

Abstract

In this research, we examine polarization as a form of rapid social change resulting from the interplay between small group processes and perceptions of society at large. Specifically, we investigate how a negative (or hostile) norm regarding minorities at the societal level can fuel polarization between majority subgroups at the local level. By employing a novel analytic approach that uses variances to capture polarization processes, we were able to study non-linear societal change. In three studies among high school and university students ($N = 347$), we manipulated the societal norm about a minority category (positive vs. negative). Subsequently, participants read about a minority member's ambiguous behavior and evaluated this target. All studies used a similar paradigm, but they varied in whether or not participants discussed the ambiguous behavior within local groups. Results showed that the societal norm affected perceptions of the minority member's behavior when people discussed this behavior in a local group but not when they reflected on it individually. Specifically, group discussions led to polarization between local groups within a broader social category, but only in the context of a negative societal norm. This effect was predicted by the a priori perception of the local group norm. Results are discussed in terms of the integration of society- and group-level processes when studying the development of intergroup attitudes, and practical implications for the coarsening climate of the societal debate about current societal issues.

Keywords: group polarization, pro-social norms, hostile norms, social interaction, multi-level integration, intergroup conflict, group processes, rapid social change

The Polarizing Effects of Group Discussion in a Negative Normative Context:
Integrating Societal, Group, and Individual Level Factors

In many Western societies, we witness a coarsening in public discourse about Muslim minorities. Despite this prevailing negative majority norm regarding immigrants, political views are often characterized by fragmentation and polarization rather than mere assimilation to the negative societal sentiment (Alba & Foner, 2017). Indeed, both extreme right-wing and extreme left-wing parties have gained support over the last decades (e.g., the Dutch Socialist Party, the German Party of Democratic Socialism; Albertazzi & McDonnell, 2008; Dunlap, McCright, & Yarosh, 2016; for an overview, see March & Mudde, 2004). Thus, there is an apparent discrepancy between a strong negative norm at the societal level, and the emergence of small groups at the local level that hold beliefs either in line with *or* against this norm (i.e., polarization).

In this paper we propose that a negative majority norm at the societal level can be considered a risk factor for catalyzing political polarization. We argue that this phenomenon, in which society rapidly divides into extreme opinion camps, is explained by communication within ‘local groups’¹: Whether majority members go along with, or take a stance against, a categorical majority norm often depends on conversations with similar others. On their own, individuals may doubt the validity of majority norms. By turning to others, they can develop a locally shared view, thereby finding and validating their position (cf. Festinger, 1950). Indeed, local group interactions, especially with fellow majority members, shape group norms and outgroup attitudes (Greijsdanus, Postmes, Gordijn, & Van Zomeren, 2015; Hewstone & Jaspars, 1982; Smith & Postmes, 2009, 2011; Smith, Thomas, & McGarty, 2015). We build on this research by investigating the impact of the valence of the societal norm (positive or negative) on this. By integrating societal and group-level processes, with

¹ We use the term local groups to refer to small groups of people that directly communicate with each other, in the current research consisting of only majority group members.

individual perceptions of these higher-level variables, we aim for a better insight in the factors underlying societal polarization.

Majority category norms

In line with social identity theory (SIT; Tajfel, 1978; Tajfel & Turner, 1986), studies have demonstrated that when an identity is salient, norms attached to this identity guide members' behaviors and attitudes in many areas (e.g., Goldstein, Cialdini, & Griskevicius, 2008; Jetten, Spears, & Manstead, 1996). When intergroup conflict is salient, isolated individuals will hence engage in behavior that favors the ingroup (e.g., Tajfel, 1978; Tajfel, Billig, Bundy, & Flament, 1971). Indeed, when opportunities to engage in interaction are minimized, as in many social psychological research paradigms and some (authoritarian) societies, the influence of categorical or societal norms will be large (Galam & Moscovici, 1991).² The reasoning behind this is that when individuals become depersonalized, they do not interact as individuals but as members of the same social category (Turner, 1985).

We argue that this focus on ingroup assimilation during intergroup conflict is valuable yet does not paint a complete picture. In real-life intergroup conflicts, local groups within the same social category often develop varying norms regarding, for instance, outgroups. Whereas some local groups develop norms promoting intergroup hostility, others develop norms for harmony (Thomas, Smith, Postmes, & McGarty, 2010). Thus, it seems crucial to distinguish between high-level collectives such as a society or broad social category (e.g., the majority of native inhabitants)³ and the lived experience of local groups with fellow majority

² Galam and Moscovici use this logic to explain why authoritarian regimes endorse social mutism, and why they encourage citizens to become informers on each other – and thus isolated from one another for fear that their neighbors might denounce them.

³ A social category includes members of a society that share a relevant characteristics for the issue at hand, in the case of immigration issues this is likely to be nationality. Hence, the majority category consists of native inhabitants of the host country (here Native-Dutch), whereas the minority category consists of immigrants (here specifically Moroccan-Dutch immigrants). In the current research we focus on a subset of these categories, namely students at universities or high schools.

members (cf. McGarty, Bliuc, Thomas, & Bongiorno, 2009, for a similar distinction between social categories and psychological groups).

To explain polarization between groups *within* a social category, attention for *intragroup* processes is pivotal. Indeed, conversations with the people around us are extremely important for the development of shared social identities and their according norms and perceptions (Greijdanus et al., 2015; Haslam et al., 1997; Koudenburg, Postmes, & Gordijn, 2013a, 2013b; for reviews see Koudenburg, Postmes, & Gordijn, 2017; Postmes, Spears, Lee, & Novak, 2005; Thompson & Fine, 1999).

Group Influence: Consensualization and Polarization

The influence of a group on its members becomes evident through two complementary processes: Consensualization and polarization (cf. Jans, Leach, Garcia, & Postmes, 2015). On the one hand, people consensualize their views when making judgments in a group, including judgements about other groups (Sherif, 1936). Research shows that group members consensualize on their *own* group's stereotypes and norms (Bliuc et al., 2007; McGarty et al., 2009; Meeussen, Delvaux, & Phalet, 2014), but also on stereotypes about, and hostility towards, *other* groups (Greijdanus et al., 2015; Haslam et al., 1997; Smith & Postmes, 2009). On the other hand, group discussions can also cause collective shifts toward more extreme positions (Moscovici & Zavalloni, 1969). Specifically, during group interaction something emerges that is not merely reducible to the aggregation of members' individual properties: Groups produce emergent normative tendencies (McGarty, Turner, Hogg, David, & Wetherell, 1992).

This second form of group influence, polarization, implies that through intragroup interaction local groups can diverge from other local groups within the broader social category. Indeed, interpretations of what the social category stands for may diverge between different local groups within that social category. For instance, feminists may have a different

view on the role of women than women in more conservative local groups (who likely endorse more traditional gender roles). Similarly, local majority groups' views of ethnic minority groups may diverge: Some local groups adopt hostile norms, whereas other local groups develop more pro-social or egalitarian norms. As such, within a social category polarization may emerge between different local groups that evolve around diverging shared opinions (Bliuc et al., 2007; Bliuc et al., 2015). We build on this research by proposing that polarization between local groups of majority members depends on the valence of the broader majority category norm.

We argue that both majority norms and local group communications are sources of social influence – and that these processes interact (cf. Postmes et al., 2005). That is, in most naturally occurring situations, majority members roughly know how the majority category in society feels about a certain issue, but they are also able to acquire, maintain, or transform these norms through interacting with each other.

Local Group Norm Formation under a Negative Majority Norm

Our key hypothesis is that the polarizing effects of local group discussions are especially likely to occur under a negative majority norm. This is because we believe a negative norm urges people to take a stance in the debate. This fits research showing that threat, or negative mood more generally, functions as an alarm signal stimulating action. Contrarily, positive mood signals that “everything is fine”, and no action or explicit stance-taking is necessary. Indeed, negative mood leads to more detailed, in-depth information processing (Bless, Bohner, Schwarz, & Strack, 1990; Forgas, 2013) and stimulates more extensive information exchange (Van Knippenberg, Kooij-De Bode, & Van Ginkel, 2010). Therefore, a negative (vs. positive) societal norm is more likely to fuel discussion. By increasing the need for a shared definition of who ‘we’ are and what ‘we’ stand for as a group

(McGarty et al., 2009), a negative majority norm should in particular facilitate polarization between local groups of majority members.

A second reason why a negative majority norm should elicit more group influence can be derived from research on the positive-negative asymmetry effect. This research shows that although group members frequently express ingroup favoritism, they are much less eager to engage in outgroup derogation (Mummendey & Otten, 2011; cf. Aaldering, Ten Velden, Van Kleef, & De Dreu, 2018). Because negative attitudes are, in essence, less morally valued than positive attitudes (Katz & Hass, 1988; Killen & Smetana, 2006; Rutland, Killen, & Abrams, 2010) and morality is a basic psychological need (Prentice, Jayawickreme, Hawkins, Hartley, Furr, & Fleeson, 2018), the development of negative attitudes about a minority category may require more social validation. Supporting this suggestion, research demonstrates that people can develop negative views about the outgroup when they engage in interaction that consensualizes around these negative stereotypes (Haslam et al., 1997, 1998; Thomas et al., 2010).

However, group discussions can also evolve towards more pro-social views and behaviors (Reicher, Cassidy, Wolpert, Hopkins, & Levine, 2006; Thomas et al., 2010). By discussing grievances, groups may consensualize around views that foster equality or pro-social behavior and this may become central to how they define themselves (Bliuc et al., 2007, 2015; Thomas et al., 2010). The alarm signal posed by the negative majority norm may thus fuel the need for social validation of what the group stands for and thereby promote the development of either negative (hostile) or positive (pro-social) views through group discussion. Individuals may thus be more strongly influenced by their local group than they would have been under a positive majority norm.

The Present Research

We propose that understanding the rapid social change process of polarization requires more advanced research paradigms, moving beyond studying societal-level or individual-level factors in isolation. In three studies, we examined the simultaneous influences of majority norms and local majority group discussions. We hypothesized that attitudes would polarize more under a negative (vs. positive) majority norm (Hypothesis 1). Secondly, we hypothesized a key role for perceptions of the local group norm in predicting the direction of polarization in a specific local group. We reasoned that majority members might tailor their contributions to the discussion to how they expect their local majority group to view the minority category. The result may be a collective shift in interpretations and according attitudes, towards the perceived local group norm (Hypothesis 2). Importantly, we compared the consequences of discussion with the consequences of individual rumination and predicted a unique catalyzing role of local group discussion in attitude polarization (Hypothesis 3). The current research tested these hypotheses with a relatively novel analytic approach, using variance in individual- and multi-level designs as a statistical proxy for polarization (see also Jans et al., 2015; Kuppens & Yzerbyt, 2014).

General Procedure

We tested our hypotheses among students at a Dutch university (Study 1) and high schools (Studies 2, 3). Each study followed the same procedure: After manipulating a majority norm, participants (i.e., majority members) reflected on an ambiguous intergroup scenario and completed a questionnaire. There was however one important difference between the studies: In Study 1, participants received the norm information and reflected on the scenario individually, whereas in Study 2 participants discussed the scenario in local groups. Comparing the finding of Study 1 with those of Study 2 allows us to assess whether a majority norm influences polarization processes directly (i.e., by influencing individuals in Study 1), or indirectly (i.e., by shaping group discussions, as in Study 2). In Study 2, one

group member (the teller) received the norm manipulation, read the scenario, and then discussed this within their local group. The other members were assigned the role of listener, and did not receive the norm information or scenario. They were encouraged to react and ask questions about the tellers' experience. This set-up allowed us to isolate the tellers' perceptions of local group norms as a potential predictor of polarization between local groups, without inference of potential listener motives or behaviors that might be instigated by a negative majority norm.⁴ Study 3 combined these individual (Study 1) and group (Study 2) settings, and served to replicate the findings of Study 1 and 2 in one design.

Majority norm manipulation. Participants were randomly allocated to either a positive or a negative majority norm condition. The norm concerned native-Dutch students' (i.e., university students or high school students) perceptions of Moroccan-Dutch students. For the native-Dutch, Moroccan-Dutch people are currently a highly relevant minority, of which especially the young men are stereotyped as aggressive, unadjusted, and criminal (e.g., Greijdanus, Postmes, Gordijn, & Van Zomeren, 2014). The *negative* [*positive*] majority norm condition presented a bogus newspaper article stating that "large-scale research shows that native-Dutch students *prefer not to* [*like to*] work together with Moroccan-Dutch students, because they feel that they are *unreliable* [*accommodating*], and therefore native-Dutch students feel that they *always need to check* [*can always receive help from*] Moroccan-Dutch students." The manipulation further stated that, "in general, native-Dutch students are *negative* [*positive*] about cooperation with Moroccan-Dutch students."

Ambiguous scenario. Subsequently, participants reflected on an ambiguous scenario involving a Moroccan-Dutch student. The scenario described a collaboration between the

⁴ We regard this a conservative test of the polarization hypothesis, as we consider it likely that listeners in the group are also motivated to develop a firm position on the issue when confronted with a negative societal norm. They could do so, for instance, by confirming the information given by the teller (which may be either positive or negative about the minority target), rather than providing potential alternative motives for minority target's behavior. It would be likely that such behaviors may amplify the polarization processes reported in this paper.

participant (“you”; participants imagined being the protagonist), Mirjam, Ernst (typical Dutch names), and Abdul (typical Moroccan name; minority target). In the scenario, the protagonist forgets his/her mobile phone at the faculty. Upon return at the faculty, they see Abdul holding their mobile phone in his hand. The scenario is ambiguous on Abdul’s intentions i.e., did he try to make sure the protagonist would receive his/her phone back, or to steal it?.

Next, participants reflected on this scenario on their own (Study 1, Study 3 individual conditions), or within a local group of Dutch students (Study 2, Study 3 group conditions). Afterwards, all participants wrote about ‘their experience’. Tellers and people who reflected on the scenario alone took the protagonist perspective, whereas listeners described the scenario as if it had occurred to the teller. All participants were asked to report both factual descriptions and their thoughts and feelings. We expected polarization only for participants who reflected on the scenario in groups.

Dependent variables (DVs) and analytic strategy. We measured effects on individual participants’ attitudes regarding the outgroup target both explicitly using Likert-scales, and less intrusively, by content-coding their written reports of the scenario. In Studies 2 and 3, we also assessed participants’ attitudes towards Moroccan-Dutch students in general. We assessed polarization on each scale by examining whether differences between participants, (Study 1, Study 3 individual conditions), or between local groups (Study 2, Study 3 group conditions) were larger in the negative than in the positive majority norm condition.

Study 1

Study 1 served to explore whether a negative (vs. positive) majority norm, in itself, would foster polarization. We reasoned that if polarization results from local group discussions, exposure to a negative majority norm without local group discussion should not increase polarization (i.e., variance between individuals).

Method

Participants and design. Fifty undergraduates (43 women, seven men; $M_{\text{age}} = 19.96$, $SD = 1.93$; 48 native-Dutch, two German) participated for partial course credit or a small monetary reward. As part of a different study (Koudenburg et al., 2013a, Study 2) participants had a 5-minute dyadic conversation via headsets.⁵ Afterwards, they were introduced to the current study, presented as an unrelated study.

Study 1 used a between-subjects manipulation of the majority norm regarding Moroccan-Dutch students (positive: $n = 25$ vs. negative: $n = 25$). After receiving the majority norm manipulation and the ambiguous scenario, participants wrote about ‘their experience’ to the person they spoke with in the previous study.

Measures. Single-item measures assessed participants’ trust in the target: “I trust Abdul” (1 = strongly disagree, 7 = strongly agree), and checked the majority norm manipulation: “Dutch psychology students in general experience interaction with Moroccan-Dutch students as negative (1) - positive (7)”. Participants reported their demographics and frequency of intergroup contact.^{6, 7}

Two independent coders rated participants’ reported experiences on *trust* in the target (1 = strong distrust, 5 = strong trust) and the participant’s *negative affect* about the situation (1 = positive, 3 = indifferent, 5 = negative).⁸ The codings had sufficient interrater reliability,

⁵ The conditions of this study did not influence the DVs of the current study, nor did controlling for condition change the pattern of results. In reporting the present results, we therefore do not include or control for these conditions.

⁶ For exploratory reasons, Study 1 also included other measures. Because none of the variables were administered in the following studies, we describe only the most important variables in the present article. Full information is available from the first author on request.

⁷ Participants’ age, gender, and intergroup contact frequency did not influence the DVs, and the results of the norm manipulation did not change when we added these variables as covariates to the analyses. We report the results without these covariates.

⁸ Initially, we used additional codings. We did not include these in the final analyses because they had very little variation in their scores or had low interrater reliabilities.

and we calculated the average between coders as indicators of trust ($ICC2^9 = .72$) and negative affect ($ICC2 = .86$).

Results

Majority norm. Table 1 displays means and standard deviations. The manipulation was effective; participants in the positive condition reported that native-Dutch students experienced intergroup interaction more positively than participants in the negative condition, $F(1,48) = 8.49, p = .005, \eta_p^2 = .15$. We found no effects on self-reported trust nor on the codings of trust or negative affect, $F_s < 1.27, p_s > .264$.¹⁰

Polarization. To investigate between-participant polarization, we also tested whether the majority norm manipulation influenced variances in the DVs. We used the Bartlett-Box Test to test heterogeneity of variances across all three DVs. As anticipated, this test indicated no significant between-condition differences, Box's $M = 5.47, p = .532$. Accordingly, Levene's test of equality of variances per variable revealed no significant between-condition differences on negative affect, $F(1,48) = 1.64, p = .207$, or self-reported and coded trust, $F_s < .256, p_s > .615$.

Discussion

We found no effect of a negative (vs. positive) majority norm manipulation on outgroup attitudes when majority participants reflected *individually* on an ambiguous scenario involving a minority member. Importantly, we also found no evidence for differences in variances. This suggests that a negative majority norm in itself does not lead to more polarized interpretations of an ambiguous scenario than a positive majority norm. Study 2 aimed to examine whether the majority norm affects majority members' interpretation of an

⁹ Because both coders rated all cases, we used the ICC2 for the average measures to estimate interrater reliability.

¹⁰ There were no condition effects on participants' perceived overlap between themselves and the target, or between the target and the Moroccan-Dutch student group, suggesting that the target was equally likely to be categorized as an outgroup member in both conditions.

ambiguous intergroup scenario differently when they discuss this scenario within a local ingroup.

Study 2

The set-up was similar to Study 1, but in Study 2 students discussed the scenario in a local group instead of reflecting on it individually. Afterwards, we assessed participants' attitudes regarding the minority target using both self-reports and codings as in Study 1 but using a more elaborate and validated scale (Leach, Ellemers, & Barreto, 2007). Additionally, we assessed whether effects would extend beyond evaluations of the target outgroup member to attitudes regarding Moroccan-Dutch students in general.

We hypothesized that local group discussion would catalyze polarization between local groups under a negative (vs. positive) majority norm (Hypothesis 1). An important additional question was what steers the direction of this polarization. We proposed that members' a priori perception of the local group norm plays an important role. That is, local group members who entered the discussion anticipating a positive local group norm would likely reinforce this norm through group discussion (and vice versa for a negatively perceived norm). We hypothesized that local group norm perceptions by the tellers (i.e., the students who introduced the scenario in their local groups) would be particularly influential under a negative majority norm. Indeed, tellers might tailor their descriptions of what happened to how they expect their local group to view the minority category. For instance, they should be more likely to share information or express opinions that confirm rather than contradict this perceived norm (Kashima, 2000; Lyons & Kashima, 2003). And a negative majority norm would boost this effect. The result may be a collective shift in interpretations of the scenario and according attitudes, towards the teller's perception of the local group norm – especially under a negative majority norm (i.e., interaction between teller's perception of the local group norm and the majority norm, Hypothesis 2).

Method

Participants and design. As part of an introduction to psychology workshop, 159 high school students (128 women, 27 men; $M_{\text{age}} = 15.52$, $SD = 1.07^{11}$) participated.¹² Most participants had a native-Dutch ($n = 146$), a different Western, or a dual nationality ($n = 5$). The data of three participants with non-Western backgrounds (Middle Eastern, Kosovar/Albanian) and their local group members ($n = 7$) were excluded from analyses.

The study used a between-subjects manipulation of majority norm (positive: $n = 77$, vs. negative: $n = 68$). Students entered in classes of 15-25 and were randomly separated into local groups of three or four students. We assigned one student in each local group to be the *teller* and the others to be *listeners*. Only group tellers received the majority norm manipulation, which was similar to Study 1, but now concerned native-Dutch *high school* student norms regarding Moroccan-Dutch *high school* students.

After the norm manipulation, participants who were assigned the role of teller read the scenario, but instead of writing about it, they were asked to tell ‘their experience’ to the listeners in their local group. Listeners were encouraged to react and ask questions about the tellers’ experience. Finally, all participants completed a set of questionnaires.

Measures. We assessed the same DVs for all participants. First, participants rated the target “Abdul” in terms of warmth, competence, and morality (Leach et al., 2007; e.g. “I think that Abdul is friendly”, 1 = not at all applicable, 7 = very applicable). The nine items were combined into one target evaluation scale ($\alpha = .92$). Perceived majority and local group norms and personal attitudes regarding Moroccan-Dutch students were assessed with three bi-polar items each: [native-Dutch students]/[my group members]/[I] view Moroccan-Dutch

¹¹ Four students did not indicate their gender, age or background.

¹² In Studies 2 and 3, sample size was determined by the number of students enrolled in the introductory workshop. We conducted a post-hoc power analysis indicating that our sample sizes afforded 98% power in Study 2 and 87% power in Study 3 to detect between-condition differences in variance of a medium effect size ($w = .30$).

students as: 1 = mostly negative – 7 = mostly positive, 1 = unreliable – 7 = reliable, 1 = unhelpful – 7 = helpful. Combining the three items gave reliable scales for perceived native-Dutch student norms ($\alpha = .81$, i.e., majority norm manipulation check), perceived local group norms ($\alpha = .90$), and attitudes ($\alpha = .90$).¹³ Finally, all participants described the scenario on paper and reported their demographics and intergroup contact frequency.

Four independent coders coded the written experiences on the same indicators as in Study 1. One coder's trust ratings had low correlations with the other three coders. Because removing this coder increased reliability ($ICC2_{\text{trust}} = .87$, $ICC2_{\text{negative affect}} = .79$), we present results based on the remaining three coders.

Results

We used multilevel analyses with participants (Level 1) nested in local groups (Level 2) to conduct two tests: First, we examined whether polarization between local groups would be increased in the negative (vs. positive) majority norm condition (i.e., comparing Level 2 variances between conditions). Second, we attempted to explain polarization between local groups by the perceived local group norm (i.e., examining interaction effects at Level 2). In the second analysis, we also tested for main effects of the majority norm manipulation and local group norm perceptions.

Manipulation check. The majority norm manipulation succeeded; tellers in the positive majority norm condition perceived native-Dutch students to view Moroccan-Dutch students more positively ($M = 3.97$, 95%CI [3.60;4.35]) than tellers in the negative majority norm condition ($M = 3.29$, 95%CI [2.86;3.71]), *mean difference* = 0.68 ($SE = 0.27$), $t(44) = 2.52$, $p = .015$.

Polarization between local groups. Comparing the ICCs between conditions, the local group explained on average 40% of the variance in all DVs in the negative majority

¹³ In Studies 2 and 3, participants also completed pre- and post-measures of their identification with the local group in the study, as well as with native-Dutch students.

norm condition, but only 19% of the variance in the positive majority norm condition.¹⁴ ICCs higher than zero can reflect decreased individual-level variability within groups (consensualization) or increased differences between groups (polarization). In the current paper, we chose not to focus on the ICC, but tease the theoretically and empirically distinct processes of consensualization and polarization apart. Because we are primarily concerned with polarization we focus on comparing *between*-group variances (τ) per condition (see Table 2; cf. Jans et al., 2015).

To this end, we estimated for each dependent variable a Model 1 with only the intercept and a main effect for condition, and a random intercept for local group. Model 2, which added a parameter to explain the difference in local group-level variance between the two conditions. This additional parameter models the covariance between the level-2 intercept (the variance at the local group level), and the majority norm condition. It can be interpreted as the difference in local group-level variance (i.e., τ) between the two majority norm conditions. Wald Z tests indicated that for all dependent variables, more variance was explained by the local group in the negative norm condition than in the positive norm condition: Target evaluation, *Wald Z* = -2.95, *p* = .008, trust, *Wald Z* = -2.85, *p* = .004, negative affect, *Wald Z* = -2.86, *p* = .004, and outgroup attitudes, *Wald Z* = -3.17, *p* = .002. This supports Hypothesis 1: More polarization occurred under a negative majority norm. We also examined whether adding this parameter improved the model fit by comparing the 2-restricted log-likelihoods of Model 1 and Model 2 with a $\Delta\chi^2$ -test.¹⁵ Results of these tests can be found in Table 2.

¹⁴ Intraclass correlation is calculated by dividing the group level variance by the total variance (individual and group level), using formula: $\tau^2 / (\tau^2 + \sigma^2)$

¹⁵ Because all chi-square difference tests were used to test variance components, we divided the p-values by 2, as recommended by Berkhof and Snijders (2001).

Teller's perception of local group norm. To explain the polarization effects, we subsequently examined whether tellers' perceptions of the local group norm would be more likely to shape group discussion, and thus influence the DVs, under a negative than under a positive majority norm (interaction effect, Hypothesis 2). Each DV was regressed onto *majority norm* (negative (0) vs. positive (1)), standardized values of *teller's perception of the local group norm*, and their interaction. A main effect for the majority norm manipulation revealed, paradoxically, higher trust in the negative (vs. positive) norm condition, $\gamma = -0.20(0.09)$, $t(41) = -2.21$, $p = .033$. Other majority norm main effects were not statistically significant, $|t|s < 1.08$, $ps > .284$. We found main effects of teller's perception of the local group norm on target evaluation, $\gamma = 0.40(0.12)$, $t(41) = 3.29$, $p = .002$, negative affect, $\gamma = -0.12(0.06)$, $t(41) = -2.11$, $p = .041$, and outgroup attitude, $\gamma = 0.36(0.10)$, $t(41) = 3.42$, $p = .001$. All effects revealed assimilation to the perceived local group norm. The effect on trust was in the same direction but not statistically significant, $\gamma = 0.16(0.10)$, $t(41) = 1.64$, $p = .109$.

Importantly, these main effects were qualified by interactions on trust, $\gamma = -0.13(0.10)$, $t(41) = -3.09$, $p = .004$, and negative affect, $\gamma = 0.22(0.06)$, $t(41) = 3.83$, $p = .001$, suggesting that assimilation to the perceived local group norm was stronger under a negative (vs. positive) majority norm (see Figures 1a-b). We found similar yet non-significant interaction patterns on target evaluation, $\gamma = -0.15(0.12)$, $p = .214$, and outgroup attitude, $\gamma = -0.09(0.10)$, $p = .374$. The pattern of interactions is in line with Hypothesis 2. That is, under a negative majority norm, tellers were more likely to tailor their message to what they expected to be the local group norm, resulting in more trust in the minority target and less negative affect among local group members when tellers expected the local group norm to be positive than when they expected it to be negative. In the positive majority norm condition, tellers' perceptions of the local group norm were less likely to shape group discussion because here,

outcomes of the discussion in terms of trust and negative affect did not align with the tellers' perceived local group norm.

Discussion

The results of Study 2 revealed that local majority group discussion about a minority target's ambiguous behavior instigated collective attitude shifts: The local group either collectively distrusted the minority target, or collectively shifted away from the negative majority norm. Thus, discussion of such events increased polarization between local groups. Importantly, we found evidence for polarization especially when the majority norm regarding the minority was negative. A negative majority norm urges local majority groups to take a stance and is therefore more likely to cause rapid social change.

Additionally, Study 2 identified prior perceptions of local group norms as a predictor of the direction in which the local group will shift. Specifically, local group members' attitudes (i.e., target evaluation, negative affect, general attitude towards the outgroup) collectively shifted towards the teller's perception of the local group norm. That is, tellers' perceptions of local group norms acted as a self-fulfilling prophecy. The data furthermore provided some evidence that a negative majority norm boosted this effect. Indeed, for trust and negative affect, assimilation to the perceived local group norm was stronger in the negative than in the positive majority norm condition.

Study 3

Study 3 served two main purposes. First, we aimed to replicate Studies 1-2. Second, we tested whether the stronger polarization effects under a negative majority norm resulted from local group processes rather than individual reflection (Hypothesis 3). Although the combined findings of Studies 1 and 2 provided an initial answer to this question, a valid test of the hypotheses should include both individual and local group conditions in the same design. Therefore, Study 3 was designed similar to Study 2, but added two conditions in

which participants reflected individually on the scenario under a negative or positive majority norm.

Method

Participants and design. Participants were 138 high school students (118 women, 18 men; $M_{age} = 15.51$, $SD = 0.64$ ¹⁶) participating as part of an introduction to psychology workshop. Most participants had native-Dutch ($n = 122$) or different Western or dual nationalities ($n = 6$). Data of eight participants with non-Western backgrounds (Latin-American, Asian, African, Middle Eastern) and their local group members ($n = 10$) were removed before analyses.

The study had a 2(majority norm: positive vs. negative) x 3(role: teller vs. listener vs. individual) between-subjects design. Students entered the room in classes of 20-27. In the local group conditions, they were randomly separated into groups of two ($n = 9$), three ($n = 21$), or four students ($n = 1$). The negative norm conditions included 16 tellers, 27 listeners, and 18 individuals; the positive norm conditions included 15 tellers, 27 listeners, and 17 individuals. We followed the procedure of Study 2, except for a minor change: We assessed the perceived local group norm (only for *tellers*) and the perceived native-Dutch student (i.e., majority) norm (only for *tellers* and *individuals*) directly *after* the majority norm manipulation, before they read the scenario or started talking with their local group members. To ensure that all local group members would be ready for the group task around the same time, *listeners* simultaneously completed a filler task (neutral word search).

In the individual conditions, participants sat behind separate tables. The procedure for *individuals* was similar to that for tellers, with two exceptions: 1) individuals did not answer questions about local group norms because there was no group to refer to, and 2) rather than telling their experience of the scenario to their local group, they reflected on it individually.

¹⁶ Two students did not indicate their gender, age, or background.

Both local groups and individuals were given 5 minutes to discuss/reflect on the scenario.

Afterwards, all participants completed questionnaires.

Measures. We used the same measures as in Study 2 to assess native-Dutch student norms (majority norm manipulation check; $\alpha = .94$), perceived local group norms (only for tellers; $\alpha = .91$), target evaluation ($\alpha = .93$), and outgroup attitudes ($\alpha = .90$). Subsequently, all participants described the ambiguous scenario in writing and indicated their demographics and intergroup contact frequency.

Two independent coders coded all scenario reports. Their scores were averaged to create indicators of *trust* in the target (ICC2 = .90), and *negative affect* (ICC2 = .72).

Results

The data were analyzed in two parts, each focusing on a different comparison. Part 1 compared local groups (including both tellers and listeners) to replicate Study 2: We used the same multilevel approach to compare polarization between local groups in the negative vs. positive majority norm conditions (testing Hypothesis 1) and to examine the role of perceived local group norms in this process (testing Hypothesis 2). Part 2 tested whether local group discussions *catalyzed* polarization processes (Hypothesis 3). We compared attitude polarization among tellers with attitude polarization among those who reflected individually (i.e., the individuals). We exclude listeners in part 2, for two reasons. First, tellers and individuals were exposed to exactly the same information (i.e., majority norm manipulation and scenario; whereas listeners did not read these documents) and thus the only difference between tellers and individuals was that they discussed this scenario in a local group or reflected on it individually (i.e., the difference of interest). Second, comparing the tellers and individuals allowed us to compare equally sized groups, thereby excluding the possibility of differences in variance due to between-condition sample size differences.

Manipulation check. Because the majority norm manipulation was displayed only to tellers and individuals, the manipulation check was also conducted on only those participants. The intended effect on majority norms was revealed: Participants in the negative norm condition perceived the native-Dutch student norm regarding Moroccan-Dutch students to be more negative ($M = 3.15$, 95%CI [2.83; 3.48]) than those in the positive norm condition ($M = 5.32$, 95%CI [4.90; 5.73]), *mean difference* = 2.17 ($SE = 0.26$), $t(63) = 8.42$, $p < .001$. Because neither of the confidence intervals around the manipulation check means includes the midpoint of the scale (i.e., 4), we are confident to have manipulated a truly positive and a truly negative majority norm.

Part 1: Polarization between local groups. To examine polarization, we compared between-local-group variances per condition as in Study 2 (see Table 3). Supporting Hypothesis 1, Wald Z tests indicated that for all dependent variables, more variance was explained on the local group level in the negative majority norm condition than in the positive majority norm condition: Target evaluation, $Wald Z = -2.46$, $p = .014$, trust, $Wald Z = -2.01$, $p = .044$, negative affect, $Wald Z = -2.06$, $p = .040$, and outgroup attitudes, $Wald Z = -1.97$, $p = .049$.

Teller's perception of local group norm. We then examined whether the teller's perception of the local group norm could explain this polarization process (Hypothesis 2). Using multi-level modeling as in Study 2, we entered the Level-2 predictors (majority norm manipulation, teller's perceived local group norm, and their interaction) to predict DVs at Level 1 (individual local group members).

First, we found convincing evidence for a main effect of the teller's perceived local group norm on target evaluation, $\gamma = 0.72(0.22)$, $t(26) = 3.24$, $p = .003$, and outgroup attitudes, $\gamma = 0.52(0.16)$, $t(26) = 3.29$, $p = .003$. The main effects on trust, $\gamma = 0.36(0.19)$, $t(26) = 1.91$, $p = .068$, and negative affect, $\gamma = -0.17(0.12)$, $t(26) = -1.45$, $p = .159$ did not

reach statistical significance. All patterns consistently showed assimilation to the local group norm.

The majority norm manipulation affected trust, but as in Study 2 this effect was counterintuitive: Trust was higher under a negative (vs. positive) majority norm, $\gamma = -.65(0.27)$, $t(26) = -2.44$, $p = .022$.¹⁷ No other main effects were found, $|t|s < 1.00$, $ps > .430$.

Importantly, the main effect on outgroup attitude was qualified by a statistically significant interaction, $\gamma = -0.48(0.22)$, $t(26) = -2.14$, $p = .042$. The interaction effect on target evaluation did not reach statistical significance, $\gamma = -0.61(0.31)$, $t(26) = -1.94$, $p = .064$ (see Figures 2a-b). The pattern of interactions suggests that under a positive majority norm, the teller's perceived local group norms did not influence local group members' target evaluations or outgroup attitudes. However, under a negative majority norm, the tellers' perceptions of the local norm guided local group members' evaluations of the outgroup target and their attitudes towards the outgroup in general. No interactions on trust or negative affect were found, $|t|s < 1.00$, $ps > .558$.

Part 2: Comparing Local Groups and Individuals. We analyzed the differences in variances between the positive and negative majority norm conditions for tellers and individuals separately. We first conducted a one-sided Bartlett-Box test for equality of variances on the DVs assessing attitudes towards the target (trust, negative affect, and target evaluation).¹⁸ We subsequently used one-sided Levene's tests to test the equality of variances of each DV separately, including outgroup attitudes, to test for potential generalization of the target evaluations to evaluations of the entire outgroup (see Table 4).

¹⁷ The main effect of trust in Study 2 and 3 suggests that local group members shift their judgments away from the societal norm. However, because the main effect in both studies only appeared when we included the interaction term in the model, it should be interpreted with caution.

¹⁸ We used one-sided tests, because we had clear hypotheses for the direction of the effects in the negative norm condition (following guidelines by Snijders and Bosker, 2012). To afford comparability, tests in the positive norm condition were also conducted one-sided.

As expected, for tellers, comparing the variances in the positive versus negative norm conditions revealed a significant overall difference, Box's $M = 12.72$, $p = .040$. Accordingly, Levene's tests for equality of error variances revealed statistically significant between-condition differences for trust, $F(1,29) = 9.88$, $p = .002$, and negative affect, $F(1,29) = 2.92$, $p = .049$. The effect on target evaluation was not statistically significant, $F(1,29) = 2.48$, $p = .063$. The pattern on all variables was consistent with more variance in the negative norm condition. We subsequently tested whether the polarized perceptions of the outgroup target generalized to attitudes towards the outgroup as a whole. The difference between the majority norm conditions was in the predicted direction but not significant, $F(1,29) = 1.11$, $p = .151$.

For individuals, the Bartlett-Box test did not reveal a statistically-significant difference in variances, Box's $M = 11.57$, $p = .054$. In line with this, the pattern was less clear: Whereas the standard deviation of target evaluation was higher in the positive majority norm condition, $F(1,33) = 3.35$, $p = .038$, no differences were found on trust and negative affect, $F_s < .704$, $p_s > .204$. Unexpectedly, the variance in outgroup attitudes was higher in the positive norm condition, $F(1,33) = 5.65$, $p = .012$.

Discussion

Replicating Study 2, Study 3 provided evidence that the valence of a broad social category's norm influences the degree of polarization between local groups within that broader social category. Across variables, the variance between local-groups was significantly larger in the negative than in the positive majority norm condition (except for outgroup attitude, which revealed a trend in this direction).

Moreover, Study 3 replicated the finding that perceptions of local group norms drive this effect, again supporting Hypothesis 2. Indeed, the teller's perception of local norms predicted local group members' attitudes towards the individual minority target and towards the minority category as a whole. Tellers' perceptions of what is normative in their local

groups act as self-fulfilling prophecies, enforcing this presupposed norm on local group members. Interestingly, the interaction effect suggests that this steering occurred especially under a negative majority norm. Here, evaluations of the minority target and general attitudes towards the minority category were predicted by the teller's perceptions of the local group norm, whereas in the positive norm condition, the teller's perceptions of the local group norm did not predict these intergroup perceptions.

Finally, Study 3 aimed to examine whether polarization occurred only in local group settings (as suggested by the null-findings of Study 1), or was a direct effect of the manipulation of a negative majority norm. To this end, Study 3 combined the designs of Study 1 and 2 to enable a comparison between participants who discussed a scenario in their local group (the tellers) and participants who reflected on it individually. Supporting Hypothesis 3, a negative majority norm only polarized tellers' attitudes regarding a minority target. For those reflecting on the scenario individually, a negative (vs. positive) majority norm did not increase polarization. If anything, evidence suggested somewhat more polarization between individuals under a positive (vs. negative) norm.

General Discussion

The current research aimed to explain contemporary observations of polarization processes in Western societies, in which local groups shift to both the extreme right- and extreme left-wing parties, rather than assimilating to the negative societal sentiment about immigrants. To this end, we employed a novel analytic approach, using variance in individual- and multi-level designs as a statistical proxy for polarization (cf. Jans et al., 2015; Kuppens & Yzerbyt, 2014). Three studies provided converging evidence that discussion within local groups catalyzes polarization of majority members' views towards minorities. Moreover, this process is influenced by perceptions of norms at both the local and the categorical level. The studies concentrated on native-Dutch high school or university students

(i.e., majority members), who discussed or individually interpreted an ambiguous scenario describing an encounter with a Moroccan-Dutch student (outgroup target). The results consistently show that majority members take on more extreme views (both positive and negative) concerning a minority when 1) they perceive that, at the categorical level, the majority endorses negative views about this minority, and 2) they engage in local-group discussion with fellow majority members. The direction in which each local majority groups' views polarized was predicted by the local norm perceptions of the majority member who introduced the scenario (i.e., the teller) to the other local group members (i.e., the actively reacting "listeners"). That is, the tellers' perceptions of the local norms regarding the target minority acted as self-fulfilling prophecies, steering their local groups towards either negative or positive outgroup perceptions and, hence, causing polarization between local groups. Such polarization did not occur under a positive categorical majority norm or when majority members engaged in individual reflection rather than small-group discussions. These findings suggest that a negative, category-level majority norm urges majority members to take a stance in the debate, but that they need local-level interaction with fellow majority members to validate their viewpoints to be able to develop a firm stance.

Theoretical Implications

Relating these findings to existing literature on social influence, a number of observations can be made. First, we find no evidence that individuals passively conform to majority norms. It is noteworthy, however, that in the majority of studies reporting such conformity, participants did not engage in any form of social interaction (e.g., "minimal group paradigm"; Crandall & Stangor, 2005; Monteith, Deneen, & Tooman, 1996; Tajfel, 1970). In our studies – mimicking more real-life contexts – we do not replicate these norm compliance effects. Instead, we show that when people *interact* about norm-relevant information, they tend to construct norms and stereotypes more dynamically. That is, by

turning to local groups people can negotiate and validate a collectively shared viewpoint at the local level that could conform to *or* clash with the existing negative majority norm at the categorical level.

Second, it is important to distinguish the present research from established findings that group polarization occurs in the context of intergroup conflict (e.g., Abrams, Wetherell, Cochrane, Hogg, & Turner, 1990; Doise, 1969). Previous work has shown that collective judgments of group members become more extreme when confronted with a rival group. In our research, the negative norm manipulation may have made the intergroup conflict between native-Dutch and Moroccan-Dutch students more salient. Literature on intergroup conflict suggests that when such conflict is salient, one is likely to minimize differences within the ingroup, and distinguish oneself from the outgroup (Doise, 1969; Turner, 1985). One would thus expect ‘polarization’ between the majority category (here native-Dutch students) and minority category (here Moroccan-Dutch students). In the present studies, however, we find polarization at a different level. We demonstrate that when a negative norm about the outgroup is established at the level of a majority category in society, individual judgments do not shift to be in line with this norm, but polarization occurs between small local groups *within* that majority category. This is a crucial difference, because it implies that intergroup tension does not motivate group members to straightforwardly support the majority category position (thereby minimizing differences within the majority group), but instead motivates them to actively construe meaning regarding the minority category with the people in their direct environment. Thus, the direction of polarization results from the complex dynamic interactions within psychologically relevant local groups of individuals rather than from abstract social category membership (cf. McGarty et al., 2009). It appears that when discussing the issue in a local group, not all groups are inclined to distance themselves from the outgroup, but they develop either more hostile or more pro-social views about the

outgroup within in their local majority subgroups. Although intergroup contrasting may have been more likely with an outgroup that is actually present or expected to enter the situation at any moment (i.e., even more salient; cf. Greijdanus et al., 2015), the current situation in which group members construed meaning *within* their own groups about an absent outgroup closely mirrors daily life experiences (Dixon, Durrheim, & Tredoux, 2005; Dixon, Tredoux, Durrheim, Finchilescu, & Clack, 2008).

Third, in our research we found evidence for group influence in the form of polarization *between* local groups. That is, rather than converging their opinions to an average of opinions, local group members collectively shifted to a more extreme point on either side of the opinion spectrum. Interestingly, these effects occurred in absence of a clear intergroup context between local groups. Previous theorizing has explained polarization as resulting from the motivation to distinguish oneself positively in intergroup comparisons (Turner, 1985; Turner et al., 1987). In our experiment, although group members were aware of other groups participating in the experiment, they had no indication of how to distinguish themselves from other local groups on this topic, as no information regarding the norms of other groups was provided. The only information given was the existence of a negative (or positive) overarching majority norm. We suggest that this negative majority norm could foster polarization in two ways: First, a negative majority norm may serve as an alarm signal, similar to threat or negative mood, that the current situation requires action (Bless et al., 1990; Van Knippenberg et al., 2010). This may stimulate more in-depth group discussion of what their group stands for, thereby facilitating the emergence of a polarized view. Second, the development of negative attitudes about a minority category may require more social validation, because negative attitudes are, in essence, less morally valued than positive attitudes (Katz & Hass, 1988; Killen & Smetana, 2006; Rutland, Killen, & Abrams, 2010).

Individuals may therefore be more strongly influenced by their local group than they would have been under a more positive majority norm.

Predictors of Group Polarization

In line with these ideas, Studies 2 and 3 provide converging evidence that the direction in which each group polarizes is predicted by the tellers' perceptions of the local group norm. It appears that when communicating to other group members about a controversial issue (e.g., an ambiguous interaction with a minority category member), people estimate the local group norm and tailor their descriptions of the issue accordingly. In this way, intragroup communication is likely to reinforce—rather than challenge—presupposed local group norms, which may differ from majority norms. Although we did not directly analyze the content of group discussions, this explanation is in line with research suggesting that group members prefer to transmit stereotype-consistent information (Clark & Kashima, 2007), and also with research on the social tuning of racial attitudes to listeners (Sinclair, Lowery, Hardin, & Colangelo, 2005), and on the consensualizing nature of group discussions (Festinger, 1950; Haslam et al., 1997).

Extending this literature, the present research provides two important novel insights; we demonstrate 1) that local polarization occurs more strongly under a negative majority norm, and 2) that norm conformity is a complex process in which perceived majority and local group norms interactively influence individuals. Specifically, our studies suggest that under a negative norm, tellers may be more likely to take position, explaining the polarization processes between groups. On the one hand, the negative norm may provide leeway for negative opinions to arise: When tellers believe their group shares the majority norm, they may feel disinhibited to derogate an ambiguously behaving minority member. On the other hand, this can lead to contrast from the majority norm: Tellers may believe that their group does not share the negative norm and feel a need to tailor their descriptions positively to

avoid appearing prejudiced. Interestingly, these processes are much less pronounced when a positive majority norm is set. It seems that the positivity of the norm – and potentially, of the debate – reduces the need to take position but allows for more nuanced opinions.

A strength of the current set of studies is that we tested polarization processes under more or less naturalistic conditions, i.e., in discussions that might actually take place at school yards or in university canteens. This stress on external validity might have gone somewhat at the expense of examining the more fine-grained cognitive processes that took place when processing and discussing the information. These cognitive processes, e.g., more thorough information processing in the negative norm condition, can be tested under more controlled lab conditions in future studies. Future research could also further unpack the impact of majority norms by closely analyzing discussion dynamics and examining the role of moral considerations in the polarization process.

Concluding thoughts

We believe that the present research provides important insights into the complex influence of norms on shaping intergroup perceptions and attitudes - insights that would not have been obtained from examining either societal-level norms, small-group interactions, or individual processes alone. Indeed, current polarization (and radicalization) patterns such as the increasing support for political parties with an extreme left- or right-wing orientation are difficult to explain when examining individual-, group-, or societal-level factors separately. Indeed, in our research we depart from classical intergroup literature to explain how polarization can occur at a different level: between local subgroups of the same majority category (here, native-Dutch students). Our findings demonstrate that majority members do not passively conform to majority norms about minority categories, but actively construe meaning about these stereotypes with the people around them. The integration of wider societal contexts, group-level processes, and individual perceptions of—and interactions

with—these higher-level variables provides a unique insight in the ways in which local majority subgroups in society can become divided on important issues such as immigration. Therefore, we argue that cross-level examinations enable a better (i.e., more realistic) prediction of behavioral patterns that are currently witnessed in society.

References

- Aaldering, H., Ten Velden, F. S., van Kleef, G. A., & De Dreu, C. K. (2018). Parochial cooperation in nested intergroup dilemmas is reduced when it harms out-groups. *Journal of Personality and Social Psychology, 114*, 909-923. doi:10.1037/pspi0000125
- Abrams, D., Wetherell, M. S., Cochrane, S., Hogg, M. A., & Turner, J. C. (1990). Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization. *British Journal of Social Psychology, 29*, 97-119. doi:10.1111/j.2044-8309.1990.tb00892.x
- Alba, R., & Foner, N. (2017). Immigration and the geography of polarization. *City and Community, 16*(3), 239-243. doi:10.1111/cico.12241
- Albertazzi, D., & McDonnell, D. (2008). *Twenty-first century populism: The spectre of Western European democracy*. New York: Palgrave Macmillan.
- Berkhof, J., & Snijders, T. A. (2001). Variance component testing in multilevel models. *Journal of Educational and Behavioral Statistics, 26*, 133-152. doi:10.3102/10769986026002133
- Bless, H., Bohner, G., Schwarz, N., & Strack, F. (1990). Mood and persuasion: A cognitive response analysis. *Personality and Social Psychology Bulletin, 16*, 331-345. doi:10.1177/0146167290162013
- Bliuc, A., McGarty, C., Reynolds, K., & Muntele, D. (2007). Opinion-based group membership as a predictor of commitment to political action. *European Journal of Social Psychology, 37*(1), 19-32. doi:10.1002/ejsp.334
- Bliuc, A., McGarty, C., Thomas, E. F., Lala, G., Berndsen, M., & Misajon, R. (2015). Public division about climate change rooted in conflicting socio-political identities. *Nature Climate Change, 5*(3), 226-229. doi:10.1038/nclimate2507

- Clark, A. E., & Kashima, Y. (2007). Stereotypes help people connect with others in the community: A situated functional analysis of the stereotype consistency bias in communication. *Journal of Personality and Social Psychology, 93*, 1028–1039. doi:10.1037/0022-3514.93.6.1028
- Crandall, C. S., & Stangor, C. (2005). Conformity and prejudice. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 295–309). Malden, MA: Blackwell.
- Doise, W. (1969). Intergroup relations and polarization of individual and collective judgments. *Journal of Personality and Social Psychology, 12*(2), 136–143. doi:10.1037/h0027571
- Dunlap, R. E., McCright, A. M., & Yarosh, J. H. (2016). The political divide on climate change: Partisan polarization widens in the US. *Environment: Science and Policy for Sustainable Development, 58*(5), 4-23. doi:10.1080/00139157.2016.1208995
- Festinger, L. (1950). Informal social communication. *Psychological Review, 57*(5), 271–282. doi:10.1037/h0056932
- Forgas, J. P. (2013). Don't worry, be sad! On the cognitive, motivational, and interpersonal benefits of negative mood. *Current Directions in Psychological Science, 22*, 225–232. doi:10.1177/0963721412474458
- Galam, S., & Moscovici, S. (1991). Towards a theory of collective phenomena: consensus and attitude changes in groups. *European Journal of Social Psychology, 21*, 49–74. doi:10.1002/ejsp.2420210105
- Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research, 35*, 472–482. doi:10.1086/586910

- Greijdanus, H., Postmes, T., Gordijn, E. H., & Van Zomeren, M. (2014). When abstraction does not increase stereotyping: Preparing for intragroup communication enables abstract construal of stereotype-inconsistent information. *Social Cognition, 32*, 505–527. doi:10.1521/soco.2014.32.6.505
- Greijdanus, H., Postmes, T., Gordijn, E. H., & Van Zomeren, M. (2015). Steeling ourselves: Intragroup communication while anticipating intergroup contact evokes defensive intergroup perceptions. *PloS one, 10*, e0131049. doi:10.1371/journal.pone.0131049
- Haslam, S. A., McGarty, C., & Turner, J. C. (1996). Salient group memberships and persuasion: The role of social identity in the validation of beliefs. In J. L. Nye & A. M. Brower (Eds.), *What's social about social cognition? Research on socially shared cognition in small groups* (pp. 29–56). Thousand Oaks, CA: Sage.
- Haslam, S. A., Turner, J. C., Oakes, P. J., McGarty, C., & Reynolds, K. J. (1997). The group as a basis for emergent stereotype consensus. *European Review of Social Psychology, 8*, 203–239. doi:10.1080/14792779643000128
- Hewstone, M., & Jaspars, J. (1982). Explanations for racial discrimination: The effect of group discussion on intergroup attributions. *European Journal of Social Psychology, 12*, 1–16. doi:10.1002/ejsp.2420120101
- Jans, L., Leach, C. W., Garcia, R. L., & Postmes, T. (2015). The development of group influence on in-group identification: A multilevel approach. *Group Processes and Intergroup Relations, 18*, 190–209. doi:10.1177/1368430214540757
- Jetten, J., Spears, R., & Manstead, A. S. (1996). Intergroup norms and intergroup discrimination: Distinctive self-categorization and social identity effects. *Journal of Personality and Social Psychology, 71*, 1222–1233. doi:10.1037/0022-3514.71.6.1222

Kashima, Y. (2000). Maintaining cultural stereotypes in the serial reproduction of narratives.

Personality and Social Psychology Bulletin, 26, 594–604.

doi:10.1177/0146167200267007

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict:

Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology*, 55, 893–905. doi:10.1037/0022-3514.55.6.893

Killen, M., & Smetana, J. G. (2006). *Handbook of moral development*. Mahwah, NJ:

Erlbaum.

Koudenburg, N., Postmes, T., & Gordijn, E. H. (2013a). Conversational flow promotes

solidarity. *PLoS one*, 8, e78363. doi:10.1371/journal.pone.0078363

Koudenburg, N., Postmes, T., & Gordijn, E. H. (2013b). Resounding silences: Subtle norm regulation in group interaction. *Social Psychology Quarterly*, 76, 224–241.

doi:10.1177/0190272513496794

Koudenburg, N., Postmes, T., & Gordijn, E. H. (2017). Beyond content of conversation: The

role of conversational form in the emergence and regulation of social structure. *Personality and Social Psychology Review*, 21(1), 50–71.

doi:10.1177/1088868315626022

Kuppens, T., & Yzerbyt, V. Y. (2014). Predicting variability: Using multilevel modelling to assess differences in variance. *European Journal of Social Psychology*, 44, 691–700.

doi:10.1002/ejsp.2028

Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality

(vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, 93, 234–249. doi:10.1037/0022-3514.93.2.234

Lyons, A., & Kashima, Y. (2003). How are stereotypes maintained through communication?

The influence of stereotype sharedness. *Journal of Personality and Social Psychology*, 85, 989–1005. doi:10.1037/0022-3514.85.6.989

March, L., & Mudde, C. (2005). What's left of the radical left? The European radical left after 1989: Decline and mutation. *Comparative European Politics*, 3, 23–49.

doi:10.1057/palgrave.cep.6110052

McGarty, C., Bliuc, A. M., Thomas, E. F., & Bongiorno, R. (2009). Collective action as the material expression of opinion-based group membership. *Journal of Social Issues*, 65, 839–857. doi:10.1111/j.1540-4560.2009.01627.x

McGarty, C., Turner, J. C., Hogg, M. A., David, B., & Wetherell, M. S. (1992). Group polarization as conformity to the prototypical group member. *British Journal of Social Psychology*, 31, 1–19. doi:10.1111/j.2044-8309.1992.tb00952.x

Meeussen, L., Delvaux, E., & Phaet, K. (2014). Becoming a group: Value convergence and emergent work group identities. *British Journal of Social Psychology*, 53, 235–248.

doi:10.1111/bjso.12021

Monteith, M. J., Deneen, N. E., & Tooman, G. D. (1996). The effect of social norm activation on the expression of opinions concerning gay men and Blacks. *Basic and Applied Social Psychology*, 18, 267–288. doi:10.1207/s15324834basp1803_2

Moscovici, S., & Zavalloni, M. (1969). The group as a polarizer of attitudes. *Journal of Personality and Social Psychology*, 12, 125–135. doi:10.1037/h0027568

Mudde, C. (2004). The populist zeitgeist. *Government and Opposition*, 39, 541–563.

doi:10.1111/j.1477-7053.2004.00135.x

Mummendey, A., & Otten, S. (1998). Positive–negative asymmetry in social discrimination. *European Review of Social Psychology*, 9(1), 107–143.

doi:10.1080/14792779843000063

- Postmes, T., Spears, R., Lee, A. T., & Novak, R. J. (2005). Individuality and social influence in groups: inductive and deductive routes to group identity. *Journal of Personality and Social Psychology, 89*, 747–763. doi:10.1037/0022-3514.89.5.747
- Prentice, M., Jayawickreme, E., Hawkins, A., Hartley, A., Furr, R. M., & Fleeson, W. (2018). Morality as a basic psychological need. *Social Psychological and Personality Science*. Advance online publication. doi:10.1177/1948550618772011
- Reicher, S., Cassidy, C., Wolpert, I., Hopkins, N., & Levine, M. (2006). Saving Bulgaria's Jews: An analysis of social identity and the mobilisation of social solidarity. *European Journal of Social Psychology, 36*, 49–72. doi:10.1002/ejsp.291
- Rutland, A., Killen, M., & Abrams, D. (2010). A new social-cognitive developmental perspective on prejudice the interplay between morality and group identity. *Perspectives on Psychological Science, 5*, 279–291. doi:10.1177/1745691610369468
- Sherif, M. (1936). *The psychology of social norms*. New York: Octagon.
- Sinclair, S., Lowery, B. S., Hardin, C. D., & Colangelo, A. (2005). Social tuning of automatic racial attitudes: the role of affiliative motivation. *Journal of Personality and Social Psychology, 89*, 583–592. doi:10.1037/0022-3514.89.4.583
- Smith, L. G. E., & Postmes, T. (2009). Intra-group interaction and the development of norms which promote inter-group hostility. *European Journal of Social Psychology, 39*, 130–144. doi:10.1002/ejsp.464
- Smith, L. G. E., & Postmes, T. (2011). The power of talk: Developing discriminatory group norms through discussion. *British Journal of Social Psychology, 50*, 193–215. doi:10.1348/014466610X504805
- Smith, L. G. E., Thomas, E. F., & McGarty, C. (2015). “We must be the change we want to see in the world”: Integrating norms and identities through social interaction. *Political Psychology, 36*, 543–557. doi:10.1111/pops.12180

- Snijders, T. A. B., & Bosker, R. (2012). *Multilevel analysis: An introduction to basic and advanced modeling (2nd edition)*. London, UK: Sage.
- Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, 223(5), 96–103.
- Tajfel, H. (1978). Interindividual behavior and intergroup behavior. In H. Tajfel (Ed.), *Differentiation between social groups* (pp. 27–60). London: Academic Press.
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behaviour. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (2nd ed., pp. 7–24). Chicago, IL: Nelson-Hall.
- Thomas, E. F., Smith, L. G. E., McGarty, C., & Postmes, T. (2010). Nice and nasty: The formation of prosocial and hostile social movements. *Revue Internationale de Psychologie Sociale*, 23(2), 17–55.
- Turner, J. C. (1985). Social categorization and the self-concept: A social cognitive theory of group behavior. In E. J. Lawler (Ed.), *Advances in group processes* (Vol. 2, pp. 77–122). Greenwich, CT: JAI Press.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford, UK: Blackwell.
- Van Knippenberg, D., Kooij-de Bode, H. J. M., & Van Ginkel, W. P. (2010). The interactive effects of mood and trait negative affect in group decision making. *Organization Science*, 21, 731–744. doi:10.1287/orsc.1090.0461
- Wittenbaum, G. M., & Moreland, R. L. (2008). Small-group research in social psychology: Topics and trends over time. *Social and Personality Psychology Compass*, 2, 187–203. doi:10.1111/j.1751-9004.2007.00065.x

Tables and Figures

Table 1. Means (*SDs*) per condition of the categorical majority norm in Study 1.
Different subscripts indicate that means per condition did significantly differ at $p < .05$.

	Negative norm ($n = 25$)	Positive norm ($n = 25$)
	<i>M (SD)</i>	<i>M (SD)</i>
Manipulation check	4.00 _a (1.00)	4.92 _b (1.22)
Reported trust target	3.16 (1.49)	3.04 (1.45)
Trust	2.46 (0.83)	2.72 (0.89)
Negative affect	3.50 (1.03)	3.14 (1.22)

Table 2. Means, group influence statistics, and test results for differences in group level variance between conditions (negative vs. positive categorical majority norms about the minority; Study 2).

	Negative norm (<i>n</i> = 68)			Positive norm (<i>n</i> = 77)			Wald Z test for between condition difference in group level variance		Chi-square comparing Model 1 and 2.and mo		
	<i>M</i>	Variance		ICC	<i>M</i>	Variance		ICC	Wald Z	p	$\Delta\chi^2$
		Individual level (σ^2)	Group level (τ^2)		Individual level (σ^2)	Group level (τ^2)					
Evaluation target	3.75	.66 (.14)***	.80 (.32)**	.55	4.01	.46 (.09)***	.32 (.14)*	.41	-2.95	.008	3.16
Trust	2.47	.77 (.16)***	.54 (.25)*	.41	2.09	.99 (.18)***	.00 (.00) ^a	.00	-2.85	.004	5.14

Negative affect	3.25	.31 (.06)***	.22 (.10)*	.42	3.28	.38 (.07)***	.00 (.00) ^a	.00	-2.86	.004	5.01
Outgroup attitude	4.65	.80 (.17)***	.48 (.24)*	.38	4.56	1.04 (.20)***	.08 (.13)	.07	-3.17	.002	1.82

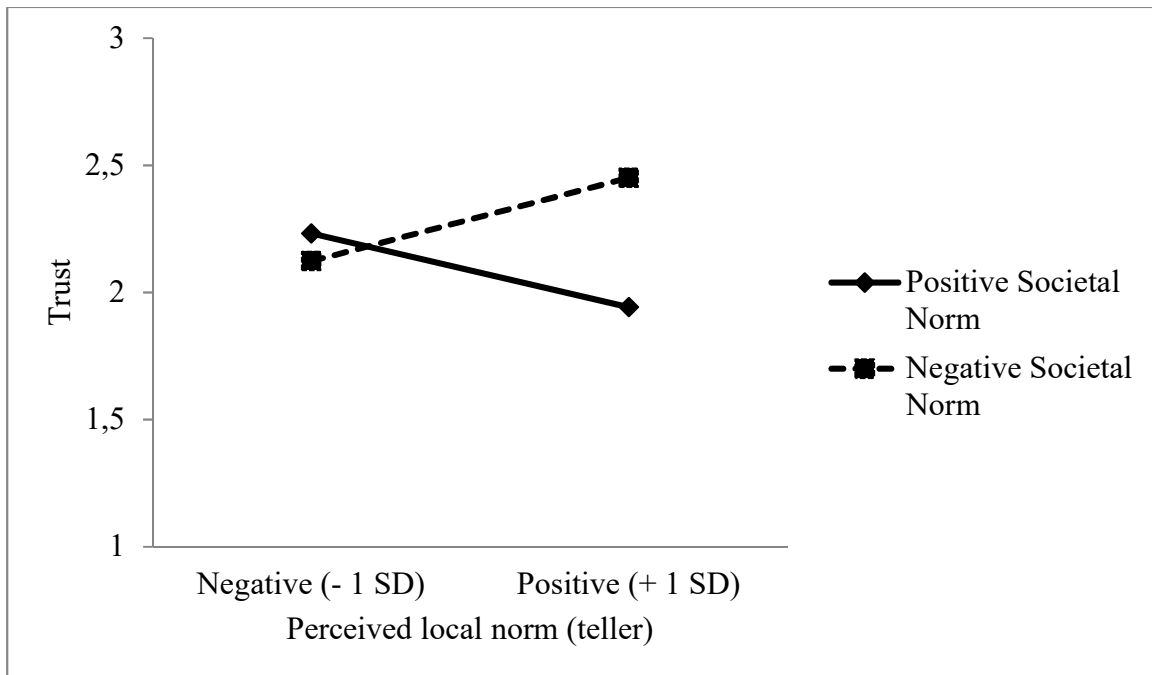
*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

^a This parameter is redundant, τ cannot be calculated.

Note. Standard errors for the variances are reported between brackets. Because chi-square difference tests were used to test variance components we divided the p-values by 2, as recommended by Snijders and Bosker (2012, p.98).

Figure 1a-b. Graphical displays of the interaction effects on (a) trust and (b) negative affect in Study 2.

a.



b.

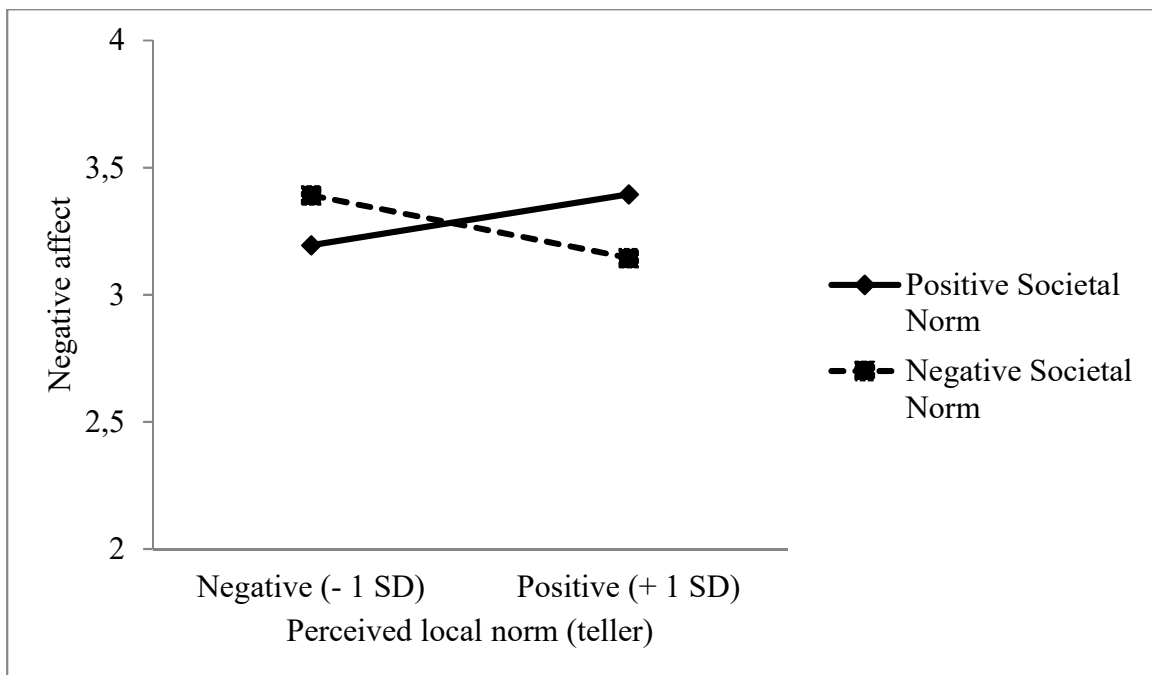


Table 3. Means, group influence statistics, and test results for between condition differences in group level variance in Study 3. Results display only data from participants in groups.

	Negative norm (<i>n</i> = 43)			Positive norm (<i>n</i> = 42)			Wald Z test for between condition difference in group level variance		Chi-square test comparing fit i Model 1 and Model 2.			
	<i>M</i>	Variance	ICC	<i>M</i>	Variance	ICC	w	p	$\Delta\chi^2$			
		Individual level (σ^2)	Group level (τ^2)		Individual level (σ^2)	Group level (τ^2)						
Evaluation target	4.04	.57 (.16)***	1.09 (.48)*	.66	4.32	.37 (.10)***	.11 (.10)	.24	-2.456	.014	7.62	.00
Trust	2.79	1.23 (.34)***	.27 (.31)	.18	2.50	.62 (.14)***	.00 ^a (.00)	.00	-2.014	.044	3.39	.00

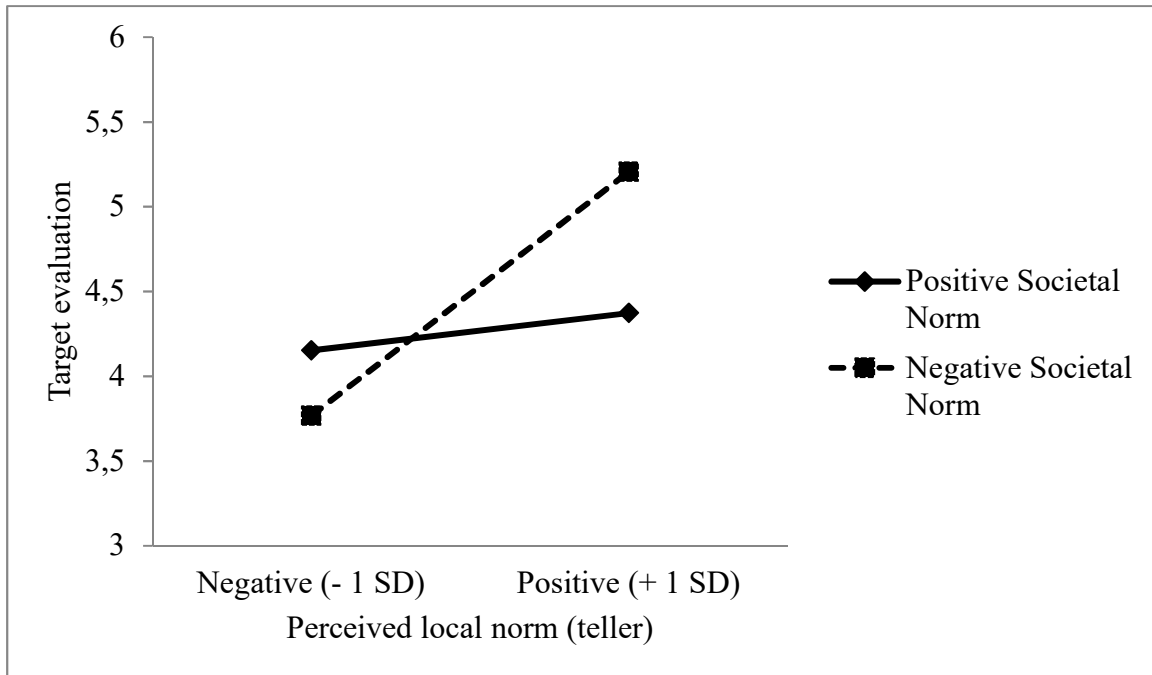
Negative affect	3.23	.51 (.14) ***	.05 (.10)	.10	3.20	.21 (.06) ***	.00 (.03)	.01	-2.055	.040	2.67	.0
Outgroup attitude	4.53	.92 (.27) ***	.17 (.24)	.15	4.98	.68 (.15) ***	.00 ^a (.00)	.00	-1.970	.049	1.44	.1

Note. Standard errors for the variances are reported between brackets. Because chi-square difference tests were used to test variance components we divided the p-values by 2, as recommended by Berkhof and Snijders (2001).

^a This covariance parameter is redundant.

Figure 2a and b. Interaction effects between majority norms and local group norms for (a) target evaluation and (b) outgroup attitude in Study 3.

a.



b.

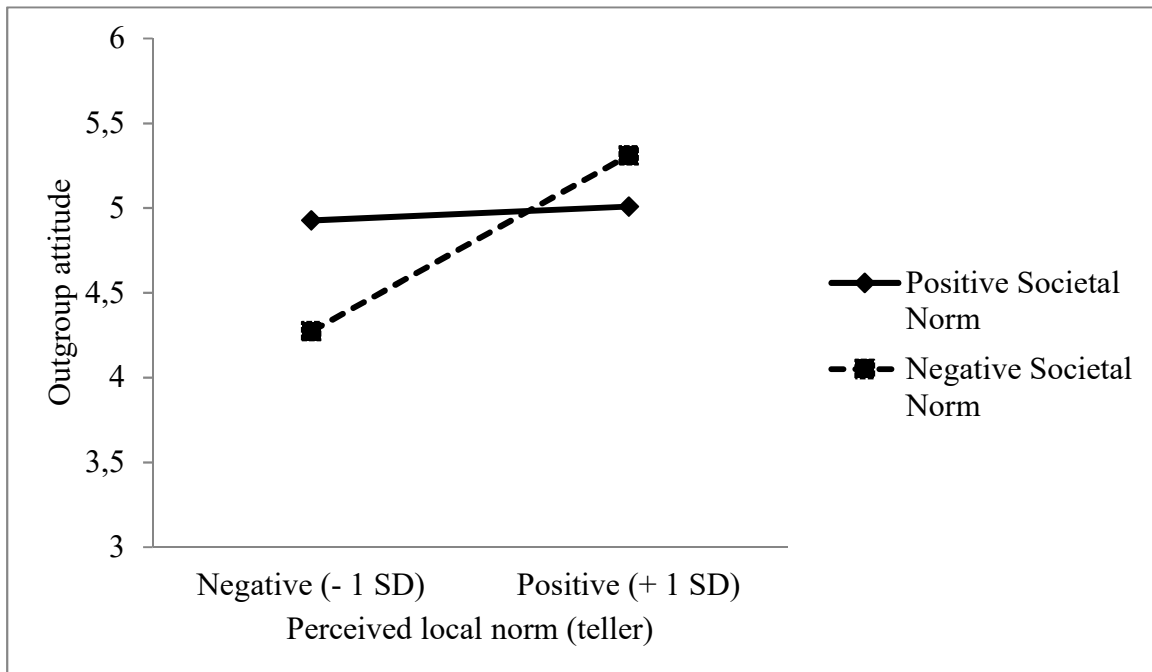


Table 4. Means and standard deviations of tellers in the group conditions, and participants in the individual conditions (Study 3).

	Group teller		Individual	
	Negative majority norm (<i>n</i> = 16)	Positive majority norm (<i>n</i> = 15)	Negative majority norm (<i>n</i> = 18)	Positive majority norm (<i>n</i> = 17)
Evaluation target	4.05 (1.35)	4.26 (0.80)	4.58 (0.77)	4.29 (1.22)
Trust	2.84 (1.41)	2.84 (0.63)	3.22 (1.15)	2.56 (1.29)
Negative affect	3.25 (0.98)	3.43 (0.50)	3.14 (0.94)	3.44 (0.83)
Outgroup attitude	4.75 (1.21)	5.16 (0.98)	4.85 (0.69)	4.69 (1.34)