The power of regression to the mean: A social norm study revisited

3 AUTHORS:

Kirsten T Verkooijen
Wageningen University
20 PUBLICATIONS   **118** CITATIONS

Marijn Stok
Universität Konstanz
22 PUBLICATIONS   **185** CITATIONS

Saar Mollen
University of Amsterdam
7 PUBLICATIONS   **16** CITATIONS

FAST TRACK REPORT

The power of regression to the mean: A social norm study revisited

Kirsten T. Verkooijen*, F. Marijn Stok†,‡ & Saar Mollen§

* Health and Society Group, Wageningen University and Research Centre, Wageningen, The Netherlands
† Department of Psychological Assessment and Health Psychology, University of Konstanz, Konstanz, Germany
‡ Department of Clinical and Health Psychology, Utrecht University, Utrecht, The Netherlands
§ Amsterdam School of Communication Research, University of Amsterdam, Amsterdam, The Netherlands

Correspondence
Kirsten T. Verkooijen. Health and Society Group, Wageningen University and Research Centre, P.O. Box 8130, 6700 EW Wageningen, The Netherlands.
E-mail: kirsten.verkooijen@wur.nl

Abstract

This research follows up on a study by Schultz et al. (2007), in which the effect of a social norm intervention on energy consumption was examined. The present studies included control groups to examine whether social norm effects would persist beyond regression to the mean. Both studies had a 2 (baseline consumption: below mean versus above mean) × 2 (message condition: no-message control versus norm message) design. Based on baseline fruit (Study 1) or unhealthy snack (Study 2) consumption, students were classified as above mean or below mean for consumption. One week later, half of the students in the above-mean and below-mean groups received normative feedback; control groups did not. Neither study showed an effect of norm messages on behavior relative to control, providing evidence for regression to the mean as an alternative explanation. Findings highlight the importance of control groups to distinguish social norm intervention effects from mere regression to the mean.

THE POWER OF REGRESSION TO THE MEAN: A SOCIAL NORM STUDY REVISITED

Social norms have been found to affect a number of different types of health behaviors, ranging from fruit consumption (Stok, De Ridder, De Vet, & De Wit, 2012) to cancer screening (Sieverding, Decker, & Zimmerman, 2010), exercise (Burger & Shelton, 2011), and alcohol consumption (Larimer, Turner, Mallett, & Geisner, 2004). The use of social norm messages to reduce alcohol consumption in US colleges has become so popular that about 48% of all four-year residential colleges and universities in the USA have tried this approach (Wechsler, 2003). The rationale behind the "social norms approach" (Perkins & Berkowitz, 1986) to reduce alcohol consumption is that people tend to overestimate the extent to which their peers consume alcohol (Borsari & Carey, 2003). By disseminating information about the actual prevalence of alcohol consumption, norm perceptions are adjusted to a lower, more realistic standard. These new, more realistic norm perceptions should, in turn, guide behavior and reduce alcohol consumption (Perkins &

Berkowitz, 1986). Results from these interventions, however, have been mixed; while some have proven to be successful (e.g., DeJong et al., 2006; Turner, Perkins, & Bauerle, 2008), others have failed to encourage significant changes in alcohol consumption (e.g., Clapp, Lange, Russell, Shillington, & Voas, 2003; Granfield, 2005).

A reason offered for the mixed results of social norms campaigns is that while they may have a positive effect on the people who consume more alcohol than the average college student, they may cause those who consume less than average to increase their intake, resulting in an unwanted boomerang effect (Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007). Schultz et al. (2007) researched this assumption. They conducted a field study in which normative messages were used to promote household energy conservation. In short, households received information about how much energy they had used in the previous week along with information about the energy consumption of the average household in their neighborhood (i.e., descriptive norm; see, e.g., Reno, Cialdini, & Kallgren, 1993). Energy readings in the following weeks showed a

decrease in energy consumption for households that consumed more than average at baseline, and as expected, an increase in energy use was found in households that were below the mean on baseline energy consumption. The latter finding indicates a boomerang effect. This boomerang effect did not occur when an injunctive norm, i.e., approval or disapproval of others for low or high energy consumption, respectively, was added to the descriptive norm message. According to the authors, the findings provide evidence that descriptive norms can cause unfavorable deviants to behave in a more favorable way, but may also cause favorable deviants to act less favorably (i.e., "boomerang effect").

However, the no-control group design of the study conducted by Schultz et al. (2007) leaves room for an alternative explanation, more specifically regression to the mean (Krueger, 2011). Regression to the mean is the phenomenon that if a variable is extreme on its first measurement, it will tend to be closer to the average on its second measurement (Stigler, 1997). Hence, the decrease in energy consumption among households that were categorized as high energy consumers and the increase in energy consumption among previously low-energy-consuming households might have been due to normal fluctuations in energy consumption. One way to rule out this explanation is to include control groups in one's study design that receive no intervention. Given the significant theoretical and applied impact of the results, a conceptual replication and extension of the study by Schultz et al. (2007) is deemed desirable. The present research aimed to examine whether social norm effects would persist beyond mere regression to the mean by including two control conditions (i.e., above-mean control and below-mean control) in the study design. Instead of energy consumption, the current studies focused on eating behavior, which aligns with a large body of work on social norms and health behavior (e.g., Mollen, Rimal, Ruiter, & Kok, 2013; Rimal, 2008; Scholly, Katz, Gascoigne, & Holck, 2005; Stok, de Ridder, de Vet, & de Wit, 2014). While Study 1 pertains to healthy eating behavior (i.e., fruit consumption), Study 2 focuses on unhealthy eating behavior (i.e., snack consumption).

STUDY 1

To study whether the effects of normative feedback interventions can potentially be ascribed to statistical regression to the mean, university students, in this study, will initially report daily fruit consumption for three days (i.e., baseline measure). Subsequently, half of the participants will receive a social norm message, indicating that their consumption is either below or

above average (i.e., depending on their baseline consumption), whereas the other half of the participants receive no such message. Fruit consumption will then again be assessed on three consecutive days.

We expect that fruit consumption during this follow-up will show a regression to the mean for the control group participants (i.e., those who are below the mean at baseline will increase consumption, and those who are above the mean at baseline will decrease consumption), and aim to establish whether there is an effect of the social norm messages above and beyond this regression to the mean effect.

Method

Participants

Eligible for participation were Dutch-speaking students enrolled at an applied or regular university in the Netherlands, who were not restricted in their fruit consumption for medical reasons. Participants were recruited via Facebook (Facebook, Inc., Menlo Park, CA, USA) and e-mail. As an incentive, participants could win one of 15 gift vouchers of 15 euro. Baseline measures were filled out by 173 students. Four students did not complete the follow-up measurement, and six students were excluded because of an exceptionally high fruit consumption at baseline and/or at follow-up (>3 standard deviations (SD) above the mean), leaving 163 students in the final sample[1] (127 women; Mage = 21.90; SDage = 5.46). A power analysis using G*power 3.0.10 (Faul, Erdfelder, Lang, & Buchner, 2007) revealed that, given the observed correlation between fruit consumption at baseline and follow-up (r = .60), a sample of 60 participants would be needed to detect interactions for a medium effect (f = .25) and an alpha and beta error of .05 each. Ethical approval was obtained from the ethical committee of the first author's institution.

Design

The design was a 2 (fruit consumption: below average versus above average) × 2 (message condition: descriptive norm versus control) × 2 (time: baseline versus

---

[1]Retaining outliers in the sample did not change any outcomes of the analyses in Study 1. For Study 2, retaining outliers in the sample did not change the outcomes of the analysis of variance, but did change the outcome of one pair-wise comparison; the decrease in snack consumption among students in the control condition who scored initially above the mean for unhealthy snack consumption was no longer significant when outliers remained in the sample (t(26) = 1.89, p = .070 vs t(23) = 2.24, p = .035).

follow-up) mixed-factorial design. Fruit consumption and message condition were between-participants factors, and time was a within-participant factor. Participants were randomly assigned to either the descriptive norm or no-message control condition.

Procedure

Data collection took place within a period of 4 weeks and consisted of a series of eight online questionnaires. The first questionnaire was distributed over a period of 2 weeks in order to recruit sufficient participants. All other questionnaires were administered on the same days for all participants. Participants had to provide informed consent before being able to continue with the first questionnaire. This questionnaire involved demographic variables (i.e., sex, age, and educational program) and socio-cognitive variables (i.e., attitude, self-identity, self-efficacy, social identity, and intention towards fruit consumption)[2]. Directly after the 2-week recruitment period, three questionnaires followed that assessed participants' baseline fruit consumption on three consecutive days. Six days later, the fifth questionnaire followed, which contained the intervention. Participants were randomly assigned to either the experimental or control condition.

Participants in the experimental conditions were given a descriptive norm message, which consisted of information on the fruit consumption of the average study participant (based on actual reports in the baseline questionnaires) together with the information that the participant's own fruit consumption was above or below this average (also based on actual reports). The text reads: "Experience has taught us that study participants like the opportunity to compare their own 'scores' with those of others. That is why we provide some more information on this. Our results of last week show that the average student consumes 2.56 portions of fruit per day. Based on your specific answers, it appears that you eat [more/less] than the average student." To ensure that the norm message was read, participants were subsequently asked the question: "Do you find it surprising that your fruit consumption is [above/below] the consumption of the average student?" (1 = *yes, very surprising* to 5 = *no, not at all surprising*). Hereafter, participants were asked some questions measuring socio-cognitive variables (i.e., intention and self-efficacy)[2]. Participants in the control conditions did not receive any descriptive norm information and were only asked to answer the questions measuring socio-cognitive variables. Six

days after this questionnaire, the same three questionnaires followed for all participants, assessing fruit consumption on three consecutive days at follow-up.

Measurement of Fruit Consumption

Fruit consumption at baseline and at follow-up was assessed in the same way. On three consecutive week-days, starting on Tuesday, participants were asked to report their fruit consumption of the previous day. A list of 26 fruits was provided on which participants could indicate which type of fruit they had consumed, along with the amount (in pieces for larger fruits such as apples and in handfuls for very small fruits such as grapes). In addition, three "other" options were provided where participants could indicate the type and amount of fruits that were not on the list. Consumption was calculated by computing the total amount of portions of fruit consumed, following nutritional guidelines. For very small fruits (e.g., berries), one handful constitutes one portion. For smaller fruits (e.g., prunes), two or three pieces constitute one portion. For normal-size fruits (e.g. apples), one piece constitutes one portion. For large fruits (e.g., melons), parts of the fruit constitute one portion. Hereafter, total consumption was divided by the number of reported days in order to obtain average daily consumption. When reports were missing for one ($n = 8$) or two ($n = 1$) days, fruit consumption was calculated on the basis of the remaining days.

Mean baseline fruit consumption was computed and used to divide the sample into a group that scored below and a group that scored above the fruit consumption mean. Because of time constraints, this division was based on a rather rough calculation of the fruit consumption mean ($M = 2.56$). Because we used a more precise calculation of fruit consumption in the results, the final baseline mean ($M = 2.31$) differed somewhat from this mean. Despite the fact that fruit consumption was positively skewed (skewness = 1.39), the division resulted in two almost equal sample sizes ($n = 83$ and $n = 80$). Hereafter, a random selection of almost half of the students in each consumption group were assigned to the control conditions ($n_{above-mean} = 36$ and $n_{below-mean} = 38$), while the other students were assigned to the experimental conditions ($n_{above-mean} = 44$ and $n_{below-mean} = 45$).

Results

Baseline Fruit Consumption

The average baseline fruit consumption of the final sample was 2.31 portions ($SD = 1.71$) per day. A

---

[2]Because the main focus was on the behavioural outcomes, no further analyses were conducted on the socio-cognitive variables that were assessed.

randomization check revealed no significant differences in demographics, socio-cognitive variables, or baseline fruit consumption between students in the experimental and in the control conditions ($ps > .43$). Further, students in the experimental condition who received the message that they consumed less fruit than the average student were equally (un)surprised by this information as students who received the message that they ate more fruit than the average student ($M_{\text{below-mean}} = 3.33$, $SD = 1.26$, $M_{\text{above-mean}} = 3.55$, $SD = .98$, $t(89) = 0.89$, $p = .378$).

Change in Fruit Consumption

A repeated measures analysis of variance (ANOVA) revealed a significant main effect of baseline fruit consumption (below versus above mean at baseline) ($F(1, 159) = 98.97$, $p < .001$, $\eta_p^2 = .38$), and an interaction effect between time and baseline fruit consumption ($F(1, 159) = 37.03$, $p < .001$, $\eta_p^2 = 0.19$), showing that those who consumed below average at baseline increased consumption at follow-up, and those who consumed above average at baseline decreased consumption at follow-up. If the social norm messages would affect behavior above and beyond mere
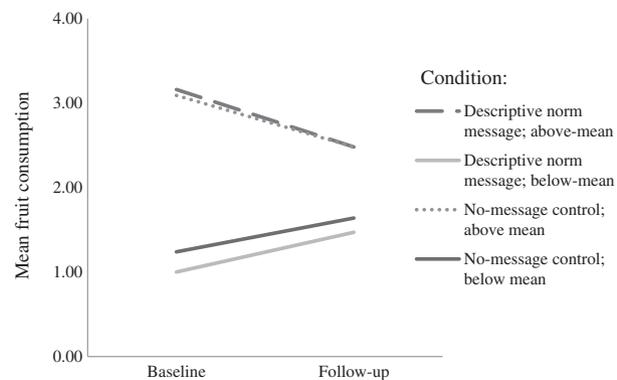

Fig. 1: Fruit consumption (portions/day) at baseline and at follow-up among the four conditions

regression to the mean effects, a significant interaction between norm message and time and/or baseline consumption would be expected. However, message condition (descriptive norm versus no-message control) showed no main effect ($F(1, 159) = 0.17$, $p = .679$), nor interaction effect with either baseline fruit consumption or time ($ps > .5$), nor was there a three-way interaction between message condition, time, and baseline fruit consumption ($F(1, 159) < 0.02$, $p = .885$). This indicates that there was no effect of the social norm messages beyond mere regression to the mean.

Figure 1 indicates that, both in the descriptive norm and in the control condition, average fruit consumption decreased among students who scored above the mean for fruit consumption at baseline, whereas it increased among students who scored below the mean for fruit consumption at baseline. Pair-wise comparisons confirmed these observations (Table 1). A significant decrease in fruit consumption, relative to baseline, was found among students whose fruit consumption was initially above the mean and who either received the descriptive norm message or no message. In contrast, students who at baseline consumed less fruit than the reported average showed a significant increase in fruit consumption at follow-up. Again, the same changes in fruit consumption were apparent in the descriptive norm message condition and the no-message control condition.

Additional Analysis

To test whether more extreme scores at baseline moderated the effect of time, an observation that would strengthen the regression to the mean explanation, an extra between-subject variable was added to the model that splits the sample up into four, instead of two, baseline consumption groups, namely those scoring more than 1 SD above the mean ($n = 22$), those scoring less than 1 SD above the mean ($n = 49$), those scoring less than 1 SD below the mean ($n = 64$), and those scoring more than 1 SD below the mean ($n = 28$). The results indeed showed a significant interaction between

Table 1. Pair-wise comparisons of fruit consumption for all four conditions

Condition	Baseline		Follow-up		Mean difference	95% CI of mean difference	t	p	Cohen's d
	M	SD	M	SD					
Norm message—above-mean	3.16	1.33	2.48	1.40	0.680	0.257–1.103	3.24	.002	0.47
No-message—above mean	3.09	0.96	2.48	1.17	0.605	0.285–0.925	3.84	< .001	0.66
Norm message—below mean	1.10	0.63	1.47	1.00	−0.374	−0.645 to −0.103	−2.78	.008	0.44
No-message—below mean	1.24	0.84	1.64	0.94	−0.40	−0.712 to −0.088	−2.60	.013	0.42

Note: M, mean; SD, standard deviation; CI, confidence interval.

time and extremity at baseline ($F(3, 156) = 9.53$, $p < .001$, $\eta p^2 = .160$). Extremity at baseline also showed a main effect on fruit consumption ($F(3, 156) = 22.18$, $p < .001$, $\eta p^2 = 0.31$), and a three-way interaction with time and baseline fruit consumption ($F(3,156) = 4.48$, $p = .013$, $\eta p^2 = 0.06$). Message condition did not show a main nor an interaction effect. A graphic representation of the significant interactions revealed that the more extreme scores showed a greater shift towards the mean compared with the less extreme scores, which was particularly the case for the above-mean consumption group.

Discussion

Results of Study 1 showed that, while descriptive norm messages produced reductions in consumption for high fruit consumers, and an increase among low consumers, an identical pattern was evident for high and low consumers in the no-message control conditions. Hence, the changes found in consumption could not be attributed to the norm messages. Instead, they appear to be the result of statistical regression to the mean. This interpretation was strengthened by our finding that people with more extreme consumption levels at baseline showed a stronger regression to the mean than people with moderate baseline scores. However, it is possible that our results deviate from previous findings because the aim of the norm messages used in our study was to increase the performance of desirable behavior (i.e., fruit consumption), while in the study by Schultz et al. (2007), the norm messages were aimed at reducing undesirable target behavior. In Study 2, we therefore aim to replicate our findings using an unhealthy target behavior.

STUDY 2

Promoting healthy behavior is not necessarily the same as discouraging unhealthy behavior (Adriaanse, Vinkers, de Ridder, Hox & De Wit, 2011; Holland, Aarts, & Langendam, 2006). Previous research has shown that the associations between social norms and eating behavior differ according to whether the norm promotes healthy eating behavior or discourages unhealthy eating behavior (Stok et al., 2015). Because the effects of a normative feedback intervention might thus be different depending on whether they target healthy or unhealthy behavior, a second study was conducted that targets unhealthy behavior, namely unhealthy snack consumption. We tested whether, like in Study 1, there would be no difference between the experimental and control conditions in terms of changes in unhealthy snack consumption from baseline to follow-up.

Methods

Participants

Participants for Study 2 were recruited in a similar vein to those in Study 1. This resulted in 144 students who filled out the baseline measures. Hereof, 20 students did not complete the follow-up measures. In addition, five students were excluded because their unhealthy snack consumption exceeded the mean by 3 *SD* at baseline and/or follow-up[1]. Hence, the final sample consisted of 119 students (103 women, $M_{age} = 21.59$, $SD = 3.76$). According to a power analysis (G*Power 3.0.10; Faul et al., 2007), a sample of 92 participants would be required, given the correlation between unhealthy snack consumption at baseline and at follow-up ($r = .37$), to detect medium-size interactions ($f = .25$) with an alpha and beta error of .05 each.

Design and Procedure

The design and procedures were identical to Study 1, except for the socio-cognitive variables, attitude, self-identity, and self-efficacy, which were not assessed in Study 2 (intention, however, was assessed again in this study[2]). In addition, a question was included at the end of the survey that assessed whether participants correctly recalled the contents of the feedback they had received earlier (response options were "I consumed more unhealthy snacks than the average student"; "I consumed fewer unhealthy snacks than the average student"; and "I don't know"). Finally, unhealthy snack consumption was assessed, and communicated in the norm message, instead of fruit consumption.

Measurement of Unhealthy Snack Consumption

First, a brief definition of a snack was given (i.e., "food that is eaten between meals, and not as part of one's breakfast, lunch or dinner"). Hereafter, unhealthy snacking was self-reported using a previously developed snack diary that has been used in earlier social psychological research (Adriaanse, De Ridder, & De Wit, 2009). Participants were asked to indicate the type of snack(s) that they had eaten and the amount. For this purpose, two lists were presented, one captioned "unhealthy snacks," which included 13 types of unhealthy snacks (e.g., small cookie, candy bar, and chips), and one captioned "healthy snacks" including 10 types of healthy snacks (e.g., yogurt, cracker, and bread

product). Both lists also included "other" options where types and consumed amounts of snacks that were not listed could be provided. Only unhealthy snack consumption was analyzed and communicated in the norm message. The measure of healthy snack consumption was only included to maintain the cover story (the research was introduced as a survey on snack consumption in general). Also, by including a list of healthy snacks, participants could obtain a better idea of what snacks were considered as healthy or unhealthy.

Like in Study 1, participants were asked to report their consumption of the previous day for three consecutive weekdays at baseline and at follow-up. The total number of unhealthy snacks consumed was calculated in a similar way as in Study 1. For instance, a small cookie counted as 0.5 portion and a big cookie or slice of cake as 1 portion. Average consumption at baseline and at follow-up was computed by dividing the total snack consumption by the number of reported days. Eighteen participants showed missing data for 1 day, and 12 participants had missing data for 2 days.

Again, an initial rough calculation of the mean baseline unhealthy snack consumption was performed to divide the sample into a group that scored below and a group that scored above this mean ($M = 2.05$). A more precise calculation of unhealthy snack consumption resulted in a final baseline mean of 1.67 snacks per day. Unhealthy snack consumption was positively skewed (skewness $= 1.32$), which resulted in 51 students who scored above the mean at baseline, and 68 students who scored below the mean at baseline. Hereafter, half of the students in each consumption group were randomly assigned to the control conditions ($n_{\text{above-mean}} = 24$ and $n_{\text{below-mean}} = 31$), while the other half were assigned to the experimental conditions ($n_{\text{above-mean}} = 27$ and $n_{\text{below-mean}} = 37$).

Results

Baseline Unhealthy Snack Consumption

The average baseline unhealthy snack consumption of the final sample was 1.67 portions per day ($SD = 1.04$). A randomization check revealed no significant differences in age ($p = .901$), gender ($p = .191$), or baseline snack consumption ($p = .512$) between students in the experimental and in the control conditions. Also, students who received the message that they consumed fewer unhealthy snacks than the average student were equally surprised as those who received the message that they consumed more unhealthy snacks ($M = 3.49$, $SD = 1.02$ and $M = 3.19$, $SD = 1.18$, respectively; $t(62) = 1.10$, $p = .278$). Most students (89%) in the
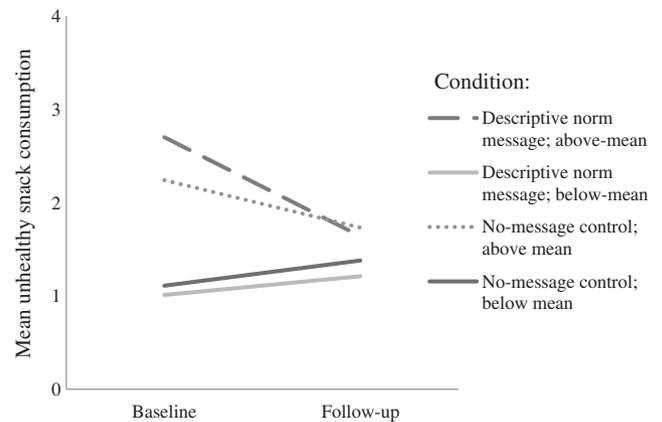

Fig. 2: Unhealthy snack consumption (portions/day) at baseline and at follow-up among the four conditions

experimental conditions correctly recalled the norm message they had received.[3]

Change in Unhealthy Snack Consumption

An ANOVA for repeated measures revealed a significant main effect of baseline unhealthy snack consumption ($F(1, 115) = 40.62$, $p < .001$, $\eta p^2 = .26$), and an interaction between time and baseline unhealthy snack consumption ($F(1, 115) = 26.79$, $p < .001$, $\eta p^2 = 0.19$). Again, there was no main effect of message condition (descriptive norm versus no-message control; $F(1, 115) = 0.03$, $p = .861$), no interaction of message condition with either baseline unhealthy consumption or time ($ps > .12$), nor a three-way interaction ($F(1, 115) = 1.44$, $p = .233$).

Figure 2 and Table 2 show that, both in the descriptive norm and in the control condition, average snack consumption decreased among students who scored above the mean for unhealthy snack consumption at baseline. In contrast, students, both in the descriptive norm condition and in the control condition, who at baseline consumed fewer unhealthy snacks than the average student, did not show a significant change in unhealthy snack consumption.

Additional Analysis

Again, an additional ANOVA was conducted to test whether initial extremity moderated the effect of time. Splitting the sample into four baseline snack

[3]Three students recalled the message incorrectly, one student did not recall the contents of the message, and a further three did not answer this question. Excluding these seven students from the data analyses did not change any of the outcomes.

Table 2. Pair-wise comparisons of unhealthy snack consumption for all four conditions

Condition	Baseline		Follow-up		Mean difference	95% CI of mean difference	t	p	Cohen's d
	M	SD	M	SD					
Norm message—above-mean	2.70	0.93	1.64	1.26	1.062	0.483–1.519	4.45	.000	0.88
No-message—above mean	2.24	0.63	1.73	0.97	0.511	0.038–0.984	2.24	.035	0.47
Norm message—below mean	1.01	0.79	1.21	0.91	−0.198	−0.525–0.025	−1.58	.123	0.27
No-message—below mean	1.11	0.61	1.38	1.16	−0.275	−0.702–0.152	−1.32	.198	0.25

Note: *M*, mean; *SD*, standard deviation; CI, confidence interval.

consumption groups resulted in 19 students who scored more than 1 *SD* above the mean, 34 students who scored less than 1 *SD* above the mean, 47 students who scored less than 1 *SD* below the mean, and 19 students who scored more than 1 *SD* below the mean. The results showed again a significant main effect of extremity at baseline ($F(3, 116) = 15.21$, $p < .001$, $\eta p^2 = 0.31$) and an interaction effect of time with extremity ($F(3, 116) = 3.72$, $p = .014$, $\eta p^2 = .10$). Students with more extreme scores showed a greater shift towards the mean compared with those with a less extreme score. No other main or interaction effects were significant.

Discussion

Study 2 again showed no effect of norm messages on students' eating behavior. Students with an above average unhealthy snack consumption at baseline showed a lower consumption at follow-up, irrespective of whether students received a descriptive norm message or not. The experimental and control conditions also showed the same pattern for students whose snack consumption was below average at baseline. Students in the descriptive norm condition, as well as the no-message control condition showed a nonsignificant increase in unhealthy snack consumption. Similar to Study 1, participants with more extreme consumption scores at baseline showed a larger regression to the mean than participants who scored closer to the mean at baseline.

GENERAL DISCUSSION

This research aimed to conceptually replicate and extend a study by Schultz *et al.* (2007), in which observed changes in energy consumption over time were attributed to a social norm intervention. The interpretation of their results has been challenged, as statistical regression to the mean could provide an alternative explanation for the change in energy consumption (Krueger, 2011). To be able to distinguish social norm

effects from regression to the mean effects, the present two studies included no-message control conditions.

The current studies did not find an effect of descriptive norm messages on behavior, as changes in fruit consumption and unhealthy snack consumption were observed, irrespective of whether participants were in the descriptive norm or no-message control condition. Moreover, effect sizes of these changes were mostly comparable across the experimental and control conditions. Additional support for interpreting the results as regression to the mean was found in the larger regression to the mean effects in participants whose baseline consumption was far away from the mean, as compared with participants with baseline scores that were closer to the mean. These findings underscore the importance of using strong research designs, such as adding control groups to longitudinal study designs. Although this does not automatically mean that regression to the mean fully explains the results by Schultz *et al.* (2007), results from that study should be interpreted with caution. Our findings suggest that communicating descriptive norms and comparing individual fruit or unhealthy snack consumption to average norms may not be an effective way to influence these types of eating behavior among students.

Other studies that have also used no-message control group designs, however, have found support for descriptive norm messages on behavior (e.g., Allcott, 2011; Mollen, Rimal, Ruiter & Kok, 2013; Schultz, 1999; Stok *et al.*, 2012; Stok *et al.*, 2014). This shows that the effect of social norms may depend on the context, such as the behavior and target group under study, or on the exact formulation of the norm message that is communicated. The lack of an effect of descriptive norms on behavior in the current studies might be related to the formulation of the norm that was communicated. In the current studies, as well as in the study by Schultz *et al.* (2007), the norm message contained a social comparative element (i.e., "you consume [more/less] [fruit/unhealthy snacks/energy] than the average [student/household]"), whereas in many studies, only a descriptive norm (i.e., "most people do X")

is communicated (e.g., Goldstein, Cialdini, & Griskevicius, 2008; Mollen, Rimal, Ruiter, Jang, & Kok, 2013; Nolan, Schultz, Cialdini, Goldstein, & Griskevicius, 2008; Stok et al., 2012). According to Cialdini and Goldstein (2004), conformity to social norms may serve three goals: accuracy, affiliation, and the maintenance of a positive self-concept. Whereas a simple descriptive norm message may influence behavior mainly through people's desire to be accurate, a social comparative norm message might activate thoughts related to social relationships and self-concept. Because the behaviors that result in social approval might vary between different referent groups, the effects of descriptive norms with a social comparative element may not be as clear-cut as descriptive norm messages that merely communicate what most others do. This may also explain why adding an injunctive norm to a social comparative element in the study by Schultz et al. (2007) had a favorable impact on energy consumption (i.e., the tendency for low-energy-use households to increase their energy consumption over time disappeared). Injunctive norms clearly state what behavior is socially approved of. Although the present research did not address injunctive norm effects, future research is needed to explore possible differences in effects between descriptive norm messages that do include and do not include social comparative aspects, as well as injunctive norms.

LIMITATIONS

Fruit and snack consumption in our research were assessed with a self-report measure, which is less reliable than an objective measure like the household energy reader that was used by Schultz et al. (2007). Self-report measures may also induce a question–behavior effect—asking questions about behavior may change the behavior in question (Dholakia, 2010; Sherman, 1980). However, this usually leads to bias in a socially normative direction: an increase in performance of socially desirable actions, and a reduction of performance of socially undesirable behavior (e.g., Sprott, Spangenberg, & Fisher, 2003). Instead, our studies showed increases as well as decreases in both desirable and undesirable behavior, depending on participants' baseline consumption levels and irrespective of whether norm information was communicated or not. Furthermore, a meta-analysis found the question–behavior effect on health-related behavior to be rather small (Rodrigues, O'Brien, French, Glidewell, & Sniehotta, 2014).

Another issue is that the present results rely on the absence of a difference (i.e., no difference between experimental and control condition), which is usually more difficult to defend than the presence of a difference, because absence of evidence does not necessarily mean that there is no effect. One could argue that more participants or a stronger experimental manipulation could cause statistically unreliable differences in our study to become reliable. Yet, according to the power calculations, the sample sizes of both studies were sufficient to detect meaningful differences. In addition, although the experimental manipulation may not have been very strong (i.e., consisting of one sentence presented at one time), most participants did remember the message at time of the final questionnaire.

CONCLUSION

The present results illustrate the importance of including control groups in longitudinal designs in order to distinguish social norm effects from statistical artifacts. Given the many unsolved questions regarding the social norm approach, combined with the high practical relevance of this approach for health and other societally relevant behaviors, we suggest it is imperative that additional research on the social norm approach and the potential for so-called boomerang effects is conducted. Future research should employ methodologically strong, controlled, and preferably longitudinal designs, allowing a clear distinction between experimental effects and statistical artifacts.

ACKNOWLEDGEMENTS

REFERENCES

Adriaanse, M. A., De Ridder, D. T. D., & De Wit, J. B. F. (2009). Finding the critical cue: Implementation intentions to change one's diet work best when tailored to personally relevant reasons for unhealthy eating. *Personality and Social Psychology Bulletin, 35*, 60–71.

Adriaanse, M. A., Vinkers, C. D. W., de Ridder, D. T. D., Hox, J. J., & De Wit, J. B. F. (2011). Do implementation intentions help to eat a healthy diet? A systematic review and meta-analysis of the empirical evidence. *Appetite, 56*, 183–193.

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics, 95*, 1082–1095. http://dx.doi.org/10.1016/j.jpubeco.2011.03.003.

Borsari, B., & Carey, K. B. (2003). Descriptive and injunctive norms in college drinking: a meta-analytic integration. *Journal of Studies on Alcohol, 64*, 331–341.

Burger, J. M., & Shelton, M. (2011). Changing everyday health behaviors through descriptive norm manipulations. *Social Influence, 6*, 69–77.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology, 55*, 591–621.

Clapp, J. D., Lange, J. E., Russell, C., Shillington, A., & Voas, R. B. (2003). A failed norms social marketing campaign. *Journal of Studies on Alcohol, 64*, 409–414.

Granfield, R. (2005). Alcohol use in college: Limitations on the transformation of social norms. *Addiction Research and Theory, 13*, 281–292.

DeJong, W., Schneider, S. K., Towvim, L. G., Murphy, M. J., Doerr, E. E., Simonsen, N. R., Mason, K. E., & Scribner, R. A. (2006). A multisite randomized trial of social norms marketing campaigns to reduce college student drinking. *Journal of Studies on Alcohol, 67*, 868–79.

Dholakia, U. M. (2010). A critical review of question-behavior effect research. *Review of Marketing Research, 7*, 147–199.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

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. http://dx.doi.org/10.1086/586910.

Holland, R. W., Aarts, H., & Langendam, D. (2006). Breaking and creating habits on the working floor: A field experiment on the power of implementation intentions. *Journal of Experimental Social Psychology, 42*, 776–783.

Krueger, J. (2011). Return of the regressed: A thought and a field experiment. Psychology Today. Retrieved from http://www.psychologytoday.com/blog/one-among-many/201110/return-the-regressed-thought-and-field-experiment

Larimer, M. E., Turner, A. P., Mallett, K. A., & Geisner, I. M. (2004). Predicting drinking behavior and alcohol-related problems among fraternity and sorority members: Examining the role of descriptive and injunctive norms. *Psychology of Addictive Behaviors, 18*, 203–212.

Mollen, S., Rimal, R. N., Ruiter, R. A. C., Kok, G. (2013). Healthy and unhealthy social norms and food selection. Findings from a field-experiment. *Appetite, 65*, 83–89. http://dx.doi.org/10.1016/j.appet.2013.01.020

Mollen, S., Rimal, R. N., Ruiter, R. A. C., Jang, S. A., & Kok, G. (2013). Intervening or interfering? The influence of injunctive and descriptive norms on intervention behaviours in alcohol consumption contexts. *Psychology & Health, 28*, 561–575. http://dx.doi.org/10.1080/08870446.2012.752827.

Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2008). Normative social influence is underdetected. *Personality and Social Psychology Bulletin, 34*, 913–923. http://dx.doi.org/10.1177/0146167208316691.

Perkins, H. W., & Berkowitz, A. D. (1986). Perceiving the community norms of alcohol use among students: Some research implications for campus alcohol education programming. *Substance Use and Misuse, 21*, 961–976.

Reno, R. R., Cialdini, R. B., & Kallgren, C. A. (1993). The transsituational influence of social norms. *Journal of Personality and Social Psychology, 64*, 104–112. http://dx.doi.org/ 10.1037/0022-3514.64.1.104

Rimal, R. N. (2008). Modeling the relationship between descriptive norms and behaviors: A test and extension of the theory of normative social behavior (TNSB). *Health Communication, 23*, 103–116. http://dx.doi.org/10.1080/10410230801967791.

Rodrigues, A. M., O'Brien, N., French, D. P., Glidewell, L., & Sniehotta, F. (2014). The question–behavior effect: Genuine effect or spurious phenomenon? A systematic review of randomized controlled trials with meta-analyses. *Health Psychology*, E-print ahead of publication. http://dx.doi.org/10.1037/hea0000104

Sherman, S. J. (1980). On the self-erasing nature of errors in prediction. *Journal of Personality and Social Psychology, 39*, 211–221.

Scholly, K., Katz, A. R., Gascoigne, J., & Holck, P. S. (2005). Using social norms theory to explain perceptions and sexual health behaviors of undergraduate college students: An exploratory study. *Journal of American College Health, 53*, 159–166.

Schultz, P. W. (1999). Changing behavior with normative feedback interventions: A field experiment of curbside recycling. *Basic and Applied Social Psychology, 21*, 25–36.

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*, 429–34. Article Stable URL: http://www.jstor.org/stable/40064634

Sieverding, M., Decker, S., & Zimmermann, F. (2010). Information about low participation in cancer screening demotivates other people. *Psychological Science, 21*, 941–943. http://dx.doi.org/10.1177/0956797610373936.

Sprott, D. E., Spangenberg, E. R., & Fisher, R. J. (2003). The importance of normative beliefs to the self-prophecy effect. *Journal of Applied Psychology, 88*, 423–31.

Stigler, S. M. (1997). Regression toward the mean, historically considered. *Statistical Methods in Medical Research, 6*, 103–114. DOI:10.1191/096228097676361431.

Stok, F. M., De Ridder, D. T. D., De Vet, E., & De Wit, J. B. F. (2012). Minority talks: The influence of descriptive social norms on fruit intake. *Psychology & Health, 27*, 956–970. http://dx.doi.org/10.1080/08870446.2011.635303.

Stok, F. M., De Ridder, D. T. D., De Vet, E., & De Wit, J. B. F. (2014). Don't tell me what I should do, but what others do: The influence of descriptive and injunctive peer norms on fruit consumption in adolescents. *British Journal of Health Psychology, 19*, 52–64. http://dx.doi.org/10.1111/bjhp.12030.

Stok, F. M., De Vet, E., De Wit, J. B. F., Luszczynska, A., Safron, M., & De Ridder, D. T. D. (2015). The proof is in the eating: Peer social norms are associated with adolescents' intended and actual eating behavior. *Public Health Nutrition, 18*, 1044–1051. http://dx.doi.org/10.1017/S1368980014001268.

Turner, J., Perkins, H. W., & Bauerle, J. (2008). Declining negative consequences related to alcohol misuse among students exposed to a social norms marketing intervention on a college campus. *Journal of American College Health, 57*, 85–93.

Wechsler, H. (2003). Perception and reality: A national evaluation of social norms marketing interventions to reduce college students' heavy alcohol use. *Journal of Studies on Alcohol, 64*, 484–94.