2023 통일기반구축사업 결과보고서
사회과학대학
김용균 (정치외교학부)

<구성>
I. SSCI급 국제학술지 제출 예정 논문 초고: “공통 정체성 호소를 통한 집단 간 사회통합 가능성: 실험연구”
II. 파일럿 조사 분석 결과: “네트워크 상에서의 정보전달: 설문실험연구”

I. 공통 정체성 호소를 통한 집단 간 사회통합 가능성: 실험연구

Civil wars have become the most common form of conflict since the end of colonialism in the mid-20th century. Once the active phase of fighting ends, societies enter on a precarious path of post-conflict reconciliation. At times, efforts at peace building break down and violence reignites, as has happened in Rwanda, Indonesia, and Yemen. Even in countries where civil conflict ends in a seemingly stable peace settlement or a decisive victory by one side – as in Colombia, Lebanon, or Northern Ireland – economic, ethnic, or sectarian conflict often continues to threaten lasting peace.

At the root of persistent post-conflict grievances might be unresolved disputes over the distribution of economic resources but often also an inability or unwillingness to admit the opposing side to full membership in the political community because of lingering mistrust and trauma of conflict (Lupu and Peisakhin 2017; Rozenas, Schutte, and Zhukov 2017). Put simply, the Humpty Dumpty of a unified national community is often difficult to piece together in the aftermath of a civil war or frozen conflict.

An oft-cited hypothesis, originating in sociology and social psychology, suggests that intergroup contact can reduce prejudice and reconcile groups in conflict (Williams 1947; Allport 1954). Much research in psychology has confirmed that intergroup contact successfully reduces prejudice (Pettigrew and Tropp 2006), but the efficacy of intergroup contact in the context of entrenched conflict stemming from racial and ethnic divisions remains in question (Paluck, Green, and Green 2019). Because a lot of empirical work on the effectiveness of intergroup contact is observational – mostly, studies in psychology with university student samples – it remains contested whether the findings hold in randomized control trials, whether they travel well to adult samples in real-world settings and extend to behaviors as well as attitudes, and how well changes in attitudes toward specific individuals in controlled interactions extend to behaviors toward entire groups (Dixon et al. 2005; Paolini et al. 2021).

In this study – set in the context of an attempted reconciliation among different communities of ethnic Koreans divided by the Korean War (1950-1953) – we examine the effectiveness of intergroup contact in engendering cooperation across group lines. Our study is designed to address the weaknesses in existing research

on intergroup contact. We study behaviors – altruistic giving, trust, and cooperation – that stem from prejudice reduction.[1] Ours is a randomized control trial (RCT), a laboratory-in-the-field experiment, in Korea with an adult sample.

There are three hypothesized mechanisms for how intergroup contact reduces prejudice: decategorization, recategorization, and longitudinal reformulation. In this study, we test the effectiveness of the recategorization mechanism, which is an illustration of common ingroup identity theory (Gaertner and Dovidio 2000). The recategorization hypothesis posits that inter-group contact allows members of different groups to unite around their commonalities and bond around these as an alternative and supraordinate identity. We focus on this mechanism in part because of the popularity of the common ingroup identity theory, and, in part, because it lends itself well to the context of the Korean conflict where there is an historical supraordinate identity – that of common pre-war cultural and linguistic Koreanness.

South Koreans, North Koreans, and Chinese Koreans – the ethnic Korean communities that came to be divided in the aftermath of the Korean War – share a common pre-conflict political and cultural identity and are of the same race and ethnicity. In the present day, all three groups agree that Korean customs and culture are important to their sense of the self (Jung and Choi 2018; Lee et al. 2020; Sun 2016). The primary identity cleavage in this population runs along economic and political lines that have been forged over the course of conflict. Generalizing to the context of other civil wars, different Korean communities can be thought of as conflicting groups, and the common Korean cultural identity is then equivalent to a unifying national identity. The Korean case should be an easy case for the intergroup contact theory because the three Korean communities have more in common, including ethnicity, language, and religion, than most groups in conflict.

Taking the intergroup contact hypothesis, and, within that, the common ingroup identity theory, as starting points, we set out to examine the effectiveness of activation of a strong pre-conflict supraordinate identity, built around shared Korean history and culture, in increasing altruistic giving, trust, and cooperation

1) Existing work shows an association between lower prejudice and higher trust; trust, in turn, is a pre-requisite for cooperation (Tam et al. 2009).

across groups that have been in conflict. We also ask whether the majority group – here, South Koreans – might be more susceptible to the supraordinate identity appeal because a unitary post-conflict identity would be more likely to reflect the majority's values and, as a corollary, minorities (Chinese and North Koreans) might be less taken in by such appeals out of fear of losing their cultural distinctiveness. In addition, we test whether the majority's susceptibility to common ingroup appeals might vary depending on the status of the minority group that participates in the identity construction project, given that the majority might be concerned with a potential loss of status in a highly unequal exchange. South Koreans generally consider Chinese Koreans to be lower status and less desirable as migrants than North Koreans (Ha, Cho, and Kang 2016; Park 2018). All of the study's hypotheses had been pre-registered in a pre-analysis plan filed with EGAP's Registry.[2)]

In a novel turn in the experimental literature on post-conflict reconciliation, we set out to activate a supraordinate group identity by appealing to common pre-conflict history of statehood and anti-colonial resistance. We examine the effectiveness of common ingroup identity appeals through three interventions: a control, a video intervention with narration and images about Korean statehood and anti-colonial resistance against the Japanese in the first half of the 20th century, and the same video intervention followed by group discussion designed to intensify cross-group interactions. The effectiveness of common group ingroup identity theory is tested against other intergroup contact – participants in the control condition do interact over the course of behavioral tasks but are not subject to messaging about the supraordinate identity. Thus, the intensity and duration of intergroup contact increase across treatments. The outcomes – altruistic giving, trust toward, and cooperation with the outgroup – are measured through incentivized behavioral games.

Overall, 252 South Korean citizens, 126 North Korean defectors, and 126 Chinese Korean economic migrants participated in the study in Seoul for a total of 504 participants. Participants were assigned randomly to mixed six-person teams, and the teams were in turn assigned at random to one of the three experimental conditions. Half the teams (42) had only South Korean and North Korean participants, and the other half only South Korean and Chinese Korean

2) [Link to PAP redacted for reasons of preserving author anonymity].

participants.[3)]

We find that, contrary to expectations, exposure to the supraordinate identity video message failed to increase intergroup altruism, trust, or cooperation. Where discussion followed from the video, South Korean participants tended to exhibit higher trust in their North Korean counterparts, but not toward Chinese Koreans. Even in interactions with North Korean refugees, an increase in trust was not sufficient to raise the propensity to cooperate among South Korean participants. Neither the video nor discussion had any discernible effect on the minority groups.

These findings suggest that intergroup contact, at least when it occurs around appeals to a common supraordinate identity, is not always successful at bringing about cooperation in post-conflict settings. Such appeals might be selectively persuasive for the majority group but are not as effective in resonating with minorities. And even majorities seem to balk at embracing groups that they consider lower status. As a matter of policy, our findings suggest that for intergroup contact and common supraordinate identity to be useful as post-conflict reconciliation tools, the groups participating in the identity construction project must be similar in status, and substantive engagements like cross-group discussion are likely to have a higher chance of success than passive consumption of pre-prepared messaging.

Our finding harken back to the original conditions that Allport (1954) posited as necessary for the success of intergroup contact, and especially equality of group status.[4)] Empirically, this study advances the experimental literature on the strategies of post-conflict reconciliation (Chang and Peisakhin 2019; Paler, Marshall, and Atallah 2020; Scacco and Warren 2018).

Context: Divided Korea

Korea as we know it today divided between the North and the South is a product of geopolitical perturbations of the 20th century. Since the 11th century, the Korean peninsula has been under the de facto control of a unitary state

3) Due to limited resources and difficulty of recruiting Chinese and North Korean participants, we were unable to measure outcomes in the Chinese Korean – North Korean dyad.

4) Allport's other conditions are common goals and a friendly institutional setting.

under the Goryeo (918-1392AD) and Joseon (1392-1897) dynasties. Although influenced by large and powerful neighboring states, inhabitants of the Korean kingdom had distinctive linguistic, cultural, religious, and administrative traditions. In an attempt to resist assimilationist pressures from the Mongolian, Chinese, and Japanese empires, by the late 19th century the inhabitants of the Korean peninsula developed a strong sense of a national community (Pratt 2006).

In 1910 Korea was fully annexed by Japan. The Japanese occupation provoked country-wide uprisings (Cumings 1999). The largest of these was the March 1 movement in 1919. Over ten percent of the Korean population is estimated to have participated in these anti-Japanese protests, and over 7,000 people were killed by colonial police (Han 1988). Violent protests broke out also in the Jiandao region of northeastern China (present-day Chinese province of Jilin) where there were many Korean migrants. Violent mass protests shook Korea once again in 1926 (June 10 Movement) and 1929 (Gwangju Student Independence Movement).

Following Japan's defeat in World War II, Korea was carved into two spheres of influence, similar to the contemporaneous partition of Germany. The competing governments aspired to unify Korea, and with encouragement from the Great Powers war broke out in June 1950. The Korean War was a civil conflict fought over an ideological cleavage with support in human and material resources from abroad. By January 1951, the conflict stabilized along the 38th parallel. The toll of the war was a mass loss of civilian life, estimated at some 2.5 million (Jager 2013). A formal peace treaty was never signed, and North and South Korea remain officially at war.

As a result of that conflict, ethnic Koreans are today scattered across several countries. The three dominant Korean communities in and around the Korean peninsula are the ethnic Koreans of South and North Korea and those residing in northeastern China. Those in North Korea live in a closed totalitarian state that sets up South Korea as an implacable enemy. The South Korean state portrays the North Korean political establishment in a similarly unfavorable light. Korean residents of China – the population that we term here Chinese Koreans and whose ancestors migrated to China mostly in the early decades of the 20th century – are also subject to state propaganda that seeks to foster loyalties

toward China (Seol and Skrentny 2009). Many Chinese Koreans were not involved in the fighting during the civil war but, like their brethren across the peninsula, they became victims of the frozen conflict.

The three Korean communities now view each other with suspicion and hostility not only due to the high intensity of violence in the civil conflict but also because of the efforts of state propaganda. Nevertheless, existing studies indicate that Korean culture and customs are of high importance to all three groups, which speaks to the high salience of the supraordinate Korean identity. In surveys, close to 90% of South Koreans say that being of Korean ancestry, sharing Korean culture and customs, and speaking Korean are important traits of Koreanness (Lee et al 2020). North Korean defectors stress the importance of Korean cultural identity to their definition of the self (Cho 2010), and Chinese Koreans maintain that they preserved Korean culture in exile (Sun 2016).

There are approximately one million Chinese Koreans currently in South Korea, and they make up about half of all non-native-born residents. Chinese Koreans began to enter South Korea as labor migrants when diplomatic relations between China and South Korea were re-established in 1992, and many have since settled in their historic homeland. The number of North Koreans in South Korea is lower at around 30,000-40,000 refugees, most of them in Seoul. North Korean citizens in South Korea are defectors who fled at a substantial risk to themselves and their families in search of economic opportunities and political freedom. Close to half of all North Koreans in the south report having experienced discrimination (Jung et al. 2016; Park et al. 2005). However, North Koreans are held in considerably higher esteem than Chinese Koreans, and Chinese Koreans are often perceived by South Koreans as being outside of the ethnic Korean community (Choi 2006; Park 2018). Chinese Koreans and North Koreans are clearly identifiable because of their distinctive dialects.

Theory and Hypotheses

The intergroup contact hypothesis informs much scholarship on prejudice reduction and cooperation building among groups in conflict. Allport (1954) hypothesized that tensions between conflicting groups are most likely to be reduced when two groups of equal status work toward common goals with the

aim of establishing closer relations in a cooperative manner and with institutional support. Evidence in support of the contact hypothesis is mixed; scholars have pointed out that evidence in its support comes mostly from observational studies of unrepresentative populations, and relatively few studies have examined the impact of contact on behavioral outcomes (Dixon, Durrheim, and Tredoux 2005; Paluck, Green, and Green 2019). Some authors have recently sought to test the contact hypothesis in the field (Mousa 2018; Scacco and Warren 2018).

Intergroup contact is thought to reduce prejudice through one of three possible mechanisms: decategorization, recategorization, and longitudinal reformulation. The decategorization hypothesis holds that through contact, members of different groups come to see one another not as stereotypical representatives of abstract communities but as unique individuals, and that this breaks down group stereotyping (Brewer and Miller 1984). The recategorization hypothesis – an application of common ingroup identity theory – posits that inter-group contact allows members of different groups to unite around their commonalities and bond around these as an alternative and supraordinate identity (Gaertner and Dovidio 2000). The longitudinal reformulation theory suggests that prolonged contact across group lines gradually allows group members to learn about mutual differences and realize that these are non-threatening (Brown and Hewstone 2005). Common ingroup identity theory has been especially prominent in this literature, and in this study we assess the validity of common ingroup identity theory as the main determinant of cooperation.

We do this by testing the effectiveness of a novel intervention – a short video reminding the groups in conflict that prior to the outbreak of hostilities they shared a history of nationhood, a strong unitary identity, and a common enemy. This message maps wells onto the realities of post-civil war reconstruction beyond the Korean case, as warring parties often have periods of cohabitation and cooperation before the onset of conflict and a prior history of shared struggle against a powerful external enemy, often a colonial power. Drawing on the insights of the common ingroup identity theory, we hypothesize that those exposed to the video message that seeks to activate a supraordinate Korean identity will exhibit higher altruism, trust, and cooperation toward other Korean communities (Hypothesis 1).

Some scholars have suggested that a particularly effective way to engender interaction and to increase cross-group trust and reduce support for sectarianism is through group discussion across ethnic and sectarian lines (Svensson and Brouneus 2013; Paler, Marshall, and Atallah 2020). We incorporate this insight about the potential importance of deliberation for overcoming group differences into our design and, in a different treatment, we supplement the video with a subsequent group discussion about the video's contents. We hypothesize that inter-group discussion that follows from watching the video about common Korean history will increase cross-group altruism, trust, and cooperation against the baseline of the video on its own (Hypothesis 2).

Allport cautions that for intergroup contact to effectively reduce prejudice the interacting groups must be equal. However, in the aftermath of many conflicts groups seeking reconciliation are not equal in status. And some work in psychology has claimed that even when the condition of group equality is not met contact reduces prejudice, albeit by a lower amount, (Hopkins and Kahani-Hopkins 2006; Pettigrew and Tropp 2006).

In order to advance the theoretical understanding of the effectiveness of intergroup contact and to formulate functional policies it is important to establish how groups of different relative status respond to cross-group contact. There is broad agreement that when choosing between several group identities individuals are motivated by the twin needs for differentiation and self-esteem (Brewer 1991). On the one hand, individuals wish to hold onto the values and cultural traits of their community; on the other, they would like to be associated with the higher-status identity. When a supraordinate identity option is presented to groups that are unequal in size or status, the desires for differentiation and self-esteem might come into conflict. For the larger and higher status group, supraordinate identity should be easy to adopt because the values of that identity will likely be similar to existing traits of the numerically dominant group. As a corollary, minority groups might be less open to cooperation with a dominant minority. Studies in psychology caution that the need for differentiation is very powerful in minorities, and, therefore, that minorities will tend to resist the appeal of the supraordinate identity when interacting with a hostile majority (Hewstone and Brown 1986; Hall and Crisp 2005).

We hypothesize that South Koreans, as the majority, will be more susceptible to the appeals of the common ingroup identity, and that Chinese and North Koreans, by virtue of being minorities, will be less drawn to the supraordinate identity (Hypothesis 3). We expect majorities to interact with higher and lower status minorities differently. For the majority, there is an important caveat that incorporation of lower status groups into the umbrella identity might make the supraordinate identity less attractive (Waldzus et al. 2003). This leads us to hypothesize that in interactions with Chinese Koreans, South Koreans will be less likely to embrace the supraordinate group identity and therefore will be more cooperative when engaging with North Koreans (Hypothesis 4).

Research Design

(i) Participants

To test these hypotheses, we recruited 504 ethnic Koreans to for the study. Of these, 126 individuals were of Chinese Korean descent, 126 were North Korean, and 252 were of South Korean descent. Participants were randomly assigned to South Korean-Chinese Korean and South Korean-North Korean mixed six-person teams; in total, 84 teams were formed. Individuals were randomly assigned to teams, and teams were randomized into experimental interventions. The study took place in Seoul in August 2019.

Participants are described in Table 2; the data are from survey responses that were collected before interventions had been administered. Participants were recruited with the help of a South Korean public opinion firm; the firm recruited South Korean participants with the view to representativeness and approached Chinese Koreans and North Koreans in districts where they are known to reside compactly. Power calculations for these participant numbers are reported in Appendix A.

Table 2. Participants' Characteristics

	South Koreans (n=252)	North Koreans (n=126)	Chinese Koreans (n=126)
Age, yrs.	38 (12)	48 (18)	41 (15)
Male, %	48.80 (0.50)	49.21 (0.50)	41.27 (49.42)

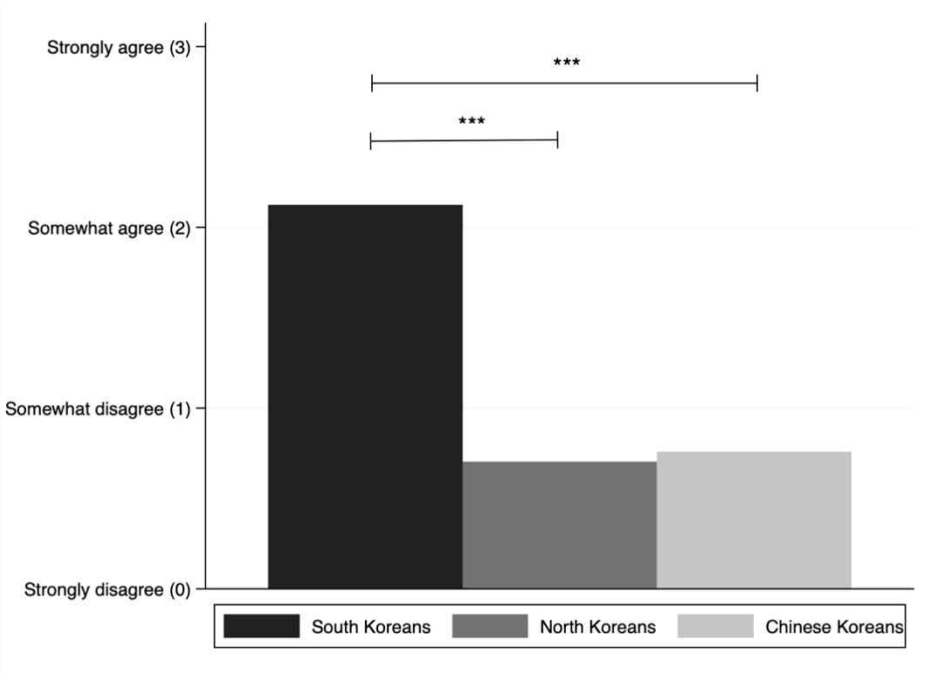
Monthly household income level, 0-6	3.06 (1.64)	0.85 (1.20)	1.55 (1.49)
Education level, 0-5	3.92 (0.54)	3.26 (0.79)	3.63 (1.00)
Employed, %	0.56 (0.50)	0.33 (0.47)	0.49 (0.50)
Knowledge of Korean history, 0-3	1.59 (0.67)	1.50 (0.81)	1.15 (0.78)
Importance of tradition and language to self-identification, 0-3	1.94 (0.64)	2.39 (0.64)	2.08 (0.74)

Note: Standard deviations in parentheses. Survey question wording in Appendix B.
Monthly Household Income level: 0 = less than 1,500,000 Won; 1 = 1,500,001 to 3,000,000 Won; 2 = 3,000,001 to 4,500,000 Won; 3 = 4,500,001 to 6,000,000 Won; 4 = 6,000,001 to 7,500,000 Won; 5 = 7,500,001 to 9,000,000 Won; 6 = 9,000,001 Won or more
Education level: 3 = high school, 4 = university.
Knowledge of Korean history: 0 = not knowledgeable at all; 1 = not very knowledgeable; 2 = quite knowledgeable; 3 = very knowledgeable.

South Korean participants in our study, by virtue of being the largest group, are quite similar to an average South Korean in the general population in terms of household income, although they are a little younger and better educated.[5] It is difficult to judge how representative the Chinese Korean migrant and North Korean defector participants are because there are no large representative surveys of these groups. Our assessment is that women are underrepresented in our pool of North Korean defectors, and while both Chinese Koreans and North Koreans in the study are better educated than an average Chinese or North Korean in Seoul, they are underemployed.

Figure 1: South Koreans' perception of their economic status as high relative to North Koreans and Chinese Koreans

5) For a comparison between our sample and the 2018 Korean General Social Survey on these variables, see Appendix C.



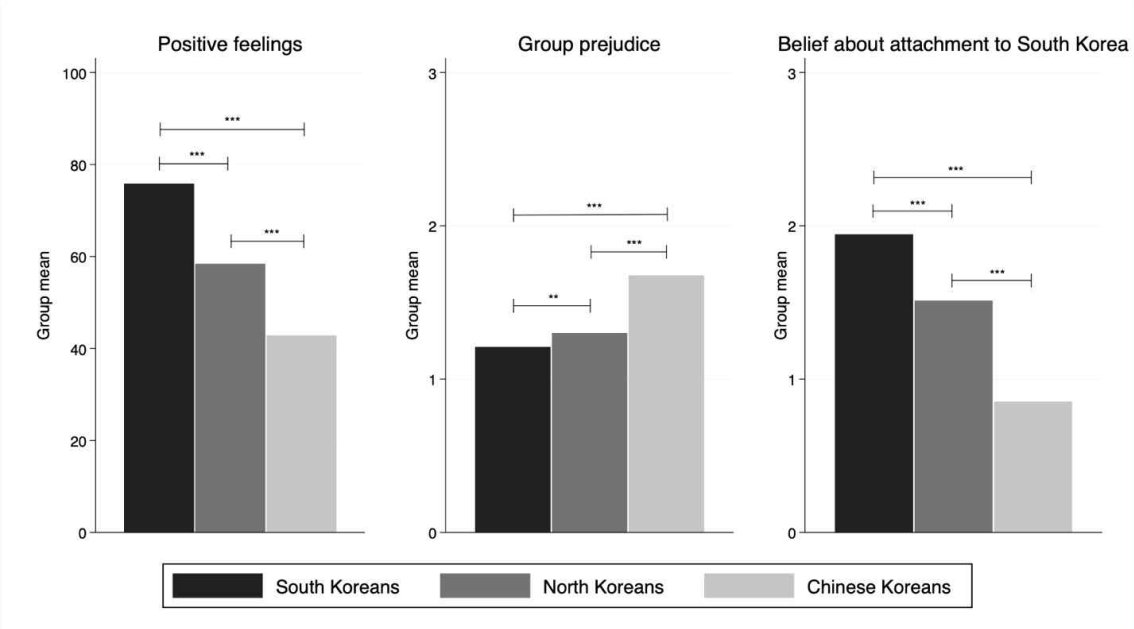
Note: The question measures agreement only among South Korean respondents (n=252) with the statement that a group's relative economic status is high by comparison to other groups. P-values report the level of difference from group comparison t-tests. * p<0.05, ** p<0.01, *** p<0.001.

Chinese and North Korean participants are considerably less wealthy than their South Korean counterparts and less likely to be employed. Consistent with this, South Korean participants consider themselves to be much better off than members of the Chinese Korean and North Korean communities as shown in Figure 1. South Korean subjects believe Chinese Koreans and North Koreans to be equally poorly off. Almost all South Korean and Chinese Korean participants have at least some university education, whereas many North Korean defectors only completed high school. The self-reported level of knowledge of Korean history is not high across the board.

One of the premises of this study is that one of the sub-groups is held in lower regard than the other by the majority. This is confirmed in Figure 2. South Koreans have less positive feelings toward Chinese Koreans relative to North Koreans, are more prejudiced toward them as a group, and believe that Chinese Koreans have a weaker attachment to South Korea.[6)]

6) Variables behind Figure 2 are explained in Appendix B.

Figure 2. South Koreans' self-reports of positive feelings, prejudice toward, and belief in the level of attachment to South Korea among other South Koreans and the two outgroups.



Note: Data are from self-reports in pre-manipulation survey (252 South Korean participants). The whiskers indicate the assessment of which two groups are being compared, and p-values report the level of difference from group comparison t-tests. * p<0.05, ** p<0.01, *** p<0.001.

All three groups agree that respect for Korean tradition and ability to speak Korean are important markers of Koreanness. North Koreans are especially likely to stress these identity traits, and there is no difference between South and Chinese Koreans in how important they consider them. This suggests that despite the economic and status differences, there might be a strong cultural foundation for a common identity between the three groups.

Participants were successfully randomized across the three interventions; see Appendix D for results of the randomization checks. As to the study's external validity, Chinese Korean migrants and North Korean defectors are unlikely to be representative of Chinese Koreans and North Koreans in general. The migrants made a conscious choice to relocate to South Korea and are therefore likely close in their values and aspirations to their South Korean co-ethnics. This creates a bias in favor of finding treatment effects given that migrants should be particularly keen to cooperate with their hosts.

(ii) Treatments: Video and Discussion

This study's design is summarized in Table 1. There are three experimental interventions. In the control condition, after being assigned to their teams, participants went straight through to outcome measurements. In the common identity video intervention, prior to playing the behavioral games, participants were shown a brief video designed to activate a common supraordinate identity. In the video + discussion intervention, participants were asked to discuss the video's content and how it resonated with them. The effectiveness of the common ingroup identity video are measured against the untreated baseline (control condition). Effects of team discussion are calculated in a comparison against the video baseline (common identity video condition). To test the effect of status differences on the willingness to cooperate we will compare interactions in mixed South and North Korean teams (baseline group status) to those teams where South Korean participants interact with Chinese Koreans (lower group status).

Table 1. Experimental Interventions

	Control	*Common Identity Video*	*Common Identity Video + Discussion*
Baseline group status	14 teams (3 South Koreans + 3 North Koreans per team)	14 teams (3 South Koreans + 3 North Koreans per team)	14 teams (3 South Koreans + 3 North Koreans per team)
Lower group status	14 teams (3 South Koreans + 3 Chinese Koreans per team)	14 teams (3 South Koreans + 3 Chinese Koreans per team)	14 teams (3 South Koreans + 3 Chinese Koreans per team)

The treatments were designed to activate a supraordinate group identity. We appealed to the idea of a common Korean political community forged in a violent struggle against the colonial Japanese oppressor. Anti-Japanese sentiment remains salient in China, North, and South Korea (Ching 2019), and therefore an appeal to anti-colonial struggle resonates with contemporary political messaging in Korean communities under study.

The video intervention reminds viewers that Korea used to be a powerful kingdom until colonization by Japan, describes the three largest waves of violent nation-wide anti-colonial protests, touches on the migration of Korean political and military elites to China, and closes with a discussion of the Cold War

partition of the Korean peninsula and an appeal to the idea of a unified Korean political nation. Throughout, the narrative stresses that many ethnic Koreans participated in the national struggle irrespective of whether they lived in the south or the north of the peninsula or on Chinese territory. The video is 11 minutes in duration, uses historical footage and maps, and is narrated by three men of similar age of Chinese Korean, North Korean, and South Korean extraction. The script of the video in English and the link to the original video file in Korean are in Appendix E.

The video intervention is a composite of multiple components. It reminds viewers about the history of common statehood but also activates the idea of a common enemy in the past. This design does not allow us to disentangle the messaging about common statehood from one about Japan as a common enemy. We do not consider this to be of concern given that the two issues are historically intertwined. The text of the intervention is also accompanied by images and patriotic music. In opting for this composite of text, images, and music we consciously favored realism in messaging and modeled the treatment on the type of messaging one might see on television. One might be reasonably concerned that a short video of this kind is a weak treatment. However, brief videos have been shown to be effective in eliciting strong effects. Exposure to short videos, 2.5 to 15 minutes in duration, has been shown to increase the intent to vote (Brader 2005), lead to opinion polarization (Druckman, Levendusky, and McLain 2017), and change respondents' preferences over government's spending priorities (Gadarian 2010).

A third of the six-person teams were randomized into the team discussion intervention. Inter-personal discussions after consuming thought-provoking material are commonplace, and the discussion intervention in this study was designed to mimic these types of conversations. Discussions lasted on average thirty minutes and were lightly moderated by trained facilitators, who intervened minimally to encourage everyone to speak and made sure that the discussion remained focused on Korean history and politics.[7] Most commonly, participants shared stories of older relatives who participated in anti-colonial protests or the Korean War, reminisced about family conversations about the Korean national community prior to the war, reflected on what they learned about Korean history

7) Guidelines for team moderators are in Appendix F.

in school and on gaps in their knowledge, and opined about what the future holds for the peninsula.

(iii) Outcome measures

In assessing the effectiveness of the interventions, we are particularly interested in the impact of the appeals to supraordinate identity on intergroup cooperation. Cooperation can take the form of non-reciprocal altruistic giving in single-shot interactions or of reciprocal exchanges in repeated interactions premised on mutual trust. To measure these different aspects of cooperation we asked study participants to play three behavioral games that measure ingroup/outgroup altruism through the other-other allocation game, relative trust in the outgroup through the trust game, and cooperation through the public goods game. Before each task team moderators read out instructions and illustrated possible outcomes through several pre-set examples; these instructions are in Appendix G.[8)]

In the other-other allocation game, participants playing in their six-person teams each received 10 tokens (around $3.50) and were asked to divide that amount between an anonymous ingroup member and an anonymous outgroup member on their team; they could not keep any of the money for themselves. In this game, higher amounts given to outgroup members (e.g., a South Korean donating to a North Korean on their team) signal higher outgroup altruism. As with the trust game that follows, payoffs were not announced until the end of the session.

The trust game was played between a Sender and a Receiver in two stages. At the outset the Sender was endowed with 10 tokens. In the first stage, she decided how many of the 10 tokens to send to Receiver. She could keep all the tokens for herself or transfer all. Whatever tokens she chose to send were doubled, and the resultant amount was transferred to Receiver. In the second stage, Receiver decided how many of the received tokens to keep and how many to send back to Sender; whatever the Receiver sent back was doubled and given to Sender. At the end of these two stages Receiver kept as earnings whatever he did not send back to Sender, and Sender earned whatever she kept back in the first stage and also the amount that she received back in the second

8) Participants recorded their preferences in each game on decision sheets, which are in Appendix H.

stage. This game measures trust insofar as Senders are likely to send larger amounts to Receivers in the first stage when Senders trust Receivers to reciprocate in the second stage, otherwise it is more rational to keep all or most of the initial endowment for oneself at the outset. We measured one's relative trust in the outgroup by subtracting the number of tokens sent as Sender to an ingroup member from the amount sent as Sender to an outgroup member. Each participant was asked to play twice as Sender and twice as Receiver, once in each role being matched with an anonymous ingroup member in the six-person team and once with an anonymous outgroup member in the team.

The final task was a public goods game, a standard measure of cooperation in behavioral experiments. The game started with every team member receiving an initial endowment of 10 tokens. Participants could then decide how much of that amount to contribute to the public pool in the knowledge that the resultant public endowment would be doubled and divided equally between each team member. Individual earnings consisted of whatever one kept for oneself initially and the payout once the public pool had been doubled and divided out. We measure willingness to cooperate as the number of tokens that a participant donates to the public pool, given that players might be tempted to keep back most of their endowment while hoping to free-ride on others' contributions. This game was repeated over five rounds. Individual contributions were disclosed anonymously at the end of every round and payoffs announced.

(iv) Compensation and Ethics

With the three behavioral games completed participants were asked to fill out a post-manipulation survey gauging their level of understanding of the tasks and attitudes toward the outgroup. Comprehension levels of the various tasks were good at 2.14, on average, on a four-point scale (0-3). Then earnings were tallied by adding up contributions in the other-other allocation game, one randomly selected interaction in the trust game, and one randomly selected round of the public goods game. Every South Korean and Chinese Korean participant received a show-up fee of $35 and between $1 and $26 in additional earnings from the three games. The show-up fee for North Korean participants was $53, as that population was especially difficult to recruit given the small size of the North Korean defector community.[9] On average, South Korean and Chinese

Korean participants earned around $46, and North Korean refugees $63; this is equivalent to 50-70% of a daily wage in South Korea. The study did not entail any deception or major discomfort. Participants were free to leave the study at any point in time. Those in discussion teams were given an opportunity to listen to the audio recordings of the discussion and have any of their comments deleted.

Results and Discussion

The paper's main results are presented in Table 3.[10] The three behavioral outcomes are the dependent variables in the horizontal panels of Table 3. The first two models capture the effect of the video intervention on cooperation in comparison to the untreated control. The last two models estimate the effect of the discussion by comparing the common identity video and the common identity video + discussion conditions.

For every dependent variable, the first and third models are parsimonious regressions that control for having a specific team moderator but do not include any other controls. In the second and fourth models we also control for participants' age, income, education level, employment status, knowledge of Korean history, cultural group (Chinese/North/South Korean), and whether they were on a team with Chinese Korean participants.[11] For the analyses of outcomes in the public goods game we also control for contribution to the common pool in previous rounds. On average, participants contributed a little over six tokens of ten toward the common pool; the pattern of contributions across the five rounds is in Appendix I.

Table 3. Effect of Treatments on Altruism, Trust, and Cooperation

Other-Other Allocation Game (Task 1): dependent variable is number of tokens sent to outgroup member (0-10)				
Model (OLS)	(1)	(2)	(3)	(4)
Video	0.12	0.11		
	(0.17)	(0.17)		
Discussion			0.28	0.27

9) Participants learned about the amount of show-up fee for the other groups only at the very conclusion of the study; they had no prior knowledge of this information.

10) The data and program code will be uploaded to a public database upon publication.

11) The number of observations varies across models because of missing survey answers that go into constructing the control variables.

			(0.18)	(0.19)
Constant	4.65***	5.96***	4.78***	4.61***
	(0.19)	(0.71)	(0.17)	(1.05)
Observations	336	304	336	307
R^2	0.01	0.06	0.02	0.06

Trust Game (Task 2): DV is tokens sent to outgroup member minus tokens sent to ingroup member (-10~10)				
Model (OLS)	(5)	(6)	(7)	(8)
Video	0.04	-0.04		
	(0.25)	(0.27)		
Discussion			0.61*	0.74**
			(0.26)	(0.27)
Constant	-0.56*	0.30	-0.35	0.26
	(0.27)	(0.98)	(0.36)	(1.12)
Observations	335	303	336	307
R^2	0.03	0.12	0.02	0.11

Public Goods Game (Task 3): DV is average contribution to common pool across five rounds (0-10)				
Model (Linear random effects)	(9)	(10)	(11)	(12)
Video	0.01	-0.04		
	(0.16)	(0.19)		
Discussion			-0.14	-0.07
			(0.19)	(0.20)
Constant	4.51***	3.83***	3.47***	3.02***
	(0.35)	(0.69)	(0.34)	(0.63)
Observations	1343	1215	1344	1228
R^2	0.33	0.33	0.29	0.30

Moderator fixed effects controlled for	Yes	Yes	Yes	Yes
Additional control variables included	No	Yes	No	Yes
Experimental conditions	1,2,4,5	1,2,4,5	2,3,5,6	2,3,5,6

Note: Standard errors, clustered at team level, in parentheses; additional control variables are *Age, Male, Monthly Household Income, Education Level, Employment Status, Knowledge of Korean History, North Korean, Chinese Korean, and Chinese Korean Session*; *Lagged Contribution* is additionally included in all models for Task 3; * p<0.05, ** p<0.01, *** p<0.001. Full regression results for all models, including coefficients for the controls, are in Appendix J.

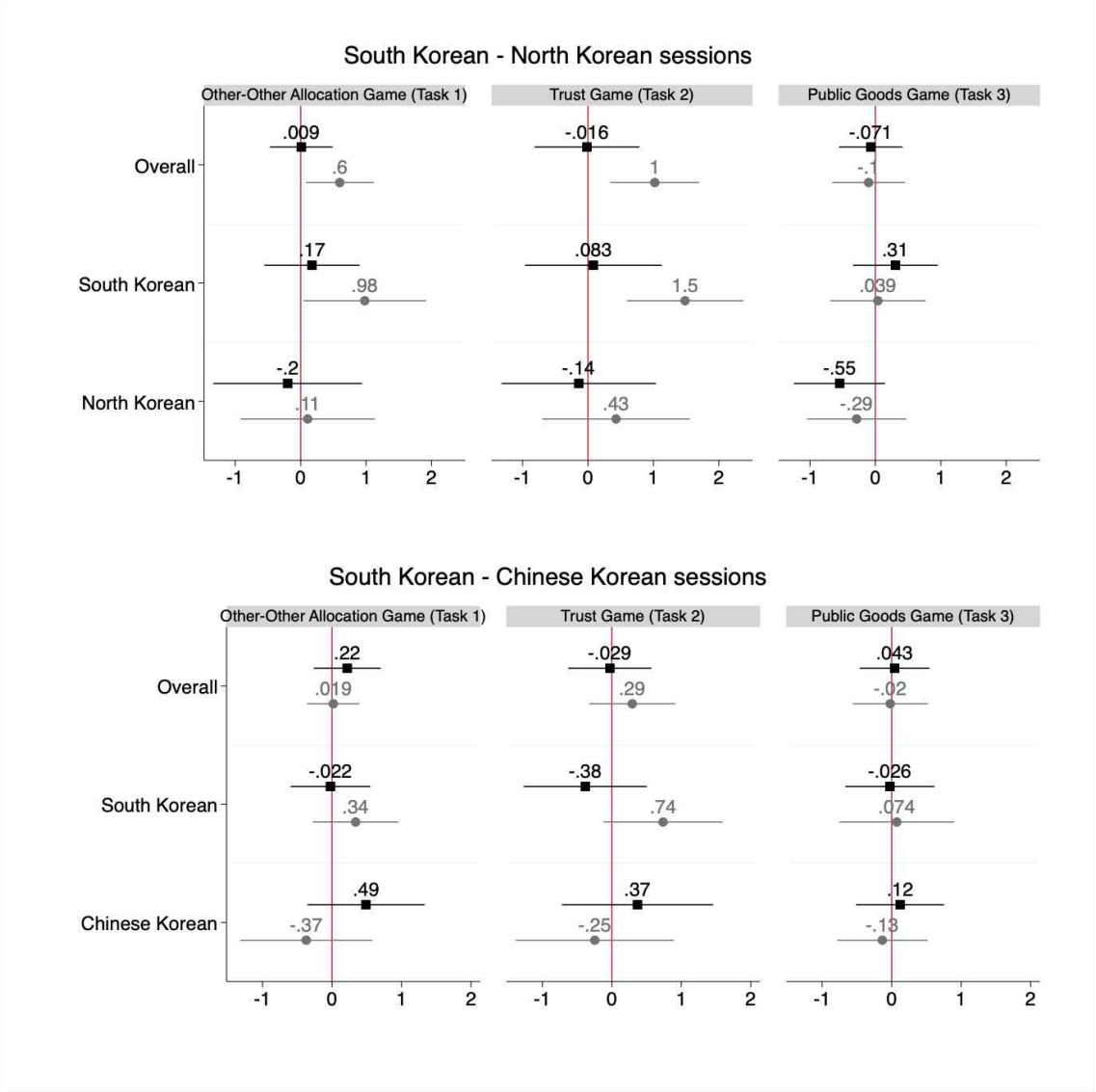
Overall, watching a video that seeks to activate a supraordinate Korean identity appears to have no discernible effect on altruistic giving, trust, or cooperation among participants. Across all three games, coefficients for the *Video* intervention are very small substantively, statistically indistinguishable from zero, and change sign depending on specification. Given the small size of the coefficients and changing sign it seems unlikely that small sample size and

therefore insufficient statistical power is to blame for null effects. This suggests that an appeal to a common pre-conflict identity – in the form delivered in this study – fails to increase cooperation in groups that are divided by conflict. Thus, hypothesis 1 is invalidated.

The effectiveness of post-video team discussion is mixed. Team discussion appears to increase altruistic giving and trust – coefficients for *Discussion* are positive and quite large in the other-other allocation and trust games. However, the positive effect of team discussion is statistically significant only for the trust game where participants send on average 0.6-0.7 of a token more (of 10) to outgroup members. The effect of team discussion on cooperation as measured in the public goods game is statistically indistinguishable from zero and has a negative sign. On balance, there is little supporting evidence for hypothesis 2.

We also hypothesized that majority and minority groups are likely to be affected differently by the treatments. The susceptibility of each of the three groups to the interventions is explored in Figure 3. There we report the marginal effects of the treatments first for all the participants combined and then separately for the majority group (South Koreans) and each minority group (North Koreans, Chinese Koreans) in the three behavioral games. The treatment effect of *Video* is denoted by a solid black line (95% confidence interval) with a small square representing the coefficient estimate and of *Discussion* by a grey line with a circular dot. Results for the South Korean-North Korean sessions are plotted in the top panel and for the South Korean-Chinese Korean sessions at the bottom. The underlying regression results are in Appendix K.

Figure 3: Contributions in behavioral games disaggregated by group



Note: Black lines and square dots are marginal effects of the *Video* intervention; gray lines and round dots are marginal effects of the *Discussion* intervention. Whiskers around point estimates are 95% confidence intervals.

The most important finding from Figure 3 is that the aggregate effect of the *Discussion* intervention reported earlier comes exclusively from within the South Korean group and only in sessions that involve interactions with North Koreans. In the altruism game (task 1), South Korean participants who had been randomized into the discussion treatment donate 1 token more (of 10) on average to North Korean counterparts than South Koreans in the control group. In the trust game (task 2), South Koreans in the *Discussion* intervention give 1.5 tokens more (of 10) to North Koreans by comparison to South Koreans assigned to the *Video* treatment. Both interventions are ineffective for the Chinese and North Korean participants. This pattern of effects provides confirmation for

hypothesis 3: the majority is more susceptible to supraordinate identity appeals than minorities. As before, neither appeal has an effect on either the majority or minority in the public goods game.

The comparison of the top and bottom panels of Figure 3 suggests that South Korean participants engage with the Chinese Korean population differently than with North Koreans. While the effects of the *Video* intervention are statistically equivalent to zero throughout, they are consistently positively signed in the interactions with North Koreans and have a negative sign when South Koreans interact with Chinese Koreans. More tellingly, the large positive effect of the *Discussion* treatment in the trust and altruism games is present only when South Korean participants interact with North Koreans, but not in interactions with Chinese Koreans. [12] This provides support for hypothesis 4, which posits that the majority group is less susceptible to a supraordinate identity appeal in interactions with a lower status. An alternative explanation is that the majority might be less open to cooperating with a minority that is culturally distant. We cannot rule this out because South Koreans consider Chinese Koreans *both* lower status and more culturally distant than North Koreans.

Returning to the main results in Table 3, a reasonable concern might be that the video treatment fails because the study participants might already be knowledgeable about the historical episodes described in the video and that the treatments did not add anything to an already high floor effect. To address this concern, we re-ran the models with an interaction term between self-reported knowledge of Korean history and assignment to the video treatment. The results are reported in Appendix M. The interaction term between treatment and history knowledge is not statistically significant, suggesting that one's level of knowledge of Korean history does not account for the video's ineffectiveness. A related concern is that older people might be more susceptible to the treatment because historical events in the 20th century are more salient for them. Interacting a continuous measure of age with the treatment, we do not find any support for the hypothesis that the treatment is more effective among older participants (see Appendix N).

12) These results are confirmed in the context of regression analyses in Appendix L that include interaction terms between treatments and session types.

Another possible explanation for the null results is that the willingness to cooperate was so high from the outset that it was impossible to measure an increase in cooperation because of high initial contributions. One plausible reason for high willingness to cooperate might be that the experience of being a team member in itself creates a desire to be cooperative quite apart from any substantive information (on team bonding effects see Scacco and Warren 2018). There is little evidence for high initial floor effects in this study. For instance, in the public goods game with average contributions at 6 tokens of 10, supraordinate identity interventions should still have become manifest given that higher contributions were possible.

The Hawthorne effect is another possible reason for null findings, as participants might have understood that the experiment was designed to measure inter-group cooperation and over the course of the study exhibited highly cooperative behavior across all treatment conditions to please the experimenters (Landsberger 1958). For this explanation to be true we should expect increasing contributions across multiple rounds of the interactions; instead, we find that contribution levels are quite constant in the public goods game (see Appendix J). We also have an alternative outcome measure, collected in the exit survey after the experiment had been completed. This measure captures participants' feelings toward the outgroup via a feeling thermometer (on correlation between outgroup feeling thermometer and prejudice see Tropp and Pettigrew 2004). This measure supports the earlier findings; feelings toward the outgroup were improved only in cross-group discussions between South Koreans and North Koreans, and only South Koreans reported warmer feelings toward the outgroup. The results of these analyses are in Appendix O.

Conclusion

We set out to establish whether intergroup contact, specifically in the form of an attempt to activate a supraordinate group identity through an appeal to common history, might help increase cooperation between groups that have been divided as a result of a frozen conflict. In a laboratory-in-the-field experiment, individuals from communities torn by conflict were randomly assigned to watch a short video appeal designed to increase the salience of a unitary pre-conflict political identity and, separately, to a cross-group discussion following on from

the video; others were randomized into the control group. The treatments' effects on altruistic giving to members of the opposite group, trust toward, and cooperation with the outgroup were measured in behavioral games.

Co-ethnics from Chinese Korean, South Korean, and North Korean communities – separated as a result of the Korean War – were the participants in this study. Ethnic Koreans are a relatively easy case for intergroup contact theory because they share the same ethnicity, language, and long pre-conflict history; what divides them is a recent history of conflict, decades of physical separation, and present-day inequality in wealth and status.

Contrary to expectations, the video intervention was ineffective in engendering higher altruism, trust, or cooperation among members of the different Korean communities. Effects of the video intervention were small, inconsistent in sign, and statistically indistinguishable from zero. Likewise, the discussion treatment was also largely ineffective, with the exception of a finding that cross-group discussion increased trust and altruism toward the outgroup only in the majority community and only when the majority interacted with the higher status group of the two minorities. Effect heterogeneity analyses indicated that the two minority groups – Chinese Koreans and North Koreans – were unaffected by either of the treatments across any of the outcomes.

There are several potential explanations for the lack of treatment effects. One theoretical possibility is that interactions between participants were negative in tone and therefore failed to engender trust and cooperation. This is not plausible given that participants were observed by team moderators, who did not report negative interactions. Another possibility is that the supraordinate historical Korean identity was not sufficiently salient a priori. However, we know from the baseline survey that all three groups prize Korean culture and traditions. It is likewise difficult to argue that the experiment might have resulted in prejudice reduction but failed to affect behavior. We reported that the treatments had little impact on feelings toward other groups.

The two most likely possibilities that remain are either that the recategorization channel, which operates by activating a supraordinate identity, is not a very effective mechanism behind intergroup contact or that, as Allport (1954) argued,

for intergroup contact to work in the field interacting groups must be equal in status. It bears noting that we attempt to activate the supraordinate identity via a video and related discussion about Korea's colonial history. It is plausible that the contents of this message, premised as they are on historically negative relations with Japan, might not be sufficiently relevant to the way the participants think about their common Korean identity. Another possibility is that a video intervention might be a weak delivery channel for this type of messaging. These potential limitations that affect the study's external validity require further attention in future work.

For now, we cautiously conclude that even in relatively favorable circumstances – like in this study, where groups divided by conflict have a rich common history and share many traits – appeals to supraordinate group identity can fail to reunite communities torn by conflict. This is, in part, because the majority group might be unwilling to associate closely with a low-status minority. And, in part, because minority groups might be reluctant to give up their cultural distinctiveness in exchange for higher status or economic opportunities. Our study adds an important wrinkle to the insights of the common ingroup identity theory and suggests that appeals to supraordinate group identity might be effective at post-conflict reconciliation in the field only if values and interests of minority groups are explicitly recognized and secured.

II. 네트워크 상에서의 정보전달: 설문실험연구

가) 하반기 사업 <Information Transmission via Network> 파일럿 실험 완료

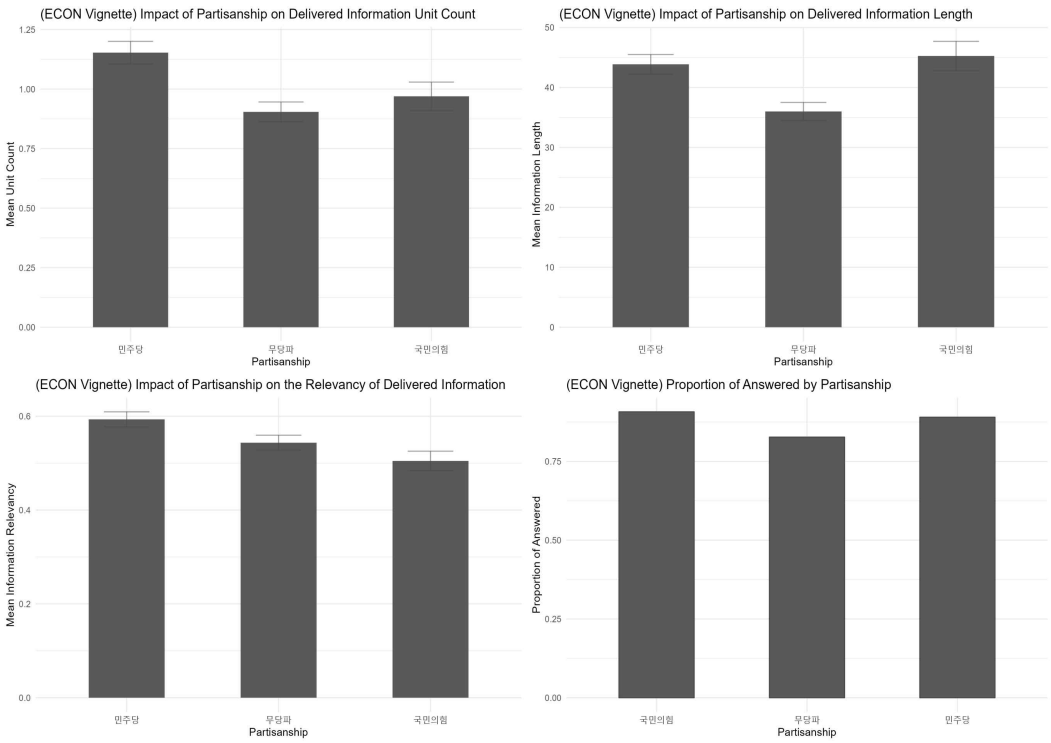
(1) 설문 실험을 통해 ①당파성에 따른 정치적 정보 전달 양상 ②당파성과 정보전달 네트워크에 따른 정보 전달 양상 ③Accuracy Nudge, Directional Nudge 등 정보 전달 양상에서의 넛지 실험 처치 효과 확인

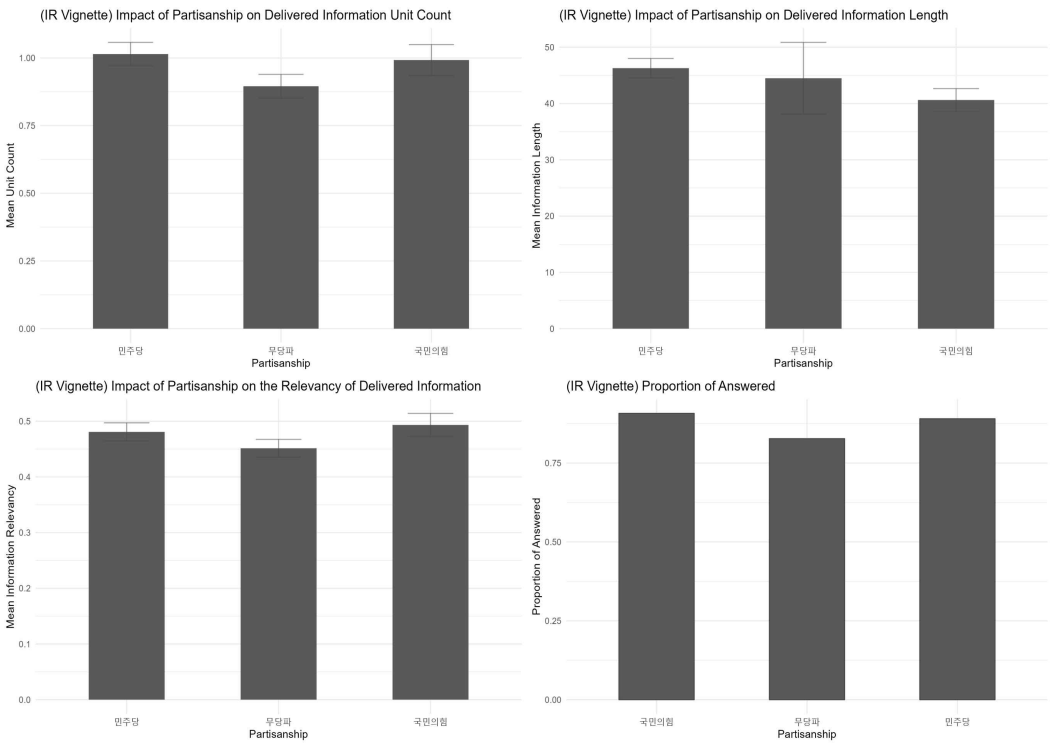
(2) 설문 실험에 활용된 지문은 실제 신문기사를 바탕으로 재구조화한 '사실에 기반한 객관적 내용 + 상충된 두 정당의 입장'으로 구성. 정치 이슈별 차이를 고려하여 국제정치 관련 지문과 경제 관련 지문을 랜덤한 순서로 제시

(3) GPT4를 활용한 Computational Text Analysis. ①정보 단위 수 information unit count ②정보 길이 information length ③관련성 relevancy ④당파성 partisan bias ⑤응답 여부 answered 등 다양한 변수를 통해 정보 전달 양상의 차이를 세부적으로 분석함

(4) 한국인 2,494명을 대상으로 한 파일럿 설문 실험 완료. 다음은 파일럿 실험의 주요 결과를 정리한 것임

나) 질문 1. 당파성에 따른 정치적 정보 전달 양상의 차이가 있는가?





(1) 경제 지문의 경우, 관련성 변수를 제외한 나머지 세 변수 - 정보 단위, 정보 길이, 응답 여부 - 에 있어 무당파에 비해 민주당, 국민의힘 지지자가 더 큰 값을 보임. 즉, 이들은 “정치적 정보 전달 과정에 높은 참여도”를 가지고 “더 많은 정보 단위의,” “더 긴” 정보를 전달하는 경향이 있음

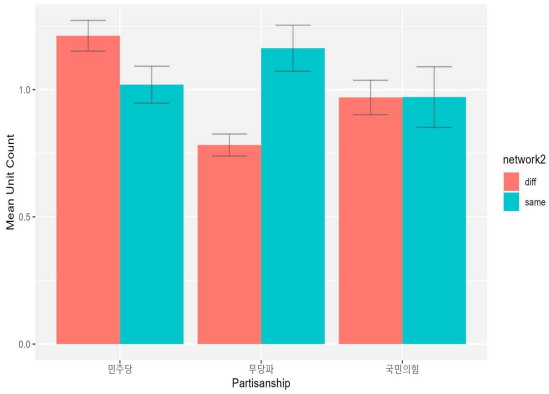
(2) 외교 지문의 경우 정보 길이 변수를 제외한 나머지 세 변수 - 정보 단위, 관련성, 응답 여부 - 에 있어 무당파에 비해 민주당, 국민의힘 지지자가 더 큰 값을 보임. 즉, 이들은 “정치적 정보 전달 과정에 높은 참여도”를 가지고 “더 많은 정보 단위의,” “더 관련성이 높은” 정보를 전달하는 경향이 있음

(3) 민주당과 국민의힘 지지자 사이에는 주목할 만한 차이는 발견되지 않음

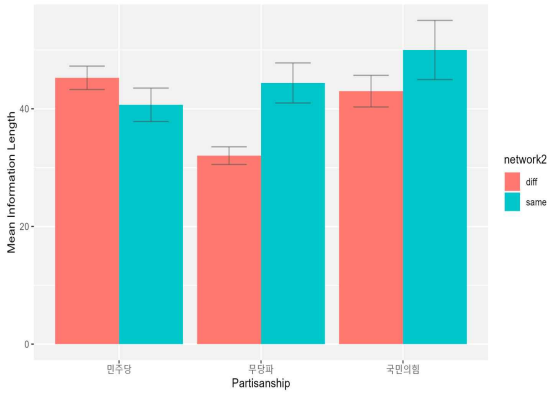
(4) 지문별 차이가 있으나, 대체로 당파성을 지닌 사람의 경우 무당파에 비해 자신이 접한 정치 관련 이슈를 타인에게 전달하려는 경향이 있으며, 또 이들이 전달하는 정보 메시지가 더 많은 수의 정보 단위를 포함하고 있는 경향이 있음을 확인함

다) 질문 2. 당파성과 정보 전달 네트워크에 따른 정치적 정보 전달 양상의 차이가 있는가?

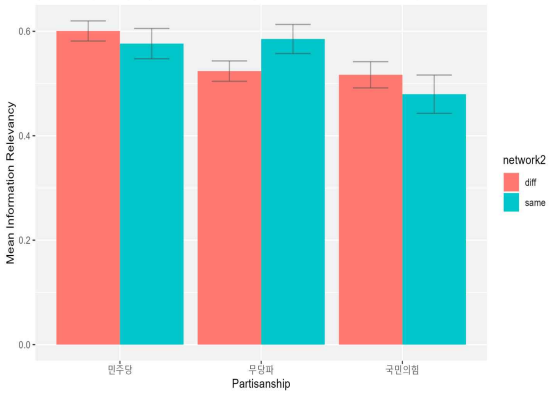
(ECON Vignette) Impact of Partisanship x Network on Delivered Information Unit Count



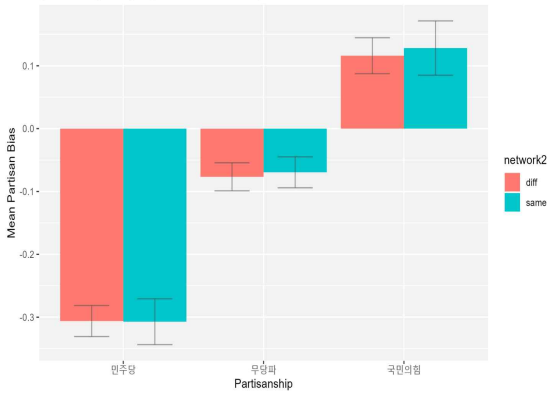
(ECON Vignette) Impact of Partisanship x Network on Delivered Information Length



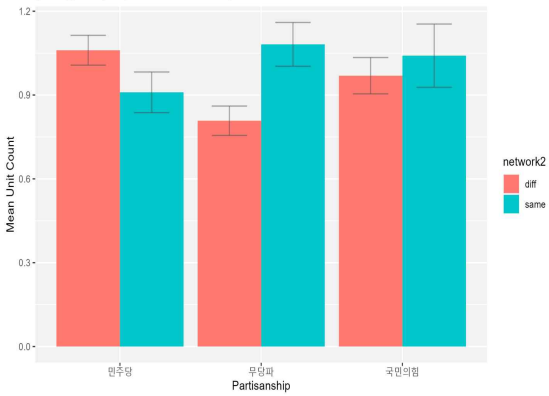
(ECON Vignette) Impact of Partisanship x Network on the Relevancy of Delivered Information



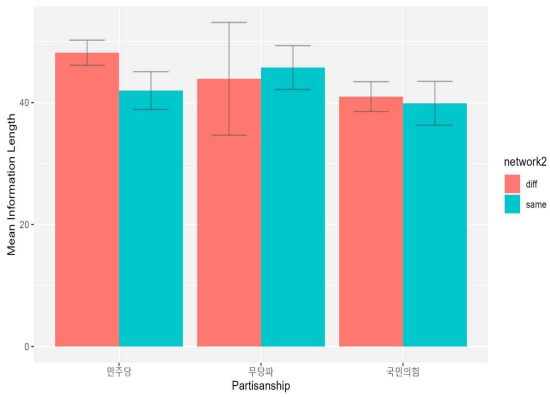
(ECON Vignette) Impact of Partisanship x Network on Partisan Bias



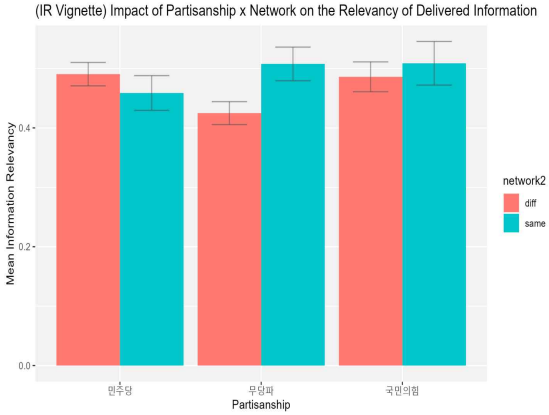
(IR Vignette) Impact of Partisanship x Network on Delivered Information Unit Count



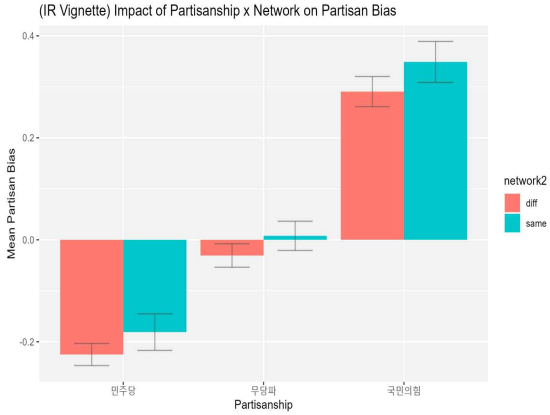
(IR Vignette) Impact of Partisanship x Network on Delivered Information Length



(IR Vignette) Impact of Partisanship x Network on the Relevancy of Delivered Information



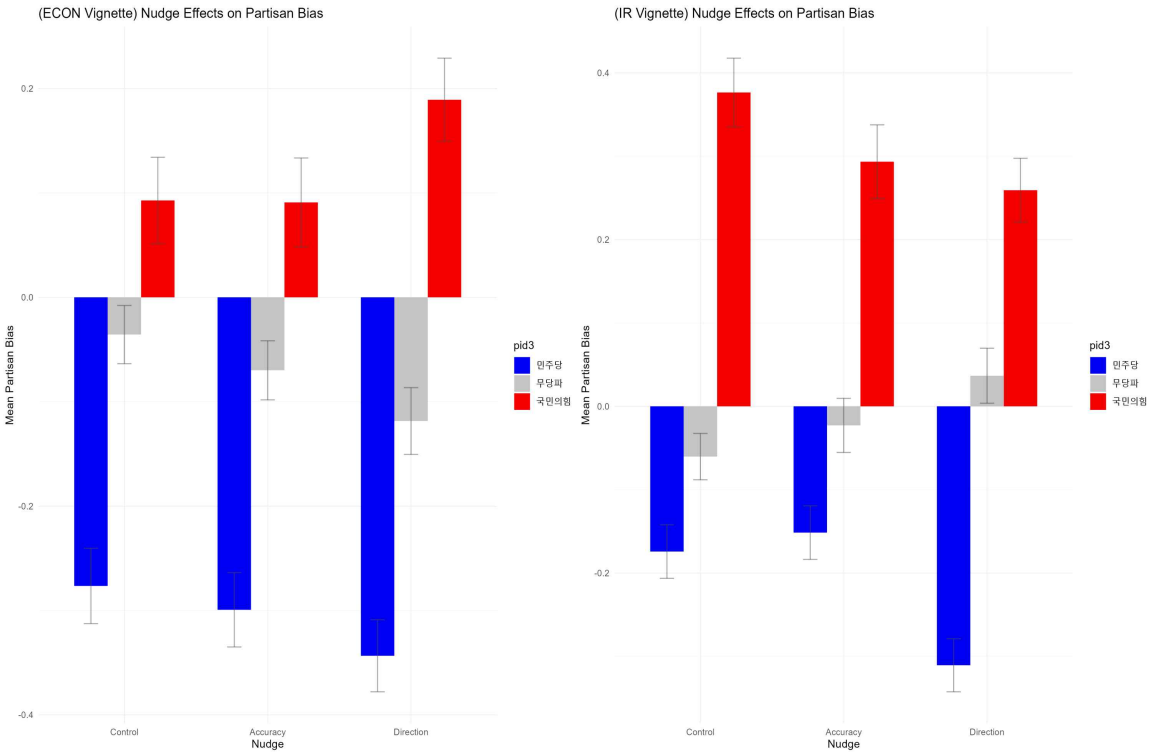
(IR Vignette) Impact of Partisanship x Network on Partisan Bias



(1) 경제 지문의 경우, 무당파는 자신과 같은 무당파 사람에게 정보를 전달할 때 “더 많은 정보 단위의,” “더 긴,” “더 관련성이 높은” 정보를 전달하는 경향이 확인됨. 반면, 민주당과 국민의힘 지지자의 경우 대체로 그와 다른 경향성이 있음을 확인. 특히 민주당 지지자에서 그와 상반된 경향성이 두드러지게 발견되는데, 이들은 오히려 그들과 상반된 정치적 성향을 가진 사람에게 정보를 전달할 때, “더 많은 정보 단위의,” “더 긴,” “더 관련성이 높은” 정보를 전달함. Partisan Bias에 있어선 세 당파성 그룹 모두 네트워크에 따른 큰 차이를 보이지 않음

(2) 외교 지문에서는 Partisan Bias 변수를 제외하고 경제 지문과 대체로 비슷한 결과를 보임. 무당파는 자신과 같은 무당파 사람에게 정보를 전달할 때 “더 많은 정보 단위의,” “더 긴,” “더 관련성이 높은” 정보를 전달하는 경향이 확인됨. 반면, 민주당 지지자는 그와 상반된 경향, 즉 오히려 자신과 다른 정치적 성향을 가진 사람에게 정보를 전달할 때에 “더 많은 정보 단위의,” “더 긴,” “더 관련성이 높은” 정보를 전달하는 경향이 있음. 한편, Partisan Bias에 있어서 민주당 지지자는 그들과 다른 당파성을 지닌 사람에게 정치 관련 메시지를 전할 때에 “당파적 색이 짙은” 정보를 전달하려는 경향이 있지만, 국민의힘 지지자는 그와 반대로 오히려 당파성이 같은 경우에 “당파적 색이 짙은” 정보를 전달하려는 경향이 있음을 발견

라) 질문 3. 넛지가 정보 전달 양상에 유의미한 변화를 가져오는가? 당파성과 네트워크에 따른 정보전달 양상을 약화, 혹은 강화시키는가?



(1) Accuracy Nudge, Directional Nudge 각각은 정보 전달 양상에 개별적으로 유의미한 효과를 가져오지 못함. 하지만, 당파성에 따라 넛지가 전달된 정보의 Partisan Bias에 미치는 영향을 주목할 만함. 위의 표는 그 결과를 정리한 것임

(2) 특히, 경제 지문에서는 본 연구의 예상과 부합하는 Directional Nudge의 효과를, 외교 지문에서는 Accuracy Nudge의 효과를 확인할 수 있었음

(3) 경제 지문에서는 세 당파성 그룹 모두에서 Directional Nudge가 전달된 정보의 Partisan Bias 정도를 강화함을 발견함. 한편, Accuracy Nudge가 Partisan Bias를 줄여줄 것이라는 본 연구의 예상과 달리 민주당과 무당파 그룹에서 오히려 Partisan Bias를 강화함을 확인

(4) 외교 지문에서는 세 당파성 그룹 모두에서 Accuracy Nudge가 전달된 정보의 Partisan Bias 정도를 완화함을 확인. 한편, Directional Nudge는 민주당 그룹에서만 본 연구의 예상에 부합하는 Partisan Bias 강화 효과를 확인

(5) 또 다른 흥미로운 지점은 외교 지문과 경제 지문의 결과가 다르다는 것으로, 정치 이슈별 차이를 이해할 필요가 있음

마) 파일럿 실험 결과, 당파성과 네트워크에 따라 상이한 정치적 정보 전달 양상이 있음을 확인함. 또한 제한적이긴 하나, Accuracy Nudge를 통해 당파적으로 편향된 메시지 전달을 완화할 수 있다는 것을 발견함. 향후 본 실험에서 실험적 처치를 강화할 방안과 설문 문항의 구성과 배치, 정치 이슈별 차이를 밝힐 수 있는 방안 등을 고려해 수정·보완하는 방향을 논의 중임

2) 관련분야 기여도

가) 본 연구는 정치적 정보 전달 양상을 개인적 수준에서 중요하게 작동하는 당파성과, 사회적 수준에서 고려되는 네트워크의 효과를 복합적으로 살펴볼 수 있는 기반을 마련함

나) 또한 Rand 등의 실험 연구 그룹에서 정보 전달 양상을 이해하기 위해 많이 사용하는 실험 처치인 넛지를 Accuracy Nudge, Directional Nudge 등으로 다변화하여 활용

다) 넛지 효과를 기존 연구에서 많이 살펴본 온라인 세팅이 아닌, 일상생활에서 실제 가까운 지인에게 전달하는 일반적인 의사소통 세팅에서 확인함

라) 더불어 Carlson (2018, 2019) 의 연구를 바탕으로 정치 커뮤니케이션 연구에서 아직 구체적으로 다뤄지지 않은 부분인 어떤 종류의 정보가 소실되고 살아남는지, 수신자와 발신자 사이의 관계가 어떤 영향을 미치는지 등에 대한 답을 제시함으로써 기존 연구의 공백을 채워넣을 수 있을 것으로 기대

마) 궁극적으로, 본 연구는 당파성과 네트워크에 따른 정치적 정보 전달 양상을 추적 및 분석하고 그에 대한 넛지 효과를 확인함으로써, 정보 전달 과정에서 나타나는 양극화 현상을 줄일 수 있는 방안에 대한 함의 제공