Spiral of silence on social media and the moderating role of disagreement and publicness in the network: Analyzing expressive and withdrawal behaviors

Hsuan-Ting Chen
The Chinese University of Hong Kong, Hong Kong

Abstract
Using two-wave panel data from Hong Kong, this study examines the spiral of silence process on social media. It extends the theoretical framework by including both supporting and disagreeing opinion expression and examining not only expressive but also withdrawal behaviors on social media. This study also investigates the moderating roles of disagreement and publicness as two affordances on social media that influence the spiral of silence process. Results from the moderated mediation model with a panel lagged and autoregressive analysis suggest that fear of social isolation (FSI) has an indirect effect on discouraging disagreeing opinion expression but not supporting opinion expression and on encouraging withdrawal behaviors through enhancing willingness to self-censor (WTSC) on social media. This indirect effect is contingent on the levels of disagreement and publicness in one’s network. Higher levels of disagreement and publicness promote the spiral of silence. Implications of the findings are discussed.

Keywords
Political disagreement, political expression, publicness, self-censorship, social media, spiral of silence

Corresponding author:
Hsuan-Ting Chen, School of Journalism and Communication, The Chinese University of Hong Kong, Room 206-207, Humanities Building, New Asia College, Shatin, Hong Kong.
Email: htchen@cuhk.edu.hk
The spiral of silence theory posits that individuals constantly monitor their opinion environment, such as mass media and interpersonal networks, because they fear being socially isolated (Noelle-Neumann, 1974). Through observing the opinion environment, people grasp a sense of whether the majority shares their own opinion. If their opinions are in line with the majority position, they tend to speak out, but if their opinions are incongruent with the perceived majority, they remain silent. As Noelle-Neumann (1974) argues, “the tendency of the one to speak up and the other to be silent starts off a spiraling process, which increasingly establishes one opinion as the prevailing one” (p. 44). In the spiraling process, perceived majority opinions are more likely to be shared and become more dominant, while perceived minority opinions become increasingly silent over time (Matthes, 2015). With the increasing popularity of social media as an online space for social interaction and political engagement (Kushin and Yamamoto, 2010), whether and how the spiral of silence occurs on social media becomes an important question. Thus, this study examines factors from two dimensions in influencing self-censorship and refraining from expression: (1) fear of social isolation (FSI) as an individual personality trait at an internal level and (2) political disagreement and publicness as two digital affordances on social media at an external level. Examining these factors will help to demonstrate how the spiral of silence occurs on social media and to what extent the spiral of silence affects people’s expressive and withdrawal behaviors.

This study extends the literature in several ways. First, it adopts a process-oriented approach to examine the indirect effect of FSI on refraining from expression through willingness to self-censor (WTSC) on social media. As Hayes et al. (2005b) stated, it is important to “conceptualize the spiral of silence as a process that is facilitated and or perpetuated by self-censors” (p. 317). Although researchers have strived to understand the antecedents and consequences of the spiral of silence on social media (Gearhart and Zhang, 2015b; Kwon et al., 2015), the underlying mechanism has not yet been fully explored. The spiral of silence on social media also warrants further investigation as most studies consist of convenience samples or cross-sectional data. Using two-wave panel data, this study can investigate how individual personality traits interact with digital affordances in affecting the spiral of silence process on social media (with WTSC conceptualized as the mediating variable) and assess the change in expressive and withdrawal behaviors.

Second, this study extends the theoretical framework by including the concept of publicness, drawing on the theory of impression and privacy management (Leary and Kowalski, 1990; Petronio, 2002). Social media have provided a novel sociality through offering opportunities for and controllability of interaction. More specifically, social media not only increase opportunities to be exposed to political disagreement and diverse perspectives (Kim et al., 2013) but also afford the controllability of information disclosure through changing privacy settings on social media (Leary and Kowalski, 1990; Petronio, 2002). That is, although social media users’ political expression can be broadcasted simultaneously to a large network of friends, users can adjust the level of publicness in their network through privacy settings to allow and limit those who see their information. This is a unique affordance of social media that has not yet received enough attention in the literature on the spiral of silence in the social media context.
Third, this study differentiates between expressive and withdrawal behaviors to understand how and to what extent people stay silent. In a face-to-face interaction, it is hardly possible for people to take back what they have expressed. Nevertheless, withdrawal behaviors such as deleting and editing posts can happen on social media, and it is important to acknowledge that not posting is not equal to deleting a post. People may remain quiet not only by passively lessening their political expression but also by actively deleting or editing what they have posted before to appeal to the opinion climate. In addition, while the literature has mostly paid attention to expressing supporting opinions as the outcome, this study examines expressive behaviors separately as expressing supporting opinion that promotes a favored candidate or party and expressing disagreeing opinion against an unfavored candidate or party. Different tones and affects are embedded in supporting and disagreeing opinions, so they should not be treated without differentiation. Clarifying the types of expressive behaviors can provide a clearer understanding of the spiral of silence process.

Setting this study in Hong Kong answers the calls raised by scholars asking for more research to examine the spiral of silence process in a cross-cultural context (Matthes et al., 2012). Hong Kong also provides a suitable context to examine the spiral of silence on social media due to its political system and technological environment. Hong Kong has a semi-democratic political system, and its political environment has grown more tense in the past few years under China’s sovereignty. There is a growing concern from scholars, the public, and reporters about the increasing self-censorship in media due to indirect pressure from the Chinese central government (Freedom of the Press, 2017). Moreover, many political issues, such as Universal Suffrage in the 2017 Chief Executive election, have polarized political opinion, which leads to a confrontational political environment that is likely to induce self-censorship and produce a spiral of silence. In addition to the political system, social media platforms have highly penetrated Hong Kong society and are significantly related to Hong Kong citizens’ political engagement (Chen et al., 2016). Accordingly, understanding whether and how the spiral of silence occurs on social media is a pertinent and timely question.

**FSI, WTSC, and refraining from expression**

The spiral of silence theory considers public opinion to be a process of social control. The theory proposes that individuals monitor what media convey about public opinion as well as what people discuss with others in everyday life, and they will refrain from speaking their opinions if their opinions are not congruent with those held by the perceived majority (Noelle-Neumann, 1974). In this spiral of silence process, FSI was found to play a significant role; however, it has been conceptualized in different ways. Some scholars consider FSI a transitory fear that is triggered by an opinion expression situation. Thus, the perception of an unfavorable opinion climate would trigger a transitory FSI (Hayes et al., 2013; Neuwirth et al., 2007). Other scholars have argued that the spiral of silence process is largely triggered by FSI as a trait-based characteristic that motivates continuous environmental surveillance for cues about the majority position. There are people who do not fear being isolated in general, while there are also people who always worry about it. Those who have an innate fear of becoming socially ostracized and
perceive themselves as having the minority views in an incongruent opinion climate will be less willing to speak out. Although in this line of research FSI plays a vital role in triggering the spiral of silence, it has been either assumed as a factum without empirical analysis (Matthes and Hayes, 2013) or inconsistently measured in previous studies (Neuwirth et al., 2007). Only recently have researchers captured FSI in a more consistent and reliable way. Matthes et al. (2012) took a trait-based individual difference perspective and developed a measure of FSI including psychometric properties which they validated using data from eight countries. In line with previous scholarship (e.g. Chan, 2017; Matthes et al., 2012), this study considers FSI as an antecedent that affects individuals’ inclination to withhold an opinion when they encounter a hostile opinion climate and set the spiral in motion.

Although researchers have mostly examined willingness to speak out as an outcome of the spiral of silence (Scheufele and Moy, 2000), they have also argued for attention to another variable: willingness to self-censor, or the likelihood of withholding one’s true opinion from an audience with a disagreeing opinion (Hayes et al., 2005a, 2005b). Hayes et al. (2005a, 2005b) suggested examining WTSC from a psychological perspective as an individual difference that affects people’s decisions to express their opinion because willingness to speak out is typically operationalized as the extent to which people are likely to speak an opinion in a hypothetical situation. Moreover, self-censorship implies active consideration of the opinion climate, which is different from simply inhibiting one’s opinion expression. According to the spiral of silence theory, FSI as an emotional fear of being excluded by social contacts triggers the spiral of silence process by prompting self-censorship. An individual’s likelihood to self-censor his or her opinion to a person with disagreeing viewpoints inhibits him or her from speaking out.

However, recent studies have shown mixed findings regarding refraining from opinion expression as a result of FSI and encountering disagreement on social media. For instance, Gearhart and Zhang (2015b) found that although people decided against posting political content that may risk offending others, they posted positive comments on political posts when encountering agreeable posts. Similarly, Kwon et al. (2015) found that although FSI prompts individuals’ WTSC, they expressed their opinions with actual posting behavior. To provide a clearer understanding of the spiral of silence process on social media, in the following sections, this study first proposes that there is a distinction between expressive and withdrawal behaviors and differentiates between supporting and disagreeing opinion expression. Second, this study emphasizes the significant role of social media affordance (i.e. disagreement and publicness) in moderating the spiral of silence process.

**Refraining from expression: distinction between expressive and withdrawal behaviors**

The distinction between expressive and withdrawal behaviors has been applied to social media studies to understand strategies in impression and privacy management (Dienlin and Metzger, 2016; Petronio, 2002). The collapsing of a wide variety of social contexts on social media (i.e. context collapse) has raised the question of whether people will be willing to share private information on social media. Some studies found that a more
diverse imagined audience leads to less willingness to share information online (e.g. Child and Starcher, 2016), while others suggested that increased context collapse prompts greater online disclosure and information sharing (Beam et al., 2017). According to the communication privacy management theory, people set up their privacy rules and boundaries depending on contextual influences and individual motivational factors (Child and Petronio, 2011; Petronio, 2002). When people are concerned about issues of privacy management, they are less likely to share information within a diverse network—more importantly, revealing and concealing information work as a dialectic (Petronio, 2002). People constantly negotiate between information dispersion and information retention on social media to manage their impression and privacy. It is, therefore, important to recognize that expressive and withdrawal behaviors are two different actions.

However, most studies involving the spiral of silence on social media have focused on expressive behavior only or simply asked the likelihood of refraining from commenting on a post in a cross-sectional survey (Gearhart and Zhang, 2015a, 2015b; Kwon et al., 2015). It is important to re-conceptualize refraining from expression in the social media context. Without tapping into active withdrawal behavior, it is difficult to grasp a comprehensive picture of refraining from expression on social media. This study conceptualizes refraining from expression in two ways: (1) decreasing or limiting “expressive behaviors” and (2) actively deleting and untagging posts as “withdrawal behaviors.” WTSC should constrain expressive behaviors and enhance withdrawal behaviors.

It is also equally important to differentiate the types of expression, namely, supporting and disagreeing opinion. Expressing a supporting opinion means that people express their support for the preferred candidate and party, while expressing a disagreeing opinion refers to expressing views that are against the oppositional candidate and party. However, most of the studies on spiral of silence in social media have not yet incorporated this differentiation. They either test only supporting opinion expression or measure expressive behavior in general (Chan, 2017; Gearhart and Zhang, 2015a, 2015b; Kwon et al., 2015). Studies have documented that disagreement is likely to produce conflict and prompt uncivil online debate when people perceive those who disagree with them as biased (Kennedy and Pronin, 2012). It can also trigger negative emotional arousal, such as anger and anxiety (Bushman, 2002; Chen et al., 2017). When people are afraid of being isolated in their social network and tend to self-censor their opinion, they should be reluctant to express a disagreeing opinion. As a result, the extent to which people refrain from expression regarding supporting and disagreeing opinion may be different. This study assumes that WTSC would be negatively related to expressive behaviors and examines the two types of expression separately.

The direct effect of FSI on WTSC and the direct effect of WTSC on expressive and withdrawal behaviors, as discussed above, also imply a basic mediation model: FSI will indirectly influence expressive and withdrawal behaviors through WTSC. Although not explicitly hypothesized, the relationships Kwon et al. (2015) proposed suggest a mediating effect of WTSC on the relationship between FSI and political posting activities. Chan (2017) also documented that WTSC mediates the relationship between FSI and willingness to express public support for a political candidate or party. Accordingly, the following hypotheses are proposed regarding how FSI affects WTSC, which, in turn, influences expressive and withdrawal behaviors separately:
H1. FSI is positively related to WTSC on social media.

H2. WTSC on social media is negatively related to expressive behaviors (i.e. supporting and disagreeing opinion expression) on social media.

H3. WTSC on social media is positively related to withdrawal behaviors on social media.

H4a. FSI indirectly influences expressive behaviors through WTSC on social media.

H4b. FSI indirectly influences withdrawal behaviors through WTSC on social media.

The moderating role of disagreement and publicness in the network

Social media have played a significant role in influencing the democratic process by providing an online space for social interaction and political engagement. Many of the associated behaviors take the form of political expression (Yamamoto et al., 2015). While social media have become mainstream channels for information and opinion sharing, some social media characteristics may affect the extent to which people self-censor their action and express their opinion. This study examines two affordances—political disagreement and publicness—that matter to the occurrence of the spiral of silence on social media. Rice et al. (2017) conceptualize media affordances as “relationships among action possibilities to which agents perceive they could apply a medium (or multiple media), within its potential features/capabilities/constraints, relative to the agent’s needs or purposes, within a given context” (p. 109). Affordances symbolize the multidimensional relationship between the object/technology and the user and how that relationship leads to potential outcomes (Evans et al., 2017). Meeting Evans et al.’s (2017) threshold criteria for confirming the proposed affordances, political disagreement and publicness are not the objects, features, or outcomes, but do vary in the dynamic relationship between social media and the users. They neither belong to the social media (i.e. the object) nor the individual (i.e. the user), but rather to the relationship between individuals and their uses and perceptions of social media. Political disagreement highlights the opportunities people have to be exposed to incongruent opinion climates on social media, while publicness represents the extent to which people control their profiles to allow themselves to be exposed to the opinion climate on social media. Accordingly, these two affordances should play significant roles in influencing the process of the spiral of silence.

Political disagreement. Exposure to counterattitudinal viewpoints is essential to reasoned public opinion and deliberative democracy (Habermas, 1989). Scholars have found that exposure to counterattitudinal viewpoints not only facilitates political tolerance, informed opinion, and a better understanding of a political issue, but it also leads to political ambivalence that makes people uncertain of their own positions and further discourages them from political engagement (Chen, in press; Mutz, 2002a, 2002b).
Building on the inadvertency thesis that suggests a structural force in the online environment that facilitates a less-than-perfect selective exposure strategy, non-avoidance of exposure to political difference, and a weakened social boundary which drives people to expose themselves to cross-cutting political views (Brundidge, 2010), social media have been found to provide great opportunities for exposure to political disagreement (Kim, 2011). They help people to expand their networks with greater heterogeneity (Lee et al., 2014). The coexistence of political and non-political characteristics on social media also prompts exposure to a variety of beliefs and perspectives (Wojcieszak and Mutz, 2009).

Exposure to political disagreement, however, may raise concern regarding to what extent people self-censor their behaviors and are willing to express their political opinion. A recent report from the Pew Research Center (2016) questioned the seemingly optimistic role of social media use in political engagement. Although people enjoy the opportunities for political information sharing and engagement that social media facilitate, many more are exhausted by the amount of political content they encounter. They feel stressed when discussing politics on social media with others who have disagreeing viewpoints. The Pew Research Center (2014) also found that Facebook and Twitter users are less likely to speak up when they encounter disagreement and do not feel their friends or followers agree with their point of view. Kwon et al. (2015) found that social media users are cautious when making political posts when their audiences hold different political ideology.

These studies demonstrate that the spiral of silence process is carried from the face-to-face environment to interaction on social media despite the fact that social media are meaningful venues for political expression. Although FSI serves as a motivational trigger for the spiral of silence process, scholars have argued that FSI should interact with opinion climate perceptions in explaining opinion expression (Hayes et al., 2013). The extent to which people perceive themselves to be exposed to political disagreement can represent the extent of hostile opinion climate. One may argue that disagreement can be an antecedent of FSI because the incongruent opinion climate can induce a fear of isolation. However, in this study, FSI is conceptualized as a personality trait that represents individual differences instead of a transitory fear that is triggered by an opinion expression situation. Accordingly, this study suggests a moderating effect of disagreement on the proposed mediating relationship (the indirect effect of FSI on expressive and withdrawal behaviors through WTSC). When people monitor their environment for cues about the opinion climate and refrain from opinion expression because of FSI, political disagreement is likely to exacerbate the influence of FSI in triggering the spiral of silence process.

Publicness. Distinct from typical online discussion forums as “privately-public” spaces, social media are characterized as a “publicly-private” realm (Papacharissi, 2009: 199), where personal information and private activities are publicized and the boundaries between public and private are blurry. Therefore, there is a lower perception of privacy on social media compared with other online forums. Many studies on the spiral of silence on social media have assumed that communication such as posting and responding
to content is seen publicly in the network. However, the development of privacy settings on social media has allowed users to adjust their profile visibility by, for example, limiting status updates and sharing information only to a specific group of people. Accordingly, social media users can decide whether a post is broadcasted to a large network of friends or narrowcasted to specific others. This is the level of publicness afforded by social media technology.

Publicness refers to “the probability that one’s behavior will be observed by others and the number of others who might see or learn about it” (Leary and Kowalski, 1990: 38). A post that is visible to everyone has a higher level of publicness than a post that is limited to a small group of people. This digital affordance has been studied mostly in the impression management literature (Leary and Kowalski, 1990). People refrain from actions that could give negative impressions in the minds of their audience, but when out of the gaze of their audience, they will be freer to behave as they wish (Leary, 1996). In this sense, when social media users allow their profile to be visible to everyone (i.e. high publicness), which could include a wide variety of people, they may exhibit higher self-censorship in limiting what they post. In a self-presentational context, studies have found that Facebook users self-censor the posts they make to avoid predicaments (Das and Kramer, 2013; Lang and Barton, 2015). Similarly, in a political context, Marder et al. (2016) found that people who are concerned about the impression they might leave on others through liking political pages will reduce their liking behaviors and choose to exercise “secret like” to increase gateway interaction (p. 282).

One may argue that a more public profile indicates that social media users are more open to their audience and are less concerned about their privacy. However, this is not necessarily the case on social media as the level of publicness relates to many different factors, such as skills in managing privacy setting (Park and Jang, 2014) and privacy self-efficacy (Dienlin and Metzger, 2016). Moreover, the level of publicness may not be able to define the level of disclosure because social media users have the need to be both private and social simultaneously (Petronio, 2002). As Baym and boyd (2012) argue, navigating collapsed contexts for social-mediated publicness requires a wide variety of strategies. Social media users need to manage multiple layers of audience and negotiate between visibility and obscurity of their mediated acts. They may change their privacy settings to maintain a low level of publicness of their profiles but interact intensively with certain sub-networks. On the contrary, those who have a high publicness of profiles may be subject to others’ attention and scrutiny and further refrain from taking action that may affect others’ impression of them. This is very likely to occur in a political expression context because politics often involves conflicting topics and strong emotional arousal.

Taken together, these two affordances should play a significant role in moderating the indirect effect of FSI on expressive and withdrawal behaviors through WTSC (see Figure 1). The following hypotheses are proposed:

\[ H5a. \] The indirect effect of FSI on expressive behaviors through WTSC is conditionally affected by political disagreement and publicness in the network. In other words,
political disagreement and publicness enhance (moderate) the direct effect of FSI on WTSC, which in turn limits expressive behaviors.

H5b. The indirect effect of FSI on withdrawal behaviors through WTSC is conditionally affected by political disagreement and publicness in the network. In other words, political disagreement and publicness enhance (moderate) the direct effect of FSI on WTSC, which in turn enhances withdrawal behaviors.

Method

The data were drawn from a two-wave panel study in Hong Kong conducted by Survey Sampling International (SSI), a Web survey panel company. Both waves of the survey were administered online. The first wave (W1) was conducted a week before the election for the 2016 Hong Kong Legislative Council in September. A total of 1141 participants completed the survey. The second wave (W2) of data collection took place in October 2016. In W2, 813 original respondents completed the survey questionnaire, yielding a retention rate of 71.3%. To proportionally represent the Hong Kong population, SSI employed a stratified quota sampling method based on gender, age, and income so that the sample would tend to match the Hong Kong Census as closely as possible. Given the continuous drop of landlines which makes it difficult to collect data using random digit dialing to reach a probability sample, stratified quota sampling is a feasible way to collect national samples and a popular method in communication research. The matched sample using census data to provide a more accurate representation of the population has been validated by previous research (Iyengar and Hahn, 2009). Some evidence also documents that the stratified quota sampling of an Internet panel produces quality responses that are not very different from a telephone survey (Ansolabehere and Schaffner, 2014).
**Measurement**

FSI. Following Matthes et al.’s (2012) measurement of FSI, respondents were asked to think about their social media use and rate the following statements from 1 = *strongly disagree* to 7 = *strongly agree*: (1) “It is scary to think about not being invited to social gatherings by people I know”; (2) “One of the worst things that could happen to me is to be excluded by people I know”; (3) “It would bother me if no one wanted to be around me”; (4) “I dislike feeling left out of social functions, parties, or other social gatherings”; and (5) “It is important to me to fit into the group I am with.” The five items were averaged to form an index ($W^1 M = 4.33$, standard deviation $[SD] = 1.16$, Cronbach’s $\alpha = .88$).

WTSC. This measure examines the extent to which respondents are willing to withhold their true opinion from others (Hayes et al., 2005a). Respondents were asked to rate their agreement with the following statements in a social media context from 1 = *strongly disagree* to 7 = *strongly agree*: (1) “It is difficult for me to express my opinion if I think others won’t agree with what I say”; (2) “There have been many times when I have thought others around me were wrong but I didn’t let them know”; (3) “When I disagree with others, I’d rather go along with them than argue about it”; (4) “I’d feel uncomfortable if someone asked my opinion and I knew that he or she wouldn’t agree with me”; and (5) “I tend to speak my opinion only around friends or other people I trust.” The items were averaged to form an index ($W^1 M = 4.39$, $SD = 0.96$, Cronbach’s $\alpha = .83$; $W^2 M = 4.20$, $SD = 0.92$, Cronbach’s $\alpha = .80$).

Expressive and withdrawal behaviors. For expressive behaviors, how respondents engaged in expressing supporting and disagreeing political views about the 2016 Hong Kong Legislative Council election on social media were examined. For expressing supporting opinion, respondents answered two items about the extent to which they express opinions to support their (1) preferred candidate and (2) preferred party on social media (from 1 = *never* to 4 = *often*; $W^1$ Spearman–Brown Coefficient = .91, $M = 1.96$, $SD = 0.87$; $W^2$ Spearman–Brown Coefficient = .89, $M = 1.99$, $SD = 0.83$). For expressing disagreeing political opinion, two items asking about expressing political views against the (1) oppositional candidate and (2) oppositional party on social media were measured ($W^1$ Spearman–Brown Coefficient = .91, $M = 1.94$, $SD = 0.87$; $W^2$ Spearman–Brown Coefficient = .87, $M = 1.97$, $SD = 0.81$). For withdrawal behaviors, respondents were asked to rate on a scale of 1 = *never* to 4 = *often* how often they (1) delete or edit their posts, and (2) ask people to untag them from a post (Dienlin and Metzger, 2016; Lang and Barton, 2015). The two items were also averaged to form an index ($W^1$ Spearman–Brown Coefficient = .67, $M = 2.25$, $SD = 0.72$; $W^2$ Spearman–Brown Coefficient = .66, $M = 2.31$, $SD = 0.71$).

Political disagreement. This measure taps into opinion incongruence, which emphasizes the extent to which members of the respondents’ social media networks have dissimilar political viewpoints. Respondents were asked, “What percentage of people you have talked to regarding politics or public affairs on social media have a dissimilar opinion to you?” Responses were 1 = 0–20%, 2 = 21–40%, 3 = 41–60%, 4 = 61–80%, and 5 = 81–100% ($W^1 M = 1.96$, $SD = 1.05$).
Publicness. The measure of publicness is adapted from the measure of privacy management (Child and Starcher, 2016; Petronio, 2002). Questions in this study emphasize publicness from a network perspective (i.e. the extent to which one’s network can see the content one shares; Liu and Kang, 2017) rather than the types of content people share on social media (e.g. personal feelings or intimate things about oneself). Respondents were asked to indicate the extent to which they make their profile public to the social media network on a 7-point scale (from 1 = strongly disagree to 7 = strongly agree) using the following questions: (1) “My network can see my up-to-date profile”; (2) “My network can see my detailed profile”; (3) “My network can see the content I share”; (4) “My network can find out what kind of person I am”; and (5) “I have changed the privacy setting to limit my profile visibility.” The last item was reverse-coded and then averaged with the other four items to create the index of publicness (W1 $M=3.28$, $SD=0.96$, Cronbach’s $\alpha=.86$).

Control variables. A host of variables from W1 were controlled in the analysis (see Appendix 1) because the literature has demonstrated their relationships with the spiral of silence process, including demographics (Noelle-Neumann, 1974), news media use (Tsfati et al., 2014; Willnat, 1996), political ideology, political interest, political efficacy (Lasorsa, 1991; Tsfati et al., 2014), and social media network size (Das and Kramer, 2013).

Statistical analysis

To test the proposed hypotheses, a two-wave panel design and the PROCESS macro for SPSS (Hayes, 2013) were employed. To assess the direct effect of FSI on WTSC, the direct effect of WTSC on expressive and withdrawal behaviors, and the indirect effect of FSI on expressive and withdrawal behaviors mediated by WTSC (H1 to H4), the Model 4 template from Hayes’ (2013) PROCESS macro with 10,000 bias-corrected bootstrap samples and 95% confidence intervals (CIs) was employed. Statistical significance ($p<.05$) is achieved when lower bound (LL) and upper bound (UL) CI do not include zero. Hayes’ (2013) PROCESS macro with the Model 9 template was then conducted to examine the proposed moderated mediation models (H5).

This study also incorporated a panel lagged and autoregressive analytic approach for the analyses. That is, not only were the W2 measures regressed on the W1 measures for the panel lagged analysis (Kenny, 2005), but the autoregressive term of the key dependent variable (measured in time one) was also introduced as a control and specified as an exogenous variable (Gil de Zúñiga et al., 2015). The results can “reflect the influence of predictor variables in wave one on the outcome variable in wave two above and beyond the causal influence of prior levels of the outcome variable on itself” (Shah et al., 2005: 549). The variables of WTSC, expressing supporting opinion, expressing disagreeing opinion, and withdrawal behaviors in W1 are controlled in the analyses in addition to the control variables. Employing the panel lagged and autoregressive analysis in the model can help to isolate the possible effects not only the control variables but also the dependent variables may have over time. This helps to ascertain the causal influence between the independent and dependent variables (Shah et al., 2001).
Results

Results support H1 that FSI is positively related to WTSC ($B = .067$, standard error $[SE] = .031$, $p < .05$). In addition, WTSC is negatively related to expressing disagreeing opinion ($B = -.088$, $SE = .031$, $p < .001$), supporting H2b. However, it is not related to expressing supporting opinion (H2a: $B = .047$, $SE = .032$, $p = .14$). In addition to suppressing expression of disagreeing opinion, WTSC is positively related to withdrawal behaviors, supporting H3 ($B = .146$, $SE = .028$, $p < .001$).

Results from the mediation analysis further demonstrate the indirect lagged and autoregressive effect of FSI on expressive and withdrawal behaviors through WTSC as proposed in H4. The findings show that WTSC significantly mediates the relationship between FSI and expressing disagreeing opinion (H4a: $B = -.006$, $SE = .004$, 95% CI = -.017 to -.0003), but not between FSI and expressing agreeing opinion ($B = .003$, $SE = .003$, 95% CI = -.001 to .013), which is no surprise given that there is no significant relationship between WTSC and expressing agreeing opinion. In other words, through heightening WTSC, FSI suppresses the expression of a disagreeing opinion against an unfavored candidate or party but does not limit the expression of a supporting opinion that promotes a favored candidate or party. Results from the mediation analysis also document an indirect effect that FSI increases WTSC on social media, which in turn boosts withdrawal behaviors such as deleting posts (H4b: $B = .010$, $SE = .006$, 95% CI = .004 to .023).

H5 proposed a moderating role of political disagreement and publicness in influencing the mediating relationship posited in H4. The first moderated mediation model demonstrates that the indirect effect of FSI on expressing disagreeing opinion through WTSC is conditionally affected by political disagreement ($B = .051$, $SE = .023$, $p < .05$) and publicness ($B = .044$, $SE = .019$, $p < .05$) in the network (H5a). The moderated mediation analysis for expressing supporting opinion as the outcome was not performed because the indirect effect of FSI on expressing supporting opinion was not established in the result above (H4a). Table 1 reports varying degrees of indirect effects on expressing disagreeing opinion (Model 1) and withdrawal behaviors (Model 2) depending on the level of political disagreement and publicness. As shown in Table 1 Model 1, the indirect effect of FSI on discouraging disagreeing opinion expression through WTSC is not significant if either disagreement or publicness is at a low level. The moderated mediation effect is significant only when disagreement and publicness are at a middle or high level, and the indirect effect was stronger at the high level than at the middle level of disagreement and publicness.

As shown in Table 1 Model 2, the indirect effect of FSI on promoting withdrawal behaviors through WTSC is not significant if either disagreement or publicness is at a low level. Similar to the findings for expressing disagreeing opinion, the moderated mediation effect is significant only when disagreement and publicness are at a middle or high level, and the indirect effect is stronger at the high level than at the middle level of disagreement and publicness. Taken together, as shown in Figure 2, the results suggest that FSI will indirectly make social media users express less disagreeing opinion and exercise more withdrawal behaviors when political disagreement and publicness in the network are at the middle and high level. The stronger the political disagreement and
publicness are, the greater the indirect effect on inhibiting disagreeing opinion expression and enhancing withdrawal behaviors becomes.

**Discussion**

This study advances the understanding of the spiral of silence process on social media by examining factors from two dimensions (i.e. individual personality trait and digital affordance) and examining not only the extent to which people express but also the extent to which they withdraw from social media interaction. Moreover, the types of political expression (i.e. supporting and disagreeing opinion expression) matter to the spiral of silence process. This study found an indirect effect of FSI on refraining from expression through WTSC, and this indirect effect is conditionally affected by the level of disagreement and publicness in one’s network.

Findings from this study first resonate with the spiral of silence literature which shows that individuals who are afraid of being isolated from their social network have a higher likelihood of withholding their expression when they disagree with others (Matthes et al., 2012). This relationship should closely relate to Hong Kong’s political climate. Since the handover of Hong Kong’s sovereignty to China in 1997, there has been rising tension between people from Hong Kong and mainland China. Two polarized political ideologies have developed, with one calling for more autonomy for Hong Kong
WTSC helps us to understand the likelihood of withholding true opinions, but whether and how people refrain from opinion expression remains in question. In addition to the effect of FSI on WTSC, this study documented that WTSC in turn prevents people from expression particularly in a way that limits the expression of a disagreeing opinion but not the expression of a supporting opinion. Disagreeing opinion often contains a more negative tone to attack the oppositional candidate or party, leading to tension in the communication and triggering negative emotions, while supporting opinion has a largely positive tone which may subject the person expressing it to less pressure to conform to norms. This is a plausible reason for why people choose to refrain from expressing disagreeing opinions but not supporting ones. However, it is important to note that this action would lessen opportunities for cross-cutting discussion, which is critical for the development of deliberative democracy.

Limiting expressive behaviors of disagreeing opinion is not the only way to refrain from expression on social media. Social media platforms allow people to actively delete or edit what they have shared and untag themselves from other people’s posts. What people share may contain a record of political leaning, and being tagged in certain political posts may
also make their networks aware of their political stances. Thus, if people are afraid of being isolated and likely to self-censor their opinion, exercising withdrawal behaviors is another strategy to avoid pressure from social norms. These withdrawal behaviors are representative of the unique activities people can do on social media and differentiate the activities afforded by social media interaction from face-to-face interaction.

In addition to understanding the spiral of silence process on social media at an individual level through the perspectives of personality trait (i.e. FSI) and psychological effect (i.e. WTSC), this study examines two unique social media affordances that are related to opinion climate—political disagreement and publicness. Findings from the moderated mediation model suggest that the spiral of silence process on social media is contingent on political disagreement and publicness. Political disagreement and publicness can boost the indirect effect of FSI on limiting disagreeing opinion expression and increasing withdrawal behaviors through WTSC. This finding implies that although social media provide a great opportunity to expose oneself to political disagreement, it does not mean that all people benefit from the opportunity. Exposure to political difference or dissimilar political views has long been valued as an essential component of a healthy democracy (Price and Cappella, 2002; Robinson, 2010). However, if people are afraid of being socially isolated, they will be sensitive to the opinion climate. Then, the more frequent disagreement people encounter, the stronger the incongruent opinion climate they perceive and the stronger willingness they have to self-censor their opinion. According to the indirect relationship, this will further discourage them from expressing disagreeing opinions and stimulate withdrawal behaviors. Studies on the spiral of silence have highlighted the effect of exposure to disagreement on refraining from speaking out in face-to-face communication. This study suggests that political disagreement on social media should be cautiously promoted.

This study also suggests that the level of publicness should be taken into consideration. Given that social media are characterized by reduced anonymity, increased peer-to-peer monitoring, and highly overlapping networks with offline social contacts (Kwon et al., 2015), studies on the spiral of silence on social media have not examined the level of publicness and have assumed that one’s post on social media will be disseminated to one’s whole network. This ignores the fact that social media platforms give users various ways to adjust privacy settings and allow users to decide with whom they would like to share information. This study found that publicness does play a significant moderating role in influencing the indirect relationship. A higher level of publicness in one’s network will enhance the effect of FSI on WTSC and in turn discourage expressive behaviors and encourage withdrawal behaviors. However, when people have a low level of publicness (i.e. their behaviors will be observed by only a small number of others and/or they can narrowcast their opinion to a certain group of people), FSI does not have an indirect effect on refraining from expression. This finding demonstrates a potentially optimistic role for the ability to control one’s information on social media in alleviating the spiral of silence process on social media. However, we should cautiously note that opinion expression may lead to an echo chamber effect, as the audience to whom people narrowcast is likely to have a likeminded opinion.

The findings cannot be interpreted without limitations. First, the two-wave panel survey with stratified quota sampling that matches with the Hong Kong Census
demographic measures suggests the generalizability of the findings as well as some
causal order (Ansolabehere and Schaffner, 2014; Iyengar and Hahn, 2009). However, the
two waves are about 1 month apart, raising some concern regarding the possibility of
clearly tracing the causes and effects. A two-wave panel that is conducted over a longer
time period can be considered in future research to examine the relationship. However,
collecting the data at two closer times helped to secure a high retention rate (71.3%),
which contributes to maintaining representative and valid data.

Another limitation is the social media measurement. This study did not ask ques-
tions about specific social media because many Hong Kong people do not use only a
single channel to obtain information and share their opinion. Various social media
platforms, such as Weibo and WeChat, are popular channels for interaction. Thus, in
the questionnaire, examples were provided when asking people to think about the
social media they use most often, suggesting that social media are like Facebook,
Weibo, and WeChat that have personal profiles and ways to interact with others and
express opinions. Future researchers may consider testing the relationships in specific
social media platforms, as each may have unique features in which the spiral of silence
process might differ (Ho and McLeod, 2008). Future researchers may consider testing
this moderated mediation model in different events and different populations as this
study focuses on a specific political election context in Hong Kong. To what extent this
model can be generalized to different situations, cultural backgrounds, and political
systems needs to be further validated.

Despite the limitations, this study provides an integrative framework with which to
study the spiral of silence process by proposing a moderated mediation model that
acknowledges individual differences and digital affordances in the process on social
media. In addition, with the support of the panel lagged and autoregressive analyses, the
findings offer evidence that extends the current literature on the spiral of silence on
social media that have so far largely relied on a cross-sectional survey or student samples
to a non-Western context and to a different political system.

**Funding**
The author disclosed receipt of the following financial support for the research, authorship, and/or
publication of this article: This work was supported by grants from the Research Grants Council of
the Hong Kong Special Administrative Region, China (CUHK 24613415), and the C-Centre of the
School of Journalism and Communication, CUHK (SS16576).

**Note**
1. One may argue a total effect of fear of social isolation (FSI) on expressive and withdrawal
behaviors as the “causal steps approach” for the mediation test suggested by Baron and Kenny
(1986). More recently, Hayes (2013) suggested that the total effect should not be used as a
gatekeeper for testing mediation and recommended a more powerful test of mediation such as
bootstrapping.

**ORCID iD**
Hsuan-Ting Chen [https://orcid.org/0000-0003-3140-5169](https://orcid.org/0000-0003-3140-5169)
References


Kim Y (2011) The contribution of social network sites to exposure to political difference: the relationships among SNSs, online political messaging, and exposure to cross-cutting perspectives. *Computers in Human Behavior* 27: 971–977.


**Author biography**

**Hsuan-Ting Chen** is an assistant professor at the School of Journalism and Communication, Chinese University of Hong Kong. Her research focuses on the use of digital media and their impacts on individuals and society. Her research has been published in journals such as *New Media & Society, Journal of Broadcasting & Electronic Media, the International Journal of Advertising, and the International Journal of Public Opinion Research*, among others.

**Appendix 1**

Four demographic control variables were included in the models: gender (male = 47%), age ($M=5.58$, standard deviation $[SD]=2.36$, $6=40–44$ years old), level of education ($M=5.99$, $SD=1.59$, $6=college$ degree or professional certificate), and household income per month ($M=7.26$, $SD=2.47$, $7=HK$30,000–HK$39,999, equivalent to US$3800–US$5100). For news media use, respondents were asked to rate on a 7-point scale (1 = never to 7 = everyday) how often they used the following media to get news and information about current events: television, radio, online newspaper, print newspaper, magazine, mobile phone, desktop, tablet, and social media ($M=4.86$, $SD=1.08$, Cronbach’s $\alpha=.76$). Political predisposition includes political ideology ($M=3.65$, $SD=1.33$), political interest ($M=3.83$, $SD=1.56$), and political efficacy ($M=3.37$, $SD=1.34$). Social media network size was measured by asking approximately how many friends the respondent has on social media. Since the range of answers was highly skewed ($M=139.76$, $Mdn=50.00$, $SD=241.58$, skewness = 6.36), as could be expected, it was transformed using the natural logarithm ($W^1 M=1.76$, $Mdn=1.71$, $SD=0.62$, skewness = −0.27).