ChinaRxiv

AI translation · View original & related papers at
chinarxiv.org/items/chinaxiv-202205.00024

Exploring the influencing factors of COVID-19 vaccination willingness: A mixed-method study in China

Authors: Yue Su, Sijia Li, Feng Huang, Jia Xue, Tingshao Zhu, Jia Xue, Tingshao Zhu

Date: 2022-05-11T19:08:39+00:00

Abstract

The COVID-19 outbreak has caused tremendous damage to individuals' normal lives and societal development. Recently, with the emergence of novel coronavirus variants, people worldwide continue to face the threat of COVID-19. Vaccination is considered one of the most effective methods to combat this pandemic. However, many individuals exhibit hesitant attitudes toward COVID-19 vaccines. For vaccination promotion efforts, it is crucial to identify the factors that influence vaccination intention. With practical application in mind, this study identifies two main themes within this factor structure: non-manipulable factors and manipulable factors. This study employs semi-structured interviews to explore factors influencing vaccination willingness. For analyzing interview results, this study employs a mixed-methods approach combining qualitative thematic analysis and quantitative topic modeling. Thematic analysis identified nine key factors, while topic modeling identified seven key factors. After synthesizing and discussing these factors, this study confirmed ten key factors as the final set. More specifically, non-manipulable factors include vaccine effectiveness and safety, and vaccine feasibility, while manipulable factors include vaccine convenience, responsibility, external reference, necessity and urgency of vaccination, perceived threat, perceived benefits, scientific and objective communication, and policy requirements. This study investigates key factors influencing public vaccination intentions in China and provides specific guidelines for vaccination promotion to authorities and public health workers.

ChinaRxiv 中X

Full Text

Exploring the Influencing Factors of COVID-19 Vaccination Willingness: A Mixed-Method Study in China

Yue Su[12], Sijia Li[12], Feng Huang[12], Jia Xue[3],, *Tingshao Zhu*[12],

[1] CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China
[2] Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China
[3] Factor-Inwentash Faculty of Social Work & Faculty of Information, University of Toronto, Toronto M5S 1A1, Canada

Correspondence: tszhu@psych.ac.cn (T.Z.), jia.xue@utoronto.ca (J.X.)

Abstract

The COVID-19 pandemic has caused enormous disruption to individuals' lives and societal development. With the emergence of novel coronavirus variants, populations worldwide continue to face COVID-19 threats, and vaccine uptake remains one of the most effective methods for combating this epidemic. However, many people maintain hesitant attitudes toward COVID-19 vaccines. For vaccination promotion efforts, it is crucial to identify the factors that influence vaccination intention. With practical application as our goal, this study distinguishes two main themes in this factor structure: non-manipulable factors and manipulable factors. Using semi-structured interviews to explore factors related to vaccination willingness, we employed a mixed-method approach combining qualitative thematic analysis and quantitative topic modeling. Thematic analysis identified nine key factors, while topic modeling identified seven. After integrating and discussing these findings, we confirmed ten key factors overall. Specifically, non-manipulable factors include vaccine effectiveness and safety, and vaccine feasibility. Manipulable factors include convenience of vaccination, responsibility, external reference, necessity and urgency of vaccination uptake, perceived threat, perceived benefits, scientific and objective expression, and policy requirements. This study investigates key factors affecting public vaccination intentions in China and provides specific guidelines for vaccination promotion for authorities and public health workers.

Keywords: influencing factor, vaccination intention, vaccine uptake, vaccination willingness, thematic analysis, semi-structured interview, topic model

Introduction

Since the COVID-19 outbreak in December 2019, nearly every country worldwide has suffered tremendously over an extended period. Despite substantial

efforts from the public, professionals, and authorities to combat COVID-19 through multiple transmission prevention measures, achieving complete control of epidemic transmission has remained challenging. Presently, vaccination uptake is still considered one of the most effective methods for protecting individuals from viral infection and establishing herd immunity to deter epidemic spread.

However, many people maintain hesitant attitudes toward COVID-19 vaccines. Vaccine hesitancy represents a significant challenge in promoting vaccination uptake. The Strategic Advisory Group of Experts on Immunization (SAGE), a working group of the World Health Organization (WHO), defines vaccine hesitancy as "delay in acceptance or refusal of vaccines despite availability of vaccine services. Vaccine hesitancy is complex and context-specific, varying across time, place, and vaccines" (The SAGE Vaccine Hesitancy Working Group, 2014).

Recent research has focused on COVID-19-related vaccine hesitancy and the factors influencing vaccination intentions. Demographic variables such as age, gender, and occupation have been found to associate with vaccine hesitancy (Wake, 2021). More in-depth research has also emerged. Wang and colleagues (2021) used online questionnaires to investigate COVID-19 vaccination intentions and influencing factors among overseas and domestic university students, finding that perceived benefits, perceived barriers, and cues to action were important factors. Beyond questionnaires, interviews provide a valuable approach for deeply and specifically investigating people' s thoughts about COVID-19 vaccination uptake. Mugur and colleagues (2021) conducted interviews with women leaving jails and found that vaccine promotion requires strengthened health education and reduced mistrust, misinformation, and conspiracy theories (Geana, Anderson, & Ramaswamy, 2021). Additionally, a qualitative study based on semi-structured interviews explored perceptions of and hesitancy toward COVID-19 vaccination among older Chinese adults in Hong Kong, identifying barriers to vaccination and motivations for vaccination as two main themes (Siu, Cao, & Shum, 2022).

Researchers have struggled to reach consensus on influencing factors related to vaccination intention, likely due to differences in participants, time periods, and analytical methods. Regarding analytical methods, we employed a mixed-method approach to achieve deeper analysis of interview results, combining traditional qualitative methods with quantitative machine learning techniques. The first method is based on thematic analysis and manual coding, while the second utilizes topic modeling algorithms.

Traditionally, interview analysis relies heavily on manual coding, requiring researchers to carefully read and discuss transcripts to reach consensus. However, new text analysis methods have emerged in recent years, with topic modeling becoming particularly popular (Blei, Ng, & Jordan, 2003; Steyvers & Griffiths, 2006). Topic modeling is an unsupervised machine learning method that provides multiple word clusters as main topics, revealing the fundamental and un-

derlying dimensions of linguistic data (Atkins et al., 2012). For example, Xue and colleagues used topic modeling to analyze COVID-19-related tweets for public discourse investigation, identifying ten key themes (Xue et al., 2020). Beyond social media texts, interview transcripts are also suitable for topic modeling, as demonstrated in couple-therapy trials (Atkins et al., 2012). However, to our knowledge, topic modeling has not been applied to COVID-19-related interview studies. In our study, we used topic modeling as a supplementary quantitative method for text data analysis. By combining manual coding and topic modeling, we provide a novel and comprehensive perspective for interview transcript analysis.

Our study aims to identify factors influencing COVID-19 vaccine willingness in China and provide references for vaccination promotion for authorities and public health workers. From a practical perspective, we focus on factors with significant practical value for vaccination promotion work. Thus, we identified two main themes in the influencing factor structure: non-manipulable factors and manipulable factors. Non-manipulable factors refer to factors directly related to the COVID-19 vaccine itself, such as experimental results and objective data. These factors' core content is difficult to change and should remain consistent regardless of publication context, such as vaccination rates. Manipulable factors refer to factors indirectly related to the COVID-19 vaccine, typically focusing on external circumstances rather than vaccine information per se. These factors are generally easier to manipulate in practice to influence individuals' feelings and decisions.

Study 1

Method

We conducted semi-structured interviews to explore influencing factors related to COVID-19 vaccination, investigating what concerns people had about vaccines and what information would motivate them to get vaccinated more actively. The interview guide was developed by four psychological experts (two faculty members and two graduate students in psychology) who reached consensus through discussion and analysis. The guide contained two sections: one addressing potential vaccine concerns and the other focusing on persuasive information and influencing factors raised by participants.

Participant Recruitment. We recruited participants through purposive and snowball sampling to maximize diversity in age, gender, educational level, and geographical location. Specifically, recruitment advertisements were posted on social media, and interested individuals completed a questionnaire embedded in the advertisement, providing their phone number and available time slots. We then selected participants to interview, enriching the demographic diversity of our sample. Participants received cash compensation for their time and contribution, calculated at 2 Yuan per minute based on interview duration, plus a bonus of 0-30 Yuan depending on performance. Our study received prior

approval from the Ethics Committee of the Institute of Psychology, Chinese Academy of Sciences (approval number: H15009).

Data Collection. Six trained psychology graduate students conducted the interviews, with each interview performed by one interviewer via Tencent Meeting, a remote conferencing application providing reliable cloud video and audio services. Interviews lasted approximately 30-60 minutes. All participants provided informed consent, and interviews were audio-recorded for subsequent analysis. Data collection concluded and sample size was confirmed when interview data reached saturation, with no novel ideas emerging (Marbán-Castro et al., 2021).

Data Analysis. Audio recordings were transcribed verbatim by the interviewers. Researchers then performed thematic analysis to manually code the interview data. Following Braun and Clarke' s thematic analysis steps (Bell et al., 2020), we familiarized ourselves with the data, generated initial codes, searched for themes, reviewed themes, defined and named themes, and produced the final report (Braun & Clarke, 2006). To enhance analytical rigor, initial code development and subsequent theme generation and definition were discussed between two researchers (the first and second authors) until final agreement was reached (Bell et al., 2020). We also selected concrete examples for each theme.

Results

Demographic Information. Twelve participants were included in our study. Table 1 presents the demographic characteristics of our interview participants. The mean age was 25.83 years (SD = 2.98), with a range from 23 to 34 years. Among participants, six were male and six were female, representing a balanced gender ratio. Regarding education level, one participant held a college diploma while eleven had bachelor' s degrees or higher. Geographical distribution was diverse, with five participants from eastern China, four from central China, and three from western China.

Interview Results. After completing all interviews, we analyzed the transcripts and generated 31 initial codes (see Appendix for details). These codes were grouped into nine key factors, which were further organized into two main themes. The final interview results are presented in Table 2.

Theme 1: Non-Manipulable Factors. One key factor was classified as non-manipulable: effectiveness and safety of vaccines. This factor meant that vaccines should not pose threats to personal safety, emphasizing detailed, clear, scientific, and rigorous explanations of vaccination conditions, applicable populations, vaccination restrictions, and possible adverse reactions. For example, participants expressed considerable concern about side effects after vaccination, and some wanted to know more about why certain people were unsuitable for vaccination.

Theme 2: Manipulable Factors. Eight key factors were classified as manipulable: convenience of vaccination, responsibility, external reference, necessity

and urgency of vaccination uptake, perceived threat, perceived benefits, scientific and objective expression, and policy requirements.

Convenience of vaccines referred to costs individuals needed to consider and pay during vaccination, such as distance and time. Participants mentioned that abundant vaccination sites and convenient access would increase their willingness to get vaccinated.

Responsibility meant that people hold the awareness, emotion, and belief that individuals should be responsible for themselves and others, their families and groups, and their country and society. They also maintain the corresponding attitude that individuals should abide by norms, accept responsibility, and fulfill obligations. Some participants indicated that vaccination uptake signifies responsibility toward others and society.

External reference meant that when forming attitudes or behaviors, individuals use specific individuals or groups as reference points for comparison. These reference objects included specific social groups, societal majorities, people in one' s immediate circle, and people with different political and cultural backgrounds. Some participants mentioned that perceptions of vaccines among people around them were important, and if friends and family chose vaccination, they would also approve of it.

Necessity and urgency of vaccination uptake emphasized the intensity and urgency of vaccination need that individuals felt under current social and epidemic conditions. Several participants mentioned that news reports of growing domestic confirmed cases would increase their vaccination intention.

Perceived threat referred to people feeling their freedom was threatened when persuasive acts interfered with individual free choice. Some participants indicated that reports stating people "must" rather than "should" get vaccinated made acceptance difficult.

Perceived benefits meant benefits individuals perceived after vaccination. For example, participants mentioned they would actively get vaccinated if it facilitated easier campus access.

Scientific and objective expression emphasized that vaccine-related information should avoid arbitrary or ambiguous expressions, provide pros and cons from an objective perspective, and use accurate, easy-to-understand language. Some participants mentioned hearing that COVID-19 vaccine effectiveness rates in Brazil were not high, and they appreciated such honest reporting rather than mere praise for vaccines.

Policy requirements meant that individuals would abandon their own opinions and comply with policy demands and group regulations. Some participants indicated that if authorities mandated COVID-19 vaccination, they would obey and get vaccinated.

Study 2

Method

Based on transcripts from Study 1, we compiled all word sets spoken by participants. Specifically, we defined each participant' s complete set of spoken words as one document, yielding 12 documents total.

Before conducting topic modeling, we used Python to preprocess all documents to ensure analysis quality and interpretability. The preprocessing included three steps:

First, we removed all special symbols except English characters, Chinese characters, and numbers. Although all interviews were conducted in Chinese and documents were written in Chinese, we retained English characters because some English expressions (e.g., "mRNA") appeared in the interviews.

Second, we used Jieba for word segmentation. Jieba is a Python package commonly used for Chinese segmentation. In our analysis, Jieba separated sentences into words and phrases. Considering that new terminologies and expressions constantly emerge, we added specific terms to Jieba' s built-in user dictionary, such as "Bilibili," "Health Code," and "Li Jiaqi."

Third, we removed stop words and words shorter than two characters from the segmentation results. This elimination reduced irrelevant terms and noise, enhancing topic model algorithm performance (Calderón, de la Vega, & Herrero, 2020).

After preprocessing, we obtained a final dataset ready for topic modeling. We then used the Latent Dirichlet Allocation (LDA) algorithm to explore the main topics and structures. LDA is a common topic modeling method frequently used to investigate multiple topics in document sets (Ramage et al., 2009; Xue et al., 2020). As a probabilistic model, LDA can infer latent thematic structures from documents without prior knowledge or manual labels (Calderón, de la Vega, & Herrero, 2020). The Mallet version of LDA, implemented in the Python package Gensim, is considered to perform better than standard LDA (Wang & Li, 2020). Therefore, we used the Mallet version to build our topic model and investigate thematic patterns.

The number of topics represents an important parameter for model performance. To determine the appropriate number, we tested topic counts ranging from 2 to 30 and calculated coherence scores for each model. Coherence scores are advisable for measuring model performance, with higher scores indicating more precise models (Calderón, de la Vega, & Herrero, 2020). Based on coherence scores, we selected the most appropriate topic number and confirmed the final topic model.

Next, we employed pyLDAvis to visualize the topic model results. pyLDAvis is a Python package providing a visual interface for inter-topic distance (Chuang et al., 2012). In the inter-topic distance map, each bubble represents a topic,

with bubble area indicating topic prevalence (Hariman, Vries, & Smeets, 2018). Topic distances were multidimensionally scaled onto two axes (Sievert & Shirley, 2014). Relatively large, non-overlapping bubbles suggest a proper topic model result (Hariman, Vries, & Smeets, 2018).

The final step involved analyzing, identifying, and describing themes based on LDA results. After determining the number of topics, two authors (Y.S. and S.L.) discussed the top 20 keywords of each topic and reviewed corresponding expressions in interview transcripts. They then reached agreement on labeling and describing each theme.

Results

To determine the appropriate number of topics, we calculated coherence scores for each topic model. Figure 1 presents coherence scores for models built with different topic numbers. We selected seven as the optimal number based on two considerations. First, when the topic number was seven, coherence scores reached a high level, indicating good model performance. Second, our corpus was not very large, and too many topics would dilute the main focus of our text materials. Thus, compared to other high-coherence options like 16 or 18 topics, the number seven captured important topics while avoiding confusion about the main focus (Lyu, Le Han, & Luli, 2021).

The inter-topic distance map appears in Figure 2. Each bubble represents Topics 1 through 7, and the bubbles are relatively large and non-overlapping, suggesting a good topic model and confirming the reasonableness of selecting seven topics. Topics are ordered by descending prevalence in the corpus (Hariman, Vries, & Smeets, 2018).

Table 1 shows the topic model results with seven topics and the top 20 relevant words for each topic in the corpus. Topic numbers correspond to those shown in the inter-topic distance map.

After analyzing and discussing the seven topics generated by topic modeling, we identified these as our final topic categories, labeled each topic, and grouped them into two themes according to our proposed structure. These results appear in Table 2.

Theme 1: Non-Manipulable Factors. This theme contained two topics focusing on information directly related to the COVID-19 vaccine itself. Topic #3 was named "Side Effects of the COVID-19 Vaccine," concentrating on concerns about vaccine side effects—specifically, worries about how vaccination might affect one' s body. Topic #4 was labeled "Vaccine Feasibility," referring to explanations of vaccine applicability: what kinds of people could be vaccinated, who should wait, who should avoid vaccination, when vaccination would be useful, when full vaccination would be insufficient, and when other epidemic protection approaches (e.g., nucleic acid testing) would also be required.

Theme 2: Manipulable Factors. This theme contained five topics focusing

on external influences rather than vaccine information itself. Topic #1, "The Influences from Others' Experiences," meant that people refer to others' experiences and behaviors when deciding about vaccination, such as friends and colleagues, whose feedback held great value. Topic #2, "The Potential Costs during the Vaccination Process," included two aspects: social pressure from not getting vaccinated and the time and energy costs of getting vaccinated—essentially all costs people must consider when planning vaccination. Topic #5, "Direct and Transparent Information," emphasized that vaccine-related language and expression should be unambiguous, direct, sincere, and transparent. Topic #6, "Positive Perception of the Epidemic Situation," referred to vaccination necessity: when people felt positive about their living circumstances and perceived low infection risk, they were less willing to get vaccinated. Topic #7, "Safe and Reliable Information Related to the COVID-19 Vaccine," showed people' s attention to trustworthy vaccine safety information, specifically official media, medical professionals, and scientific reports that support and verify vaccine safety.

Discussion

This study explored influencing factors related to public vaccination willingness through semi-structured interviews. We recruited 12 participants and obtained verbatim transcripts. Study 1 used thematic analysis to manually code transcripts, identifying one key factor in the "non-manipulable factors" theme and eight key factors in the "manipulable factors" theme. Study 2 employed the LDA topic modeling method to examine semantic patterns in transcripts, confirming two key factors in "non-manipulable factors" and five key factors in "manipulable factors."

The mixed methods produced similar yet subtly different results. For non-manipulable factors, thematic analysis identified one key factor—effectiveness and safety of vaccines—while topic modeling identified two: side effects of the COVID-19 vaccine and vaccine feasibility. Topic modeling thus confirmed more factors than thematic analysis. The safety component within effectiveness and safety clearly focused on vaccine side effects, similar to the "Side Effects of the COVID-19 Vaccine" topic. However, vaccine feasibility emphasized situations and populations where vaccines should or should not be applied, revealing semantic connections between targeted groups and appropriate situations. This differed from the focus on vaccine effectiveness and safety and indeed supplemented a new perspective on non-manipulable factors.

For manipulable factors, thematic analysis identified eight key factors: convenience of vaccine, responsibility, external reference, necessity and urgency of vaccination uptake, perceived threat, perceived benefits, scientific and objective expression, and policy requirements. Topic modeling identified five: influences from others' experiences, potential costs during vaccination process, direct and transparent information, positive perception of epidemic situation, and safe and reliable information related to the COVID-19 vaccine.

Comparing these factors, external reference closely resembled influences from others' experiences, while scientific and objective expression was similar to direct and transparent information. Convenience of vaccine partially overlapped with potential costs during vaccination process, as the latter also mentioned social pressure related to vaccination, which thematic analysis included within external reference. Additionally, safe and reliable information related to the COVID-19 vaccine from topic modeling was also encompassed within external reference in thematic analysis. One interesting finding involved a factor present in both methods but expressed in different directions: necessity and urgency of vaccination uptake focused on situational pressure and risk making people feel urgent and more active about vaccination, while positive perception of epidemic situation focused on how positive feelings about the epidemic decreased public vaccination willingness. Nevertheless, both emphasized the necessity of vaccine uptake. Thematic analysis also identified responsibility, perceived threat, perceived benefits, and policy requirements—factors not captured by topic modeling. These findings suggest that for manipulable factors, thematic analysis provided more detailed and distinct results than topic modeling.

Overall, after considering both thematic analysis and topic modeling, we added vaccine feasibility to non-manipulable factors based on our thematic analysis results, ultimately identifying ten key factors related to COVID-19 vaccination intention.

Several factors in this study align with previous research. In a systematic review, Wake (2021) identified perceived risk of COVID-19, norms, perceived vaccine benefit, perceived vaccine barriers, perceived vaccine efficacy, and COVID-19 vaccine safety concerns as associated with vaccination willingness (Wake, 2021). However, to our knowledge, this is the first study to treat vaccine feasibility, responsibility, and policy requirements as distinct key factors and incorporate them into a complete factor structure, providing diverse perspectives for increasing public vaccination intention.

This study demonstrates the feasibility of using thematic analysis supplemented by topic modeling. In most cases, thematic analysis provided detailed insights, while topic modeling topics were relatively broad. However, topic modeling could identify novel relationships among participants' expressions that manual analysis sometimes overlooked.

Several models have been developed to explain vaccination willingness factors. The SAGE working group proposed the 3C model including confidence, convenience, and complacency (MacDonald, 2015). Thomson et al. (2016) suggested the 5A model including acceptance, access, awareness, affordability, and activation (Thomson, Robinson, & Vallée-Tourangeau, 2016). These models summarize influencing factors into mature theoretical systems. Unlike these models, our study emphasizes practical value and frames the factor structure into two themes: non-manipulable and manipulable factors. To our knowledge, this is the first study to propose an influencing factor structure for COVID-19 vaccines based on practical application, providing clear guidance on how these factors

should be used in real-world contexts.

Our findings deepen understanding of factors influencing public COVID-19 vaccine uptake intention in China and contribute to guidelines for targeted vaccination promotion programs. Based on the key factors identified, authorities and public health workers can focus on one or several factors according to specific social situations and epidemic conditions. For example, if a COVID-19 outbreak begins in a neighboring country, authorities could emphasize the severity and risk that country faces, helping people feel they remain threatened by COVID-19 and raising their vaccination awareness.

Conclusion

This study explored influencing factors related to COVID-19 vaccination intention in China using semi-structured interviews and mixed methods (thematic analysis and topic modeling). Results suggest thematic analysis generally produces more detailed analysis than topic modeling, though topic modeling can examine novel relationships among participants' expressions and offer new perspectives. We identified ten key factors: vaccine effectiveness and safety, and vaccine feasibility as non-manipulable factors; and convenience of vaccine, responsibility, external reference, necessity and urgency of vaccination uptake, perceived threat, perceived benefits, scientific and objective expression, and policy requirements as manipulable factors. We recommend that authorities and public health workers focus on these key factors to implement targeted vaccination promotion programs.

References

Atkins, D. C., Rubin, T. N., Steyvers, M., Doeden, M. A., Baucom, B. R., & Christensen, A. (2012). Topic models: a novel method for modeling couple and family text data. *Journal of Family Psychology*, 26(5), 816.

Bell, S., Clarke, R., Mounier-Jack, S., Walker, J. L., & Paterson, P. (2020). Parents' and guardians' views on the acceptability of a future COVID-19 vaccine: A multi-methods study in England. *Vaccine*, 38(49), 7789-7798.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan), 993-1022.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.

Calderón, C. A., de la Vega, G., & Herrero, D. B. (2020). Topic modeling and characterization of hate speech against immigrants on Twitter around the emergence of a far-right party in Spain. *Social Sciences*, 9(11), 188.

Chuang, J., Ramage, D., Manning, C., & Heer, J. (2012, May). Interpretation and trust: Designing model-driven visualizations for text analysis. In *Pro-*

ceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 443-452).

Geana, M. V., Anderson, S., & Ramaswamy, M. (2021). COVID-19 vaccine hesitancy among women leaving jails: A qualitative study. *Public Health Nursing*, 38(5), 892-896.

Hariman, N., Vries, M. D., & Smeets, I. (2018, November). Topic modeling for exploring cancer-related coverage in journalistic texts. In *Benelux Conference on Artificial Intelligence* (pp. 43-51). Springer, Cham.

Lyu, J. C., Le Han, E., & Luli, G. K. (2021). COVID-19 vaccine–related discussion on Twitter: Topic modeling and sentiment analysis. *Journal of Medical Internet Research*, 23(6), e24435.

MacDonald, N. (2015). Vaccine hesitancy: Definition, scope and determinants. *Vaccine*, 33(34), 4161-4164.

Marbán-Castro, E., Pons-Duran, C., García-Otero, L., Chen, H., Herrera, L. B., Gil, M. D. M., ···& Enguita-Fernàndez, C. (2021). Acceptability of clinical trials on COVID-19 during pregnancy among pregnant women and healthcare providers: A qualitative study. *International Journal of Environmental Research and Public Health*, 18(20), 10717.

Ramage, D., Hall, D., Nallapati, R., & Manning, C. D. (2009, August). Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing* (pp. 248-256).

Sievert, C., & Shirley, K. (2014, June). LDAvis: A method for visualizing and interpreting topics. In *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces* (pp. 63-70).

Siu, J. Y. M., Cao, Y., & Shum, D. H. (2022). Perceptions of and hesitancy toward COVID-19 vaccination in older Chinese adults in Hong Kong: A qualitative study. *BMC Geriatrics*, 22(1), 1-16.

Steyvers, M., & Griffiths, T. (2006). Probabilistic topic models. In T. Landauer, D. McNamara, S. Dennis, & W. Kintsch (Eds.), *Latent Semantic Analysis: A Road to Meaning* (pp. 427-448). Mahwah, NJ: Erlbaum.

The SAGE Vaccine Hesitancy Working Group. (2014). *Report of the SAGE Working Group on Vaccine Hesitancy*. http://www.who.int/immunization/sage/sage_{wg}{*vaccine*}{hesitancy

Thomson, A., Robinson, K., & Vallée-Tourangeau, G. (2016). The 5As: A practical taxonomy for the determinants of vaccine uptake. *Vaccine*, 34(8), 1018-1024.

Wake, A. D. (2021). The willingness to receive COVID-19 vaccine and its associated factors: "Vaccination refusal could prolong the war of this pandemic" — A systematic review. *Risk Management and Healthcare Policy*, 14, 2609.

Wang, J., & Li, C. (2020, July). Research on user satisfaction of video education application based on reviews. In *2020 IEEE 10th International Conference on Electronics Information and Emergency Communication (ICEIEC)* (pp. 340-343). IEEE.

Xue, J., Chen, J., Chen, C., Zheng, C., Li, S., & Zhu, T. (2020). Public discourse and sentiment during the COVID-19 pandemic: Using Latent Dirichlet Allocation for topic modeling on Twitter. *PLOS ONE*, 15(9), e0239441.

Appendix: 31 Codes Generated from Interview Transcripts

No.	Code Name	Operational Definition
1	Vaccine Safety	The vaccine does not pose a threat to personal safety, such as a low incidence rate of serious adverse reactions and good health conditions in the population after vaccination.
2	Vaccine Effectiveness	The effective result and time period—that is, the vaccine could protect individuals from coronavirus infection with high probability and remains effective for a specific period.
3	Convenience of Vaccination	Costs that need to be considered in the vaccination process, such as distance and time.
4	Restriction of Vaccination Uptake	The specific groups that are not suitable for vaccination.
5	General Vaccine Information	Basic information about vaccines, such as vaccine components.
6	Specific Explanation of Vaccination Requirement	Explanation of specific behaviors and requirements in vaccine promotion, such as reasons why certain people are unsuitable for vaccination.
7	Unofficial Position	The information publisher is not official media but rather media with neutral political positions and no political purpose.
8	Official Position	Confidence in public power.
9	Scarcity	The scarcer the product, the more people want it—people are more likely to desire products if told they are hard to obtain.
10	Data Exhibition	Providing specific numeric values in expression.
11	Popular Science	Providing vaccine-related scientific information for the public in accurate, rigorous, and easy-to-read language.

No.	Code Name	Operational Definition
12	Responsibility	People hold the awareness, emotion, and belief that individuals should be responsible for themselves and others, their families and groups, and their country and society. They also maintain the corresponding attitude that individuals should abide by norms, accept responsibility, and fulfill obligations.
13	Group Identification	The individual realizes they belong to a particular social group and recognizes the emotion and value that group brings, forming spiritual cohesion that regulates group members' values and moral norms.
14	Risk Perception	Individuals' subjective judgments about how safe their current environment is, how serious the epidemic is, and their infection risk.
15	Conformity	An individual tends to follow group norms to form beliefs, attitudes, and behaviors.
16	Policy Requirement	An individual abandons their own opinion and behaves according to policy demands and group regulations.
17	Specialization	Characteristics of the disseminator, including professional knowledge and rich experience.
18	Trustworthiness	Information receivers believe the communicator is able and willing to provide objective, fair, true, and effective information. The communicator should be sincere, honest, and objective, with no specific communication motives or intentions.
19	Attraction	We are more likely to follow requests from someone we admire or like.
20	Dual Role Persuasion	Both pros and cons information should be provided.
21	Perceived Threat	Acts with the aim of persuasion can be regarded as threats to freedom if they interfere with individual free choice.
22	Deadline Strategy	Increase compliance by indicating individuals have limited time to obtain certain benefits.
23	Vaccination Certification	People who take the COVID-19 vaccine receive certification in a certain form, such as a unique health code.

No.	Code Name	Operational Definition
24	Personal Involvement Degree	Individuals' subjective experience of the relationship between personal life and vaccination uptake.
25	Extra Benefits	Additional benefits beyond vaccine benefits per se.
26	Providing Comparison	Provide goods of different quality and effects for users to compare and choose.
27	Reference Group	An individual or group that serves as a reference when a person forms an attitude.
28	Vicarious Experience	Individuals can gain knowledge about vaccination by observing how others behave.
29	In-group Pressure	An influence that a group has on its members. When a group member' s thoughts or behaviors conflict with group norms, members feel mental pressure to abide by group norms to maintain their relationship with the group.
30	Certainty of Vaccination Information	Does not contain arbitrary or ambiguous expressions.
31	Psychological Distance	The emotional or psychological distance that an individual maintains with another object.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.