

Goals and Social Comparisons Promote Walking Behavior

Gretchen B. Chapman, PhD, Helen Colby, PhD, Kimberly Convery, MSW,
Elliot J. Coups, PhD

The effectiveness of a pedometer intervention was affected by manipulating the goals given to participants and by providing social comparison feedback about how participants' performance compared with others. In study 1 (n = 148), university staff members received a low, medium, or high walking goal (10%, 50%, or 100% increase over baseline walking). Participants walked 1358 more steps per day (95% confidence interval [CI], 729, 1985), when receiving a high goal than when receiving a medium goal, but a medium goal did not increase walking relative to a low goal (554 more steps; 95% CI, -71,1179). In study 2 (n = 64), participants received

*individual feedback only or individual plus social comparison feedback. Participants walked 1120 more steps per day (95% CI, 538, 1703) when receiving social comparison feedback than when receiving only individual feedback. Goals and the performance of others act as reference points and influence the effect that pedometer feedback has on walking behavior, illustrating the applicability of the principles of behavioral economics and social psychology to the design of health behavior interventions. **Key words:** behavior change; walking; physical activity; pedometer; reference point; goals; social comparison. (Med Decis Making XXXX;XX:xx-xx)*

Walking has the potential to dramatically impact public health and reduce health care costs in the United States. Its many health benefits include the prevention and management of multiple chronic medical conditions, including cardiovascular disease, diabetes, obesity, and dementia.¹ However, the vast majority of the US population does

not meet public health recommendations for engaging in regular physical activity, such as walking.² Across a variety of health-related behaviors, including diet, medication adherence, and physical activity, self-monitoring is one of the most effective methods of promoting behavior change. In the context of walking, such self-monitoring can be achieved through the use of a pedometer, which provides an objective assessment of the number of steps walked. Randomized controlled trials of walking interventions indicate that the use of a pedometer and tracking of daily steps toward a goal (e.g., 10 000 steps/d) increases walking by an average of 2500 steps/d.³ We apply principles of behavioral economics to test the effect of 2 pedometer interventions on individuals' walking levels.

Research in behavioral economics (notably, Prospect Theory⁴) indicates that decision makers code outcomes as gains or losses relative to a reference point. Goals act as reference points, with failure to attain the goal experienced as a loss.⁵ Because decision makers are averse to losses, they are highly motivated to achieve goals but less motivated to exceed them. Social comparison information can also act as a reference point, such that doing worse than others is coded as a loss, whereas doing better than others is coded as a gain. Individuals are thus especially motivated to improve behavior when they learn that they are doing worse than others.⁶

Received 2 July 2014 from the Institute for Health, Healthcare Policy, and Aging Research, Rutgers University, New Brunswick, NJ, USA (GBC, KC); Anderson School of Management, UCLA, Los Angeles, CA, USA (HC); and Rutgers Cancer Institute of New Jersey, Department of Medicine, Rutgers Robert Wood Johnson Medical School, Rutgers University, New Brunswick, NJ, USA (EJC). Financial support for this study was provided entirely by a grant from the Robert Wood Johnson Foundation's Pioneer Portfolio and the Donahue Foundation through "Applying Behavioral Economics to Perplexing Health and Health Care Challenges." The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. Revision accepted for publication 23 May 2015.

Address correspondence to Gretchen Chapman, Institute for Health, Healthcare Policy, and Aging Research, Rutgers University, 112 Paterson Street, New Brunswick, NJ 08901, USA; (848) 932-5866; e-mail: gretchen.chapman@rutgers.edu.

© The Author(s) 2015

Reprints and permission:

<http://www.sagepub.com/journalsPermissions.nav>

DOI: 10.1177/0272989X15592156

MEDICAL DECISION MAKING/MON–MON XXXX

In study 1, participants were randomly assigned to receive a goal number of steps per day that was a small (10%), medium (50%), or large (100%) increase relative to their baseline walking. We explored whether higher goals motivated more walking (as predicted by Goal-Setting Theory⁷) or whether a very high goal would be less motivating than a more realistic goal. Prospect Theory⁵ predicts that a very high goal could be counterproductive if it means that performance is always far from the reference point, where the theoretical value function is flat. In study 2, participants were randomly assigned to receive feedback about their own performance only or feedback about how their walking performance compared with that of other study participants. We predicted that those in the social comparison condition would walk more than those in the control condition because the social comparison provides a reference point.

METHOD

Subjects

The participants in study 1 were 148 university staff members (125 women) who responded to e-mail invitations. Two additional participants who started the study but dropped out after attending only the first in-person visit are not included in the analyses. Participant age ranged from 22 to 60 y (mean, 48 y), and 60% were non-Hispanic white, 16% African American, 14% Asian, 6% Hispanic, with the remainder other races or multiple races. The mean body mass index (BMI; based on self-reported height and weight) was 27.87 (range, 18.64–45.10).

The participants in study 2 were 64 university staff members (60 women; mean age, 45 y [range, 23–69 y]; 72% were non-Hispanic white, 11% Asian, 9% African American, and 8% Hispanic; and the mean BMI was 27.66 [range 18.79–48.23]). Data from an additional 13 participants were excluded from analyses (without looking at the data) because of experimenter error in executing the protocol for these participants. (Two of these participants received faulty social comparison feedback [or would have received faulty feedback, had they been in the social comparison condition] because their baseline average was not entered in the Weblog program. As a result, they were always told [or would have been told] that they did better than 99% of others. Eleven of the 13 participants received Weblog feedback as if they were in the social comparison condition but e-mail

feedback as if they were in the control condition, or the reverse.)

Procedure

The study received Institutional Review Board approval. University staff were eligible to participate if they were 18 y or older, proficient in English, and in good health as assessed by the Physical Activity Readiness Questionnaire (PAR-Q).⁸ Individuals answering in the affirmative to any PAR-Q question, apart from the item about taking prescription medication for elevated blood pressure, and those who were pregnant were required to gain clearance from their physician before they could participate in the study. Because recruitment efforts targeted staff who worked in the buildings where study sessions were held, many of the participants knew one another and saw each other regularly during the work week. One hundred thirty-eight of 148 participants in study 1 reported on the exit survey that they knew other people who were enrolled in the study.

Participants in both studies completed a 3-wk protocol that included a 1-wk baseline period followed by a 2-wk intervention period. Participants attended an initial session where they were issued a New Life-styles NL-800 pedometer with 7-d memory and were instructed to fill out a log sheet each day, recording the time they put the pedometer on in the morning and the time they took it off in the evening. Pedometers were secured with a tamper-proof tie so that participants could not open them to view their number of steps. When participants returned 1 wk later, research staff removed the tie to open the pedometers and extracted the data from the 7-d memory. A baseline average was computed for each participant as the mean number of steps walked across each of the 6 full days between the first visit and the second visit, excluding any days on which the participant wore the pedometer for fewer than 8 h or on which fewer than 1000 steps were recorded.⁹ (In study 1, 42 of 148 participants had 1 to 4 of the 6 full baseline days excluded [mean exclusions over all 148 participants was 0.43 days per participant]. In study 2, 8 of 65 participants had 1 or 2 of the 6 full baseline days excluded [mean, 0.14].)

Participants then wore the pedometers for a 2-wk intervention period, viewing the step count feedback whenever desired and filling out daily log sheets. In the middle of the 2-wk period, research staff called each participant and walked him or her through the process of reading out the 7-d memory on the pedometer, so that the researchers would have pedometer

memory step count data for each day of the study in addition to the log sheets where participants recorded their daily step totals themselves. At the end of the 2-wk period, participants returned for a final session to turn in pedometers and log sheets. Step count data were again retrieved from each participant's pedometer using the 7-d memory function.

Participants completed questionnaires at the baseline and final study sessions. The questionnaire at both time points included the International Physical Activity Questionnaire (Short Form; IPAQ, <https://sites.google.com/site/theipaq/>), which includes a self-report measure of walking. Questionnaire measures other than the IPAQ walking items are not presented here. Participants were offered lunch at each of the in-person sessions and received a \$40 gift card as a thank you after completing the study. Participants in all conditions received standard information about how to use their pedometers as well as tips for how to increase the amount they walked.

In study 1, at the end of the baseline period, participants were randomly assigned to receive a personalized steps/d walking goal calculated as a 10%, 50%, or 100% increase over their baseline average. Research staff communicated the personalized goal to each participant and encouraged her or him to achieve or come as close as possible to the goal each day. Participants were not told how the goals were calculated, nor were they told their baseline average. At the end of the baseline week, research staff erased the memory on the pedometers before returning them to participants so that they could not review their baseline week performance, thus focusing participants' attention on their goal. When keeping their log sheets during the intervention period, participants were instructed to circle the smiley face that appeared on the log sheet if they had met their goal for the day and the frowning face if they had not, as a way of focusing their attention on the relationship between their walking behavior and their goal.

Study 2 differed from study 1 in that participants were not given any goals, and they were asked to complete a daily Weblog of walking behavior (on a study-specific Web site). In the intervention phase, participants were randomly assigned to the control or social comparison feedback condition. Those in the control condition simply tracked their walking behavior and tried to increase it. Those in the social comparison condition also received social comparison feedback via 2 mechanisms. First, when completing their Weblog each day, a message was displayed that said, "You did better than xx% of other people" (if the participant was in the top half of the distribution)

or "You did worse than xx% of other people" (if the participant was in the bottom half of the distribution). Second, participants received a personalized e-mail twice per week with similar percentile information corresponding to their performance over the past 3 d. Control participants received twice-weekly e-mails indicating whether they had completed the Weblog as requested.

The social comparison feedback was based on an actual comparison of the participant's performance to the performance of participants in study 1. We used study 1 as the reference population because that database was complete and available for comparison prior to the first participant enrolling in study 2 and because the 2 studies drew from the same population of university staff members. To determine the percentile social comparison feedback, an algorithm computed the difference between the participant's steps on the current day and her mean number of steps during the baseline period. That difference score was then compared with the distribution of analogous difference scores from all the participants on all the days of the intervention phase in study 1 to retrieve the appropriate percentile. If study 2 participants were not compliant in completing the Weblog, they could not receive the social comparison feedback. Consequently, the twice-weekly e-mail to participants in the social comparison condition either gave social comparison feedback (if the participant had completed the Weblog over the previous 3 d) or issued a request to update the Weblog (if the participant had not completed the Weblog over the past 3 d). E-mails to the control participants simply indicated whether the participant's Weblog entries had been received or whether the participant needed to update the Weblog.

RESULTS

The primary outcome measure was steps walked per day as recorded in the pedometer memory. We calculated for each participant a baseline average number of steps per day during the 6 full days of the sealed pedometer baseline period and examined daily steps walked during the 12 full days of the intervention period. In study 1, most participants (99 of 148) had data for all 12 full days in the intervention period; however, 16 had intervention period data for only 1 wk or less, and 3 of these had data for fewer than 6 d. In study 2, most participants (48 of 64) had intervention period data for all 12 full days; however, 5 had data for only 1 wk. All participants in both

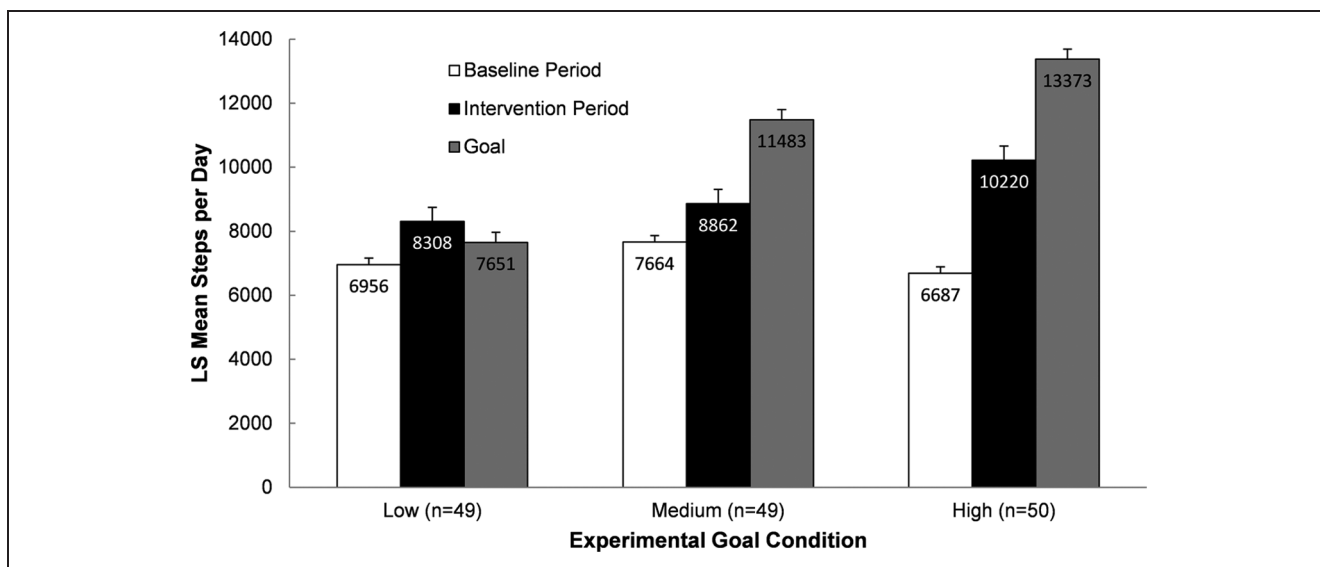


Figure 1 Least squares (LS) mean number of steps walked per day during the baseline and intervention study periods, as well as the mean goal assigned in each of 3 experimental conditions of study 1. LS mean steps per day during the intervention period control for baseline mean (centered) and day of intervention period. Error bars show 95% confidence interval.

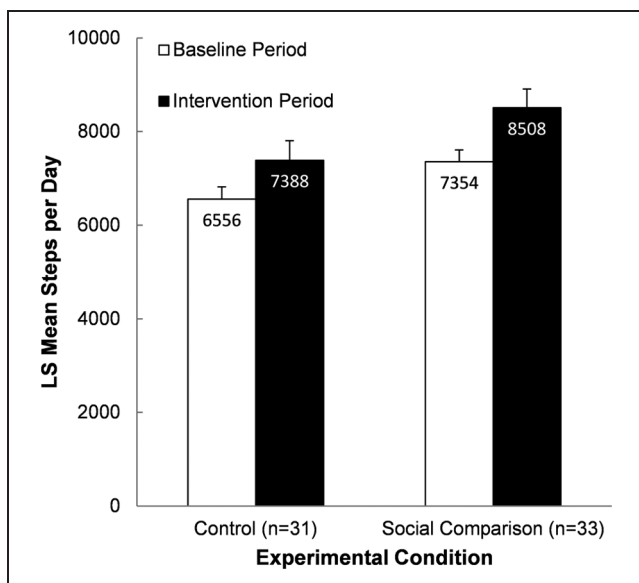


Figure 2 Least squares (LS) mean number of steps walked per day during the baseline and intervention study periods for participants in the control and social comparison conditions in study 2. LS mean steps per day during the intervention period control for baseline mean (centered) and day of intervention period. Error bars show 95% confidence interval.

studies were retained in the analyses regardless of the number of days of data they had. In study 1, mean number of steps per day during the baseline period

did not differ significantly across the 3 conditions, $F(2, 145) = 1.98, P = 0.14$ (Figure 1), although baseline walking was marginally higher in the medium goal condition compared with the high goal condition (506 step difference, 95% confidence interval [CI], -23 to 1978). In study 2, the mean number of steps per day during the baseline period did not differ significantly between conditions, $t(62) = 1.24, P = 0.22$ (Figure 2). Subsequent analyses controlled for baseline walking.

Steps per day during the intervention period were subjected to a mixed model that included experimental condition, day of the intervention period (12 levels), and baseline average (centered). In both studies, we had data on steps walked during the intervention period not only from the pedometer memory but also from the log sheets that participants kept. Analyses using the log sheet data show the same effects (data not shown). The study data sets are available from the first author upon request.

In study 1, participants assigned higher goals walked more than those assigned lower goals (Figure 1). Walking during the intervention period was correlated with baseline average, $B = 0.94, t(144) = 18.46, P < 0.0001$, and did not change systematically across days, $F(11, 1458) = 1.11, P = 0.35$. Most importantly, the experimental condition affected walking during the intervention period, $F(2, 144) = 19.71, P < 0.0001$. Specifically, those with a medium goal

walked only 554 (95% CI, -71 to 1179) more steps per day than those with a low goal, $F(1, 144) = 3.07$, $P = 0.08$, but those with a high goal walked 1358 (95% CI, 729 to 1985) more steps per day than those with a medium goal, $F(1, 144) = 18.25$, $P < 0.0001$.

We examined whether goal achievement moderated the effect of goal condition. Forty percent of participants in the high condition never met their goal on any day, compared with 16% in the medium condition and only 4% in the low condition, $\chi^2(2, n = 148) = 20.46$, $P = 0.0001$, $\phi = 0.37$. We repeated the mixed-model analysis using days 2 through 12 of the intervention period. We dropped the day variable and included instead a binary variable that indicated whether the participant had met or achieved her goal on the previous day. An interaction between experimental condition and goal achievement on the previous day, $F(2, 114) = 7.18$, $P = 0.001$, indicated that the effect of the experimental condition is strongest among people who achieved their goal on the previous day. Thus, high goals are not much more motivating than medium goals when one has failed to achieve them, but when one has achieved the very high goal, it is more motivating than lower goals.

In study 2, participants who received social comparison feedback walked more than those who did not (Figure 2). Walking during the intervention period was correlated with baseline average, $B = 0.75$, $t(61) = 13.02$, $P < 0.0001$, and did not change systematically across days, $F(11, 642) = 0.27$, $P = 0.99$. Most importantly, experimental condition affected walking during the intervention period, $F(1, 61) = 14.78$, $P = 0.0003$. Participants in the social comparison condition walked 1120 (95% CI, 538 to 1703) more steps per day than those in the control condition.

In study 2, participants in the social comparison condition were more compliant with completing the Weblog than were those in the control condition, presumably because the Weblog in the former condition provided new information (social rank) and was therefore more engaging. For each participant, we computed the percentage of e-mails during the intervention period of the protocol that informed participants that they were not up to date in completing their Weblog. That noncompliance rate averaged 25% ($s, 0.26$) for the control group, compared with 12% ($s, 0.22$) in the social comparison group, $t(62) = 2.19$, $P = 0.03$, a difference of 13% (95% CI, 1% to 25%). After statistically controlling for differences in compliance, the effect of the social comparison manipulation on steps walked remained (analyses not shown).

Previous studies have found that behavior declines when participants learn that they are performing better than others^{6,10} or that social comparison feedback has little impact on those who are performing better than others.¹¹ To examine whether the impact of the social comparison feedback in the current study was due to the effects of being told that one was doing worse than others or the effects of being told one was doing better than others, or both, we repeated the analysis using days 2 through 12 of the intervention period, dropping the day variable and including instead a binary variable that indicated whether the participant's previous day performance was in the top or bottom half of the comparison distribution (for participants in the experimental condition, this corresponds to being told that they performed better or worse than average). This analysis revealed no interaction between experimental condition and previous day's performance, $F(1, 56) = 0.00$, $P = 0.98$. (A follow-up analysis that treated the previous day's performance as a continuous variable also did not reveal a significant interaction.) That is, relative to the no-feedback control condition, the social comparison feedback increased performance to the same extent regardless of whether the feedback was that one had done better ($\beta = 893$ steps, 95% CI, -27 to 1812) or worse than average ($\beta = 880$ steps, 95% CI, 161 to 1598). Thus, unlike previous studies,^{6,10,11} study 2 revealed no evidence that the effects of social comparison feedback are limited to participants who learn that they were performing worse than average.

In both studies, participants reported on the IPAQ how many days per week they walked for at least 10 min at a time and how much time they usually spent walking on those days. We computed the log of the reported minutes per week spent walking. In study 1, at the first visit, as expected, there were no differences across the 3 goal conditions in self-reported walking, $F(2, 145) = 0.76$, $P = 0.47$. We examined self-reported walking at the end of the 3-wk protocol controlling for self-reported walking at the first visit (centered) but found no differences across conditions, $F(2, 98) = 1.12$, $P = 0.33$ (see Table 1; the IPAQ was inadvertently omitted from the exit questionnaire for 49 participants who are therefore missing responses for this data point).

In study 2, at the first visit, there was an unexpected marginal difference between conditions in self-reported walking, $F(1, 63) = 3.65$, $P = 0.06$, with participants in the social comparison condition reporting somewhat more walking, even though the experimental manipulation had not yet been applied.

Table 1 Geometric Mean Self-Reported Number of Minutes per Week Spent Walking

| Study Condition | Baseline | | End of Intervention | |
|-----------------|----------------|---------|---------------------|----------|
| | Mean | 95% CI | LS Mean | 95% CI |
| Experiment 1 | <i>n</i> = 148 | | <i>n</i> = 99 | |
| Low | 67 | 38, 118 | 213 | 150, 302 |
| Medium | 77 | 44, 136 | 260 | 183, 369 |
| High | 108 | 62, 191 | 314 | 217, 455 |
| Experiment 2 | <i>n</i> = 64 | | <i>n</i> = 62 | |
| Control | 65 | 36, 117 | 143 | 85, 240 |
| Social comp | 141 | 80, 250 | 230 | 140, 380 |

Note: LS means for end-of-intervention control for baseline levels. CI = confidence interval; LS = least squares.

We examined self-reported walking at the end of the 3-wk protocol controlling for self-reported walking at the first visit (centered) but found no differences across conditions, $F(1, 59) = 1.69, P = 0.20$ (see Table 1; 2 participants did not complete the exit questionnaire). Thus, self-reported walking did not reveal the effect of the study interventions that the objective pedometer measure did. This could be due to decreased power resulting from missing observations and lack of repeated measures in the self-report assessment, or it could reflect participants' inaccurate recollection of their walking behavior.

DISCUSSION

In study 1, participants walked more when assigned a high goal than a low or medium goal, despite the fact that the high goal was arguably unrealistically high. Research in behavioral economics indicates that goals act as reference points,⁵ and Prospect Theory⁴ predicts that extremely high goals will be less effective than realistic goals if, as a result of the high goals, performance is coded as a large loss, far from the reference point. Because the Prospect Theory value function is relatively flat far from the reference point, an increase in walking behavior at that location would result in little increase in subjective value. A more realistic goal, however, might be expected to place performance closer to the reference point, where the value function is steeper and performance increases would be more motivating. In contrast to this Prospect Theory prediction, social psychology research indicates that specific, high, challenging goals elicit higher performance than do easy goals.⁷ Our results are more consistent with the latter, although it is possible that our high goals were not sufficiently high to produce the pattern

predicted by Prospect Theory. Future research is warranted to examine whether the study results are retained with even higher step goals.

In study 2, social comparison feedback motivated participants to walk more than did simple individual feedback, a result consistent with previous findings that social comparison influences behavior by making social norms salient.⁶ Some commercially available pedometer tools provide social comparison leaderboards, and our results suggest such features may motivate physical activity. The optimal nature of social comparison feedback remains to be determined in future research. For example, in study 2, participants in the social comparison condition were told that they did better than $p\%$ of other people or worse than $(1 - p)\%$ of other people, depending on whether their percentile p was above or below 50%. It is possible that feedback that reported “better than $p\%$ ” or “worse than $(1 - p)\%$ ” for all participants would have produced different results.

Limitations

The main limitation of the current studies is that the participants were a self-selected group of mostly women who volunteered for a study that was advertised as one in which participants would wear a pedometer and try to increase their walking activity. Thus, these healthy adults were likely highly motivated to become more physically active and may have been more responsive to any intervention than would a less select group. Future research will need to address whether our interventions would have similar effects in other samples. Because so few men volunteered for the current studies, our data do not provide the power to examine gender differences. In addition, the sample size in study 2 was small, and in both studies, there were marginal tendencies for the conditions to differ in baseline walking activity. Our results found no “boomerang” effect of positive social comparison, as has been found in some previous studies.^{6,10} This could have been due to a competition effect driven by the face-to-face nature of the interactions of the participants or perhaps by subtle injunctive walking norms, as injunctive norms have been found to eliminate the boomerang effect of positive social comparison. Further research will be needed to investigate these possibilities.

CONCLUSION

Reference points to which individuals can compare their feedback motivate changes in health

behavior. The current studies add to a growing literature demonstrating that principles from behavioral economics and social psychology can be profitably applied to the science of healthy behavior change.

REFERENCES

1. Lee IM, Buchner DM. The importance of walking to public health. *Med Sci Sport Exerc.* 2008;40(7 suppl):S512–8.
2. Loustalot F, Carlson SA, Fulton JE, Kruger J, Galuska DA, Lobelo F. Prevalence of self-reported aerobic physical activity among U.S. States and territories: Behavioral Risk Factor Surveillance System, 2007. *J Phys Activity Health.* 2009;6(suppl 1):S9–17.
3. Bravata DM, Smith-Spangler C, Sundaram V, et al. Using pedometers to increase physical activity and improve health: a systematic review. *JAMA.* 2007;298(19):2296–304.
4. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica.* 1979;47(2):263–91.
5. Heath C, Larrick R, Wu G. Goals as reference points. *Cogn Psychol.* 1999;38(1):79–109.
6. Schultz WP, Nolan JM, Cialdini RB, Goldstein NJ, Griskevicius V. The constructive, deconstructive, and reconstructive power of social norms. *Psychol Sci.* 2007;18(5):429–34.
7. Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation. *Am Psychol.* 2002;57(9):705–17.
8. Thomas S, Reading J, Shephard RJ. Revision of the Physical Activity Readiness Questionnaire (PAR-Q). *Can J Sport Sci.* 1992;17(4):338–45.
9. Schmidt MD, Blizzard CL, Venn AJ, Cochrane JA, Dwyer T. Practical considerations when using pedometers to assess physical activity in population studies: lessons from the Burnie Take Heart Study. *Res Q Exerc Sport.* 2007;78(3):162–70.
10. John LK, Norton MI. Converging to the lowest common denominator in physical health. *Health Psychol.* 2013;32(9):1023–8.
11. Allcott H. Social norms and energy conservation. *J Public Econ.* 2011;95:1082–95.