The U-Turn Effect in Social Norm Interventions: Individual-to-peer Behavior Gap as Endogenous Driver of Social Identification with Peers, and its Consequence on Peer Effects

Pieter F. E. Verhallen*

Maastricht University and Netspar

Thomas Post

Maastricht University, Open University, and Netspar

This version: October 13, 2020

Abstract

Social norm interventions have been shown to be an effective tool for changing human behavior towards a desirable direction across many domains, including financial decision-making. Yet, there is substantial variation found in the effectiveness of social norm interventions, and even evidence that such interventions can backfire into oppositional reactions. In this paper, we posit that social identification with a peer group used in a social norm intervention is endogenously affected by the respective group′s behavior made salient in the intervention. Specifically, we propose, and empirically confirm, that the perceived distance between a subject′s current behavior and a peer group′s salient behavior affects the subject′s social identification with this peer group, with potential to transform a peer group into a non-peer group, with consequences – which we term the U-turn effect. As social identification positively affects convergence to norm behavior, the negative effect of perceived behavioral distance on social identification thus indirectly negatively affects convergence to norm behavior, and this can even lead to oppositional reactions. We thereby address the puzzle of heterogeneous findings generally, and oppositional reactions specifically, in social norm intervention studies.

Acknowledgements: We thank Andrea Dinneen, Anja Schanbacher, and Mariel Beasley for helpful comments. We acknowledge financial support from the Common Cents Lab, Center for Advanced Hindsight, Duke University, and from the Network for Studies on Pension, Aging and Retirement (NETSPAR).

JEL Classification D14, D91

Keywords: Social Norms, Peer Effects, Social Identification, Social Identity Theory, Consumer Behavior, Oppositional Reaction

* Corresponding author: Pieter F. E. Verhallen, Maastricht University, School of Business and Economics, Department of Marketing and Supply Chain Management, P.O. Box 616, 6200 MD, The Netherlands. Tel.: +31 43 38 83 774. E-mail: p.verhallen@maastrichtuniversity.nl. The remaining coauthors can be reached at this same address.

Introduction

Social norm interventions have been shown to be an effective tool to change human behavior towards a desired behavior. For instance, such interventions have affected alcohol consumption (Ridout & Campbell, 2014), recycling (Thomas & Sharp, 2013), voting (Panagopoulos, Larimer, & Condon, 2012), sustainability behaviors (Goldstein, Cialdini, & Griskevicius, 2008; Van Der Linden, 2015), academic performance (Azmat & Iriberri, 2010), tax compliance (Bobek, Hageman, & Kelliher, 2013), energy usage (Allcott, 2011), and financial decisions (Verhallen, Brüggen, Post, & Odekerken-Schröder, 2020). However, there has been substantial variation found in the effectiveness of social norm interventions and even evidence that such interventions can backfire. In this paper, we propose a new conceptual model and present causal empirical evidence that shows that if a peer group's behavior is too far away from the individual's own behavior, social identification with the peer group is reduced and, in turn, the peer effect is weakened or, in some cases, even converted into an oppositional reaction – a social norm intervention phenomenon, which we call the U-turn effect.

The literature on peer effects documents two scenarios in which social norm interventions are ineffective or even backfire. First, for individuals already outperforming a given norm (e.g., those already consuming less alcohol or saving more than the norm), converging towards the norm involves shifting behavior in an undesirable direction. For example, in an energy conservation intervention where households were provided with energy consumption levels of their neighbors, households that were consuming less energy than their neighbors started to consume more (Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007). Second, for some individuals, a given norm may trigger an oppositional reaction, causing individuals to move away rather than converge towards the norm. For example, Beshears et al. (2015) find that, for some individuals subjected to

a retirement account savings intervention, they were not only less likely to converge towards a norm, but they even behaved oppositional to the norm, that is, saved less than when not presented with the norm.

Social identification with peers is a key driver in convergence towards peer behavior (e.g. Hogg, 2006; Phua, 2013; Reed, Lange, Ketchie, & Clapp, 2007; Verhallen et al., 2020). That is, the more an individual identifies with a peer group, the more likely he or she is to converge towards the peer group behavior. Verhallen et al. (2020) show that low social identification with a peer group not only diminishes convergence to a peer group's behavior, but can also drive the second form of oppositional reactions to peer group behavior. In a peer behavior intervention field experiment, members of a pension fund received an e-mail highlighting a useful article for them, and some received additional peer behavior treatments stating that this article was most read by a specific peer group. For those who received the peer behavior cue from a peer group scoring high on social identification, there was a 31.2 to 42.1 per cent increase in clicking the e-mail link to the article, compared to control. For those who received the same peer behavior cue from a peer group low on social identification, there was a 24.4 per cent decrease in clicking the e-mail link, compared to control. That is, if the peer group used in an intervention provides low social identification for the individual based on observable characteristics, the intervention is not effective or even counterproductive.

But, even interventions with norms from seemingly similar peer groups, again based on observable characteristics, such as age or gender, may yield ineffective or counterproductive results. Beshears et al. (2015) ran a field experiment with a large manufacturing firm in the United States, where subjects were randomly assigned to either a control with no peer information, or social norm intervention treatments providing information on the percentage of colleagues within

the same age bracket that were contributing at least 6% of their income to (were enrolled in) the firm's 401(k) retirement account to subjects (not) enrolled in the respective retirement account. The authors found that only for subjects who had never opted into the optional retirement account by choice or through inertia, compared to subjects who had either opted into or opted out of (after automatic enrolment) the optional retirement account, mean enrolment rates were *lower* in the social norm intervention treatments compared to control. As seen in the work by Beshears et al. (2015), a social norm intervention with a seemingly identifiable peer group can lead to oppositional reactions, too. The authors discuss that discouragement from upwards social comparison may be driving this oppositional reaction. Potentially, and that is the research avenue we pursue in this paper, the perceived distance to the peer group's behavior endogenously has an effect on individuals' social identification with the respective peer group, and consequently on their convergence to the social norm intervention. In the case of the work by Beshears et al. (2015), the distance in behavior on the dimension of engagement (in this case, enrolment itself as proxy for engagement) may have instigated lower social identification with the peer group in the social norm treatment for only those who never made any active (engaged) decision, as their distance from the salient peer behavior was largest. Beshears et al. (2015) also found that, for those subjects already enrolled and who received information on the fraction of peers contributing a specific 6% of income or more to their retirement account, only those with low prior contribution rates showed a negative reaction to the peer information. This, too, aligns with our hypothesis that distance from the peer behavior negatively affects convergence to that behavior, through lowered social identification with the respective peer group. To provide another example of what we posit:

Imagine, for a moment, that you are a casual runner, who often runs at a speed of 12 kilometers per hour. Tomorrow, you will go running with peers with whom you are close and have a high

social identification with. When you start running with your peers, you notice that they all run at a speed of 13 kilometers per hour. You internalize this as the norm, and, based on either, or both, your assumption that there must be a good reason for this speed, as well as your desire to signal affiliation with these peers, you increase your speed towards theirs. This is an example of a classic peer effect from exposure to a social norm that differs from your existing behavior. Now, imagine the same scenario – running with your peers – but, instead of running at a speed of 13 kilometers per hour, your peers exhibit an incredible speed of 16 kilometers per hour. While you initially greatly identified with these peers, their running speed of 16 kilometers per hour is far above yours, and puts your high identification with them into question. Are they really my peers? Are we genuinely similar? You may, at this point, very well lower your identification with these peers. The faster these peers run, the stronger the reduction in social identification with this peer group. This reduction in social identification subsequently reduces your drive to alter your speed towards the speed of your peers. At a certain peers' speed, you may even start to judge them as an outgroup and, consequently, push your behavior away from theirs – that is, exhibit an oppositional reaction to norm exposure. This is what we call the U-turn effect.

(Alternative example: In 2020, the world went at war with a virus, the Covid-19 pandemic. Outside of parts of Asia, wearing any protective gear, such as face masks, was unusual. When the pandemic started spreading, the need for protective wear increased, while the habit to wear protective wear had not yet developed. Imagine, for a moment, that you are someone who has never worn a face mask before and are currently not wearing any mask, despite the spread of the pandemic. Tomorrow, you will meet two peers, with whom you are close and have a high social identification with, for a coffee. When you meet your peers, you notice that they are both wearing a face mask. You internalize this as the norm, and, based on either, or both, your assumption that there must

be a good reason for this face mask, as well as your desire to signal affiliation with these peers, you buy a face mask and decide to wear one moving forward. This is an example of a classic peer effect from exposure to a social norm that differs from your existing behavior. Now, imagine the same scenario – meeting two peers for a coffee – but, instead of wearing face masks, the two peers are wearing a face mask, face shield, baseball cap, and surgical gloves instead. While you initially greatly identified with these peers, the seemingly excessive protective gear that they wear is far different from your current behavior, and puts your high identification with them into question. Are they really my peers? Are we genuinely similar? You may, at this point, very well lower your identification with these peers. The more different your peers' protective outfit is, the stronger the reduction in social identification with these peers. This reduction in social identification subsequently reduces your drive to alter your behavior towards the behavior of your peers. At a certain distance in behavior, you may even start to judge the peers as an outgroup and, consequently, push your behavior away from theirs – that is, exhibit an oppositional reaction to norm exposure. This is what we term the U-turn effect.)

To capture such behavior, we develop a conceptual model and present causal empirical evidence that uncovers why even seemingly similar peer groups' norms can lead to ineffective interventions. Through an experiment, we show that providing information on a peer group's behavior affects individuals' social identification with that peer group. We show that, as the perceived distance (hereafter distance) between an individual's behavior and their peer group's behavior grows, at some boundary point the distance is perceived to be so large that the peer group's behavior endogenously disqualifies the group as peers with whom one socially identifies. In other words, as the distance in behavior grows, the individual identifies less with the peer group, based on their behavior, which first diminishes the magnitude of the peer effect (i.e., convergence

to the peer group norm), and, at larger distances, can even fuel oppositional reactions by individuals, likely based on in-group versus out-group identification of the peer group. In our experiment, we manipulate the distance between an individual's and peer group's behavior (e.g. discretionary spending on eating out) to examine the effect of distance on the individual's perceived social identification with the respective peer group and the subsequent behavioral reaction relative to the norm.

We add to extant theory by showing that when subjects are presented with information that a peer group behaves significantly differently (spends substantially less on eating out), there is a significant reduction in subjects' social identification with that peer group. Moreover, we find that both social identification and distance between an individual's and peer group's behavior have a positive main effect on norm convergence. Keeping distance constant, the higher the social identification with a peer group, the higher the convergence to peer behavior. Keeping social identification with the peer group constant, the larger the distance in behavior, the greater the convergence to the peer group's behavior. Lastly, we find that perceived distance matters most, which is influenced by both relative and absolute distance to peer behavior. Increasing the distance between a subject's behavior and a peer group's behavior thus both positively (direct effect) and negatively (through social identification) affects convergence to peer group behavior. Our findings are also important from a policymaker and managerial perspective, as they inform how to avoid ineffective, and even oppositional, reactions to social norm interventions in the field. Practitioners ought to be wary when crafting social norm interventions where salient peer behavior differs significantly from the subjects being targeted.

This paper is structured as follows. First, relevant literature in social norm and social identity theory are reviewed to establish our hypotheses. Next, we report on study 1, where we

manipulated the individual-to-peer behavior gap to test its effect on social identification and subsequent norm convergence. Lastly, the findings are discussed, and limitations as well as avenues for future research are presented.

Background and Hypotheses

Social Norms and Peer Effects

The peer effect is one of Cialdini's (2001) identified channels of persuasion, inducing changes in behavior by providing behavioral information about peers, that is, individuals sharing one or more characteristics. When behavior of a peer group is made salient (Cialdini, Reno, & Kallgren, 1990), subjects are driven to social compliance and conformity, towards what those others do or what those others approve or disapprove of doing (Cialdini & Goldstein, 2004; Schultz, 1999). Behavioral information about a peer group has been shown to indeed affect many behaviors, such as alcohol consumption (Ridout & Campbell, 2014), recycling (Thomas & Sharp, 2013), voting (Panagopoulos, Larimer, & Condon, 2012), sustainability behaviors (Goldstein, Cialdini, & Griskevicius, 2008; Van Der Linden, 2015), academic performance (Azmat & Iriberri, 2010), tax compliance (Bobek, Hageman, & Kelliher, 2013), financial decisions (Beshears, Choi, Laibson, Madrian, & Milkman, 2015), and energy usage (Allcott, 2011).

Subjects are driven to conform through their goal of accuracy and goal of affiliation (Cialdini & Goldstein, 2004; Schultz, 1999), the extent of which may differ across people. The goal of accuracy highlights the motivation "to achieve…goals in the most effective and rewarding manner possible" (Cialdini & Goldstein, 2004, pg. 592). This goal of accuracy drives the search for information on what constitutes an optimal behavior, and the behavior of a peer group can act as a data point on what this optimal behavior is. Additionally, the goal of affiliation may drive

people to conform to peer group behavior as conformity can signal affiliation to others, and drive social approval (Cialdini & Goldstein, 2004), e.g., with people one wishes to affiliate with.

Social Identity Theory and Social Identification

One influence driving convergence to a norm is normative influence, that is, influence from what are considered norms within certain peer groups – average behavior of peer groups, and behavior that peer groups may approve or disapprove of. People have an inherent goal of affiliation with others, meaning that "humans are fundamentally motivated to create and maintain meaningful social relationships with others" (Cialdini & Goldstein, 2004, p. 598). By converging to a peer group's social norm, subjects not only feel affiliated with the particular group, but they also use their own behavior to signal affiliation and membership with this peer group. People can seek to affiliate due to value expressive and utilitarian influences (Bearden, Netemeyer, & Teel, 1989). Value expressive influence refers to one's "desire to enhance self-image by association with a peer group … motivated by the individual's desire to enhance or support his or her self-concept through referent identification" (Bearden, Netemeyer, & Teel, 1989, p. 474). Utilitarian influence refers to conformity to group behavior in order to avoid punishment or receive rewards from the respective group. Bearden, Netemeyer, and Teel (1989) define the extent to which a subject is influenced by these value expressive and utilitarian influences as a subject's susceptibility to normative influence. People can vary in their susceptibility to normative influence, and it generally declines with age (Batra, Homer, & Kahle, 2001; Park & Lessig, 1977). The susceptibility to normative influence is not specific to any particular peer group. Which peer groups people are most influenced by, however, depends on their social identity. In social identity theory, one's social identity is "the individual's knowledge that he/she belongs to certain social groups" (Tajfel, 1982, pg. 31). It is founded on self-stereotyping, that is, the belief of similarity to the social group's

average or prototypical members (Leach et al., 2008). The extent to which one identifies with a group is termed social identification (Abrams & Hogg, 1990). Therefore, we hypothesize that,

H1$_a$. Higher perceived distance between an individual's and peer group's behavior lowers individual's social identification with the peer group.

When one socially identifies with a group, conformity to the group's behavior is driven by value expressive and utilitarian influences. Therefore, we hypothesize that,

H1$_b$. Social identification with a peer group increases convergence towards a salient social norm from the respective peer group.

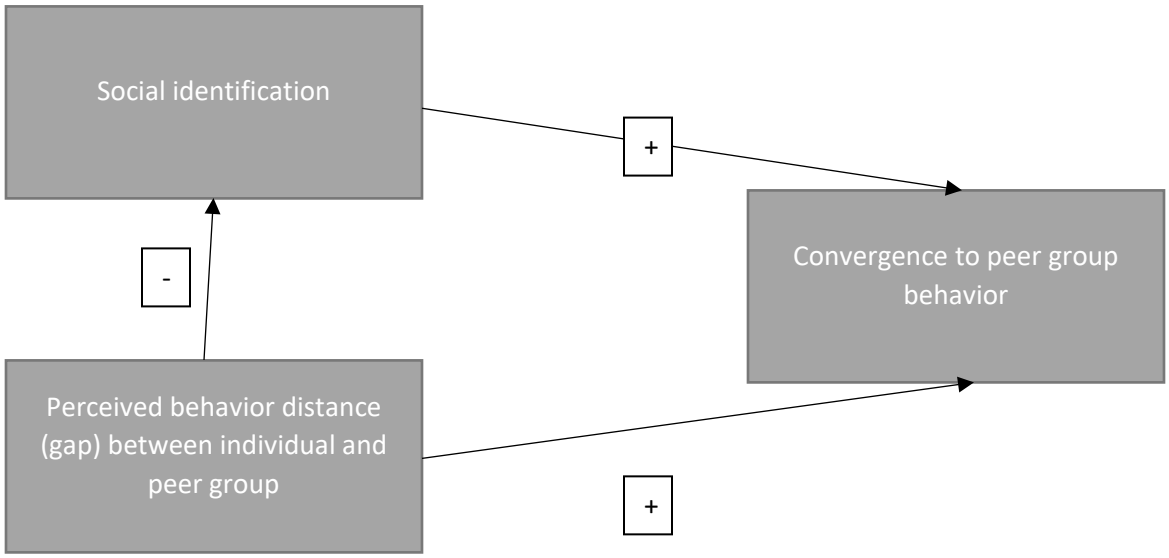
The extent to which a subject is influenced by a social norm is also contingent on the distance between the subject's current behavior and the salient peer group behavior. If there is no distance between the subject's current behavior and the salient peer group behavior, then there is no need to shift behavior to satisfy the goals of accuracy and affiliation. Generally, the larger the distance in behavior, the greater the shift in behavior to satisfy the goals of accuracy and affiliation. Therefore, we hypothesize that,

H2. The distance between an individual's and peer group's behavior has a direct positive effect on the absolute behavior change towards the peer group behavior (keeping social identification constant).

Figure 1 summarizes the conceptual model.

Figure 1

Conceptual model



Study 1: Empirical Verification of Hypotheses

Methodology

We conducted an online experiment by recruiting 372 subjects on Amazon's Mechanical Turk (mTurk) platform. mTurk is an established recruitment platform for obtaining subjects for experiments (e.g. Paolacci, Chandler, & Ipeirotis, 2010), especially within the context of decision-making heuristics (Horton, Rand, & Zeckhauser, 2011).[1] Given some concerns on mTurk subject quality from mid-2018 onwards (Ahler, Roush, & Sood, 2019; Chmielewski & Kucker, 2020), we used the Turkprime (now CloudResearch) service to preemptively circumvent bot and low quality

[1] By 2015, the percentage of papers in *Journal of Personality and Social Psychology* (*JPSP*)*, Personality and Social Psychology Bulleting,* and *Psychological Science* containing at least one mTurk study was between 20% and 45%, up from less than 10% in 2012 (Anderson et al., 2019; Zhou & Fishbach, 2016). Between 2015 and 2018, more than half of the mTurk studies in *JPSP* used experimental methods (Buhrmester, Talaifar, & Gosling, 2018). While there has been criticism, studies have shown that mTurk is suitable for experimental research where the focus is not on sample representativeness as such, but on treatment effects between groups. Undesirable subject behavior has been demonstrated to be similar between mTurk, campus, and community participants (Necka, Cacioppo, Norman, & Cacioppo, 2016).

responses. Subjects were recruited exclusively from the United States. We followed recommended procedures to exclude responses from subjects who provided unrealistic demographics (household size > 30 and bogus responses; N=4). 39.40% of subjects were female (0.54% *other*), and median age was 33 years (M_{AGE} = 34.84, SD_{AGE} = 9.84).

First, at the start of the survey, participants were asked to provide their demographics, followed by a set of questions about their recent spending on eating out ("When you eat out, how much, on average, do you spend on yourself per meal (in US dollars)?", "How many times did you eat out last week?"). Then, from responses to these questions, subjects' approximate spending in the prior week was calculated and presented to them in the survey. Subjects subsequently had the option to adapt the shown spending amount if desired.

Then, participants were randomly assigned into either a control group (receiving no peer comparison) or into one of three treatment groups with peer comparisons. Participants in the control group were told that they may be compared to peers ("people like you… who share similar characteristics with you, like income and age"), and were then elicited for perceived social identification with this peer group absent of any peer behavior (social identification baseline), using a validated 1-item measure (Postmes, Haslam, & Jans, 2013). In the three treatment groups, the social identification measure was elicited, too, but only after subjects received a comparison to the respective peer group's behavior. The three treatment groups each received a peer comparison with varying levels of relative distance between the subject's and peer group's behavior (small distance, medium distance, large distance; operationalized as a fixed lower percentage of each subject's self-reported spending; 14%, 51%, and 86%, respectively). For example, a subject in the large distance (14% spending, 86% distance) peer group's behavior treatment, who was calculated to spend $50 on eating out each week, would be presented with the

following text (each sentence on a separate line): *Your spending on eating out: $50. Spending of people like you: $7. According to this comparison, you spend $43 more than the people that you have been compared to.*

Next, all participants were asked a behavioral intention measure, that is, whether they plan to decrease, increase, or sustain their spending on eating out, on a 5-point Likert scale (from "greatly decrease", 1, to "greatly increase", 5). Finally, as our manipulation check for our perceived distance between subject's and peer group spending manipulations, subjects were asked whether the distance in spending seemed like a large, moderate, small, or negligible difference to them. This was important, as our intervention treatments utilized relative distance in peer spending from subject's spending, and subject's spending was self-reported, thus the actual absolute distance (in US dollars) may differ between subjects despite identical relative distance. For example, a 86% relative distance in spending would operationalize as a $8.60 absolute distance for a subject with $10.00 expenditure, or a $86 absolute distance for a subject with $100.00 expenditure.

This design allows us to test the causal effect of perceived behavior distance (gap) between subjects' and the peer group's behavior on social identification with the peer group as well as on behavioral intent to converge behavior towards the norm. To isolate these effects, we control for social identification in the analyses. This was necessary, as we expected both a negative effect of behavior distance treatment on behavioral intent through reduced social identification and a positive direct effect of behavior distance treatment on behavioral intent.

Results

Manipulation Check for Perceived Distance

Table 1 shows descriptive statistics for subjects' weekly expense, and the resulting absolute and perceived distance to peer comparison behavior across groups.

Table 1
Descriptives of weekly spend and manipulation check from study 1.

Treatment group	Distance	N	Mean weekly expense (S.D.)	Mean absolute distance (S.D.)	Mean perceived distance (S.D.)
Control	N/A	93	32.25 (71.68)	N/A	N/A
Small distance	14%	89	29.71 (34.79)	4.16 (4.87)	1.97 (0.80)
Medium distance	51%	92	26.71 (30.37)	13.09 (14.88)	2.58 (1.04)
Large distance	86%	94	32.47 (38.62)	27.92 (33.22)	3.06 (1.19)

Notes: This table shows the descriptives of subjects' weekly expense on eating out, and the manipulation for the peer comparison treatment groups. Weekly expense refers to the calculated average expense on eating out * frequency of eating out per week, in US dollars. Absolute distance refers to the effective US dollar difference between peer comparison expense and subject's expense, that is, percentage distance * weekly expense. Perceived distance was measured on a scale from 1 to 4 (negligible, small, moderate, large).

Analysis of variance (ANOVA) tests were conducted to compare subjects' weekly expenses on eating out across groups, and to confirm our manipulation, that is, confirm differences between peer comparison treatments in perceived distance to peer behavior across groups. First, the weekly expenses did not significantly differ between the groups ($F(3,367)=0.30$, $p = .82$). Consequently, the relative distance manipulations in our treatments led to significant differences in the absolute distance ($F(2,274)=28.83$, $p < .001$) and, more importantly, the perceived distance ($F(2,274)=26.16$, $p < .001$) between subjects' and peer comparison expenses among the treatments. Bonferroni post-hoc tests show that the perceived distance between subjects' and peer comparison expenses was indeed largest for the large distance group ($M_{PERdist} = 3.06$, $SD = 1.19$), followed by the medium distance ($M_{PERdist} = 2.58$, $SD = 1.04$) and small distance ($M_{PERdist} = 1.97$, $SD = 0.80$) groups.

Effect of Peer Behavior Distance on Social Identification With Respective Peer Group

See Table 2 for results. An ANOVA shows a significant difference in social identification between the four groups ($F(3,367)=32.82$, $p < .001$). Post-hoc tests show that social identification

was highest for the small distance group (M_{SI} = 5.08, SD = 1.39), followed by the control (M_{SI} = 4.72, SD = 1.35), medium distance (M_{SI} = 3.95, SD = 1.59), and large distance (M_{SI} = 2.97, SD = 1.87) groups. Social identification was significantly different ($p < .01$) between all groups, with the exception of the control and small distance groups ($p = 0.75$). These results support $H1_a$.

Table 2
Results study 1.

Treatment group	Distance	N	Mean social identification (S.D.)
Control	N/A	93	4.72 (1.35)
Small distance	14%	89	5.08 (1.39)
Medium distance	51%	92	3.95 (1.59)
Large distance	86%	94	2.97 (1.87)

Notes: This table shows the results from study 1, that is, the mean social identification score of the same superordinate group of peers, as affected by this peer group's weekly spending distance from subjects' weekly spending (small distance, medium distance, large distance). Social identification was measured on a scale from 1 (fully disagree) to 7 (fully agree).

Effects of Peer Behavior Distance & Social Identification With Peer Group on Behavioral Intent

An ANOVA shows no difference in behavioral intent between the four groups ($F(3,367)=1.32$, $p = .27$). This was expected, as the distance to peer group behavior manipulations should affect behavioral intent (future spending) negatively in a direct effect, and positively in an indirect effect through lowering social identification with the respective group whose behavior has been shared with the subject. Furthermore, as the distance to peer group behavior manipulations were relative, and the absolute distance in US dollars could therefore differ between subjects, the treatment effects are noisy unless we control for our included perceived distance variable. A series of regression models confirm these effects. For an overview, see Table 3. In model 1, including only subjects' perceived distance significantly predicts behavioral intent scores, $\beta = -.25$, $t(274) = -7.42$, $p < .001$; $R^2 = .16$, $F(1, 273) = 55.11$, $p < .001$. In model 2, we show that only social identification did not predict behavioral intent scores. However, in model 3, we show that by

adding subjects' perceived distance to model 2, both perceived distance, $\beta = -.28$, $t(274) = -7.73$, $p < .001$, and social identification, $\beta = -.05$, $t(274) = -2.22$, $p = .027$, significantly predict behavioral intent scores, $R^2 = .18$, $F(1, 272) = 30.42$, $p < .001$. Model 4 confirms that medium and large distance treatments, compared to the small distance treatment, do not predict behavioral intent scores. However, in model 5, by again controlling for perceived distance, we show that the medium, $\beta = .19$, $t(274) = 2.02$, $p = .044$, and large distance, $\beta = .30$, $t(274) = 3.00$, $p = .003$, treatments, which we showed earlier to negatively influence social identification, significantly predict behavioral intent scores, $R^2 = .20$, $F(1, 271) = 21.93$, $p < .001$. These results support $H1_b$ and H2.

Table 3

VARIABLES	(1) BehIntent	(2) BehIntent	(3) BehIntent	(4) BehIntent	(5) BehIntent
Perceived distance	-0.251***		-0.284***		-0.295***
	(0.034)		(0.037)		(0.036)
Social identification		0.020	-0.049**		
		(0.022)	(0.022)		
Medium distance treatment				0.011	0.191**
				(0.102)	(0.094)
Large distance treatment				-0.025	0.299***
				(0.101)	(0.099)
Constant	3.296***	2.577***	3.576***	2.663***	3.242***
	(0.094)	(0.098)	(0.157)	(0.073)	(0.097)
Observations	275	275	275	275	275
Adjusted R-squared	0.165	-0.001	0.177	-0.007	0.186

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: BehIntent = reverse-coded (1=low, 4=high). Observations (N=275) are restricted to peer comparison treatments.

Discussion and Conclusion

Our results show that a larger perceived distance between subjects' behavior and reference group behavior lowers subjects' social identification with this reference group. When controlling for perceived distance, we find that lower social identification weakens the intention to converge to reference group behavior. In terms of main effects, our treatments did not produce a net main effect. However, this was not unexpected, as each treatment produces both a positive direct effect of increased behavioral distance on intent to converge to the social norm, as well as an indirect negative effect by lowering social identification with the reference group. Indeed, when controlling for perceived distance, we do find a significant net oppositional reaction of the medium and large distance treatments on intent to converge to reference group behavior, compared to the small distance treatment (see Table 3, models 4 and 5). Our assumption is that social identification must be even lower for a net oppositional reaction to occur, as our baseline social identification is high, and, despite significant reduction through manipulations, there is room for further reduction. Additionally, it may help to study this phenomenon in a context where the optimal direction is not necessarily known. With expenses, saving is well-known to be the *correct* behavior, thus going against this trend, i.e. oppositional reaction of saving less, due to lower social identification or otherwise, is less likely. In contexts where there is more ambiguity with regard to correct behavior, a net oppositional reaction is likely easier to induce. Future studies should examine whether social norm interventions with lower baseline peer group social identification and more ambiguous desirable behaviors indeed result in net oppositional reactions (i.e. as seen in Beshears et al., 2015).

For managers and policymakers, it is crucial to be mindful of the perceived distance between actual behavior of individuals targeted for a social norm intervention and the salient (desirable) behavior of the peer group used in the intervention. If the expected behavioral distance is high,

then it is crucial to ensure a high baseline social identification with the respective peer group to preemptively counter the consequences of a likely reduction in social identification.

References (Endnote – will be manually transferred and copy-edited @ final draft)

Ahler, D. J., Roush, C. E., & Sood, G. (2019). *The Micro-Task Market for Lemons: Data Quality on Amazon's Mechanical Turk.* Paper presented at the Meeting of the Midwest Political Science Association.

Allcott, H. (2011). Social norms and energy conservation. *Journal of public Economics, 95*(9), 1082-1095.

Anderson, C. A., Allen, J. J., Plante, C., Quigley-McBride, A., Lovett, A., & Rokkum, J. N. (2019). The MTurkification of social and personality psychology. *Personality and social psychology bulletin, 45*(6), 842-850.

Azmat, G., & Iriberri, N. (2010). The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of public Economics, 94*(7), 435-452.

Batra, R., Homer, P. M., & Kahle, L. R. (2001). Values, susceptibility to normative influence, and attribute importance weights: A nomological analysis. *Journal of Consumer Psychology, 11*(2), 115-128.

Bearden, W. O., Netemeyer, R. G., & Teel, J. E. (1989). Measurement of consumer susceptibility to interpersonal influence. *Journal of consumer research*, 473-481.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., & Milkman, K. L. (2015). The effect of providing peer information on retirement savings decisions. *The Journal of Finance, 70*(3), 1161-1201.

Bobek, D. D., Hageman, A. M., & Kelliher, C. F. (2013). Analyzing the role of social norms in tax compliance behavior. *Journal of Business Ethics*, 1-18.

Buhrmester, M. D., Talaifar, S., & Gosling, S. D. (2018). An evaluation of Amazon's Mechanical Turk, its rapid rise, and its effective use. *Perspectives on psychological science, 13*(2), 149-154.

Chmielewski, M., & Kucker, S. C. (2020). An MTurk crisis? Shifts in data quality and the impact on study results. *Social Psychological and Personality Science, 11*(4), 464-473.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology, 55*, 591-621. doi:10.1146/annurev.psych.55.090902.142015

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*(3), 472-482.

Hoaglin, D. C. (2003). John W. Tukey and data analysis. *Statistical Science*, 311-318.

Hogg, M. A. (2006). Social identity theory. *Contemporary social psychological theories, 13*, 111-1369.

Leach, C. W., Van Zomeren, M., Zebel, S., Vliek, M. L., Pennekamp, S. F., Doosje, B., . . . Spears, R. (2008). Group-level self-definition and self-investment: A hierarchical (multicomponent) model of in-group identification. *Journal of personality and social psychology, 95*(1), 144.

Necka, E. A., Cacioppo, S., Norman, G. J., & Cacioppo, J. T. (2016). Measuring the prevalence of problematic respondent behaviors among MTurk, campus, and community participants. *PloS one, 11*(6), e0157732.

Panagopoulos, C., Larimer, C. W., & Condon, M. (2012). *Social pressure, descriptive norms, and voter mobilization*. Paper presented at the APSA 2012 Annual Meeting.

Park, C. W., & Lessig, V. P. (1977). Students and housewives: Differences in susceptibility to reference group influence. *Journal of consumer research, 4*(2), 102-110.

Phua, J. J. (2013). The reference group perspective for smoking cessation: An examination of the influence of social norms and social identification with reference groups on smoking cessation self-efficacy. *Psychology of Addictive Behaviors, 27*(1), 102.

Postmes, T., Haslam, S. A., & Jans, L. (2013). A single-item measure of social identification: Reliability, validity, and utility. *British Journal of Social Psychology, 52*(4), 597-617.

Reed, M. B., Lange, J. E., Ketchie, J. M., & Clapp, J. D. (2007). The relationship between social identity, normative information, and college student drinking. *Social Influence, 2*(4), 269-294.

Ridout, B., & Campbell, A. (2014). Using Facebook to deliver a social norm intervention to reduce problem drinking at university. *Drug and alcohol review, 33*(6), 667-673.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological science, 18*(5), 429-434.

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology, 33*(1), 1-39.

Thomas, C., & Sharp, V. (2013). Understanding the normalisation of recycling behaviour and its implications for other pro-environmental behaviours: A review of social norms and recycling. *Resources, Conservation and Recycling, 79*, 11-20.

Van Der Linden, S. (2015). Exploring beliefs about bottled water and intentions to reduce consumption: The dual-effect of social norm activation and persuasive information. *Environment and Behavior, 47*(5), 526-550.

Verhallen, P., Brüggen, E., Post, T., & Odekerken-Schröder, G. (2020). *Friends with benefits: Strengthening peer effects through aligning consumer traits with reference group attributes*. Maastricht University.

Zhou, H., & Fishbach, A. (2016). The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of personality and social psychology, 111*(4), 493.