# Real Name Verification Law on the Internet: A Poison or Cure for Privacy?

Daegon Cho (daegonc@andrew.cmu.edu)

School of Information Systems and Management
The H. John Heinz III College
Carnegie Mellon University, Pittsburgh, PA 15213

## Abstract

As the internet media has become more widely used over time, public opinions formed by internet discussions affect political and social issues more critically. While the internet space guarantees equal status for every participant and equal opportunity in terms of freedom of speech based on anonymity, baseless rumors, personal defamation, and privacy invasion against particular individuals and groups are more likely to happen rather than in the face-to-face offline communications. In order to prevent this undesirable effect, the South Korean government implemented Real Name Verification Law in July 2007 by which discussion participants should pass verification process in order to express their opinions in most websites.

This study examines the effects of Real Name Verification Law in several aspects. By applying content analysis to abundant data of postings in a leading discussion forum that is subject to the law, the results suggest that Real Name Verification Law has a dampening effect on overall participation in the short-term, but the law did not affect the participation in the long term. Also, identification of postings had significant effects on reducing uninhibited behaviors (swear words and antinormative expressions), suggesting that Real Name Verification Law encouraged users' behavioral changes in the positive direction to some extent. The impact is greater for Heavy User group than for Light and Middle User groups. Also, discussion participants with their real names showed more discreet behaviors regardless of the enforcement of the law.

By analyzing the effect of this policy at the forefront of internet trends of South Korea, this paper can shed light on some useful implications and information to policy makers of other countries that may consider certain type of internet regulations in terms of privacy and anonymity.

## Keywords:

anonymity, pseudonymity, deindividuation, computer-mediated communication, Real Name Verification Law, freedom of speech, content analysis

"The incident, 'Dog Shit Girl[1]' (in South Korea), involves a norm that most people would seemingly agree to clean up after your dog. But having a permanent record of one's norm violations is upping the sanction to a whole new level... allowing (internet) bloggers to act as a cyber-posse, tracking down norm violators and branding them with digital scarlet letters."

(Daniel Solove, Washington Post, 7/7/2005)

"Internet space in our country has become the wall of a public toilet." (A member of South Korean National Assembly) (…) Widespread verbal abuse on the Web has become a growing social problem, and efforts to stamp it out have become a government priority. (The New York Times, 10/13/2008)

# 1. Introduction

Long ago, Plato raised a concern in his Phaedrus: he expressed the fear that the *emerging technology of writing* would destroy the rich oral literacy that was central to his culture (Wood & Smith, 2001). Plato preferred *speech* to *writing* due to his suspicion of writing as a threatening new technology. He worried that people's thoughts would become unwavering by written documents, and this would ruin the root of democracy by preventing a variety of free discussions. Plato's concern is analogous to the present worries with regard to a newly emerged technology, *the internet or computer-mediated communication (CMC)* (Hwang, 2007).

With the growth in sociality and interaction around online discussion forums, these mediums are increasingly becoming places for communities to discuss and address common issues (Diakopoulos & Naaman 2011). Through this intermediary based on cutting-edge technology, our opinions are delivered and shared rapidly, and the history of communications via the internet is cumulated exponentially. That is, internet-based information technologies have enabled the emergence of new types of communicative practices (Karyda & Kokolakis, 2008). The other unique characteristic of the internet communication is *anonymity or pseudonymity*. People want to surf the web, purchase online, and post messages or comments without exposing their identities, interests, and activities to others (Acquisti et al., 2003). On the one hand, participants in online discussion forums are all equal under the anonymous condition, no matter what their backgrounds and circumstances are. Minorities and neglected classes of people in the real world can have equal opportunities to express their thoughts: freedom of speech is guaranteed more highly thanks to online anonymity. The immensely popular and large-scale interactions are associated with distilling expertise and opinions from thousands of participants who may or may not know each other.

However, on the other hand, anonymity mitigates the social norm and establishes conditions to neglect principles of mutual respect. Malicious profanities and groundless online rumors can proliferate quickly under this circumstance, and those online slanders can lead to serious privacy invasion and personal defamation. Not surprisingly, there have been a variety of conflicts surrounding freedom of speech and anonymity on the cyberspace, particularly in South Korea, a country that is the most highly networked in the world. According to a recent research[2], South Korea's household broadband penetration reached 95%, which was the highest rate among those of

[1] 'Dog Shit Girl' case has changed public sentiment in favor of Real Name Verification Law, as discussed more specifically in section 2.

[2] http://www.strategyanalytics.com/default.aspx?mod=PressReleaseViewer&a0=4748

all 57 surveyed countries.[3] This fact indicates that cyberspace has become an important part of daily life for South Koreans than people in any other countries. In spite of several positive impacts of this advanced network, radical changes have created several spillover effects in South Korea's society. For instance, online harassment in anonymous cyberspace has constantly occurred, such as the 'Dog Shit Girl' case[4].

This study examines the impact of policy that was firstly taken into effect in 2007 by which internet users should verify their real identification when they write on widely-used internet websites. The paper proceeds as follows: section 2 discusses the description of legislation and related literature. Section 3 lays out the research model and hypotheses and Section 4 describe the data and methods of analysis. Main results in terms of policy impacts are specified in Section 5. Finally, Section 6 provides a discussion and conclusion of the results.

# 2. Backgrounds and Relevant Literature

### 2.1. History and Description of Real Name Verification Law

The initial debate was ignited by political domain. Presidential election in 2002 was a turning point from a traditional campaign to an internet-oriented one. In the election, it was believed to some extent that young voters' assertive online activities and campaigns in favor of a particular candidate upset the election result. Some conservative members of National Assembly enacted Real Name Verification Law for the first time in 2005. Under this rule, only verified users with their real identity can leave postings and comments on the website. The suggested bill contained that the rule would be applied to election-related online discussion boards before 90 days of the Election Day. They believed that this policy would protect the privacy of candidates and promote more reliable online discussions. Due to the constant social issues caused by online slanders, the extended version of Real Name Verification Law was built in July 2007. By this law, all websites that have average daily viewership of over 300,000 went into effect. Once again, this law was strengthened to websites that have average daily viewership of over 100,000 in 2009. Accordingly, 37 websites had been subject to the law in 2007, but 153 websites[5] were concerned with the law in 2009. In order to verify a user's identity, Resident Registration Number (RRN)[6] is used at the associated websites.

The original policy goal is to prevent widespread online abuse in postings and comments that can seriously cause privacy invasion and personal defamation by legal enforcement and penalties.

---

[3] 2nd: Singapore (88%), 3rd: Netherlands (85%), USA (60%) (at the end of 2008)

[4] When a girl riding a South Korean subway refused to clean up her dog's excrement, a witness took pictures of her and posted them on a popular website, along with an account of her misbehavior. Within hours, she and her dog were recognized everywhere they went. "Within hours, she was labeled 'dog-shit-girl' and her pictures and parodies were everywhere on the cyberspace. Within days, her identity and her past were revealed. Request for information about her parents and relatives started popping up and people started to recognize her by the dog and the bag she was carrying as well as her watch, clearly visible in the original picture. All mentions of privacy invasion were shouted down with accusations of being related to the girl.
(Source: http://boingboing.net/2005/06/29/woman_doesnt_clean_u.html)

[5] It is not an exaggerated statement that almost all frequently-visited websites by South Korean people came to be laid under this restriction

[6] Resident Registration Number is equivalent to Social Security Number in the US. This is a 13-digit number issued to all residents of South Korean citizens. It is used to identify people in various private transactions such as in banking and employment. It is also used extensively for online identification purposes.

However, this strict law can discourage users' willingness to express, which is not desirable. In this context, research questions are as follows: was Real Name Verification Law effective? Did people's online behaviors change due to the law? Which factor is more influential to behaviors of users between an anonymous condition with the law and condition with revealed real identity? Relevant literature is reviewed in the following section.

### 2.2. Anonymity and Pseudonymity

As noted above, Real Name Verification Law can be regarded as a changed rule-setting in terms of anonymity. First of all, the relevant terminology should be clearly defined. According to Pfitzmann & Kohntopp (2001), *anonymity* is the state of being not identifiable within a set of subject. *Unlinkability* and *unobservability* are accompanied with anonymity. The distinction between anonymity and *pseudonymity* is another important matter. Pseudonymity indicates that the user maintains one or more persistent pseudonyms that are not connected to the user's physical identity (Goldberg, 2000). Pseudonymity comprises all degrees of linkability to a subject. For example, third parties (website operators) may have the possibility to reveal the identity of the holder (users) in order to provide means for investigation or prosecution. In online communications, pseudonyms contain a variety degree of anonymity (Froomkin, 1995). High linkable *public pseudonyms* indicate that the link between pseudonym (or nickname) and a real identity is publicly known or easy to discover. *Unlinkable pseudonyms* mean that system operators or third parties cannot detect a certain pseudonym's real identity (Pfitzmann & Kohntopp, 2001). Real Name Verification Law led to the cyberspace to switch from an unlinkable pseudonymous condition to a publicly pseudonymous condition.

### 2.3. Deindividuation

The paper is directly associated with a large body of literature of the social value of anonymous communication and deindividuation research. Social psychology has a long history of studying the effect of anonymity and its disinhibiting effect on individual's behaviors. Two main streams are existent with regard to anonymous communication: positive and negative aspects. On the positive side, anonymity communications enable minorities to express their own opinions and to maintain privacy protection (Froomkin, 1996). An anonymous environment is helpful in promoting more active involvements without revealing personal identity (Siegel et al., 1986). Also, anonymous speech helps to settle the imbalance of information through revelation of autonomous personal identity (Zarsky, 2004). According to the SIDE (Social Identity model of Deindividuation Effects) model that was built by Reicher, Spears & Postmes (1995), anonymity should accentuate the effects of the salient social identity and the dominant normative response associated with it. From Information Systems literature, anonymity led to a reduction in behavioral constraints and enabled individuals to engage in behavior they would not engage in when identified (Jessup et al., 1990).

On the other hand, opponent groups argue that it is more likely that defamation, threat, insulting words, and slander can occur under an anonymous communication environment (Cohen, 1996). According to classical deindividuation theory, anonymity in the group can lead to reduced self-awareness and influence of social norms that ultimately correspond to antinormative behaviors (Postmes & Spears, 1998).

### 2.4. CMC and Deindividuation

Finally, according to recent literature on impacts and influence of computer-mediated communication (CMC), reduction of identifiable information reduces self-awareness and stimulates

antinormative behavior (McLeod et al., 1997). However, aforementioned SIDE model proposes that certain feature of CMC can intensify social influence and enhance normative behavior (Postmes et al., 1998; Spears et al., 1990).

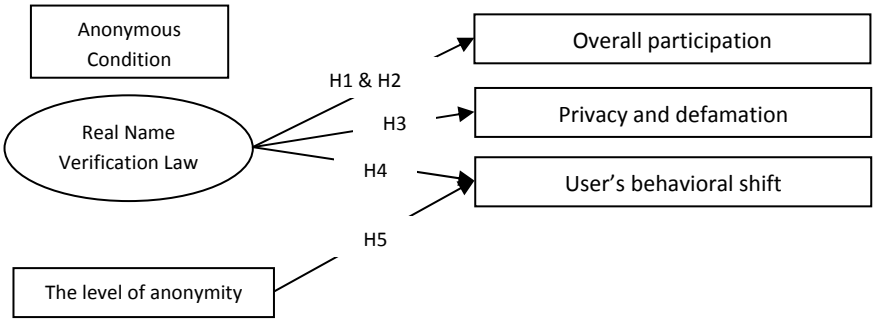
The main limitation of the previous studies is that they are based on designed experiments. None of these studies have investigated actual behavioral changes of online communications by using real-world data. Real Name Verification Law provides a proper condition to examine the actual behavioral change of users. Moreover, previous studies regarding anonymity have been conducted from legal and economic aspects. There is little research of empirical aspects of anonymity which are associated with behaviors of people and impacts of a particular policy.

# 3. Research Model and Hypotheses

The goal of Real Name Verification Law is to reduce undesirable and antinormative postings by changing the level of anonymity in which linkability and traceability are enhanced. Previous studies found that the anonymous condition is negatively associated with privacy invasion, such as more slanders and defamation (Danielson, 1996; Sproull & Kiesler, 1986). However, at the same time, if this strict legal enforcement leads to discouragement of willingness-to-express related to freedom of speech to the internet users, the law is dysfunctional and the results may yield undesirable spillover effects (Myers & Lamm, 1976; Walther 1992). For this reason, some privacy experts and groups in South Korea also criticize this policy which enables a person's real identity to be fully traceable. Some researchers tried to examine the impact of Real Name Verification Law in South Korea. Most argued the legal aspects and appropriateness of the policy (Kim, 2009; Hwang, 2007; Myung, 2005). More recently, Woo et al.(2010) pioneered the empirical study to explore the effects of the law with real world dataset, and their findings are the following: the number of postings decreased, slanderous and swear words in postings did not decrease, but those in replies(or comments) decreased in replies after the implementation of the law. However, the study only observed each 10 days before and after the law implementation and did not examine the long term impact. Not surprisingly, users would be more cautious led by excessive sense of alarm right after the law. This paper explores both short term and long term effects of the law with abundant data from several sources. The research model has been designed for this purpose, as shown in Figure 1.

Figure 1 shows the main factors to be examined and the structure of the research. Site 1 and Site 3 use nickname policies at their online discussion forums; by contrast, Site 2 uses real name policy regardless of the existence of the law. In other words, Site 2 is with more enhanced level in terms of exposing users' real identity, compared to other two sites. Site 1 and Site 2 are subjected to Real Name Verification Law, but Site 3 is not applied to the law due to its smaller number of daily visitors. Site 1 is of main interest in the study in order to see policy effects and users' behavioral change, and it is compared to those in Site 2 and Site 3 for the robustness check. By comparing Site 1 to Site 2, I can observe how users behave differently when they write messages with their real identity. In addition, Site 3 is a reference of Site 1 for verifications.

**Figure 1. Research Model.**




**Before Real Name Verification Law**

| Policy \ Level of Anonymity | Pseudonym (Nickname) | Real name |
|---|---|---|
| No regulation | **Site 1** and Site 3 | Site 2 |

**Removal of anonymity**
**Creation of traceability**

**After Real Name Verification Law**

| Policy \ Level of Anonymity | Pseudonym (Nickname) | Real name |
|---|---|---|
| Subject to Real Name Verification Law | **Site 1** | Site 2 |
| Not applied | Site 3 | -- |

Real Name Verification Law gives participants the feeling that others can observe and track their posting history. For example, government's inspection authority may be involved and punished in case that a particular user's postings lead to negative social cues. This in turn can affect the quantity of interaction and the richness of discussion in the online forums. Not surprisingly, anonymity has been found to increase actual participation over identified groups (Kahai et al. 1998). In addition, when members are originally reticent to participate, anonymity frees them from evaluation and apprehension and leads to more participation (Pinnsonneault et al. 1999). Thus, the freeing of participants from surveillance and traceability seems to be a pronounced effect on participation. Since the enforcement of Real Name Verification Law was widely reported by the mass media and websites that were applied to the law at the point of its implementation, it is assumed that most internet users were aware of it. The first two hypotheses are to measure whether the law directly discouraged the willingness-to-express at the online forum:

*Hypothesis 1: Participants decrease after the law comes into effect.*
*Hypothesis 2: The number of postings decreases after the law comes into effect.*

Next, I measure how effective the law was in terms of reducing swear words and antinormative expressions. Excessive slanders and abuse are closely associated with privacy invasion and defamation against particular individuals and groups, such as politicians and political parties.

> *Hypothesis 3: The numbers of postings that include swear words and antinormative expressions decrease after the law comes into effect.*

Next, I propose that the behavioral responses against the law would be varied according to the nature of the users. Some users may stay in the online forum longer than others and post more. Other users may participate in the discussion of topics in which they are particularly interested. I conjecture that people involved more heavily in the discussion were affected by the law rather than temporary users.

> *Hypothesis 4: The heavy users are more influenced by the law than the light users.*

Also of interest is how the operating policy of the online discussion board influences users. I assume that users behave more responsibly and discreetly with their real name, because writings with real name cannot be separated with their real-world identity.

> *Hypothesis 5: There are less abusive and antinormative postings in the real-name discussion forum than those in the pseudonym-based discussion forum.*

# 4. Data and Methods

### 4.1. Data

Data is collected from Site 1, which is the most popular online discussion forum in the South Korea's top portal website[7], using a web-crawling method. I chose a discussion board with the topic of political issues. Political and diplomatic issues, including topics relevant to North Korea, public policy, elections are more arguable in the online discussion; therefore, it provides a higher likelihood to represent users' thoughts and emotions directly. Since it was unlikely for users to change their nicknames in the given period, it is assumed that each nickname represents each separate individual.

During the period from January 2005 to December 2010, 2,024,274 postings written by 207,124 participants (nicknames) are collected, and the mean of postings and participants per month are 28,114 and 2,798 respectively. Figure 2 illustrates the trends in the number of postings and participants. A vertical solid line in the middle of the graph is the point of Real Name Verification Law, and the oval around the line is the period of interest in this study. In the long term point-of-view, it is seemingly conjectured that the number of postings and participants were more influenced by particular politically disputable issues rather than by the law. Two kinked points in the rectangular were caused by two nationwide striking issues, the recent presidential election in December 2007 and the dispute against the government's decision to import American beef respectively.

[7] http://agora.media.daum.net/ According to the official announcement of the portal operator, the total number of postings in all categories of forums was 784,107 by 74,949 participants in the period April 1, 2008 to June 18, 2008, and the number unique visitors were 7.9 million in April, 2008.

In order to see behavioral changes and policy impacts more directly caused by Real Name Verification Law, I selected two separate periods of short term and long term, which are 60 days and 6 months, as shown in Figure 3.

**Figure 2. The trend of number of postings and participants at Site 1 (2005-2010).**

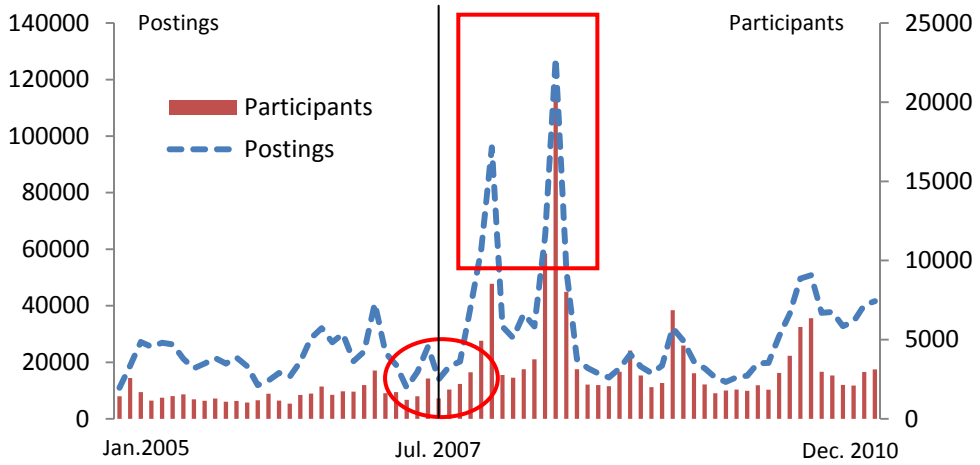
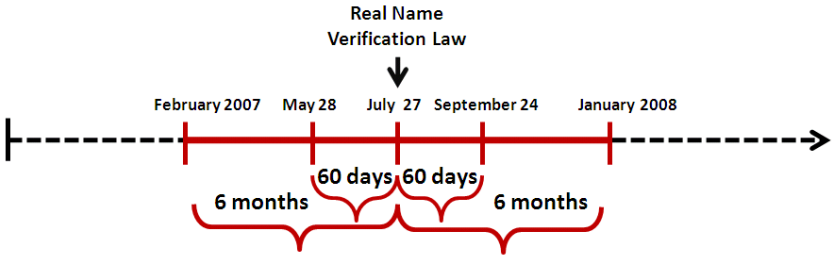



**Figure 3. Two Periods (short term and long term) of the study.**




For the supplemental study, two websites, Site 2[8] and Site 3[9], are considered: Site 2 is a discussion forum of a leading newspaper site and Site 3 is a debate forum specialized in political issues. As noted above, Site 1 and Site 2 have been subject to Real Name Verification Law since July 2007. In these Sites, only users who passed real identity verification process could leave postings in the latter period. By contrast, Site 3 is not associated with the law, because its daily visitor was less than 300,000. Anybody can express their opinions with their own nicknames without any authorization process in Site 3.

### 4.2. Methods

[8] http://www.chosun.com/
[9] http://www.freezone.co.kr/

First of all, several terms have to be clearly defined. *Posting* is a message that is written by an individual (nickname), and each posting is represented with time, date, writing number by time order, and nickname set by each user. Content analysis is based on two criteria: *swear words* and *antinormative expressions*. Even though it is certain that these two standards are closely correlated, each factor indicates salient meaning. Postings with swear words[10] can be regarded as the most aggressive behaviors. This is closely associated with privacy invasion and defamation as well as deindividuation that was discussed in the previous section. Contrastingly, postings with antinormative expressions indicate that writings include slanderous and abusive words against certain governmental policies, politicians, or other discussion participants. It is common to use irrational contractions of words and transformed insulting words which readily lead to conceive the original words. Although these expressions are less aggressive than expressing swear words directly, it is assumed that participants use these words by exploiting anonymous condition in the online forum. To do this analysis, 34 and 585 keywords are predetermined for swear words and antinormative expressions respectively. In order to examine the behavioral shift across the nature of users, six categories and three groups of users are defined as following criteria on the basis of the number of postings per month: (1) 1 posting and (2) 2 postings: light user, (3) 3 posting and (4) 4-10 postings: middle user, (5) 11-15 postings and (6) more than 15 postings: heavy user.

The following technical approach is implemented. To run and display content on the browsers, web data such as HTML and JavaScript are downloaded to the client side (Site 1). This data stream can be read and its target content can be parsed, a concept known as "crawling and parsing". A web crawler is designed to reconstruct all URLs, also known as the seeds of the target sites, and retrieve content from those pages using Java and MS-SQL. Based on extracted data from the target sites, data mining and filtering methods can be implemented to retrieve postings with specific keywords—swear words or antinormative expressions.

# 5. Results

### 5.1. The change in the number of postings and participants

Before I examine the change of postings based on 60 days and 6 months windows, the change in each 3 weeks before and after Real Name Verification Law is illustrated in Table 1 and Figure 4. Salient decrease at the point of the implementation of the law on July 27, 2007 is observed in Figure 4, and the average number of daily postings in the discussion has decreased from 477 to 379, and the difference is statistically significant. The average number of daily participant has also decreased from 124 to 115, but the difference is not significant. This can be interpreted that some users might be more cautious in posting their opinions right after the law enforcement.

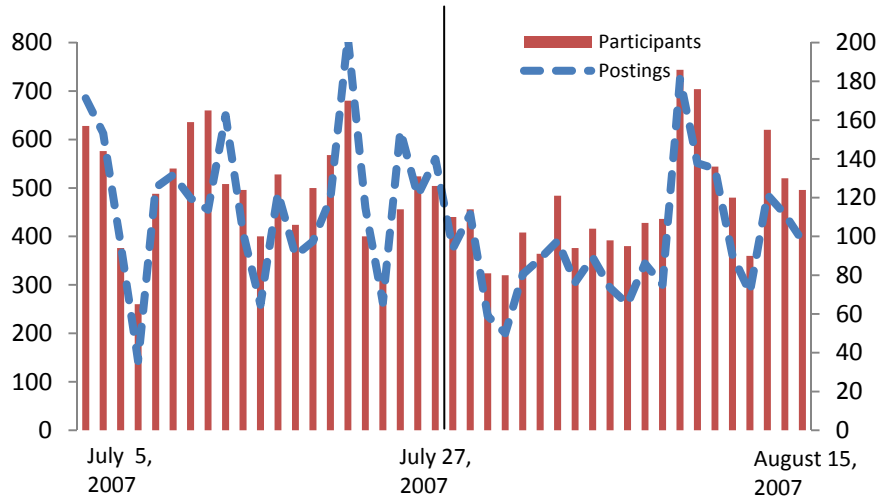
**Table 1. The comparison between before and after the law (3-week windows).**

| | Before the law | After the law | Change (%) | t-statistic (p-value) |
|---|---|---|---|---|
| Average daily posting | 477 | 379 | -98 (-20.5%) | 2.28** (0.014) |
| Average daily participants | 124 | 115 | -9 (-7.6%) | 1.08 (0.14) |
| Average posting per user | 3.78 | 3.25 | -0.53 (-14.0%) | 2.61*** (0.006) |

* significant at 10%, ** significant at 5%, *** significant at 1%

[10] Most discussion board systems automatically block writings with well-known swear words by their own filtering algorithm, so it is common for participants to use these words in a slightly different way. These words were regarded as homogeneous words to the original swear words in this paper.

**Figure 4. The trend of number of postings and participants at Site 1 (July – August 2007).**




Regardless of this result, it is difficult to conclude that users' willingness-to-express was discouraged due to the law. It was straightforwardly expected that the number of postings would decrease in this short time period, because the implementation of the law was highly disputable through online and offline and offline media at that time. Thus, I examine the impact of the law on the basis of short term (60 days) and long term (6 months) by each user group, and the result is provided in Table 2. First, while the number of postings is decreased by about 10% from 20,083 to 18,132 in the short term, the number of participants is slightly reduced by 3.7%. The difference of postings is statistically significant, but the difference of participants is not. When it comes to the long-term, the numbers of both postings and participants are strikingly increased. As noted previously, the main reason that more participants expressed their opinions was the disputable political issues in South Korea in the late 2007-2008, but, at the same time, this is clear evidence that Real Name Verification Law is not associated with users' willingness-to-express on the Internet unlike prevalent concerns by experts and activists in South Korea (Hwnag 2007; Woo et al., 2010). In other words, active discussions are maintained according to the political issues, so Hypothesis 1 and 2 are rejected based on the long term perspective. This result is somewhat conflicting with the findings by Woo et al. (2010) in which they concluded that Real Name Verification Law affected users' intention to post negatively. It should be noted that the period of their study was only 20 days around the point of the law, and their result is along with the narrow time period result that is specified earlier. As they specified that the short time period as the limitation of their study, examination based on the longer periods is required to see the effect of the law in terms of freedom of speech at the online space. However, I should be cautious to interpret the result. My conclusion is that Real Name Verification Law does not seem to be related to overall participation in the long term, according to Figure 2 and Figure 4. This means that there is no causal relationship between the law and the participation.

### Comparison with Site 2 and Site 3

I checked relatively fewer postings after the law in the short term, but no relevance to the law in the long term from Site 1. In order to make this argument more concrete, I compared the trend of the number of postings in Site 2 (a site with real-name posting policy and applied to Real Name Verification Law) and Site 3 (a site with pseudonym policy and not applied to the law) in the same period. It is possible that some users in Site 1 who were sensitive to the law implementation might

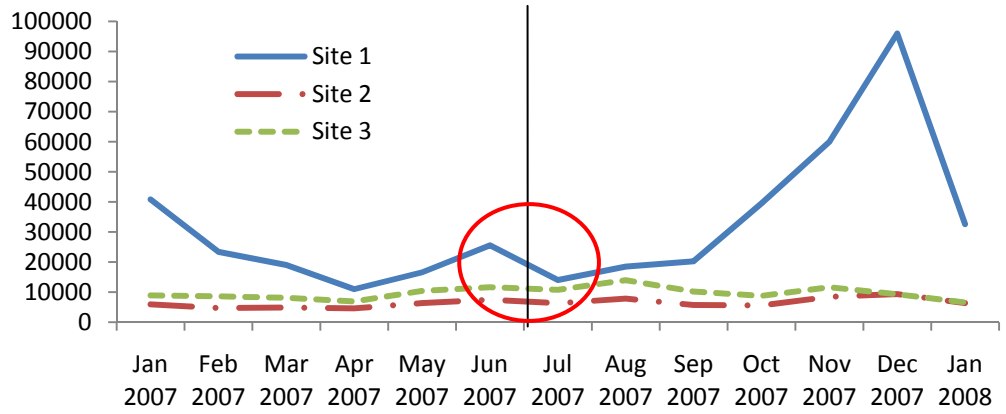
move to the other discussion forums to which Real Name Verification Law was not applied. Nevertheless, it would be difficult to create a new actively debating place, since all popular websites including newspaper websites, portals, and social networking sites are subject to the law. Yet, some debate-specialized websites, such as Site 3, can be alternative choices for some users as their shelter. Thus, I checked if there was a significant change in other relevant and equivalent websites. If there is a sign that the significant increase in the number of postings and participants, Real Name Verification Law would cause an undesirable spillover effect. As shown in Figure 5, no striking shifts in Site 2 and Site 3 unlike somewhat radical variations in Site 1 are observed around the point of the implementation of the law even in the short term.

**Table 2. The change in participants and postings by each user group.**

| | | (1) Short term (60 days) | | | | (2) Long term (6 months) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Before the law | | After the law | | Before the law | | After the law | |
| User Group | # of postings | Parti-cipants | Postings | Parti-cipants | Postings | Parti-cipants | Postings | Parti-cipants | Postings |
| Light User | 1 | 1,026 (52.80%) | 1,026 (5.11%) | 962 (51.38%) | 962 (5.30%) | 839 (51.60%) | 839 (4.59%) | 1,775 (45.86%) | 1,775 (3.99%) |
| | 2 | 251 (12.93%) | 503 (2.50%) | 265 (14.15%) | 529 (2.92%) | 219 (13.44%) | 438 (2.39%) | 572 (14.77%) | 1,143 (2.57%) |
| Middle User | 3 | 128 (6.60%) | 385 (1.92%) | 128 (6.85%) | 384 (2.12%) | 109 (6.72%) | 327 (1.80%) | 300 (7.75%) | 900 (2.20%) |
| | 4 ~ 10 | 291 (14.98%) | 1,758 (8.76%) | 281 (15.02%) | 1,682 (9.28%) | 239 (14.67%) | 1,430 (7.83%) | 649 (16.76%) | 3,894 (8.76%) |
| Heavy User | 11 ~ 15 | 59 (3.04%) | 809 (4.03%) | 60 (3.19%) | 700 (3.86%) | 51 (3.13%) | 645 (3.53%) | 146 (3.76%) | 1,859 (4.18%) |
| | 16 and more | 187 (9.65%) | 15,603 (77.69%) | 176 (9.42%) | 13,875 (76.52%) | 170 (10.44%) | 14,591 (79.86%) | 429 (11.10%) | 34,903 (78.48%) |
| Total | | 1943 (100%) | 20,083 (100%) | 1,871 (100%) | 18,132 (100%) | 1,626 (100%) | 18,269 (100%) | 3,870 (100%) | 44,473 (100%) |

Each value is the average number per month. The percentage value in parenthesis is the proportion out of total value in the last row.
Short-term: Participant: $\chi^2 = 1.67$, df = 5, p = 0.89, Postings: $\chi^2 = 13.90$**, df = 5, p = 0.016
Long-term: Participant: $\chi^2 = 15.72$***, df = 5, p = 0.007, Postings: $\chi^2 = 46.38$***, df = 5, p = 0.000
* significant at 10%, ** significant at 5%, *** significant at 1%

**Figure 5. Comparisons of the number of postings (Jan.2007 – Jan.2008).**



### 5.2. The change in the number of postings with swear words and antinormative expressions

The main interest of this study is to see how much participants' behaviors in terms of using swear words and antinormative expressions were shifted after the implementation of the law. These negative postings are associated with privacy invasion and defamation, and the reduction of these postings is the primary policy goal. The results are reported in Table 3. The proportion of each category which is specified in the parenthesis is calculated by the number of postings with swear words divided by the number of total postings in each category of the given period. While the proportions of swear words and antinormative expressions in Light User group are mostly greater than those in Middle and Heavy User groups, the actual numbers of bad postings in Heavy User group is greater than those in Light and Middle User groups across all columns. The proportions of swear words and antinormative expressions in the first row of Light User group are around 20% and 30%, compared to around 2% and 5% in the last row of Heavy User group in Table 3. There are two reasons for this finding. First, a few heavy users post far greater number of postings, several hundred in each month; therefore, the distribution is highly skewed to this group. Second, participants in Light User group usually post one or two postings according to their interests, and these postings are more likely to contain either swear words or antinormative expressions.

**Table 3. Trends of using swear words and antinormative expressions.**

| | | (1) Short term (60 days) | | | | (2) Long term (6 months) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Before the law | | After the law | | Before the law | | After the law | |
| User Group | # of postings | Swear words | Anti-normative | Swear words | Anti-normative | Swear words | Anti-normative | Swear words | Anti-normative |
| Light User | 1 | 195 (20.00%) | 282 (28.78%) | 166 (16.06%) | 261 (25.26%) | 171 (20.91%) | 244 (29.92%) | 334 (17.92%) | 509 (27.54%) |
| | 2 | 82 (17.75%) | 130 (26.37%) | 77 (12.96%) | 146 (24.87%) | 84 (19.80%) | 132 (30.71%) | 153 (13.43%) | 285 (25.29%) |
| Middle User | 3 | 71 (17.49%) | 90 (27.08%) | 53 (12.88%) | 98 (23.20%) | 61 (18.20%) | 92 (29.19%) | 104 (12.66%) | 169 (20.09%) |
| | 4 ~ 10 | 241 (13.73%) | 447 (26.22%) | 176 (9.46%) | 402 (21.29%) | 247 (17.18%) | 413 (28.93%) | 312 (9.28%) | 664 (19.53%) |
| Heavy User | 11 ~ 15 | 111 (13.45%) | 137 (16.24%) | 96 (14.30%) | 160 (20.23%) | 108 (17.84%) | 184 (30.63%) | 119 (8.93%) | 252 (16.54%) |
| | 16 and more | 323 (1.77%) | 1,184 (6.93%) | 150 (1.03%) | 589 (4.03%) | 355 (2.33%) | 987 (6.49%) | 330 (1.01%) | 1,116 (3.62%) |
| Total | | 1,022 (4.83%) | 2.269 (10.85%) | 717 (3.71%) | 1,655 (8.54%) | 1,025 (5.58%) | 2,051 (11.13%) | 1,351 (3.26%) | 2,996 (7.38%) |

Each value is the average number per month.
Short-term: Swear words: $\chi^2 = 26.99$***, df = 5, p = 0.000, Antinormative expressions: $\chi^2 = 112.65$***, df = 5, p = 0.000
Long-term: Swear words: $\chi^2 = 49.14$***, df = 5, p = 0.000, Antinormative expressions: $\chi^2 = 75.57$***, df = 5, p = 0.000
* significant at 10%, ** significant at 5%, *** significant at 1%

When it comes to the shift between two periods, the proportions of swear words and antinormative expressions are reduced in all cases, and all differences are statistically significant. In particular, behavioral shifts are observed more saliently in Heavy User group and this mainly lead to the difference of total proportions between two periods. Assuming that these heavy users' frequent postings influence the mood of discussion forum more critically, this result is meaningful in terms of the effects of Real Name Verification Law. That is, if their behavioral shift is truly maintained, this may affect behaviors of both Light and Middle User groups in positive directions. In addition, it is noted that the discrepancies in the long term are greater than those in the short term. This fact can be interpreted that the effect of policy is accelerated over time. Unlike the noticeable change in

Heavy User group, there were less significant shifts in Light and Middle User groups. In these groups, it is highly likely that users who posted swear words before the law implementation are different from those who posted after the law. That is, it seems that new participants after the law did not change their behaviors and Real Name Verification Law did not affect the fundamental shift of behaviors. There are some limitations of this study and caveats in interpreting results. Undoubtedly, postings which are involved in negative emotional expressions notably depend on the level of disputes surrounding the political issues. In addition, since the number of postings by a few heavy users is quite great, their behavioral change led to this result. In order to solve these problems, future research should explore the consistency of the results from other websites and by tracking political issues more precisely. Nevertheless, the findings of this study indicate that the governmental intervention was effective to reduce swear words and antinormative expressions that might cause privacy invasion to some extent, and Hypothesis 3 and 4 are supported. This result is also consistent to the findings from experimental designs in the previous research that argued people would behave more politely and conformingly under the identifiable condition rather than anonymous condition.

### 5.3. The behavioral changes of particular individuals

In the previous sections, I checked the trends in the number of postings and participants as well as the reduction of bad postings after Real Name Verification Law. Even if swear words and nonnormative expressions can be detected through text mining and filtering techniques, the result could be downward biased, because some slanderous postings may not contain selected keywords in the predetermined algorithm. Thus, as supplemental analysis, I conducted additional content analysis in greater details by investigating directly[11]. The same approaches for detecting slanderous and abusive postings which may or may not include swear words and antinormative expressions are conducted. By doing this, I can also check the robustness of the previous analysis and more precise behavioral changes across user groups. Since I am interested in postings around the implementation of the law, I selected participants who wrote their postings in July and August 2007. 1,289 users posted at least one message in July 2007, and 120 participants[12] are randomly selected for this study. In order to examine the behavioral shift, I only need to see users who write postings before and after the law; therefore, 41 users are taken into account.

Table 4 reports the results. In Column (1), the total number of postings increased, and this difference is not statistically significant. This finding is associated with the previous conclusions in terms of the trends in the number of postings in which the participation in the discussion is not relevant to the law enforcement. The results in terms of the number of swear words and antinormative expressions are reported in Column (2) and Column (3). Some interesting findings are as follows. First, the proportions of swear words and antinormative postings are greater than the previous results, which were the pre-coded detection algorithm. The main reason could be that more number of postings is classified to these bad postings by detailed content analysis by coders. In particular, some postings can be regarded as antinormative postings without containing significant informal expressions, such as serious slanders against the other discussion participants. Other possible reasons are due to smaller sample size and shorter period of analysis. For example, there are only 3 users in some categories. Second, the proportions of swear words and antinormative postings have decreased, and the differences are statistically significant. However, salient finding is that the differences between two

[11] Two independent coders rated each set of postings on the basis of the provided standards, and cross-checked for coding reliability. Both coders are PhD students in social science fields in the US, whose native language is Korean.

[12] After sorting by the number of postings in the given period, every tenth of participants are selected.

periods are smaller than those in the previous section (5.2). It can be interpreted that some users still behaved in aggressive ways even after the law, but the usage of explicitly abusive words might be decreased. Thus, it should be more cautious to conclude whether Real Name Verification Law has achieved its policy goal. The decrease in swear words and antinormative expressions can be positively associated with less slanders and defamation, but the finding suggests that this is not guaranteed by counting the words they used. That is, it is required to explore in more details in order to assess the policy impacts properly. Nevertheless, main results in terms of the reduction of swear words and antinormative expressions are supported and consistent with individual-level data.

**Table 4. Detailed content analysis across groups.**

| | | (1) Postings | | (2) Swear words | | (3) Antinormative | |
|---|---|---|---|---|---|---|---|
| **User Group** | **# of postings** | **Before law** | **After law** | **Before law** | **After law** | **Before law** | **After law** |
| Light Users | 1 (11) | 11 | 45 | 3 (27.27%) | 9 (20.00%) | 5 (45.45%) | 20 (44.44%) |
| | 2 (3) | 6 | 8 | 2 (33.33%) | 2 (25.00%) | 2 (33.33%) | 3 (37.50%) |
| Middle Users | 3 (3) | 9 | 17 | 2 (22.22%) | 3 (17.65%) | 3 (33.33%) | 5 (29.41%) |
| | 4 ~ 10 (10) | 78 | 102 | 12 (15.38%) | 14 (13.73%) | 18 (23.08%) | 21 (20.59%) |
| Heavy Users | 11 ~ 15 (3) | 47 | 44 | 6 (12.77%) | 4 (9.09%) | 12 (25.53%) | 7 (15.91%) |
| | 16 and more (11) | 548 | 762 | 24 (4.38%) | 26 (3.39%) | 70 (12.77%) | 72 (9.38%) |
| Total (41) | | 699 | 978 | 49 (7.01%) | 58 (5.89%) | 110 (15.74%) | 128 (13.01%) |

The percentage value in parenthesis of "# of postings" column is the number of samples in each category. The percentage value in parenthesis of column (2) and (3) is the proportion out of total postings in each row.

Postings: t value = 1.36 , df = 5 p = 0.11

Swear words: t value = 3.67 ***, df = 5 p = 0.007

Antinormative expressions: t value = 1.48 *, df = 40 p = 0.09

* significant at 10%, ** significant at 5%, *** significant at 1%

### 5.4. Postings with real name vs. postings with pseudonym

The last part of study is to examine how the level of anonymity affects users' behaviors. Regardless of whether a form of legal enforcement by a regulatory authority existed or not, encouraging normative behaviors can be achieved by the other mechanism, such as self-regulation by website operators[13]. I argue that users behave more responsibly and discreetly with their real name regardless of the existence of the law. Writings with real name cannot be separated with their real-world actual identity. Site 1 and Site 2 are used for this study: as aforementioned, both websites are subject to Real Name Verification Law, but the only difference is that the messages are posted with *pseudonym* at Site 1, but users have to post with *real name* at Site 2. Two controversial political issues are chosen; the each discussion topic was in May 2007 and in May 2009 respectively, which are before and after the law implementation. Then, 100 postings are randomly selected at each discussion topic from both Site 1 and Site 2. I followed the method and the procedure conducted by Part et al. (2008). They found that postings in the discussion board with real name showed less emotional behaviors and less grammatical errors than those in pseudonym-based online boards. I selected different topics and additional periods as well as different criteria, compared their study.

[13] Some websites including Facebook.com hold their own policy in terms of using users' real names rather than their nicknames.

Topics:

- Construction of Grand Canal Project (May 2007)
- Administrative capital relocation plan and new policies (May 2009)

The results are reported in Figure 6 and Table 5. Firstly, as shown in Figure 6, the results are clearly distinguishable between two Sites. Site 2 shows less postings that contain swear words and antinormative expressions than Site 1 for both topics in two periods. It is noted that Topic 2 seems to be more disputable than Topic 1; therefore, there are more postings that contain swear words and antinormative expressions even after the law. The salient finding in this study is that participants with their real identity show more prudent behaviors than those with pseudonyms. Thus, Hypothesis 5 is supported with this result. Since the number of bad postings is closely correlated to the nature of topics, it is not meaningful to compare both periods in this case. Secondly, the result in Table 5 shows that users participated in the discussion more actively at Site 2. Real-name online discussion board (Site 2) shows more frequent participation rates than those in Site 1. That is, one-off postings are the majority in the nickname-based board of Site 1.

**Figure 6. Comparison between Site 1 and Site 2 (Swear words and Antinormative expressions).**

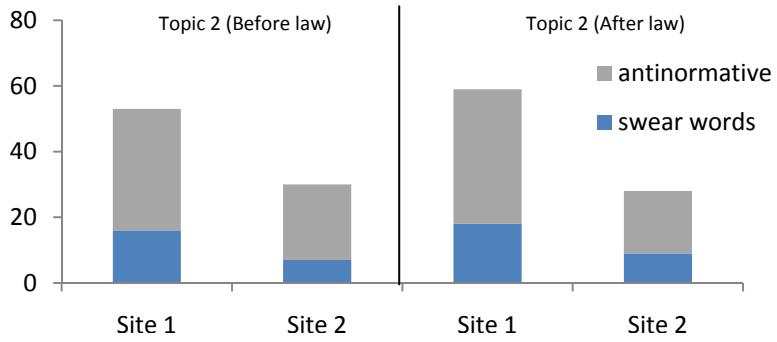



**Table 5. Participation frequency and rate in Site 1 and Site 2.**

| | Topic 1 (Before the law) | | Topic 2 (After the law) | |
|---|---|---|---|---|
| | Site 1 | Site 2 | Site 1 | Site 2 |
| 1 | 79 (90.1%) | 62 (81.6%) | 73 (86.9%) | 53 (74.6%) |
| 2 | 3 (3.4%) | 7(9.2%) | 8 (9.5%) | 10 (14.1%) |
| 3 | 2 (2.2%) | 5 (6.6%) | 2 (2.4%) | 7 (9.9%) |
| More than 3 | 3 (3.4%) | 2 (2.6%) | 1 (1.2%) | 1 (1.4%) |
| Total | 87 (100%) | 76(100%) | 84 (100%) | 71 (100%) |

The percentage value in parenthesis is the proportion out of total value in the last row.

# 6. Discussion and Conclusion

## 6.1. Discussion

There is an interesting analogy in terms of real identity between analyzed South Korean websites and famous social networking services, *Facebook and MySpace*. Users can usually be identified by their real

names on Facebook under its self-regulation policy[14]. By contrast, a constant criticism of MySpace is that real names were hidden when users see someone's profile. According to an article[15], anonymity is great when you do not want your actual identity to get in the way of whatever fantasy life you are living online. But, it is also one of the reasons Facebook, which identifies users by their real names, is gaining so quickly on MySpace. That is, people make online social connections on the basis of real-world networks. Users are willing to reveal more private stories and to share their opinions on Facebook that are consistent with their true thoughts in the real world. Abundant use of language containing abuse, slander, and aggressive slang terms on MySpace might have caused the declining use of the website. Moreover, a further study that may be highly relevant could include that of the controversial but widespread commenting systems of Facebook[16], whose advanced features were officially launched in March 2011. Even if this is helpful if users are interested in tying their non-Facebook activity to the social-networking site, this might discourage users to leave comments in sensitive topics.

### 6.2. Conclusion

This paper analyzed privacy and anonymous issues surrounding Real Name Verification Law in South Korea. The law was established in order to prevent indiscreet postings on the web; as a result, the country's cyberspace has become more traceable to users' identity as well as more vulnerable to guarantee freedom of speech.

The results suggest that Real Name Verification Law has a dampening effect on overall participation in the short-term, which is consistent with Haines et al. (2006),Kahai et al. (1998), and Woo et al.(2010); however, it is more closely associated with the nature of issues. Thus, the law did not affect the participation of the discussion in the long term. Furthermore, identification of postings had significant effects on reducing uninhibited behaviors (swear words and antinormative expressions), suggesting that Real Name Verification Law encouraged users' behavioral changes in the positive direction to some extent. The impact is greater for Heavy User group than for Light and Middle User groups. Also, discussion participants with their real names showed more discreet behaviors regardless of the enforcement of the law. It seems that users have recognized that the level of anonymity was shifted by the law, from complete dissociation of real and online identities to only visual anonymity by pseudonyms in which their real identity can be detectable. By analyzing this unprecedented policy at the forefront of internet trends of South Korea, this paper can shed light on some useful implications and information to policy makers of other countries that consider a certain type of internet regulation in terms of privacy.

This study contains some limitations. Only a few websites have been considered for analysis and comparison. If more websites are examined in each category and this significant difference is still maintained, the suggested results will be more convincing. Another limitation is *self-selection* problem.

---

[14] "Impersonating anyone or anything is not permitted. To encourage people to use their real names on Facebook, we limit the number of times names can be changed.", "Your account was disabled because the name under which it was registered was fake. We do not allow users to register with fake names, to impersonate any person or entity, or to falsely state or otherwise misrepresent themselves or their affiliations. In addition, Facebook accounts are intended for use by single individuals, so groups, clubs, businesses, or other types of organizations are not permitted to maintain accounts." *(excerpt from Help Center at Facebook.com)*

[15] http://techcrunch.com/2008/12/17/myspace-quietly-begins-encouraging-users-to-use-their-real-names/

[16] The update to Facebook's Comments Box plugin was intended to make comments more relevant. Users to a particular site see comments from Facebook friends first and Facebook added public information about users next to their names. (http://www.pcmag.com/article2/0,2817,2381630,00.asp)

Presumably, some users who are reluctant to show their real identity might use pseudonym-based online discussion forums, and they are more likely to be assertive and abusive. In other words, it is reasonable to say that more discreet users prefer to participate in real name boards, and more violent users prefer to discuss in a pseudonymous board. Notwithstanding these limitations, users' behaviors on online boards may be affected by anonymous conditions and topics rather than loosely proclaimed legal enforcement.

There are several promising future research agenda. First, future research can be conducted by collecting data from more discussion boards in the given period. People usually have their own frequently-visiting websites and participate in the discussion in those spaces. Thus, when more discussion forums are included in the study and current results are still maintained, the argument in this paper becomes highly strengthened and reasonable. Second, as noted previously, this study is linked to the next step of research in terms of the impact of recently-implemented commenting systems with real identity by popular social networking service providers in the US. Morio & Buchholz (2009) pointed out that cross-cultural differences affected interpersonal motivation (autonomy vs. affiliation) when individuals decided whether or not to remain anonymous in the online communication. Their conclusion is that Western societies will gravitate toward online communities that allow lower levels of anonymity, while individuals in Eastern societies will be more likely to seek out online communities that promote higher levels of anonymity. However, currently-implementing social comments systems in the US is more relevant to higher levels of anonymity. It would be interesting to observe whether users' behaviors are shifted when they post comments with their real personalities and how this result differs from the findings in this study.

# Reference


Acquisti, A., R. Dingledine, R. Syverson. 2003. *On the Economics of Anonymity*. In J. Camp and R. Wright, editors, Financial Cryptography (FC '03). Springer-Verlag, LNCS.

Cohen, J.E. 1996. The Right to Read Anonymously: A Closer Look at "Copyright Management" in Cyberspace, 28. *Cornell Law Review* 98.

Cohen, J.E., 2000. Examined Lives: Information Privacy and the Subject as Object, 52 *Stanford Law Review,* 1373-1398.

Danielson, P. 1996. *"Pseudonyms, MailBots, and Virtual Letterheads: The Evolution of Computer-Mediated Ethics," Ess, C. (ed.) Philosophical Perspectives on Computer-Mediated Communication.* Albany: State University of New York Press.

Diakopoulos, N., M. Naaman. 2011. Towards quality discourse in online news comments. *Proceedings of the ACM 2011 conference on Computer supported cooperative work.*

Froomkin, M.A. 1995. Anonymity and Its Enemies. *Journal of Online Law* Article 4.

Froomkin, M.A. 1996. Flood Control on the Information Ocean: Living with Anonymity, Digital Cash, and Distributed Database. *Journal of Law and Commerce* 15(2): 395-507.

Goldberg, I. 2000. A Pseudonymous Communications Infrastructure for the Internet, *PhD thesis,* UC Berkeley.

Jessup, L.M., T. Connolly, J. Galegher. 1990. The effects of anonymity on GDSS group process with an idea-generating task. *MIS Quarterly*, 14, 3: 313–321.

Hwang, Y. 2007. Critical approach to the implementation of real-name system on bulletin board of the internet: exploring right of anonymous communication, (in Korean), *Media and Society,* 15(2) 97-130.

Kahai, S. S., J. J. Sosik, B. J. Avolio. 2003. Effects of Leadership Style, Anonymity, and Rewards on Creativity Relevant Processes and Outcomes in an Electronic Meeting System Context. *The Leadership Quarterly,* 14 499-524.

Karyda, M., S.Kokolakis. 2008. Privacy perceptions among members of online communities. In Acquisti, A., Gritzalis, S., Lambrinoudakis, C. and De Capitani di Vimercati, S. (eds.), *Digital Privacy: Theory, Technologies and Practices. Auerbach Publications*, Boca Raton, 253-266.

Kim, K. 2009. The technical, legal and business issues of the real-name-verification requirement. (in Korean) *Human rights and justice*, 7: 78-97.

Marx, G., 1999. What's in A Name? Some Reflections on The Sociology of Anonymity. *The Information Society* 15(2).

McLeod, P. L., R.S. Baron, M. W. Marti, K. Yoon. 1997. The eyes have it: Minority influence in face-to-face and computer-mediated groups. *Journal of Applied Psychology*, 82: 706–718.

Morio, H., C. Buchholz. 2009. How anonymous are you online? Examining online social behaviors from a cross-cultural perspective. *AI & Society*, 23:297-307.

Myers, D. G., H. Lamm. 1976. The group polarization phenomenon. *Psychological Bulletin*, 83, 602-627.

Myung, J. 2003. Trace of the implementation of real-name system on the internet bulletin board in a public institution, (in Korean), *Center for Law & Information Society* 5:2-13.

Park, C.S, M.K. Go, S. Ki, 2008. The Impact of Anonymity on Behavior of Comments in Online Discussion Board, manuscript.

Pfitzmann, A., M. Koehntopp. 2001, Anonymity, unobservability, and pseudonymity — a proposal for terminology. In Hannes Federrath, editor, Designing Privacy Enhancing Technologies - *Proceedings of the International Workshop on Design Issues in Anonymity and Unobservability*, volume 2009 of *LNCS*. Springer.

Pinsonneault, A., H. Barki, R. B. Gallupe, N. Hoppen. 1999. Electronic Brainstorming: The Illusion of Productivity. *Information Systems Research,* 10(2) 110-133.

Postmes, T., R. Spears, M. Lea. 1998. Breaching or Building Social Boundaries: Side-Effects of Computer- Mediated Communication. *Communication Research*. 25, no.6: 689-715.

Postmes, T., R. Spears. 1998. Deindividuation and antinormative behavior: A meta-analysis. *Psychological Bulletin*, 123, 238-259.

Reicher, S.D., R. Spears, T. Postmes. 1995. A social identity model of deindividuation phenomena. *European Review of Social Psychology*, 6: 161-198.

Reidenbach, R. E., D. P. Robin, L. Dawson. 1991. An Application and Extension of a Multidimensional Ethics Scale to Selected Marketing Practices and Marketing Groups. *Journal of the Academy of Marketing Science*, 19(2) 83-92.

Siegel, J., V. Dubrovsky, S. Kiesler, T.W. McGuire. 1986. Group Processes in Computer-Mediated Communication" *Organizational Behavior and Human Decision Processes*, 37, 157-187.

Spears, R., M. Lea, S. Lee. 1990. De-individuation and group polarisation in computer-mediated communication. *British Journal of Social Psychology*, 29, 121-134.

Sproull, L., S. Kiesler. 1986. Reducing Social Context Cues: Electronic Mail in Organizational Communication, *Management Science*, 32(11), 1492-1512.

Solove, D.J. 2004. "*The digital person : technology and privacy in the information age"*, New York University Press, New York.

Solove, D.J. 2006. A Taxonomy of Privacy, *University of Pennsylvania Law Review* (154:3), pp 477-560.

Solove, D.J. 2007. *The Future of Reputation: Gossip, Rumor, and Privacy on the Internet"*, Yale University Press, New Haven, CT.

Woo, J.W., H.S. Nah, H.S., J.M. Choi. 2010. An Empirical Analysis of the Effect of Real Name System on Internet Bulletin Boards, South Korea Journal of Public Administration, 48(1), 78-96.

Walther, J. B. 1992. Interpersonal Effects in Computer-Mediated Communication. *Communication Research,* 19(1) 52-90.

Wood, A., M. Smith. 2001. *"Online communication: Linking technology, identity, and culture. Mahwah",* NJ: Erlbaum.

Zarsky T.Z., 2004. Thinking Outside the Box: Considering Transparency, Anonymity, and Pseudonymity as Overall Solutions to the Problems of Information Privacy in the Internet Society, 58 *University of Miami Law Review.*, 991, 1028-1032.