SPECIAL ISSUE ARTICLE

JOURNAL OF COMMUNITY PSYCHOLOGY | WILEY

Understanding search autocompletes from the perspectives of English and Spanish speakers during the early months of the COVID-19 pandemic

Pamela Valera[1] | David Carmona[1] | Vivek Singh[2] | Sarah Malarkey[1] | Humberto Baquerizo[3] | Nadia Smith[1]

[1]Department of Urban-Global Public Health, Rutgers School of Public Health, Newark, NJ, USA

[2]Rutgers School of Communications and Information, New Brunswick, NJ, USA

[3]Office for Diversity and Community Engagement, Rutgers New Jersey Medical School, Newark, NJ, USA

Correspondence
Pamela Valera, Rutgers School of Public Health, Community Health Justice Lab (www.chjl.org), 1 Riverfront Plaza, 10th Floor, Newark, NJ 07102, USA.
Email: pv181@sph.rutgers.edu

Abstract

The purpose of the study was to explore differences in Google search autocompletes between English and Spanish-speaking users during the first wave of the coronavirus disease 2019 (COVID-19) pandemic. Twenty-nine individuals who were in areas with shelter-in-place state orders participated in a virtual focus group meeting to understand the algorithm bias of COVID-19 Google autocompletes. The three focus group meetings lasted for 90–120 minutes. A codebook was created and transcripts were coded using NVivo qualitative software with a 95% intercoder reliability between two coders. Thematic analysis was used to analyze the data. Among the 29 participants, six self-identified as White, seven as Black/African American, five as American Indian or Alaska Native, four as Asian Indian, and three as Native Hawaiian or Pacific Islander. In terms of ethnicity, 21 participants identified as Hispanic/Latino. The themes that emerged from the study were: (1) autocompletes evoked fear and stress; (2) skepticism and hesitation towards autocomplete search; (3) familiarity with COVID-19 information impacts outlook on autocomplete search; (4) autocompletes can promote preselection of searches; and (5) lesser choice of autocomplete results

for Spanish-speaking searchers. Spanish speakers expressed concerns and hesitation due to social factors and lack of information about COVID-19.

KEYWORDS

autocomplete search, COVID-19, focus groups, health equity, health information, Hispanics/Latinos, qualitative methods

1 | INTRODUCTION

Coronavirus disease 2019 (COVID-19) is caused by severe acute respiratory syndrome called SARS-CoV-2 and is transmitted by respiratory droplets, which can induce a variety of symptoms such as fevers, shortness of breath, and even a loss of smell/taste (Centers for Disease Control and Prevention [CDC], 2021). COVID-19 has been raging worldwide since it was first declared a pandemic on March 11, 2020 (Katella, 2021). Once it was declared a national emergency, the United States began to issue stay-at-home orders for all residents beginning with California, which affected numerous businesses and their employees (Katella, 2021). As stay-at-home mandates were put into place to combat the COVID-19 pandemic, certain groups, such as the Hispanics/Latinos, were directly impacted by the pandemic at disproportionate rates.

Hispanics/Latinos are the largest racial and ethnic minority group in the United States, accounting for ~19% of the population (US Census Bureau, 2021). Hispanics/Latinos make up ~1/4 (24.8%) of all COVID-19 cases in the United States as of August 2022 (CDC, 2022a). As of July 2022, they are also 2.1 times more likely to be hospitalized from COVID-19 compared with non-White Hispanics (CDC, 2022b). Factors including social determinants of health (American Medical Association [AMA], 2020; Macias Gil et al., 2020; Pareek et al., 2020; Rodriguez-Diaz et al., 2020), inequities in information about COVID-19 (AMA, 2020; Joseph et al., 2020), and limited online search opportunities about COVID-19 in multi-languages (Makhortykh et al., 2020; Singh et al., 2007) may explain why COVID-19 heavily affected the Hispanic/Latino community.

To understand the various reasons why COVID-19 impacted Hispanics/Latinos, we must first review how people locate information about COVID-19 using online search engine queries that are used to find information. Google generates millions of autocomplete suggestions based on an algorithm defined by factors such as relevance, location, words choice in the search menu, and a user's settings on their device (Google, 2022). COVID-19 searches have been an important concept for Google searches in the United States, particularly during the first wave of the pandemic (Effenberger et al., 2020). Furthermore, with Google being the most prominent source for seeking information, health information is often automated and applied to existing knowledge, thus influencing a person's health decision and potential health outcome (Bento et al., 2020; Chen et al., 2018; Powell et al., 2003).

1.1 | Inequities in information about COVID-19

Even though search engines provide a plethora of health information, the Hispanic/Latino community faces numerous structural inequalities that may impact their health and access to COVID-19 information (AMA, 2020). In the face of growing anti-immigration rhetoric, there has been a heightened fear of deportation among undocumented Hispanic/Latinos, which may exacerbate mistrust of the United States government (AMA, 2020). Language proficiency is another reason Hispanics/Latinos may lack access to appropriate COVID-19 information (Kusters et al., 2022). Some Hispanics/Latinos do not speak or fully grasp the English language, resulting in limited access to credible COVID-19 information (AMA, 2020; Calo et al., 2020; Joseph et al., 2020). Although there seems to be no racial and ethnic

differences when owning a smartphone, Hispanics/Latinos are less likely to own a computer and home broadband than other racial and ethnic groups (Atske & Perrin, 2021).

Inequities in COVID-19 information may extend to online search opportunities in other languages. A study conducted by Singh et al. (2007) compared various health websites to measure the frequency of search result overlap when translated into different languages. Google searches in non-English languages had less than 10% of search results than English searches, indicating that most search results are not well-translated into other languages (Singh et al., 2007). In recent years, autocompletes have become important mediators of the search experience and scholars have argued for a need to understand their societal impact (Noble, 2018; Olteanu et al., 2020). Specifically, there have been an emphasis to understand the interplay between language and autocompletes searches as it relates to health, but little empirical work has been done in this direction (Loh, 2016; Rovetta, 2021). Given the lack of research in understanding health-related algorithms in the context of English and Spanish speakers, particularly during the first wave of the pandemic, the present study explores how Google autocomplete results varied across English and Spanish speakers. The research questions of this study were the following:

1. How do Spanish and English speakers use Google search autocompletes to make decisions during the early months of the COVID-19 pandemic?
2. How do perceptions and circumstances of the COVID-19 pandemic inform the autocompletes of non-Hispanic Whites and Spanish speakers when using Google searches?

2 | MATERIALS AND METHODS

The present study was part of a larger mixed-methods research project focused on understanding algorithm bias search for COVID-19 information on Google. Phase I of the study cataloged Google search autocompletes for several terms related to COVID-19 between March 2020 and September 2020, and examined the difference between the results presented via the English and Spanish search interface. In Phase II, 30 participants were recruited to participate in a focus group meeting. These focus group meetings were conducted to explore how residents living under shelter-in-place orders made decisions about health, safety, and security.

To be eligible to participate in the focus group meetings, the inclusion criteria included the following: (i) at least 18 years of age; (ii) reside in states where shelter-in-place orders were implemented (iii) must be bilingual (Spanish and English) for focus group 3; (v) provide informed consent; and (vi) agree to the focus group being digitally recorded. For one of the focus group sessions, as an additional criterion, only Spanish-speaking participants were recruited; the other two sessions did not have this criterion.

The focus group meetings lasted about 60–90 min and these meetings were facilitated by bilingual/bicultural Latino research assistants trained in qualitative research methods. Participants were compensated for their participation with an Amazon electronic gift card worth $100 and approved by Rutgers University's Institutional Review Board.

The first focus group (FG1) was conducted in English with non-Spanish speakers. The second focus group (FG2) meeting was conducted with Spanish speakers only, and the third focus group meeting (FG3) was conducted with bilingual and bicultural Latino descent participants. Study participants also completed the Household Pulse Survey developed by the US Census Bureau, a 20 minute survey, to understand the social, health, and economic impact of COVID-19 (US Census Bureau, 2020).

Although suggestions for the number of focus groups vary, the number of focus groups that have been previously recommended depends on the study design and intended sample size (Fern, 1982; Patton, 2002). That said, the team conducted three focus groups with individuals experiencing COVID-19 shelter-in-place orders with two groups predominately Spanish speaking or Bilingual in Spanish and English bias. It was determined that after

three focus groups that saturation would be achieved concerning the barriers and concerns with obtaining reliable COVID-19 information (Castel et al., 2008).

2.1 | Data collection procedures

The study was publicized widely in Spanish and English using online channels and announcements were shared through social media such as Facebook, Twitter and Instagram. To ensure broader participation, we posted an advertisement in Spanish-language newspapers and television channels. We expected to draw diverse participants from states with the highest number of COVID-19 confirmed cases. We screened 85 participants and 30 individuals were randomly selected to complete the Household Pulse Survey in English or Spanish and were invited to participate in the focus group discussions afterward. Ten people were randomly assigned to participate in FG1—if they spoke English-only. Of those 10 people invited, two did not attend the focus group discussion, resulting in eight participants for FG1. Nine Spanish-speaking only participants were identified based on whether they completed the Household Pulse Survey in Spanish. All nine completed informed consent procedures and participated in FG2. Additionally, of the 85 people who completed the screener, 30 people indicated that they were bilingual. Of those 30, 13 were randomly screened and invited to participate in FG3, but one study participant did not attend the focus group meeting, resulting in 12 participants.

The focus group discussions were conducted virtually using Microsoft Teams, a video conferencing tool approved by the authors' academic institution. Researchers tested Microsoft Teams to ensure that the technology functioned well before the interview and eligible participants were provided an electronic invitation in Microsoft Teams.

2.2 | Focus group protocol

The focus group protocol covered the following domains: (1) experience with using search engines during COVID-19; (2) description of words searched during COVID-19; (3) coping and managing life during COVID-19; (4) understanding and knowledge of COVID-19; and (5) concerns about COVID-19. In addition, participants were shown autocompletes and were asked to react to them. An example of the focus group question included: What are your thoughts on the autocompletes and which autocomplete are you most likely to choose? The semi-structured interview format was flexible for participants to respond to questions freely but structured enough to discuss relevant topics.

2.3 | Positionality of the research team

The positionality of the research team included racial and ethnic minority graduate level researchers or graduate students in public health. The research team resided in a state with shelter-in-place orders and were directly impacted by the COVID-19 pandemic either by having family members work on the frontlines, and/or identified as Hispanic/Latino.

2.4 | Qualitative data analysis

The focus group meeting was transcribed verbatim in Spanish by a native Spanish speaker from Peru and reviewed by a research team member from Colombia. The team developed a translation protocol consisting of translating the

Spanish transcript to English separately by two translators. The first translator is a native Spanish speaker from Puerto Rico and English is her working language. The second translator is a native English speaker and Spanish is her working language. Once the translation was complete, both translators reviewed their transcripts together, addressed discrepancies, and provided reasons for alternative word choice for the last stage of agreement of the final translation. Lastly, another native Spanish and English speaker from Ecuador reviewed the final version of the translation.

After comparing the transcripts to the audio files, the team isolated the portion of each focus group transcript that discussed the algorithm biases and opinions on COVID-19 Google search autocompletes. In addition, research team members then reviewed the transcripts.

2.5 | Coding process

A team trained in qualitative methods (one undergraduate public health student, one MPH student, and a lead PhD researcher) developed the codebook. The codebook was created to achieve intercoder reliability and coding consistency across the coders. The coders developed first-level codes by using the focus group protocol and reviewing the transcript and marking initial categories for the codebook. Research staff used NVIVO software to code the transcripts and determine the intercoder reliability of the coding process between the coders.

Coded data were organized into themes and examined for relevance, coherence, and consistency, and then checked against the original qualitative data set to ensure accuracy. Categories were developed and refined using passages retrieved from the transcripts and assigned a tentative code. The coding manual included the parent codes, the second-level codes, inclusion, exclusion criteria of when to use each code, and an example quote from the transcript. Internal validity was corroborated through the development of the coding manual. This stage involved identifying relationships among the codes and developing connections or relationships within the previously identified codes. The codebook was used to code the focus group transcripts. The codebook allowed the authors to highlight passages in which participants described their behaviors, attitudes, experiences, knowledge, and how participants found information regarding COVID-19. This was done in an iterative process and discrepancies between coders were discussed and addressed by the lead qualitative researcher. Reliability between coders was checked regularly through an iterative process. Frequency counts and interquartile ranges were calculated for the number of times each code was used by the coders. The inter-coder reliability yielded 0.95, suggesting the codebook was both accurate and reliable. Table 1 provides a description of the codebook.

3 | RESULTS

Table 2 describes the demographic characteristics of the study participants. Among the 29 participants across all three focus groups, 21 self-reported having Hispanic or Latin ethnic origin. All participants who identified as Hispanic, Latino, or Spanish were in FG2 and FG3. The Hispanic, Latino, or Spanish origins represented were as follows: Mexican American/Chicano ($n = 1$), Puerto Rican ($n = 2$), Colombian ($n = 5$), Peruvian ($n = 6$), Ecuadorian ($n = 3$), Dominican ($n = 2$), Venezuelan ($n = 1$), Guatemalan (1), Honduran ($n = 1$), and Other ($n = 2$). Regarding race, some selected multiple categories, but majority self-identified as White. On average, participants were 29 years old (SD = 8.67), the majority self-identified as female ($n = 20$), and self-reported as heterosexual or straight ($n = 26$). Most participants completed either a college or graduate degree, and about half of the participants ($n = 15$) reported an annual income of <\$50,000. Most participants resided in New Jersey, but four participants were from New York and two from Michigan.

Focus group participants completed the US Census Bureau Household Pulse Survey (US Census Bureau, 2020), which described how individuals located in shelter in place orders managed the early days of the pandemic

TABLE 1 Autocompletes codebook.

First Code	Second Code	Inclusion	Exclusion	Example
Thoughts on Autocomplete	1. COVID-19 Search Results. 2. General Perceptions of the autocomplete	Any discussion about first impressions, opinions, or perceptions of COVID-19 search results	No mention of first impressions, opinions, or perceptions of COVID-19 search results	"Those sorts of questions are somewhat disappointing. I guess it's...yeah, I think that's my initial thought, is that these are disappointing autocompletes."
English versus Spanish Autocompletes	Limitations of Translating Information	Any discussion about the differences between English COVID-19 search results compared with Spanish COVID-19 search results	No mention of the differences between English and Spanish COVID-19 search results	"I remember when I was Googling for my family in Spanish, it was just a lot of translating. How do I say pandemic in Spanish? How do I say certain medications? How do I say like the symptoms? Like I don't know how to say loss of taste or smell in Spanish or foreign. I was just trying to look for like, I guess the shortest, most condensed way of saying it"
Consequence of COVID-19 Misinformation	1. Noncredible Sources. 2. Misled Understanding of COVID-19	Any mention of misconstrued perceptions of COVID-19	No mention of misconstrued perceptions of COVID-19	"Seeing all of these autocompletes from that search, you know, if you're someone that doesn't have a lot of information you can start freaking out from seeing is it deadly, is it a bioweapon, contagious, things like that."
Selection of Autocomplete	1. The Current Stage of the Pandemic. 2. Earlier Stages of the Pandemic	Any mention of selecting a specific autocomplete	No mention of selecting an autocomplete	"As new information was coming out, like I've clicked on it, is it airborne before? Pretty sure I looked up is it curable?" "At this point in the stage of this virus, I would probably be interested in if it's curable."
Emotions Towards COVID-19	Emerging Feelings from COVID-19	Any mention of feelings regarding COVID-19	No mention of feelings regarding COVID-19	"I think I'm anxious because sort of what a potential second wave could look like."

TABLE 1 (Continued)

First Code	Second Code	Inclusion	Exclusion	Example
Precautions Against COVID-19	CDC Guidelines and Recommendations	Any mention of wearing a mask, washing hands, and social distancing	No mention of guidelines and policies enacted due to COVID-19	"I feel that it's wearing the mask, hand sanitizing, washing your hands—first and foremost washing your hands. But I feel the mask will protect you and protect others if you have the virus."
Response to COVID-19	1. Attitudes and Perceptions. 2. Lifestyle Changes	Any discussion about the influence of COVID-19 search results on people's response to coronavirus	No mention of how COVID-19 search results can influence a person's response to coronavirus	"If I were someone that...didn't believe in hand sanitizer working and I did this, my eyes would jump straight to, you know, hand sanitizer makes you sicker or it's not effective, it's not good for you, then I would automatically use that to validate my initial arguments."
Sources of COVID-19	Friends and Family as Source of Information. 1. Health Experts. 2. News Sources	Any mention of a particular source to obtain information related to COVID-19	No mention of any sources of information related to COVID-19	"For me, I think it was Dr. Fauci who made a very relevant point at one stage in his interviews."
Infringement of Rights Due to COVID-19		Any discussion about the infringement of rights due to government mandated COVID-19 policies	No mention of the infringement of rights	"Telling someone you have to wear a mask, like automatically no, you can't make me."
Conspiracy Theory	1. Bioweapon. 2. Manmade	Any discussion about unproven politically motivated origins of COVID-19	No mention of politically motivated origins of COVID-19	"Suggesting that it might have been like a bioweapon or things like that. Like some of these might be out of fear, some of these might be out of curiosity."
Hand Washing versus Hand Sanitizer	1. Effectiveness. 2. Overuse. 3. Appropriate Times for Use	Any discussion about the effectiveness, overuse, or appropriate times to use hand	No mention of the effectiveness or overuse of hand sanitizer in	"I remember after I started working, I started volunteering in a hospital I was using hand sanitizer a lot. I started looking into I guess the

(Continues)

TABLE 1 (Continued)

First Code	Second Code	Inclusion	Exclusion	Example
		sanitizer in comparison to that of washing hands	comparison with washing hands	negative effects of overuse of hand sanitizer". "While every time the real soap is going to remove all germs, the sanitizer is not going to be the same…you don't want people to use the sanitizer like it's what you really have to use. What you really have to be doing all the time is washing your hands."
Vaccine	1. Speculation Regarding Arrival of Vaccines. 2. Anticipating a Vaccine. 3. Expectations	Any discussion about feelings, expectations, or opinions regarding a vaccine for COVID-19	No mention of feelings, expectations, or opinions regarding a vaccine for COVID-19	"Like, just don't relax. Just wait until the cure is here. Like the vaccine. And then, you know, and then just you wait. It's just a little longer. And hopefully we'll be able to go back to not normal, but like the new normal."

Abbreviation: COVID-19, coronavirus disease 2019.

TABLE 2 Sociodemographic characteristics of participants (n = 29).

Variable	Total sample (n = 29) N (%) or Mean (SD)
Age	28.83 (8.67)
Gender	
Female	20 (69.0)
Male	9 (31.0)
Race (check all that apply)[a]	
White/Caucasian	10 (34.5)
African American/Black	7 (24.1)
American Indian or Alaskan Native	5 (17.2)
Asian	6 (19.4)
Native Hawaiian or Pacific Islander	3 (9.7)
Hispanic, Latino, Spanish origin	
No, not of Hispanic, Latino, or Spanish origin	8 (27.6)
Yes	21 (72.4)
Sexual orientation	
Heterosexual/Straight	26 (89.7)
Gay	1 (3.4)
Bisexual	2 (6.9)
Education	
Some high school	1 (3.4)
High school/GED	2 (6.9)
Some college	6 (20.7)
Associate degree	1 (3.4)
Postsecondary and beyond	19 (65.5)
Marital status	
Married	7 (24.1)
Divorced/Separated	2 (6.9)
Never married	20 (69.0)
2019 Household Income	
Less than $25,000	10 (34.5)
$25,000–$34,999	3 (10.3)
$35,000–$49,999	2 (6.9)
$50,000–$74,999	7 (24.1)
$75, 000+	7 (24)

(Continues)

TABLE 2 (Continued)

Variable	Total sample (n = 29) N (%) or Mean (SD)
Employment	
Employed	22 (75.9)
Self-employed	7 (24.1)
State of residence	
New Jersey	23 (79.3)
New York	4 (13.8)
Michigan	2 (6.9)

[a]Percent may add up to more than 100% due to check all that apply.

(see Table 3). The survey results present insights into the context in which the participants were living their lives when they were undertaking COVID-related searches using Google.

Regarding feeling anxious, nervous, or on edge, 51.7% (n = 15) of participants stated they felt nervous and anxious for several days. Similarly, 48.3% (n = 14) of participants reported feeling down, depressed, or hopeless for several days. In terms of working in the last 7 days, 44.8% (n = 13) of participants did not work for pay or profit. Of those who did not work for pay or profit, 24.1% (n = 7) noted that they did not work due to COVID-19 pandemic reduction in businesses. In addition, 41.4% (n = 12) of participants delayed getting medical care due to COVID-19 and 48.3% (n = 14) indicated that they were afraid to go out or did not want to go out to buy food.

The themes that emerged from the study were as follows: (1) autocompletes evoked fear and stress; (2) skepticism and hesitation towards autocomplete search; (3) familiarity with COVID-19 information impacts outlook on autocomplete search; (4) autocompletes can promote preselection of searches; and (5) lesser choice of autocomplete results for Spanish-speaking searchers.

3.1 | Theme 1: Autocompletes evoke fear and stress

Table 4 describes the range and frequency of sentiments (i.e., terms used to describe their feelings) that participants in the study expressed in response to the autocomplete shown. The top three emotions used to evoke feelings towards COVID-19 autocompletes were panic or fear (n = 25), hopeful or optimistic (n = 17), and cautious or vigilant (n = 12). The least frequent emotions were assurance or protection (n = 6) and exhausted or drained (n = 4). Panic or fear was the most frequent and negative emotion expressed in all focus groups.

During the first wave of the COVID-19 pandemic, these focus groups were conducted when information regarding the virus, recommended protocols to follow, and discussions about the roll-out of a possible vaccine were continuously developing. After being shown the series of autocompletes, participants were asked to describe in three words how they felt about COVID-19. Although the specific responses varied, feelings reflecting a negative state were shared among all three focus groups. However, the burden of the pandemic was heavily experienced by participants from FG2 and FG3 only. For example, one Hispanic male from FG2 (Spanish speaking group, transcript was translated from Spanish to English) described his feelings as, "for me, it would be stress, the future, and panic." Another Hispanic female from FG3 stated, "I would say concerned because we don't know when it's going to end yet like the death toll is just increasing."

The frequency of negative emotions in response to autocompletes suggests that autocompletes searches evoked feelings of anxiety and stress among Hispanics/Latinos. As the same autocompletes were shown in all of the focus groups, the feelings of panic and fear were more prevalent in the Hispanic/Latino participants than English-speaking participants in FG1.

TABLE 3 Household pulse survey.

Variable	Total sample (n = 29) N (%)
Number of people in household	
1–3	17 (58.6)
4–6	11 (37.9)
7 or more	1 (3.4)
Any work done (last 7 days)	
Yes	16 (55.2)
No	13 (44.8)
Reasons for not working	
I did not want to be employed at this time.	3 (10.3)
I did not work because I am/was caring for children/I am/was caring for an elderly person.	3 (10.3)
I am/was sick (not coronavirus related) or disabled.	1 (3.4)
I did not have work due to coronavirus pandemic related reduction in business (including furlough).	7 (24.1)
My employment closed temporarily due to the coronavirus pandemic.	3 (10.3)
Other reason, please specify:	12 (48.3)
Student	3 (25.0)
Employed	6 (50.0)
N/A	2 (16.7)
Fear of infection	1 (8.3)
Reasons for not having enough to eat	
Unable afford to buy more food	6 (20.7)
Unable get out to buy food	5 (17.2)
Afraid to go out to buy food	14 (48.3)
Unable get groceries or meals delivered to me	4 (13.8)
The stores did not have the food I wanted	12 (41.4)
Amount spent on grocery stores (last 7 days)	
$0–$50	0 (0)
$50–$74	1 (3.4)
$75–$99	28 (96.6)
Amount spent on prepared meals (last 7 days)	
$0–$24	9 (31.0)
$25–$49	8 (27.6)
$50–$74	4 (13.8)
$75+	8 (27.6)

(Continues)

TABLE 3 (Continued)

Variable	Total sample (n = 29) N (%)
Confidence to afford food (next 4 weeks)	
Not confident	2 (6.9)
Confident	27 (93.1)
Overall health	
Excellent	7 (24.1)
Very good	13 (44.8)
Good	5 (17.2)
Fair	3 (10.3)
Poor	1 (3.4)
Bothered by feeling nervous, anxious, or on edge (over 7 days)	
Not at all	6 (20.7)
Several days	15 (51.7)
More than half the days	5 (17.2)
Nearly every day	3 (10.3)
Bothered by not being able to stop or control worrying (over 7 days)	
Not at all	13 (44.8)
Several days	7 (24.1)
More than half the days	7 (24.1)
Nearly every day	2 (6.9)
Bothered by having little interest or pleasure in activities? (over 7 Days)	
Not at all	13 (44.8)
Several days	7 (24.1)
More than half the days	7 (24.1)
Nearly every day	2 (6.9)
Bothered by feeling down, depressed, or hopeless (over 7 days)	
Not at all	11 (37.9)
Several days	14 (48.3)
More than half the days	2 (6.9)
Nearly every day	2 (6.9)
Delay in medical care due to COVID-19 pandemic (last 4 weeks)	
Yes	12 (41.4)
No	17 (58.6)
No medical care received due to COVID-19 pandemic (last 4 weeks)	
Yes	10 (34.5)
No	19 (65.5)

Abbreviation: COVID-19, coronavirus disease 2019.

–WILEY⌐

TABLE 4 Frequency of sentiments concerning autocompletes search (FG1–3).

Feelings toward Covid-19 autocompletes	Frequency
Panic or fear	25
Hopeful or optimistic	17
Cautious or vigilant	12
Anticipation or uncertainty	6
Assurance or protection	6
Exhausted or drained—emotional fatigue	4

3.2 | Theme 2: Skepticism and hesitation towards autocomplete search

Participants from all three focus groups expressed skepticism and hesitation towards the autocompletes shown to them (see Table 4). Participants noted how the autocompletes seemed like advertising. They described how the search queries generated in the past were handpicked and chosen based on past search history, location, or current emotional state. One non-Hispanic male from FG1 stated,

> These were all the common questions and ideas…It's sort of a sequence herein which, you know, if one of the first things I wanted to know is wasn't serious, was deadly. So…all of these autocompletes were what was running through my brain throughout the time, especially in the beginning.

When asked about their initial thoughts on the autocompletes, another participant, a Hispanic/Latino male from FG2 stated,

> My opinion I think all these forms appear because most people in a panicked state look[ed] for more deadly information, how dangerous, that it was truth or lie, whether it was really going to be a pandemic, whether it was actually provoked by an animal and whether it was really going to bring serious consequences.

When shown a comparison of an English versus Spanish autocomplete search on COVID-19, another participant, a Hispanic/Latino male from FG3 stated,

> For me, just right off the bat, "mentira" they're basically like questioning the authenticity and the reality of it being a real thing so I feel like that's surprising cause we didn't see that on the English version—questioning if it was real or not.

The sentiments reflected above suggest how algorithms are continuously adapting based on factors that are specific to each person.

3.3 | Theme 3: Autocompletes promotes preselection of searches

Many individuals described how autocompletes could be a catalyst for spreading misinformation. A non-Hispanic male from FG1 stated,

> So, if I were someone that didn't believe in masks or didn't believe in it working and I did this, my eyes would jump straight to, you know that masks are not effective, it's not good for you. Then I would automatically use that to validate my initial arguments. And that would just validate like I said what I believed originally.

Another participant, a Hispanic female from FG2, stated,

> I decide to be aware and informed because good news, or bad news, is that we have to be informed. Many times, we do not inform ourselves and we do not also know a little ignorance and informing ourselves and being aware without entering into that panic and as they say, take all the necessary precautions that have been told to us so far.

Another individual from FG3, a Hispanic/Latina female, expressed the improper use of Google search queries on behalf of the Spanish-speaking community stating,

> There [White people] might be a little more serious about the disease, like trying to follow the right procedures, the steps to stay safe. While Spanish people are still trying to figure out what it is. Is COVID real or was a made up?

Although Google serves to be a valuable tool for both English and Spanish-speaking communities, pre-existing beliefs, biases, and comprehension of COVID-19 may cause some individuals to select the information that affirms their thoughts, opinions, and views of the disease and associated health recommendations (Hart et al., 2009).

3.4 | Theme 4: Lesser choice of autocomplete results for Spanish-speaking searchers

Most individuals from FG2 and FG3 expressed their reliance on Google as a means for receiving information on COVID-19. Due to the mistrust in news sources (e.g., Telemundo and Univisión) and routine spread of information through social media outlets (e.g., WhatsApp and Facebook), Google search autocompletes served as a primary source of information for Spanish-speaking participants. However, participants noticed a difference in the query results in English compared with those in Spanish. For example, one Hispanic/Latina female from FG3 stated,

> Just like adding one word really changes what you're going to get the results. And if you don't put in, I guess, like the right combination of words. You might be looking in like the wrong information or getting your like a resource from like not reliable websites. And I know my mom is not really tech-savvy at all. So, like, essentially, we do all her searches for her.

Another participant, a Hispanic/Latino male from FG3 noted,

> I completely agree, I feel like Google's definitely geared more towards English speakers and English writers. So, like if someone were to use it for like the Spanish language, I feel like they are already at a disadvantage just based on that single word and how it changed the results. And exactly what you said, my mom isn't tech savvy so I feel she could have not included that word or spelled something wrong and just based off of that, the results you would get are completely different than what I would get if I searched it in English.

Another Hispanic/Latina female shared the same feelings,

> Yeah, I agree with that, too. I do think the English ones are more informative or for more knowledge in English, the Spanish ones are more like, is it real? Is it false or are you lying to me? Just like not very valid. Not very useful compared to English.

Overall, limited query results are a significant barrier to making informed decisions about health and safety for Spanish-speaking users, because it contributes to the scarcity of information available to the Hispanic community and may serve as a potential catalyst for the spread of misinformation regarding COVID-19.

4 | DISCUSSION

Using search engines for health information is becoming increasingly common (Acuna et al., 2020; Pogacar et al., 2017), with autocompletes acting as "information gatekeepers" that may inform an educated search or serve as a vehicle to spread misinformation (Lukenbill & Immroth, 2009). Throughout the existing literature of autocomplete studies, bias in search engines has been observed mainly in English (Gao & Shah, 2020; Pan et al., 2007). Yet, studies have not focused on comparing English and Spanish Google autocompletes particularly during the COVID-19 pandemic.

As seen in FG2 and FG3, these findings have shown that Google autocompletes are limited in Spanish (Goldman, 2006). Although the results contribute to the existing literature of autocomplete research (Gao & Shah, 2020; Pan et al., 2007), our study extends the science further to explore how autocomplete searches impact language and health. This lack of query results in Spanish indicates the bias in language models of autocompletes (Nguyen, 2020). These inequities in COVID-19 searches are problematic as they allow health disparities to persist and further exacerbate Hispanic/Latino communities' marginalization and vulnerability. Although all participants identified misleading autocomplete searches, English and bilingual/bicultural Spanish speakers responded with curiosity, whereas Spanish speakers responded with genuine concern or confusion.

The bilingual/bicultural speakers represented in FG2 were characterized by higher educational attainment and role as interpreters for their Spanish-speaking family members. However, many participants noted their inability to translate health-specific terms and concepts related to COVID-19, such as social distancing, transmission, and droplets. The lack of query results for Spanish speakers and the challenges of translating health information about COVID-19 from English to Spanish may promote preselection of autocompletes. These barriers may have led to misinformed health-seeking behaviors, uncertain responses, poor decision-making, and subsequent unintended health outcomes.

4.1 | Limitations

Several limitations must be addressed, given the results found in this study. For instance, online presence through video conferencing during the focus group meeting may have influenced how participants responded to the questions (Anderson, 2010). Another limitation is that the findings are not representative of Hispanic/Latinos who do not have access to computers or smartphones. According to Pew Research Center, only 67% of Hispanic adults have a traditional computer or access to high-speed internet at home as of 2021 (Atske & Perrin, 2021). Single focus group meetings were also a limitation as we may not have achieved saturation of Spanish speakers experiences during the pandemic. Additional focus groups with Spanish speakers could have added more insight into how our participants were using Google autocompletes over time and whether they accessed the same information.

Despite these limitations, there are several strengths of the study. First, the codebook achieved internal validity as the intercoder reliability score was 0.95. This, in turn, provides confidence that the study findings were not plagued with research bias or contained methodological errors. Our multicultural and bilingual team was an additional strength that contributed insight and improved communication with the focus group participants.

4.2 | Implications for policy and the field of community psychology

Algorithms allow artificial intelligence to detect human emotion by analyzing vocal patterns and facial expressions in technology such as automated voice systems and facial recognition (e.g., Face ID, Memoji) (Affectiva, 2022; VoiceSense, 2022). The use of sentimental analysis is growing in popularity and used mostly on product reviews (Mäntylä et al., 2018; Medhat et al., 2014). Future sentimental analysis research of Google autocomplete may provide more insight into the emotions of Spanish speakers during the COVID-19 pandemic.

It is also essential to investigate the impact of misinformation on health during a pandemic. Given that Google is used to monitor certain behaviors such as suicide and other health behaviors (Bento et al., 2020; Gunn III & Lester, 2013). Furthermore, there is a gap in the efficiency of COVID-19 responses between the United States and other countries. Misinformation has been widespread throughout the pandemic. Many countries, such as India and countries in Sub-Saharan Africa and Latin America, struggle with varying levels of misinformation (Krishnan, 2021; Osuagwu et al., 2021). In India, the spread of misinformation escalated to the point where internet shutdowns were ordered over 100 times in the past year to combat the continuous spread of misinformation (Funke & Flamini, 2022). Furthermore, the diffusion of COVID-19 misinformation or fake news across Latin American countries has greatly influenced the decision-making, practices, and behaviors of its people across the region and is intertwined by political motives and bias (Ceron et al., 2021). For example, Venezuelan President Nicolas Maduro shared on his Twitter feed a natural remedy for COVID-19 by mixing lemongrass, ginger, and water (Lyons, 2020).

The COVID-19 pandemic continues to exacerbate existing health inequities and autocompletes have the potential to amplify bias. Furthermore, tools (e.g., artificial intelligence) used to develop these models do not account for racial and ethnic minorities' health disparities (Röösli et al., 2021). Search result algorithms during the pandemic exacerbated racial inequality, including linguistic barriers, for racial and ethnic minorities seeking reliable health information, as seen with the Spanish speakers (Röösli et al., 2021).

Autocompletes suggest words to complete the queries or searches conducted by an individual. The impact of autocomplete search varies from being able to simplify a search, the ability to help with query formulation, aids those with limited language literacy, to altering perceptions of the search (as seen in this study), and the potential to perpetuate stereotypes of certain groups. These benefits and concerns directly affect the field of community psychology, particularly as it amplifies implicit bias and the potential illusions of perceived choice. Lastly, if unchecked, autocompletes may perpetuate stereotypes by normalizing them and providing the search results that validate bias. Given these risks, community psychologists and practitioners should investigate how autocomplete searchers may be spurring group thinking that could subtly shape the experiences of people online (Loh, 2016).

5 | CONCLUSION

Spanish speakers expressed more emotional burden and hesitation due to social factors and lack of information, affecting how COVID-19 information is searched on Google. Spanish speakers expressed limited query results and translations, leading to the preselection of certain autocompletes. Search engines should be held accountable for the differences in the quality of health information provided to different sections of society.

CONFLICT OF INTEREST STATEMENT
The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID
Pamela Valera 🆔 http://orcid.org/0000-0002-0095-9209

PEER REVIEW
The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/jcop.23013.

REFERENCES

Acuna, N., Vento, I., Alzate-Duque, L., & Valera, P. (2020). Harnessing digital videos to promote cancer prevention and education: A systematic review of the literature from 2013–2018. *Journal of Cancer Education*, *35*(4), 635–642. https://doi.org/10.1007/s13187-019-01624-0

Affectiva. (2022). *Interior sensing AI*. Retrieved August 4, 2022, from https://go.affectiva.com/auto

American Medical Association. (2020). *Latinx COVID-19 health inequities report: Insights for the health care field*. Retrieved from https://cloud.e.ama-assn.org/20-1643-COVIDLatinx

Anderson, C. (2010). Presenting and evaluating qualitative research. *American Journal of Pharmaceutical Education*, *74*(8), 141. https://doi.org/10.5688/aj7408141

Atske, S., & Perrin, A. (2021). *Home broadband adoption, computer ownership vary by race, ethnicity in the U.S. Pew Research Center*. Retrieved July 16, 2021, from https://www.pewresearch.org/fact-tank/2021/07/16/home-broadband-adoption-computer-ownership-vary-by-race-ethnicity-in-the-u-s/

Bento, A. I., Nguyen, T., Wing, C., Lozano-Rojas, F., Ahn, Y.-Y., & Simon, K. (2020). Evidence from Internet search data shows information-seeking responses to news of local COVID-19 cases. *Proceedings of the National Academy of Sciences*, *117*(21), 11220–11222. https://doi.org/10.1073/pnas.2005335117

Calo, W. A., Murray, A., Francis, E., Bermudez, M., & Kraschnewski, J. (2020). Reaching the hispanic community about COVID-19 through existing chronic disease prevention programs. *Preventing Chronic Disease*, *17*, 200165. https://doi.org/10.5888/pcd17.200165

Castel, L. D., Williams, K. A., Bosworth, H. B., Eisen, S. V., Hahn, E. A., Irwin, D. E., Kelly, M. A. R., Morse, J., Stover, A., DeWalt, D. A., & DeVellis, R. F. (2008). Content validity in the PROMIS social-health domain: A qualitative analysis of focus-group data. *Quality of Life Research*, *17*, 737–749. https://doi.org/10.1007/s11136-008-9352-3

Centers for Disease Control and Prevention. (2021). *Coronavirus disease 2019 (COVID-19)*. Retrieved, December 21, 2021, from https://www.cdc.gov/dotw/covid-19/index.html

Centers for Disease Control and Prevention. (2022a). COVID-NET: COVID-19-associated hospitalization surveillance network. Retrieved from https://gis.cdc.gov/grasp/COVIDNet/COVID19_3.html

Centers for Disease Control and Prevention. (2022b). COVID data tracker-cases by race/ethnicity. US Department of Health and Human Services. Retrieved from https://covid.cdc.gov/covid-data-tracker

Ceron, W., Gruszynski Sanseverino, G., de-Lima-Santos, M.-F., & Quiles, M. G. (2021). COVID-19 fake news diffusion across Latin America. *Social Network Analysis and Mining*, *11*(1), 47. https://doi.org/10.1007/s13278-021-00753-z

Chen, Y.-Y., Li, C.-M., Liang, J.-C., & Tsai, C.-C. (2018). Health information obtained from the Internet and changes in medical decision making: Questionnaire development and cross-sectional survey. *Journal of Medical Internet Research*, *20*(2), e47. https://doi.org/10.2196/jmir.9370

Effenberger, M., Kronbichler, A., Shin, J. I., Mayer, G., Tilg, H., & Perco, P. (2020). Association of the COVID-19 pandemic with Internet search volumes: A Google TrendsTM analysis. *International Journal of Infectious Diseases*, *95*, 192–197. https://doi.org/10.1016/j.ijid.2020.04.033

Fern, E. F. (1982). The use of focus groups for idea generation: The effects of group size, acquaintanceship, and moderator on response quantity and quality. *Journal of Marketing Research*, *19*, 1–13. https://doi.org/10.2307/3151525

Funke, D., & Flamini, D. (2022). *A guide to anti-misinformation actions around the world Poynter*. Retrieved August 4, 2022, from https://www.poynter.org/ifcn/anti-misinformation-actions/

Gao, R., & Shah, C. (2020). Toward creating a fairer ranking in search engine results. *Information Processing & Management*, *57*(1), 102138. https://doi.org/10.1016/j.ipm.2019.102138

Goldman, E. (2006). Search engine bias and the demise of search engine Utopianism. In A. Spink, & M. Zimmer (Eds.), *Web Search* (Vol. 14, pp. 121–133). Springer. https://doi.org/10.1007/978-3-540-75829-7_8

Google. (2022). *Google search - Discover how Google search works*. Retrieved August 4, 2022, from https://www.google.com/search/howsearchworks/

Gunn III, J. F., & Lester, D. (2013). Using Google searches on the Internet to monitor suicidal behavior. *Journal of Affective Disorders*, *148*(2–3), 411–412. https://doi.org/10.1016/j.jad.2012.11.004

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin*, *135*(4), 555–588. https://doi.org/10.1037/a0015701

Joseph, N. P., Reid, N. J., Som, A., Li, M. D., Hyle, E. P., Dugdale, C. M., Lang, M., Betancourt, J. R., Deng, F., Mendoza, D. P., Little, B. P., Narayan, A. K., & Flores, E. J. (2020). Racial and ethnic disparities in disease severity on admission chest radiographs among patients admitted with confirmed coronavirus disease 2019: A retrospective cohort study. *Radiology*, *297*(3), E303–E312. https://doi.org/10.1148/radiol.2020202602

Katella, K. (2021). *Our pandemic year—A COVID-19 timeline*. Yale Medicine. Retrieved March 9, 2021, from https://www.yalemedicine.org/news/covid-timeline

Krishnan, M. (2021). COVID: Misinformation complicates India's fight against pandemic. DW.COM. Retrieved May 3, 2021, from https://www.dw.com/en/india-covid-misinformation/a-57414876

Kusters, I. S., Gutierrez, A. M., Dean, J. M., Sommer, M., & Klyueva, A. (2022). Spanish-Language communication of COVID-19 information across US local health department websites. *J Racial Ethn Health Disparities*, *13*, 1–8.

Loh, L. C. (2016). Autocomplete: Dr Google's "helpful" assistant? *Canadian Family Physician Medecin de Famille Canadien*, *62*(8), 622–623.

Lukenbill, B., & Immroth, B. (2009). School and public youth librarians as health information gatekeepers: Research from the lower Rio Grande valley of Texas. *School Library Media Research*, *12*.

Lyons, K. (2020). *Twitter removes tweets by Brazil, Venezuela presidents for violating COVID-19 content rules*. The Verge. Retrieved March 30, 2022, from https://www.theverge.com/2020/3/30/21199845/twitter-tweets-brazil-venezuela-presidents-covid-19-coronavirus-jair-bolsonaro-maduro

Macias Gil, R., Marcelin, J. R., Zuniga-Blanco, B., Marquez, C., Mathew, T., & Piggott, D. A. (2020). COVID-19 pandemic: Disparate health impact on the Hispanic/Latinx population in the United States. *The Journal of Infectious Diseases*, *222*(10), 1592–1595. https://doi.org/10.1093/infdis/jiaa474

Makhortykh, M., Urman, A., & Ulloa, R. (2020). How search engines disseminate information about COVID-19 and why they should do better. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-017

Mäntylä, M. V., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Computer Science Review*, *27*, 16–32. https://doi.org/10.1016/j.cosrev.2017.10.002

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

Nguyen, G. (2020). Biased language models can result from internet training data. Search Engine Land. Retrieved December 29, 2020, from https://searchengineland.com/biased-language-models-can-result-from-internet-training-data-345077

Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press; JSTOR. https://doi.org/10.2307/j.ctt1pwt9w5

Olteanu, A., Diaz, F., & Kazai, G. (2020). When are search completion suggestions problematic. *Proceedings of the ACM on Human-Computer Interaction*, *4*, (CSCW2) 1–25. https://doi.org/10.1145/3415242

Osuagwu, U. L., Miner, C. A., Bhattarai, D., Mashige, K. P., Oloruntoba, R., Abu, E. K., Ekpenyong, B., Chikasirimobi, T. G., Goson, P. C., Ovenseri-Ogbomo, G. O., Langsi, R., Charwe, D. D., Ishaya, T., Nwaeze, O., & Agho, K. E. (2021). Misinformation about COVID-19 in Sub-Saharan Africa: Evidence from a cross-sectional survey. *Health Security*, *19*(1), 44–56. https://doi.org/10.1089/HS.2020.0202

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication*, *12*(3), 801–823. https://doi.org/10.1111/j.1083-6101.2007.00351.x

Pareek, M., Bangash, M. N., Pareek, N., Pan, D., Sze, S., Minhas, J. S., Hanif, W., & Khunti, K. (2020). Ethnicity and COVID-19: An urgent public health research priority. *The Lancet*, *395*(10234), 1421–1422. https://doi.org/10.1016/S0140-6736(20)30922-

Patton, M. Q. (2002). *Qualitative research and evaluation Methods*. Sage Publications.

Pogacar, F. A., Ghenai, A., Smucker, M. D., & Clarke, C. L. A. (2017). The positive and negative influence of search results on people's decisions about the efficacy of medical treatments. *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval*, 209–216. https://doi.org/10.1145/3121050.3121074

Powell, J. A., Darvell, M., & Gray, J. A. M. (2003). The doctor, the patient and the world-wide web: How the Internet is changing healthcare. *Journal of the Royal Society of Medicine*, *96*(2), 74–76. https://doi.org/10.1258/jrsm.96.2.74

Rodriguez-Diaz, C. E., Guilamo-Ramos, V., Mena, L., Hall, E., Honermann, B., Crowley, J. S., Baral, S., Prado, G. J., Marzan-Rodriguez, M., Beyrer, C., Sullivan, P. S., & Millett, G. A. (2020). Risk for COVID-19 infection and death among latinos in the United States: Examining heterogeneity in transmission dynamics. *Annals of Epidemiology*, *52*, 46–53. https://doi.org/10.1016/j.annepidem.2020.07.007

Röösli, E., Rice, B., & Hernandez-Boussard, T. (2021). Bias at warp speed: how AI May contribute to the disparities gap in the time of COVID-19. *Journal of the American Medical Informatics Association*, *28*(1), 190–192. https://doi.org/10.1093/jamia/ocaa210

Rovetta, A. (2021). Reliability of Google trends: Analysis of the limits and potential of web infoveillance during COVID-19 pandemic and for future research. *Front Res Metr Anal*, *6*, 670226. doi:10.3389/frma.2021.670226

Singh, P. M., Wight, C. A., Sercinoglu, O., Wilson, D. C., Boytsov, A., & Raizada, M. N. (2007). Language preferences on websites and in Google searches for human health and food information. *Journal of Medical Internet Research*, *9*(2), e18. https://doi.org/10.2196/jmir.9.2.e18

US Census Bureau. (2020). *Household pulse survey*. Retrieved from https://www.census.gov/data/experimental-data-products/household-pulse-survey.html

US Census Bureau. (2021). *020 Census Redistricting Data (Public Law 94-171) Summary File. Table P2 [Dataset]*. https://data.census.gov/cedsci/table?q=Race%20and%20Ethnicity&tid=DECENNIALPL2020.P2

VoiceSense. (2022). *Predictive behavioral voice analytics*. Retrieved August 4, 2022, from https://www.voicesense.com/

How to cite this article: Valera, P., Carmona, D., Singh, V., Malarkey, S., Baquerizo, H., & Smith, N. (2023). Understanding search autocompletes from the perspectives of English and Spanish speakers during the early months of the COVID-19 pandemic. *Journal of Community Psychology*, 1–19. https://doi.org/10.1002/jcop.23013