimates. We normalized the volumes of search queries for one year, dividing the total aggregate volume by 12 months. This provides rates of search volumes comparable to one another, while the raw total counts can be sensitive to the fluctuations of particular months.

4.2 Measures of dependent variable: HIV and AIDS diffusion rates

Data were imputed from the database of the Center for disease Control (CDC), 2011 HIV Surveillance Report in US metropolitan cities. The CDC database reports on US cities with over 1,000,000 populations and tracks the epidemiology of HIV/AIDS in the US. The CDC Report includes US dependent areas like San Juan, Puerto Rico, which we excluded for our analysis (n = 103).

A series of CDC Reports had been published since 1982, but this study's analysis was designed to match the 2011 Report, within which the year 2010 was the most updated and comprehensive dataset that had both HIV and AIDS infection rates of metropolitan cities in the US and this choice made it possible to use temporally valid and corresponding socioeconomic datasets based on the 2010 US census. Accordingly, 2010 estimates of US city-level data were used as the baseline, but the latest 2015 CDC data were added to consolidate the findings across Time 1 and Time 2 as reported in over-time comparisons. The reported rates for total infections each city were per 100,000 population. Health outcome variables of HIV infection (M = 15.89, SD = 8.99; Min = 1.40, Max = 49.70) and AIDS (the stage 3 HIV) (M = 10.53, SD = 6.49; Min = 0.70, Max = 33.70) were analyzed in two separate models.

4.3 Contextual variables

In assessing effects of information attention, we were interested in two contextual moderations: (1) socio-economics (SES) and (2) broadband access penetration. For this, publicly accessible large-scale extant data were in order. In general, the city-level SES estimates are not widely available. Yet, fortunately for our purpose, the 2010 US census that matches the baseline year of Google Trends search data provides the most complete and updated dataset of following variables: population size (M = 532068.61, SD = 929831.11; Min = 32,736, Max = 8,175,133), median household income (M = 43747.92, SD = 9900.40; Min = 24,421, Max = 81,349), the percentage of college education (25 ages and up) (M = 29.86, SD = 10.63; Min = 10.90, Max = 56.80), age (M = 33.53, SD = 3.08; Min = 23.30, Max = 42.80), and minority populations (%) of Blacks (M = 24.65, SD = 18.80; Min = 0.70, Max = 82.70), Asians (M = 5.28, SD = 7.28; Min = 0.40, Max = 54.80), and Latinos (M = 20.57, SD = 18.33; Min = 1.60, Max = 84.60).

For broadband rate, a measure of urban broadband penetration rate (%) of the state, in which each city was located, was used. Figures are based on the National Telecommunications and Information Administration (NTIA) Broadband Report (2011), which published the 2010 estimates (HH %) (M = 70.26, SD = 4.66; Min = 56.10, Max = 79.50). The selection of these contextual variables was concurrently triangulated but was also in line with this study's theoretical contour (Castells, 2005; Rogers, 2003). Furthermore, we were informed by AIDS studies (Stall *et al.*, 2003), individual-level survey data in health information seeking (Ramirez *et al.*, 2013), and digital disparity literature that consistently indicated the significance of infrastructural access and socioeconomic conditions.

5. Results

The key concept under investigation is the function of information search attention and the extent of its differential effects at the aggregate city-level. We first focused on a big picture by showing main effects of information attention on health outcomes of AIDS and HIV (RQ1). Figure 1 describes baseline measures of AIDS and HIV information search volumes in their

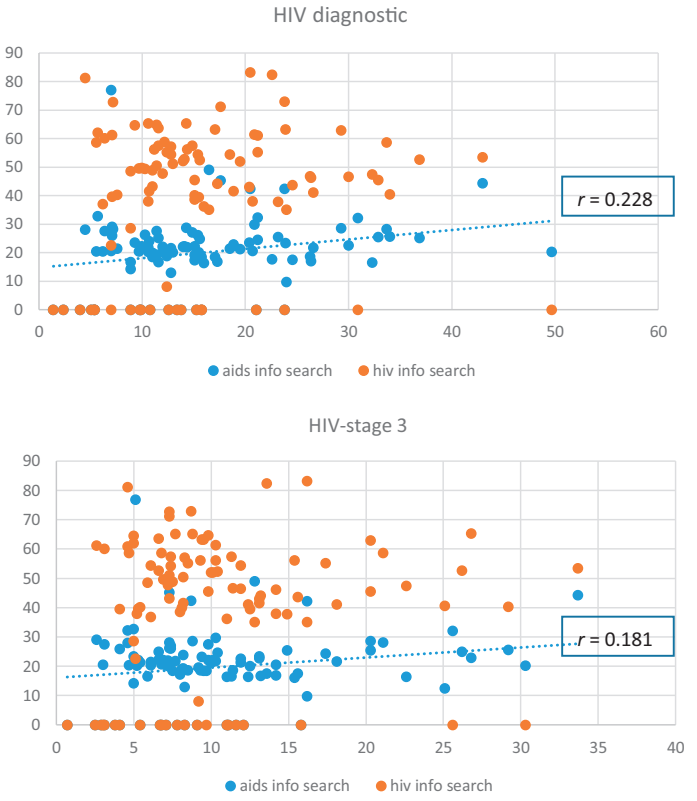


Figure 1.
Information attention and HIV rates, HIV-diagnostic and stage 3 (AIDS)

Note(s): Information search volumes are normalized to be 0 to 100 prior to analysis

respective relationships with AIDS and HIV infection rates. Significant correlations were found between AIDS search queries and each of AIDS and HIV rates ($r = 0.23, 0.17$, respectively). However, the effect sizes were relatively modest and the significance for AIDS (HIV-stage 3) infection remained only marginal at 0.10 level. Further, no correlation was found for HIV search queries, indicating that there was no simple direct relationship between information attention and HIV/AIDS diffusion rates.

To provide more detailed descriptive overviews, Table 1 dissects HIV/AIDS diffusion rates by high and low volumes of information attention, and presents ratios between overall volumes and HIV/AIDS rates. The findings reveal important patterns. First, when broken down into high and low information attention volumes, there was no discernible difference among cities in terms of HIV/AIDS diffusion rates. Second, even high levels of information had no significant relationship with any of HIV/AIDS infection rates, whereas correlations were significant only for those cities with low AIDS search volumes ($r = 0.285, 0.281$, HIV and AIDS rates, respectively). Combined, these indicate that significant correlations were driven by the lower end of HIV and AIDS information search attention, suggesting that the relationships were distinctively non-linear. Finally, the ratios between diffusion rates and information search volumes were tighter for AIDS (1.54) than HIV (3.14, that is, a ratio of about 3 to 1), which hints on a possibility that information attention may be more readily accessible in terms of AIDS than HIV in public minds.

Table 2 estimates the function of search attention in predicting the disparity of health outcomes, as indicated by HIV and AIDS diffusion rates (RQ2). First, in both HIV-diagnostic and stage 3 (AIDS) models, we found no effect of information search (of either AIDS or HIV), controlling all variables. However, it was found that the contextual predictors played significant roles in their respective effects, as an environment condition such as the city population size was found to be significant for HIV/AIDS diffusion rates ($\beta = 0.19$, $\beta = 0.20$, $p = 0.05$). Racial and ethnic backgrounds of Blacks and Latinos were also positively associated in both models ($\beta = 0.67$, $\beta = 0.68$, $p = 0.000$, for Blacks; $\beta = 0.35$, $p = 0.00$, $\beta = 0.24$, $p = 0.05$, for Latinos). Interestingly, we found a significant negative effect of

	Information search attention		Average diffusion rate	SD	Bivariate Correlation	*p*-value	Ratio
HIV Diagnostic	High Volume	HIV Info	16.45	8.55	−0.137	ns	2.50
	Low Volume		15.26	9.42	0.221	ns	
	High Volume	AIDS Info	17.36	9.07	0.049	ns	3.78
	Low Volume		14.29	8.65	0.285	0.050	
							Overall: 3.14
Stage 3 (AIDS)	High Volume	HIV Info	10.53	6.32	−0.062	ns	1.22
	Low Volume		10.54	6.73	0.246	0.085	
	High Volume	AIDS Info	11.25	7.25	0.050	ns	1.85
	Low Volume		9.79	5.57	0.281	0.049	
($n = 103$)							Overall: 1.54

Note(s): Information search volumes are normalized to be 0 to 100 prior to analysis

Table 1. Correlations and ratio of information search volume to HIV rates

Table 2.
Estimates of influence of information attention on HIV-AIDS diffusion

	Predictors	HIV-diagnostic model β	SE	R Square	Stage 3 (AIDS) model β	SE	R Square
Step 1	HIV Info	0.06	0.03	0.06	0.05	0.02	0.02
	AIDS Info[a]	0.16	0.07	0.08	0.13	0.05	0.04
	Broadband Rate	−0.23*	0.19		−0.14	0.13	
Step 2	City Population Size	0.19*	0.00		0.20*	0.00	
	Age	0.14	0.26		0.10	0.19	
	Black (%)	0.67***	0.05		0.68***	0.03	
	Asian (%)	−0.03	0.18		−0.04	0.09	
	Latino (%)	0.35**	0.04		0.24*	0.03	
	HH Income	0.19	0.00		0.20	0.00	
(n = 103)	Education	0.21	0.09	0.41	0.17	0.06	0.37

Note(s): $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; Intercept not reported; H = HIV info search, A = AIDS info search
[a]A separate regression model was run for AIDS information. Cell entries are before-entry standardized coefficients for Step 2. R square reported in Step 2 is based on HIV info, taking all variables into account

broadband rate in HIV-diagnostic model, indicating that those cities with a higher broadband penetration tended to display the lower HIV infection rate ($\beta = -0.23$, $p = 0.05$).

To further investigate the extent to which the access to Internet and socioeconomic conditions (SES) in urban settings moderate the effect of information search attention, interaction terms between search volumes, each of seven socioeconomic variables, and broadband rate were constructed. Table 3 shows interactive relationships from separate hierarchical regression analyses. We found significant but discrete patterns of interactions, which displayed highly conditional impacts of information attention. The moderating roles of median household income remained consistent in HIV-diagnostic and stage 3 (AIDS) models (for HIV information search: $\beta = 0.43$, $p = 0.00$; $\beta = 0.29$, $p = 0.05$, respectively; for AIDS information search: $\beta = 0.40$, $p = 0.05$, in HIV-diagnostics) as a higher level of information tended to benefit those cities with higher level of household income. We also found a significant interaction between education and HIV information search in HIV-diagnostic, which indicates that the cities with a higher education level were more likely to have a lower HIV infection rate when exposed to a higher level of information search ($\beta = -0.22$, $p = 0.05$). In the stage 3 (AIDS) model, however, we found no such significant interaction term.

6. Discussion

Renowned scholars (Rogers, 2003) have proposed a socio-technological model, the premise of which lies in the dynamic forces of socioeconomic (SES) disparities in harnessing the potential of the use of information technologies. We have applied this premise to understand health disparities in HIV/AIDS diffusion rates among US cities and their relationships with

Table 3.
Estimates of interactive influences of information attention on HIV-AIDS diffusion

	Interactions	HIV-diagnostic model β	SE		Stage 3 (AIDS) model β	SE	
HIV Info	x Income	0.43**	1.15		0.29*	0 0.91	
	x Education	−0.22*	0.86	R square: 0.53			R square: 0.43
AIDS Info	x Income	0.40*	1.26	R square: 0.50			R square: 0.42[a]

Note(s): $^{*}p < 0.05$; $^{**}p < 0.01$; Intercept not reported; prior steps include all variables in main effects (n = 103) Only significant interactions shown in the table, with no interaction found between AIDS info and education [a] Change in R square results from the inclusion of all variables and interaction terms in the final model

information search attention on the internet. Recent scholarship (Kelly *et al.*, 2010; Niederdepee, 2008; Valente and Pumpuag, 2007) supports this study's broad conceptual contention that the use of information technology may exacerbate disparities in urban environments as the distinctive influences from socioeconomic conditions are likely to produce different benefits and outcomes (Chen *et al.*, 2014; DiMaggio *et al.*, 2001). Most importantly, we applied these insights at the aggregate city-level to put forth our thesis regarding the powerful influence of social determinants in conditioning the effects of online information search attention. Overall, we examined the extent to which the function of information can be seen as conditional upon existing socioeconomics in metropolitan cities and reached two conclusions.

First, there were positive bivariate correlations between information search attention and diffusion rates. However, the bivariate relationships remained relatively small and confined for AIDS search queries. Furthermore, only a limited impact was found, as any direct effects of information search attention were conspicuously absent in the multivariate models that incorporated socioeconomic conditions. Second, with no apparent benefit of information *per se*, the strong direct effects of SES indicated how deeply disparities related to HIV/AIDS were embedded in existing social structures. For instance, racial and ethnic backgrounds were strong social predictors (Rice *et al.*, 2012) as the cities with more black and Latino populations tended to display higher HIV and AIDS diffusion rates. This finding supports the conclusion of earlier studies (Duh, 1991; Whitehead, 1997) that existing disparities complicate the problem of unmet health resources, particularly for underserved communities in urban environments.

In this regard, the findings concerning the interactions indicate the function of information search attention may not be uniform across all cities. Illustrative is discrete patterns in which health gaps may be exacerbated (see Figure 2). The upper panel of Figure 2 shows that the gap between high and low levels of information tends to be magnified among the cities with lower household income levels. Even when cities benefit (through a lower HIV rate) from information attention, the benefit for lower-income cities happens with the increased or exacerbated gap, whereas the same pattern does not exist for higher-income cities [1]. The lower panel displays a similar pattern in which education exacerbates the disparity between levels of high and low information, while benefiting the cities with higher education levels for having a lower HIV rates. These results carry significant implications in that certain types of urban environments may harness information to their benefits more effectively than others, far from producing a monolithic effect of information. That is to say, the simplistic proposition that 'the more information, the more benefit' may require qualifications. Collectively, the predictability of information attention for aggregate health outcomes turned out to be more complex (not simple correlational) and contextual (environmentally-sensitive) than expected.

The findings concerning descriptive statistics of search volumes merit further discussion, because the results shed lights on subtle dynamics in which the function of information attention is far from linear. Note the difference between high and low information search conditions. The fact that positive correlations existed only among those cities with lower AIDS search volumes indicates that the linear relationship (i.e., the more search, the more AIDS) occurred only at the lower end of AIDS diffusion. This finding is probably different from robust direct impacts of health information found in prior studies at the individual level (Laurent and Vicker, 2009; Ramirez *et al.*, 2013). In other words, the impact of information may differ by different stages of AIDS diffusion, as found in the difference between high and low information conditions. Fundamentally, this suggests that the role of information attention is not easily predictable in a linear manner as one might hope in simply detecting epidemiological diffusion, informing that the increase in information search volume must be carefully interpreted in conjunction with urban contextual variables such as education.

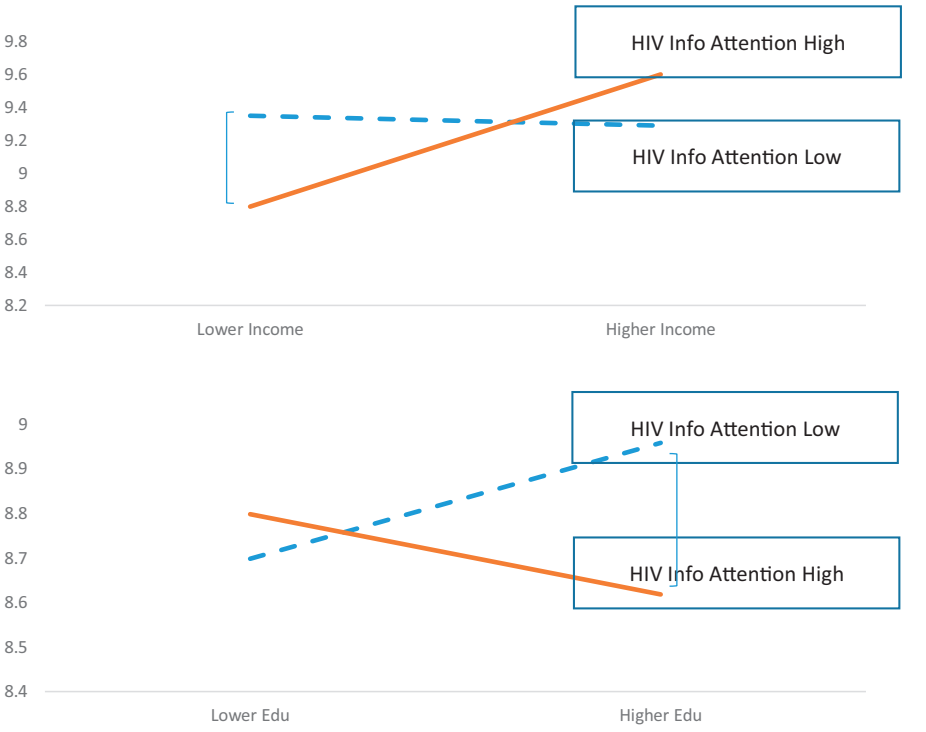


Figure 2.
Moderating role of income and education on HIV-diagnostic

Note(s): Standardized coefficients from the models with all interaction terms are graphed

Finally, the negative effect of broadband penetration in HIV diffusion deserves serious attention. First, the large effect of broadband is in sharp contrast to no impact of information found in any of the HIV-AIDS models. This negative effect is to stress that cities, in which people are exposed to higher broadband penetration, tend to enjoy a lower HIV infection rate, while the benefits of information remain subtle or even conditional. Here we cannot determine any causal directionality; however, this finding suggests that the broader access to Internet infrastructure may be more likely to incubate favorable conditions associated with a wide range of health-related resources (Rice, 2006; Rogers, 2003; also see Benda *et al.*, 2020) than highly-specific information attention alone. This was mapped in Figure 3, which delineates patterns in which broadband may help to reduce the HIV infection among the cities with a higher level of HIV information search attention. The upper-left quadrant (low broadband) represents the cities with high HIV-diffusion rate, whereas the lower-right quadrant (high broadband) displays those cities standing out for better health outcomes. The resulting four dimensions produce a rather cohesive picture in which the cities tend to cluster around broadband rates, creating a health disparity of uneven HIV diffusions even among the cities with higher information attention. Here our thesis is that information alone is not a panacea (nor is it a symptom), but its effect on health outcomes may be a reflection of the broader supporting environments of existing resources.

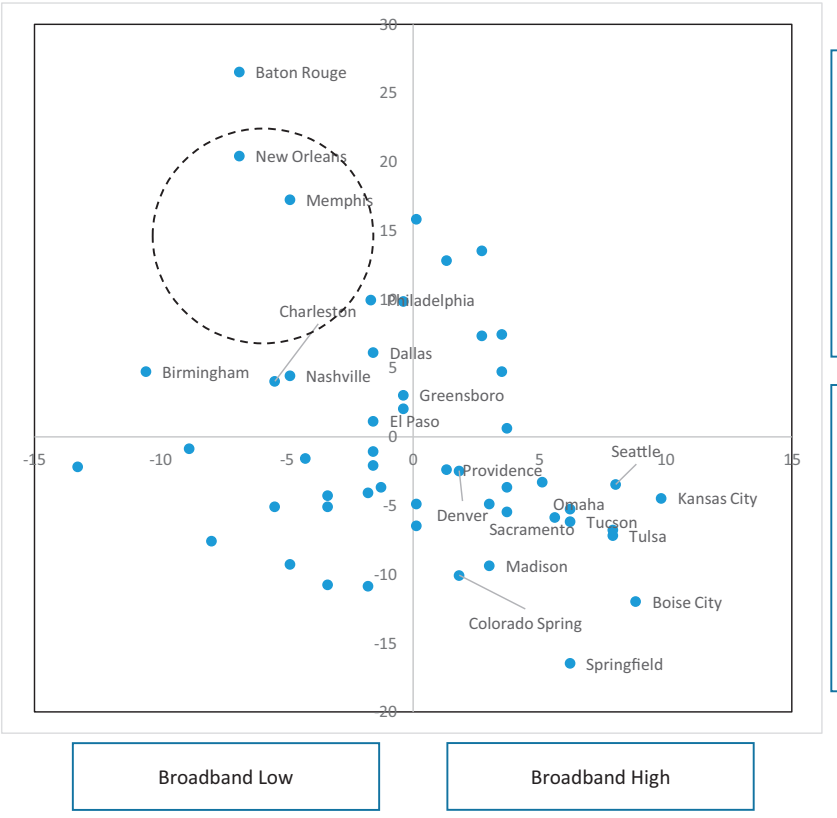


Figure 3.
Mapping of HIV-diagnostic, among high HIV-info cities

Note(s): High (both broadband and HIV rates) denotes those cities above a median value

7. Conclusion

7.1 Practical policy implications

This study's findings carry significant ramifications for municipal health administrators and policymakers who seek effective interventions to address a specific case of HIV/AIDS diffusion, as well as a global health crisis like Coronavirus pandemic in 2020. First, this study demonstrates that access to health-related information can function as a critical element especially when other social structures may be well in place. Note that not all cities remain equipped with infrastructural systems that are capable of translating information to the benefits. This gives rise to a disparity among the cities – the ones with wide supporting resources and the others without such robust socioeconomic systems. In this regard, there seems to be no definite answer to developing the precise mechanism of informational intervention that curtails HIV/AIDS diffusion. Still, the carefully-patterned arrangements of available resources will give stability to the system that can encourage the development of financial and human capitals to fight off the spread of any epidemic. That is to say, the focal point of this study is that the focus on information (or the content of the message) and public attention alone would not be particularly helpful when there are no concerted efforts to allocate resources to nurture the right environments that can translate information into tangible benefits. This point is precisely in line with prior studies that found a relationship

between population health, on the one hand, and information patterns on the other. But these studies also argued that if there were possibilities to develop intervention, the measure to improve the quality of health information should be complemented by supplying the public with appropriate infrastructure, which may in turn incubate related information skills (Ma *et al.*, 2019; Park and Chung, 2017; Percheski and Hargittai, 2011; Wamuyu, 2017).

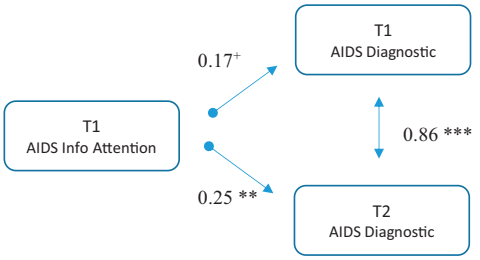
A key challenge is, then, for health administrators to invest in the supportive environments and tools to address or buffer the disparities deriving from the macro-social structure, such as the racial and ethnic characteristics of cities. Notably, the impact of broadband suggests the significance of broad availability of associated informational resources, instead of developing particular lines of mediated information or promoting episodic knowledge of HIV/AIDS-related facts. Practical examples may include urban initiatives that support the quality of broadband and this municipal effort to eradicate divide among different communities within cities. This is particularly the case with sexually transmitted diseases (STDs), such as the AIDS/HIV, which are likely to diffuse through geographically bounded, community-based social networks. The insight will be applicable to fight off a global epidemic crisis like COVID-19, given its geographical diffusion and associated needs for sanguine informational sources. Resulting lessons are not to seek an immediate solution for the socioeconomic divisions through an extravagant health campaign, but to recognize and harness holistic influences from supportive social environments to narrow the gap.

7.2 Direction of future studies

Given these, future research should move forward to expand the application of socio-technological framework to understand social uses of information technology in a holistic urban ecology. The usher into the digital era has brought online health information to the forefront of public attention. Analytically, this will entail the continuous extension of real-time field data, calling for the improvement of their measures to capture aggregate-level public health behavior.

In this vein, more complex field study designs should replace simple correlational "big data" analytics (Hargittai, 2015) as this study attempted to put forth predictive multivariate models that place search query data in social contexts. Still, one question that this study could not explore is whether a specific intervention measure would generate sufficient information attention over time, which can in turn engender substantial changes in health outcomes. A quasi-experimental setup, with a specific stimulus in a natural field, will produce a meaningful real-time dataset to inform policymakers and health administrators with regards to (1) precise volumes and (2) the locales of attention that a particular type of campaign can generate. Moreover, we also do not know the precise routes through which information can be translated into the benefits of health knowledge and desired actions by individuals. To tease out this underlying process, individual-level survey or experimental studies would best complement the macro-level analysis of systematic patterns of HIV/AIDS disparities that this study presented.

Granted this design setup provides a valid snapshot of the year and its urban ecology, warranted is the criticism about triangulated datasets in cross-sectional nature, which eliminates our confidence about data novelty and temporal causality. We cannot tell if search attention precedes AIDS occurrence, or vice versa. In the absence of updated longitudinal datasets over a considerable period, one can speculate correlation between 2010 search queries and 2015 AIDS occurrence, for instance, will be weaker than what's found in this study's concurrent 2010 datasets. Still, additional bi-variate analyses between Time 1 and Time 2 reveal interesting new patterns, displaying the strong effect of T1 info search on T2 AIDS but with no effect on T1 AIDS, as shown in Figure 4. This is fascinating, given that the correlation between T1 and T2 AIDS is close to 1. This result holds, taking account other main SES variables (population, education, and income) ($\beta = 0.25$, $p < 0.01$), which suggests that information effect may be even delayed over time before it becomes sensitive to



Note(s): $^{+}p < 0.10$; $^{**}p < 0.01$; $^{***}p < 0.001$; correlations reported

Figure 4.
Information attention (time 1) and health outcomes (time 1 and 2)

variations of health-related incidents. A longitudinal study of prospective or retrospective city- or state-level design will improve on predictive validity that this study could not achieve. Future studies should test these suggestions – for instance, in understanding the 2020 COVID-19 diffusion and prevention as this study's findings serve as a departure point to recognize discrete functions of different dimensions of information. For those studies, remaining tasks will be to investigate the extent to which even information access that is highly voluntary, rather than the one sitting to be received passively, may not operate independent of constant influences from social environments.

Note

1. The interaction reveals complex dynamics in which preexisting SES conditions shape effects of information. Here the gap between high and low levels of information is exacerbated greatly (although higher information still benefits those cities through lower HIV), displaying effects of information remain non-linear and contingent upon other social conditions.

References

Benda, N.C., Veinot, T.C., Sieck, C.J. and Ancker, J.S. (2020), "Broadband Internet access is a social determinant of health!", *American Journal of Public Health*, Vol. 110 No. 8, pp. 1123-1125.

Bollier, D. (2010), *The Promise and Peril of Big Data*, Aspen Institute, Washington, DC.

boyd, D. and Crawford, K. (2012), "Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon", *Information, Communication, and Society*, Vol. 15 No. 5, pp. 662-679.

Castells, M. (2005), *The Network Society: From Knowledge to Policy*, Johns Hopkins Center for Transatlantic Relations, Washington, DC.

Chen, W., Lee, K.H., Straubhaar, J. and Spence, J. (2014), "Getting a second opinion: social capital, digital inequalities, and health information repertoires", *Journal of Association for Information, Science, and Technology*, Vol. 65 No. 12, pp. 2552-2563.

DiMaggio, P., Hargittai, E., Neuman, W.R. and Robinson, J.P. (2001), "Social implications of the internet", *Annual Review of Sociology*, Vol. 27 No. 1, pp. 307-336.

Duh, S.V. (1991), *Blacks and AIDS: Causes and Origins*, Sage, New York, NY.

Ettema, J.S. and Kline, F.G. (1977), "Deficits, differences, and ceilings contingent conditions for understanding the knowledge gap", *Communication Research*, Vol. 4 No. 2, pp. 179-202.

Eysenbach, G. (2009), "Infodemiology and Infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the internet", *Journal of Medical Internet Research*, Vol. 11 No. 1, p. e11.

Fox, S. and Duggan, M. (2013), *Health Online 2013*, Pew Internet and American Life Project, available at: http://www.pewinternet.org/files/old-media/Files/Reports/PIP_HealthOnline.pdf.

Freimuth, V.S. (1992), "Theoretical foundations of AIDS media campaigns", in Edgar, T., Fitzpatrick, M.A. and Freimuth, V.S. (Eds), *AIDS: A Communication Perspective*, Lawrence Erlbaum Associates, Hillsdale, pp. 91-110.

Gordon, M. and Pathak, P. (1999), "Finding information on the world wide web: the retrieval effectiveness of search engines", *Information Processing and Management*, Vol. 35 No. 2, pp. 141-180.

Hargittai, E. (2015), "Is bigger always better? Potential biases of big data derived from social network sites", *The Annals of the American Academy of Political and Social Science*, Vol. 659 No. 1, pp. 63-76.

Kelley, M.S., Su, D. and Britigan, D.H. (2016), "Disparities in health information access: results of a county-wide survey and implications for health communication", *Health Communication*, Vol. 31 No. 5, pp. 575-582.

Kelly, B.R., Hornik, A., Romantan, J.S., Schwartz, K., Armstrong, A. and Wong, N. (2010), "Cancer information scanning and seeking in the general population", *Journal of Health Communication*, Vol. 15 No. 7, pp. 734-753.

King, G., Keohane, R.O. and Verba, S. (1994), *Designing Social Inquiry: Scientific Inference in Qualitative Research*, Princeton University Press, Princeton, New Jersey.

Krieger, N., Chen, J.T., Waterman, P.D., Rehkopf, D.H. and Subramanian, S.V. (2005), "Painting a truer picture of US socioeconomic and racial/ethnic health inequalities: the public health disparities geocoding project", *American Journal of Public Health*, Vol. 95 No. 2, pp. 312-323.

Laurent, M.R. and Vickers, T.J. (2009), "Seeking health information online: does wikipedia matter?", *Journal of the American Medical Informatics Association*, Vol. 16 No. 4, pp. 471-479.

Ma, S., Zhang, S., Li, G. and Wu, Y. (2019), "Exploring information security education on social media use", *Aslib Journal of Information Management*, Vol. 17 No. 5, pp. 618-636.

Mavragani, A. and Ochoa, G. (2018), "The internet and the anti-vaccine movement: tracking the 2017 EU measles outbreak", *Big Data and Cognitive Computing*, Vol. 2 No. 1, p. 2.

Meekers, D. (2000), "The effectiveness of targeted social marketing to promote adolescent reproductive health: the case of Soweto, South Africa", *Journal of HIV/AIDS Prevention and Education for Adolescent and Children*, Vol. 3 No. 4, pp. 73-92.

Melkote, R., Sundeep, R. and Goswami, S. (2000), "Social and economic factors in an integrated behavioral and societal approach to communications in HIV/AIDS", *Journal of Health Communication*, Vol. 5, pp. 17-27.

Neuman, W., Guggenheim, L., Jang, S.M. and Bae, S.Y. (2014), "The dynamics of public attention: agenda-setting theory meets big data", *Journal of Communication*, Vol. 64 No. 2, pp. 193-214.

Niederdeppe, J. (2008), "Beyond knowledge gaps: examining socioeconomic differences in response to cancer news", *Human Communication Research*, Vol. 34 No. 3, pp. 423-447.

NTIA (2011), "Exploring the digital nation", *Broadband Adoption Report*, available at: https://www.ntia.doc.gov/headlines/2011/ntia-releases-new-broadband-adoption-report.

Ogburn, W.F. (1922), *Social Change with Respect to Culture and Original Nature*, BW Huebsch, New York.

Park, Y.J. and Chung, J.E. (2017), "Health privacy as sociotechnical capital", *Computers in Human Behavior*, Vol. 76, pp. 227-236.

Park, Y.J. and Shin, D.D. (2020), "Contextualizing privacy on health-related use of information technology", *Computers in Human Behavior*, Vol. 105, p. 106204.

Park, Y.J., Jang, S.M., Lee, H. and Yang, G.S. (2018), "Divide in Ferguson: social media, social context, and division", *Social Media + Society*, Vol. 4 No. 3, 2056305118789630.

Park, Y.J. (2018a), "Explicating net diversity in trend assessment", *Communication Research*, Vol. 45 No. 5, pp. 783-809.

Park, Y.J. (2018b), "Social antecedents and consequences of political privacy", *New Media and Society*, Vol. 20 No. 7, pp. 2352-2369.

Percheski, C. and Hargittai, E. (2011), "Health information-seeking in the digital age", *Journal of American College Health*, Vol. 59 No. 5, pp. 379-386.

Ramírez, A.S., Freres, D., Martinez, L.S., Lewis, N., Bourgoin, A., Kelly, B.J. and Hornik, R. (2013), "Information seeking from media and family/friends increases the likelihood of engaging in healthy lifestyle behaviors", *Journal of Health Communication*, Vol. 18 No. 5, pp. 527-542.

Rice, R.E. and Katz, J.E. (2000), *The Internet and Health Communication: Experiences and Expectations*, Sage, New York, NY.

Rice, R.E., Wu, Z., Li, L., Detels, R. and Rotheram-Borus, M.J. (2012), "Reducing STD/HIV stigmatizing attitudes through community popular opinion leaders in Chinese markets", *Human Communication Research*, Vol. 38 No. 4, pp. 379-405.

Rice, R.E. (2006), "Influences, usage, and outcomes of internet health information searching: multivariate results from the pew surveys", *International Journal of Medical Information*, Vol. 75 No. 1, pp. 8-28.

Rogers, E.M. (2003), *Diffusion of Innovations*, Free Press, New York, NY.

Scheerder, A., van Deursen, A. and van Dijk, J. (2017), "Determinants of Internet skills, uses and outcomes. A systematic review of the second-and third-level digital divide", *Telematics and Informatics*, Vol. 34 No. 8, pp. 1607-1624.

Scharkow, M. and Vogelgesang, J. (2011), "Measuring the public agenda using search engine queries", *International Journal of Public Opinion Research*, Vol. 23 No. 1, pp. 104-113.

Schwarz, N. (1994), "Judgment in a social context: biases, shortcomings, and the logic of conversation", *Advances in Experimental Social Psychology*, Vol. 26, pp. 123-162.

Stall, R., Mills, T.C., Williamson, J., Hart, T., Greenwood, G., Paul, J. and Catania, J.A. (2003), "Association of co-occurring psychosocial health problems and increased vulnerability to HIV/AIDS among urban men who have sex with men", *American Journal of Public Health*, Vol. 93 No. 6, pp. 939-942.

Valente, T.W. and Pumpuang, P. (2007), "Identifying opinion leaders to promote behavior change", *Health Education and Behavior*, Vol. 34 No. 8, pp. 881-896.

van Lent, G.G.L., Sungur, H., Kunneman, A., van de Velde, B. and Das, E. (2017), "Too far to care? measuring public attention and fear for ebola using twitter", *Journal of Medical Internet Research*, Vol. 19 No. 6, p. e193.

Wamuyu, P.K. (2017), "Bridging the digital divide among low income urban communities. Leveraging use of community technology centers", *Telematics and Informatics*, Vol. 34 No. 8, pp. 1709-1720.

Whitehead, T.L. (1997), "Urban low-income African American men, HIV/AIDS, and gender identity", *Medical Anthological Quarterly*, Vol. 11 No. 4, pp. 411-447.

Zhu, X. and Bao, Z. (2018), "Why people use social networking sites passively: an empirical study integrating impression management concern, privacy concern, and SNS fatigue", *Aslib Journal of Information Management*, Vol. 70 No. 2, pp. 158-175.

Corresponding author
Yong Jin Park can be contacted at: yongjinp@hotmail.com