Should All Steps Count When Using a Pedometer as a Measure of Physical Activity in Older Adults?

Alison L. Marshall

Purpose: The aim of this study was to determine if feedback on step counts from a pedometer encourages participants to increase walking. Methods: Randomly recruited older adults (n = 105) were asked to wear a pedometer for 2 wk. Half the participants were asked to monitor and record daily step counts during week 1 (feedback), then seal the pedometer shut during week 2 (no feedback). Half completed the study in reverse order. Self-reported walking was assessed via telephone interviews. Results: Significantly more steps were recorded per day (approximately 400 steps per day) when participants (n = 103, 63% women; mean BMI 25 ± 4) monitored their daily step count [t(102) = –2.30, P = 0.02] compared to the no feedback condition. There was no statistically significant difference in self-reported walking (P = 0.31) between feedback conditions. Conclusion: The difference in daily step counts observed between conditions, while statistically significant, may not be considered clinically significant. Further, the non-significant difference in self-reported walking between conditions suggests that feedback on daily step counts from a pedometer does not encourage participants to increase their walking.

Key Words: randomized trial, measurement, walking, mailed materials

The accurate measurement of physical activity has long presented a challenge to researchers and practitioners. Self-report measures of physical activity have been the instrument of choice for both population monitoring and intervention evaluation. Self-report measures can suffer from social desirability and recall biases, but perhaps their greatest limitation is their inability to accurately assess unstructured and incidental ambulatory physical activity.

Ironically, unstructured lifestyle-related activities are now routinely promoted to the public as they are encouraged to accumulate 30 min of moderate-intensity physical activity on most days of the week. Thus, the accurate assessment of these types of unstructured activities is very important. The failure of self-report measures to accurately account for unstructured incidental activity has lead to the use of objective measures of physical activity such as the pedometer.

The pedometer is gaining credibility as an accurate measure of ambulatory physical activity because it can provide a direct measure of incidental and
unstructured physical activity. Using the pedometer as a measure of physical activity has many other advantages. They have been shown to be a reliable measure of physical activity, are sensitive to change, and unlike self-report measures, can detect subtle changes in an individual’s incidental physical activity. Pedometer data can also be accurately compared between studies.

Some researchers and practitioners are also using the pedometer to facilitate behavior change. Pedometers can help individuals increase their physical activity level by providing immediate feedback and acting as a constant reminder to be active. In addition, pedometers can help individuals to set concrete goals.

Much research has been published on the role of the pedometer as both a measure and motivator for physical activity. In a simple search of the Medline advanced database (1996 to 2005) using search terms “pedometer” and “measurement,” 22 papers were identified. Twenty-one papers were identified when the search term “intervention” was substituted for “measurement.” Interestingly, some studies that used the pedometer as an intervention strategy also used the step count data as an outcome measure.

While some researchers attempted to “blind” their participants’ to pedometer feedback by sealing the pedometer shut during the measurement period, there are two potential measurement biases that may confound the results: 1) social desirability bias, and 2) reactivity bias. If participant behavior changes as a result of monitoring step counts, then the pedometer’s usefulness as an accurate measure of physical activity is compromised.

Four studies have evaluated pedometer reactivity bias. Two studies were conducted with children, and two with adults. Both studies with children reported no reactivity bias following either 8-d or 4-d blinded monitoring periods. Using adult volunteers Matevey et al. and Eastep et al. both concluded that pedometer feedback did not influence their participants’ activity levels. However, Matevey et al. had instructed their participants “to maintain their usual activity levels during each monitoring period” (p. 8). This instruction may have been sufficient to attenuate their participants potential reactivity to self-monitoring. Further research in this area is warranted to overcome this limitation and to explore potential reactivity in other populations and settings.

The aim of this study was to determine whether feedback on daily step counts from a pedometer would encourage older Australian adults (age 50 to 75 y) to increase their walking and thus answer the question “Should all steps count when using a pedometer as a measure of physical activity in older adults?” Based on the findings of previous research which suggested that reactivity may account for about 420 steps per day increase in physical activity, the clinical significance of any change in physical activity observed in this study will also be considered. The hypothesis under investigation was that there would be no difference in step counts recorded when participants were asked to monitor and record their daily step count, compared to not monitoring and recording daily step counts.

Method

A randomized crossover study design was used to assess whether the feedback from a pedometer significantly changes a participant’s step count over a 1-wk period. The study protocol was approved by the university human research ethics committee.
All participants consented to wearing a pedometer for 2 wk. During one of the 2 wk, participants were asked to wear a pedometer and record steps taken on a daily basis (“feedback” condition), thus participants were able to review their step count whenever they wanted. During the second week of the trial the pedometer was sealed (with a cable tie), so that participants were blinded to the pedometer’s feedback (“no-feedback” condition). Step counts under the “no-feedback” condition were recorded by the participant on the last day of the week. To control for a potential ordering effect, participants were stratified by age and gender, then randomly allocated via a computer-generated randomization sequence in blocks of 10 into one of two groups; 1) complete the trial in the order described above, or 2) complete the trial in the reverse order (“no-feedback” week 1 and “feedback” week 2).

Participants

Potential participants \((n = 199)\), over 50 y of age, were randomly selected from a registry of research volunteers and sent an invitation to participate in this study. Of these people 108 (56%) returned the informed consent form and agreed to participate \((n = 57\) were randomly assigned study group 1, and \(n = 51\) were assigned to group 2). One-hundred and five of those who gave written informed consent were contacted by telephone to confirm participation and complete the baseline assessment.

Measures and Procedures

Data were collected by trained telephone interviewers on three separate occasions. During the baseline telephone interview, data on standard demographic variables (age, sex, height, weight, education, marital status, and employment status) were collected along with physical activity data. Physical activity was assessed using the Active Australia Questionnaire (AAQ).\(^{19,20}\) The AAQ includes eight questions to assess the frequency and duration of walking (for transport and recreation), vigorous gardening/chores, and moderate- and vigorous-intensity leisure time physical activities done in the past week.\(^{19,20}\) The AAQ shows good reliability \([\text{intraclass correlation } 0.59 (0.52 \text{ to } 0.65)]\)^{21} and validity against a CSA accelerometer \((\text{Spearman’s rho range } 0.28 \text{ to } 0.33)\).\(^{22}\)

Following baseline assessment, participants were mailed a physical activity logbook and a Yamax model SW-700 pedometer (Yamasa Corp., Tokyo, Japan). The Yamax pedometer has been repeatedly shown to be one of the most accurate and reliable pedometers for research purposes.\(^5\) Participants were asked to record their daily or weekly step count in the logbook depending on which phase of the study they were in. The logbooks included detailed instructions on how to wear the pedometer, to ensure consistent pedometer placement between groups, and had space to record daily or weekly step counts. The participants were not aware of the aim of the study, and unlike the study by Matevey et al.\(^{17}\) they were not given any advice on how much physical activity they should or shouldn’t engage in during the study period.

Each participant was contacted at the end of the first week via telephone to re-administer the AAQ. They were also asked to either seal or unseal the pedometer for the following week. Participants were telephoned again at the end of the second
week to complete the AAQ questionnaire one more time, record weekly step counts, and reminded to return the pedometers in the postage-paid envelopes provided.

**Data Analysis**

Data were entered into a Microsoft Excel database and analyzed using SPSS version 11 (SPSS, Inc., Chicago, IL). Differences between participants randomized to each study condition were assessed by comparing sociodemographic variables and baseline self-reported physical activity (AAQ data) using chi square statistics and unpaired *t*-tests where appropriate. Body mass index (BMI: weight in kg/height in m²) was calculated from self-reported height and weight.

Consistent with recommendations for analyzing the AAQ, only the data from the walking, moderate- and vigorous-intensity (weighted by two) physical activity questions were used to estimate total time spent in physical activity in the past week.¹⁹,²¹ Total physical activity data were then categorized according to whether the participants reported activity was “sufficient” based on current national physical activity recommendations (i.e., 150 min on most days of the week).³,¹⁹-²² Data from the AAQ walking items were analyzed separately, to more closely approximate data collected by the pedometer. Differences between physical activity data between the two study conditions were assessed using paired *t*-tests.

To determine if participants’ step counts were influenced by pedometer feedback, a four-step data analysis plan was followed. First, the weekly step data were described using descriptive statistics and box plots. Second, to determine if there was an ordering effect, a two sample *t*-test was conducted on the difference scores (“no feedback” step count minus “feedback” step count). If there was no statistically significant order effect then the data from the two groups could be merged to conduct the final analyses. The final analysis which aimed to test the main hypothesis was a one sample *t*-test conducted on the difference scores between the two study conditions. If there was an ordering effect in the second step then only the data from the first week should be compared.

As previous research has demonstrated gender differences in pedometer-measured physical activity,⁴ separate analyses were conducted to examine if there was a differential effect of pedometer feedback between genders. Further, because participants categorized as “inactive” at baseline may have more opportunity and motivation to react to physical activity monitoring, the data were also stratified by activity status at baseline and compared across study conditions. Alpha was set at 0.05 for statistical significance.

**Results**

Ninety-eight percent of participants (*n* = 103) completed the study [*n* = 56 from study group 1 (“no feedback” week 1) and *n* = 47 from study group 2 (“no feedback” week 2)]. There were no significant differences between the two groups of participants in terms of baseline self-reported physical activity (AAQ data) or sociodemographic variables (see Table 1). Further, there was no significant difference in the average amount of time participants wore the pedometer per day (see Table 2).
Table 1  Demographic Characteristics of Participants

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group 1a n = 56</th>
<th>Group 2b n = 47</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean ± SD</td>
<td>mean ± SD</td>
</tr>
<tr>
<td><strong>BMI (M ± SD)</strong></td>
<td>25±3</td>
<td>26±5</td>
</tr>
<tr>
<td>Gender (M)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>n (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender (M)</td>
<td>20 (36%)</td>
<td>18 (38%)</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-64 y</td>
<td>29 (52%)</td>
<td>24 (51%)</td>
</tr>
<tr>
<td>65+ y</td>
<td>27 (48%)</td>
<td>23 (49%)</td>
</tr>
<tr>
<td><strong>Married (yes)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married (yes)</td>
<td>38 (68%)</td>
<td>29 (62%)</td>
</tr>
<tr>
<td><strong>Educated (tertiary)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Educated (tertiary)</td>
<td>32 (57%)</td>
<td>23 (49%)</td>
</tr>
<tr>
<td><strong>Employed (full- or part-time)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed (full- or part-time)</td>
<td>19 (34%)</td>
<td>20 (43%)</td>
</tr>
<tr>
<td><strong>Sufficient total physical activity (≥ 150 min over 5 sessions)</strong></td>
<td>37 (66%)</td>
<td>33 (70%)</td>
</tr>
</tbody>
</table>

a Participants not blinded to pedometer feedback during week 1 and blinded week 2.
b Participants blinded to pedometer feedback during week 1 and not blinded week 2.
c Includes sum of time spent in walking, moderate-and vigorous-intensity physical activity in past week

Table 2  Differences in Pedometer Step Counts and Self-Reported Physical Activity Between “No-Feedback” and “Feedback” Conditions (n = 103)

<table>
<thead>
<tr>
<th>Variable</th>
<th>“No-feedback” from pedometer (mean ± SD)</th>
<th>“Feedback” from pedometer (mean ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time wearing pedometer (h/d)</td>
<td>14.4 ± 1.5</td>
<td>14.1 ± 1.7</td>
</tr>
<tr>
<td>Pedometer steps/wk</td>
<td>58,517 ± 23,056</td>
<td>61,373 ± 25,932a</td>
</tr>
<tr>
<td>Self-reported walking (min/wk)</td>
<td>187 ± 192</td>
<td>201 ± 176</td>
</tr>
<tr>
<td>Self reported total physical activity (min/wk)b</td>
<td>329 ± 305</td>
<td>347 ± 301</td>
</tr>
</tbody>
</table>

a Statistically significant difference between conditions (t = −2.30 [102]; P = 0.02).
b Includes sum of time spent in walking, moderate-intensity, and weighted vigorous-intensity (x2) physical activity in past week

Descriptive analyses of the pedometer data suggested that there were no differences between mean weekly step counts within groups or between conditions (see Figure 1). Box plots revealed one extreme outlier in group 2, however, it was not excluded from the analysis.

A two sample t-test revealed no statistically significant ordering effect between the difference scores (“no feedback” step count minus “feedback” step count) of the two groups [t = 0.18 (101); P = 0.86]. The one sample test conducted on the merged data set revealed a statistically significant treatment effect (see Table 2).
The mean difference in weekly step counts between the “no-feedback” condition and the “feedback” condition was –2856 (95% CI = –5321 to –391) steps/wk. This equates to participants recording –408 (95% CI = –760 to –56) fewer steps/d when they were not able to monitor their accumulating step count.

Based on the data from the AAQ, there were no statistically significant differences between self-reported walking (min/wk: \( t = –1.00 (102); P = 0.31 \)) or overall physical activity (min/wk: \( t = –0.77 (102); P = 0.45 \)) between the two treatment conditions (see Table 2).

**Secondary Analyses**

For women, the two sample \( t \)-test revealed no statistically significant ordering effect between the difference scores (“no-feedback” minus “feedback” step counts \( [t = 0.49 (63); P = 0.62] \). The one sample \( t \)-test conducted on the merged data set revealed a statistically significant treatment effect \( [t = –2.3 (64); P = 0.03] \). The mean difference in weekly step counts between the “no-feedback” and “feedback” conditions was –3126 (95% CI = –5874 to –378) steps/wk. This equates to the women taking approximately 447 (95% CI = –839 to –54) extra steps/d when they were able to monitor, or have “feedback” on their accumulating step count. Again, however, there were no statistically significant differences in the women’s self-reported AAQ data between weeks.

---

**Figure 1**—Box plots created from weekly step counts recorded by the two study groups across the two treatment conditions (“no-feedback” vs. “feedback”).
For the men, the two sample t-test revealed no statistically significant ordering effect between the difference scores (“no-feedback” minus “feedback” step counts \( t = -0.22 \) (36); \( P = 0.83 \)). The one sample t-test conducted on the merged data revealed no statistically significant treatment effect \( t = -0.98 \) (37); \( P = 0.33 \). The mean difference in weekly step counts between “no-feedback” and “feedback conditions was -2392 (95% CI = -7328 to 2542) steps/wk. There were no statistically significant differences in the men’s self-reported AAQ data between weeks.

In terms of activity status at baseline, there were no statistically significant differences in terms of reactivity among those participants classified as “sufficiently active” or “inactive” at baseline. For example, the mean difference in time spent walking (measured by AAQ) between the “no-feedback” and “feedback conditions among the inactive participants at baseline was -30 (95% CI = -75.5 to 15.4) min/wk \( t = -1.35 \) (32); \( P = 0.19 \). While the mean difference in pedometer step counts between conditions among inactive participants at baseline was -1978 (95% CI = -6420 to 2465) steps/wk \( t = -0.91 \) (32); \( P = 0.37 \).

**Discussion**

This randomized cross-over study compared the step counts and self-reported walking of older adults under two study conditions, either wearing a sealed pedometer which eliminated feedback on accumulating step counts or an unsealed pedometer that allowed feedback. Contrary to the hypothesis, there was a statistically significant difference in the number of steps recorded between the two study conditions. On average, participants recorded 2859 more steps per week when they were able to monitor (have feedback) their accumulating daily step count compared to the week they received “no feedback” from the pedometer. However, this is the equivalent of only 500 steps per day, and since 500 steps is estimated to be the equivalent of about 5 min walking, the clinical significance of this finding is questionable.

Even though 5 min bouts of moderate-intensity physical activity have been shown to significantly increase aerobic fitness, the 5 min bout is only relevant when the bouts add up to at least 30 min per day. Since the participants in this study were only accumulating the equivalent of an extra 5 min activity per day, the findings of this study cannot be considered clinically significant. The 400 step difference observed in this study is just under a 5% increase in the participants average daily step counts (i.e., 8400 to 8800 steps/d) and is similar to the difference observed in the study by Matevey et al.\(^{17}\).

Hatano (1997) reported that an acceptable level of within-instrument error for the Yamax pedometer was around 3%.\(^{25}\) Schneider and colleagues (2004) suggested an even higher level of error (up to 10%) is acceptable in free living conditions.\(^{7}\) Thus, the average increase in step count observed in this study between conditions was within the limits of acceptable measurement error and may not be a result of measurement bias.

If measurement bias is non-differential, then it is of little or no concern. In this study, the significance of the difference between “no-feedback” and “feedback” step counts was different for men and women. The women in this study accumulated significantly more steps per week when they were able to monitor their step count
under the “feedback” condition compared to the “no-feedback” condition, but the men did not. Other research has also shown that women record more steps per day compared to men,\textsuperscript{5,26} and that they are also more likely to adopt the pedometer as a motivational tool to increase their physical activity (OR = 1.52 95% CI: 1.10, 2.12).\textsuperscript{27} Again, although the difference in steps counts for women recorded in this study may not be considered clinically significant, further investigation of the distinction between genders in terms of physical activity measurement and motivation is warranted.

It has been reported in the literature that pedometers underestimate the step counts at slow speeds.\textsuperscript{8,28} As the participants in this study were older adults, it is possible that some participants did increase their incidental physical activity, but that it was not registered on the pedometer due to the slow pace. However, this type of activity also would not normally be reported in self-report assessments either and is of little consequence to health or intervention.

Caution should be exercised when interpreting the finding of this study. It is virtually impossible to implement the perfect study design to answer the research question posed, as it is not possible to have a non-pedometer-wearing control group and still obtain comparable physical activity data. However, the lack of effect on the self-report measures between the two groups included in this investigation suggests that a control group was not necessary. Further, the non-significant difference in self-reported physical activity between study conditions supports the overall finding that daily monitoring of step counts using a pedometer does not significantly influence wearers to increase their physical activity.

It is also important to note the subtle difference between participants being “personally aware” and being “instructed” by researchers to either change or not change their behavior. The aim of this study was to assess the impact of wearing a pedometer on step counts and physical activity behavior. The fact the participants in this study were not giving any specific instruction on what the purpose of this study was, or about how much physical activity they should or should not do while participating in the study means the participants were “free” to behave or react in whatever way they wanted. Thus, this study is an important addition to the work of Matevey and colleagues who instructed their participants to not change their behavior.\textsuperscript{17}

A potential limitation of this study is that most participants were considered to be “sufficiently active” at baseline, thus less likely to “react” to activity monitoring. However, stratification and analysis of the data by activity status at baseline revealed no significant differences in terms of the self-reported AAQ data or the pedometer step counts over time or between conditions.

This study also had a number of methodological strengths. First, the randomized cross-over study design enabled within- and between-participant variability between the two test conditions to be assessed. Second, as most studies in this area have used reactively recruited volunteer participants with an apparent interest in physical activity, the high recruitment rate (56%) from a proactively recruited sample (not selected because of their interest in physical activity), and low participant drop-out rate enhance the generalizability of the findings. Thirdly, this is the first study of its kind to include older adults (almost half the sample was over
65 y of age). Finally, although a self-report measure of physical activity was used to substantiate the overall conclusions, the AAQ has been shown to have acceptable measurement properties.19,20

The findings of this study support previous research on participant reactivity to pedometer-based assessment of physical activity in children and middle-aged adults15-18 by demonstrating that feedback on daily step counts from a pedometer does not encourage older adults to substantially increase their physical activity. Therefore, all steps should count when using a pedometer (and obligatory logbook) to assess physical activity in free living older adult populations.

Acknowledgments

This study was supported by a start-up grant from the University of Queensland awarded to the author. Participants were recruited from the “50+ Registry” of the Australasian Centre on Ageing, the University of Queensland. Thanks to Prof. Annette Dobson for her advice on the statistical analyses.

References


