Physical Activity Measurement in Older Adults: Relationships With Mental Health

Sarah J. Parker, Scott J. Strath, and Ann M. Swartz

This study examined the relationship between physical activity (PA) and mental health among older adults as measured by objective and subjective PA-assessment instruments. Pedometers (PED), accelerometers (ACC), and the Physical Activity Scale for the Elderly (PASE) were administered to measure 1 week of PA among 84 adults age 55–87 (mean = 71) years. General mental health was measured using the Positive and Negative Affect Scale (PANAS) and the Satisfaction With Life Scale (SWL). Linear regressions revealed that PA estimated by PED significantly predicted 18.1%, 8.3%, and 12.3% of variance in SWL and positive and negative affect, respectively, whereas PA estimated by the PASE did not predict any mental health variables. Results from ACC data were mixed. Hotelling–William tests between correlation coefficients revealed that the relationship between PED and SWL was significantly stronger than the relationship between PASE and SWL. Relationships between PA and mental health might depend on the PA measure used.

Keywords: accelerometers, pedometers, affect, exercise, elderly

The aging process is associated with an increased prevalence and number of both mental and physical health concerns and disabilities (Blazer, Burchett, Service, & George, 1991; Ko & Coons, 2005; Mroczek & Spiro, 2005). Poor mental health is an important consideration for the older adult population, because it appears to be a substantial component of perceived quality of life (Kahn, Hessling, & Russell, 2003), can affect physical health (Diefenbach, Leventhal, Leventhal, & Miller, 1996), and is likely to impede engagement in positive health behaviors (Cohen-Mansfield, Marx, & Guralnik, 2003). Given the high and rising cost of health care and prescription medications, it is important to consider relationships between modifiable behaviors and lifestyles that might affect mental health, such as physical activity (PA).

The investigation into the relationship between PA and mental health is not new, but questions remain as to what frequency, intensity, and duration of PA is most feasible and effective for affecting mental health (Dunn, Trivedi, & O’Neal, 2001). An important distinction has emerged in the literature between studies of structured exercise programs and studies promoting lifestyle PA such as walking, gardening, or cycling. An extensive body of literature suggests that participating...
in structured exercise programming has a beneficial effect on global (e.g., satisfaction with life, general psychological well-being) and specific (e.g., depression, anxiety) indices of mental health in older adults (Arent, Landers, & Etnier, 2000; Netz, Wu, Becker & Tenenbaum, 2005). Despite the known benefits of exercise participation, long-term habitual adherence to structured exercise programs remains a major limiting factor in its effectiveness as a public health approach.

Partially driven by the adherence limitations of exercise programs, researchers in the physical health domain have shown that engagement in regular, moderate-intensity PA (e.g., walking, gardening) is adequate for older adults to reap many of the protective health benefits associated with an active lifestyle (Pate et al., 1995). This recommendation might prove more palatable to older individuals, who might dislike exercise, have little exercise experience, or are otherwise anxious about or averse to structured exercise participation (Cohen-Mansfield et al., 2003).

There is a small but expanding body of research examining the relationship between lifestyle (i.e., not necessarily exercise related) PA and mental health in older adults. Meeting the minimum PA recommendation is correlated with better general psychological health scores (Mummery, Schofield, & Caperchione, 2004). Furthermore, recent studies have shown that accumulating more PA is related to higher satisfaction with life (Fox, Stathi, McKenna, & Davis, 2007; McAuley et al., 2000; Stubble, de Moor, Boomsma, & de Geus, 2007), although others have failed to detect such a relationship (Menec, 2003; Netz et al., 2005). In regard to discrete indices of mental health, PA has been found to negatively correlate with depression (Fukukawa et al., 2004; Lindwall, Rennemark, Halling, Bergelund, & Hassman, 2007; Yoshiuchi et al., 2006) and anxiety (Fox et al.; Stewart, Mills, Sepsis, et al., 1997), whereas others have not detected this relationship (Cairney, Faught, Hay, Wade, & Corra, 2005). Furthermore, in a review of cross-sectional studies, a greater amount of PA was associated with preferable levels of both positive and negative affect (Arent et al., 2000). Although this evidence points toward a general beneficial effect of PA on mental health, conflicting results as to the overall benefit of PA on mental health variables have been found.

One explanation for these conflicting results pertains to the instrumentation used to measure PA (Fox & Stathi, 2002). Specifically, studies in this area tend to use self-report PA-recall questionnaires (Fox & Stathi; Mummery et al., 2004), broadly labeled as subjective-recall instruments. Although these instruments are reasonably accurate and valid for measuring vigorous activities or structured exercise, studies have shown that self-report, subjective PA monitoring might underestimate time spent in moderate-intensity activities (Ainsworth, Bassett, Strath, Swartz, O’Brien, et al., 2000; Strath, Bassett, & Swartz, 2004; Wareham & Rennie, 1998). The minimizing effect of self-report PA measurement on overall estimated activity level might then attenuate relationships between PA and measures of mental health.

Partially driven by the aforementioned limitations of PA surveys, objective PA-measurement instruments such as pedometers and accelerometers are becoming increasingly prominent in research as tools to measure relationships between PA and physical health (Sesso, 2007). Pedometers are used to count the number of steps taken by participants, which is used as an index of overall activity level. Accelerometers not only detect total activity but also measure the magnitude of
vertical accelerations and decelerations that are collected over a user specified time interval and can be used to provide a summary of time spent in different PA intensities (i.e., sedentary, light, moderate, or vigorous activity). To date, only two studies have examined the relationship between objectively measured PA and mental health in older adults, both reporting that total PA energy expenditure, measured by accelerometers, is positively correlated with indices of global psychological well-being (e.g., satisfaction with life; Fox et al., 2007), as well as lower negative mood scores (e.g., anger, anxiety, depression; Fox et al.; Yoshiuchi et al., 2006). Although these results are encouraging for the use of objective monitors currently, to our knowledge no studies have compared the relative effectiveness of subjective versus objective monitors in identifying relationships between PA and mental health. Given that objective PA instruments are expensive and more technically demanding than subjective measures, a comparison of the instruments as they relate to mental health is warranted. As such, the purpose of the current study was to examine the impact of PA-measurement instruments (two objective and one subjective measure) on the observed relationship between PA level and mental health (affect and satisfaction with life) among older adults.

**Methods**

**Study Participants**

Participants were 84 men \((n = 28)\) and women \((n = 56)\) between the ages of 55 and 87 (mean = 71) years recruited from the local community through word of mouth and posted announcements. Individuals were eligible to participate if they had no major limitations to daily physical activities, including severe health ailments or orthopedic limitations that prevented them from ambulatory activity. Participant characteristics can be viewed in Table 1.

**Study Protocol**

Study procedures were approved by the university’s institutional review board, and informed consent was obtained before any data collection. Study participants underwent measures of height and mass and general screening measures of resting blood pressure and resting heart rate. Screening measures were carried out to determine study eligibility. Participants were then asked to complete the Positive and Negative Affect Scale (PANAS) and a Satisfaction With Life Scale (SWL) to assess general affect or mood and psychological well-being, respectively. After physical- and psychological-data collection, participants were shown how to wear a pedometer and an accelerometer. These objective monitoring devices were worn on the midline of the right thigh for 7 consecutive days during all waking hours except when bathing or swimming; the electronic devices used in the current study were not waterproof. As such, ambulatory activities (i.e., walking, jogging) were exclusively measured by the current study. Although mental health measures were collected before PA measurement, participants were asked not to change their typical activity levels during the course of monitoring, so as to provide PA data representative of their current typical lifestyle. This enabled direct comparison between typical PA measures and current typical affect and SWL assessments.
After the seventh day of PA monitoring, participants returned the objective devices and were administered the Physical Activity Scale for the Elderly (PASE) to assess self-reported PA over the same period as the objective monitors were worn.

**Study Measures**

*Positive and Negative Affect Scale.* The PANAS (Watson, Clark, & Tellegen, 1988) is a 20-item scale that measures frequency of various emotional states over a 1-month time period as a predictor of psychological well-being (Watson, Clark, & Carey, 1988). Participants are asked to estimate to what extent they experienced a given feeling or emotion on a scale of 1 (*slightly/not at all*) to 5 (*extremely*). Responses to each of 20 items are scored and averaged across 10 positive-affect (PAF) and 10 negative-affect (NAF) subscales to yield final PAF and NAF subscale scores, each with a minimum score of 10 and a maximum score of 50. Good internal reliability for the PANAS has been reported for both PAF (α = .87) and NAF (α = .87), which has been shown to be consistent in an older adult sample (Crawford & Henry, 2004).

*SWL.* The SWL (Diener, Emmons, Larsen, & Griffin, 1985) is a 5-item questionnaire designed to assess global perceptions of well-being and life satisfaction. Items reflect positive judgments of accomplishments and life satisfaction, and participants indicate the degree to which they agree with each item on a scale of 1 (*strongly disagree*) to 7 (*strongly agree*). The SWL has been shown to have good internal reliability (α = .87) and 2-month test–retest reliability (r = .82) in an older adult validation sample (Diener et al.).

*Estimation of Physical Activity via Pedometry.* Each participant wore a Yamax SW-200 pedometer (Yamax Corp., Tokyo, Japan) positioned over the midline of
the right thigh. All pedometers underwent a 20-step calibration procedure to confirm accurate step counting. Step values of 19–21 were deemed acceptable. All pedometers passed the calibration test before implementation. Daily steps accrued were recorded in a step log provided to the participant. At the end of the 7-day monitoring period daily step totals were averaged to yield mean steps per day. The Yamax SW-200 pedometer has been reported to accurately estimate steps taken during walking at 3.0 mph within 1% of actual steps taken (Bassett et al., 1996).

**Estimation of Physical Activity via Accelerometry.** Each participant wore an accelerometer (Actigraph 7164, Manufacturing Technology Inc., Fort Walton Beach, FL) positioned over the midline of the right thigh. Each participant was given written instructions to refer to each morning to aid compliance in accurate wear. Raw signals (i.e., counts) from accelerometers were converted from a voltage to a series of digital analog signals, which were then integrated for a user-determined time interval (1 min for the current study) via a computer-algorithm-derived calculation of the area under the counts-per-minute curve. Specified accelerometer cut points (Matthew, 2005) were used to determine time spent in selected physical activity categories: inactivity (<260 counts/min), light activity (261–759 counts/min), moderate general activity (760–5,924 counts/min), and vigorous activity (>5,925 counts/min). Each accelerometer was calibrated pre- and poststudy use and was found to meet manufacturer calibration standards (± 5%). Accelerometer-based estimates of PA correlated well with actual metabolic assessment of walking ($r = .77$) and correlated reasonably well with most other activities of daily living such as gardening and household chores ($r = .59$ for mixed activities; Hendelman, Miller, Bagget, Debold, & Freedson, 2000).

**Estimation of Physical Activity by Self-Report Questionnaire.** The Physical Activity Scale for the Elderly (PASE) was administered as a subjective estimate of PA. The PASE is a brief survey designed specifically for older adults to estimate PA recalled over the course of 1 week (Washburn, Smith, Jette, & Janney, 1993). Participants are asked to indicate how often (i.e., seldom, 1–2 days; sometimes, 3–4 days; or often, 5–7 days) in the previous week they had engaged in a variety of activities. These included walking; activities of light (e.g., stretching), moderate (e.g., pairs tennis), or strenuous intensity (e.g., jogging); and muscle-conditioning activities. Duration spent in these activities was categorized as less than 1 hr, 1–2 hr, 2–4 hr, or more than 4 hr/day. Furthermore, time spent in paid or volunteer work involving at least some standing or walking was recorded in total hours per week. Finally, engagement in household activities such as light housework, yard work, and caring for others was categorized as yes or no. The final PASE score is derived by multiplying activity participation by established item weights (Washburn et al.) to generate an overall PA score for the week. The PASE correlated well with a reference composite consisting of a PA diary, PA motion-sensor counts, and a global activity self-assessment ($r = .84$; Washburn et al.). The PASE also has reported criterion validity for predicting differences in some physiological measures including heart rate, maximal oxygen uptake, and body composition among older adults (Washburn et al.).
Data Analyses

Accelerometer data were summarized by established and previously described methods (Masse et al., 2005) according to whether each individual wore the monitoring device for a minimum of 10 hr/day and a minimum of 4 days/week, including 1 weekend day. Accelerometer data that were uninterrupted for 60 continuous min or more were removed from analyses to control for episodes in which participants might have forgotten to wear the device (Masse et al.). Outliers were identified by accelerometer counts greater than 20,000 counts/min, which represents a magnitude of acceleration outside the range of human movement patterns. All objectively determined PA data (pedometer and accelerometer) were averaged to yield a mean daily value. Mean steps per day were calculated to estimate average daily ambulatory activity. Accelerometer-measured average counts per day (ACC CTS), as well as average minutes of moderate and vigorous PA (ACC MV) per day, were calculated as indicators of PA volume and intensity, respectively. Body-mass index (kg/m²) was calculated by dividing weight (kg) by height (m) squared.

Pearson’s bivariate correlations were calculated to examine the relationships between the different PA measures. Linear-regression analyses were then conducted to determine unique variance in mental health measures accounted for by each type of PA assessment. For each of the three psychologically dependent variables (i.e., PAF, NAF, SWL), the following regression models were tested against PA measures as independent variables: Model 1 included the PASE only, Model 2 included steps per day only, Model 3 included ACC CTS only, and Model 4 included ACC MV only. Three separate independent linear regressions were chosen over a multiple-regression or a hierarchical/stepwise design because the PA measures were correlated with each other to prevent the influence of multicollinearity and to allow the evaluation of the independent relationship between each of the different PA and mental health measures. Hotelling–Williams calculations were used to determine the statistical significance of any differences observed in the correlation coefficients between the different PA measures and a given mental health measure. This calculation generates a t value between two dependent correlation coefficients that have been generated from the same sample (i.e., PASE and PAF versus steps per day and PAF), which can then be analyzed for a significant difference based on the sample degrees of freedom (Howell, 2002). Hotelling–Williams calculations were performed manually in an Excel spreadsheet, and all other statistics were calculated using the Statistics Package for the Social Sciences, version 12.0 (SPSS; Chicago, IL). Significance was set at \( p < .05 \).

Results

Physical health status of the participants indicated that the average body-mass index was 27.9 kg/m², with 29% classified in a normal-weight category, 46% classified in the overweight category, and 25% classified as obese. Average heart rate was normal, with a normal average diastolic blood pressure and an average systolic blood pressure classified as prehypertensive (Chobanian, Bakris, Black, Cushman, Green, et al., 2003).
Significant positive correlations \((r = .34-.85, p < .01)\) were recorded between steps per day, ACC CTS, ACC MV, and PASE PA measures (data not shown). Notably, objective measures were more highly correlated with each other than with the subjective measure.

To explore the strength of each measurement technique in predicting relationships between PA and mental health variables, regression analyses were performed (see Table 2). For each dependent variable (PAF, NAF, and SWL), separate linear-regression equations were computed for PASE, steps per day, ACC CTS, and ACC MV. Results indicated that steps per day was the strongest predictor of all three mental health indicators, accounting for 8.3% of the variance in PAF scores, 12.3% of the variance in NAF scores, and 18.1% of the variance in SWL scores. All three models incorporating steps per day were significant. In addition, ACC CTS accounted for 8.5% and 8.9% of the variance in NAF and SWL, respectively, but did not significantly predict variance in PAF. ACC MV minutes significantly predicted 6.1% of variance in NAF scores but failed to reach significance in predicting PAF or SWL. Finally, PASE scores did not significantly predict levels of PAF, NAF, or SWL.

Hotelling-Williams calculations (Table 3) yielded a significant difference \((t = -2.3, p < .05)\) between the correlation coefficient calculated for steps per day and SWL \((r = .395)\) versus the coefficient calculated for PASE and SWL \((r =

### Table 2 Linear-Regression Analysis of Predictive Validity for Physical Activity Measurement Techniques

<table>
<thead>
<tr>
<th>Model</th>
<th>(r)</th>
<th>(R^2)</th>
<th>(F)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positive affect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Activity Scale for the Elderly</td>
<td>.054</td>
<td>.003</td>
<td>.22</td>
<td>.640</td>
</tr>
<tr>
<td>steps/day*</td>
<td>.288</td>
<td>.083</td>
<td>7.13</td>
<td>.009</td>
</tr>
<tr>
<td>moderate and vigorous physical activity (min/day)</td>
<td>.215</td>
<td>.046</td>
<td>3.70</td>
<td>.058</td>
</tr>
<tr>
<td>accelerometer counts/day</td>
<td>.140</td>
<td>.020</td>
<td>1.46</td>
<td>.230</td>
</tr>
<tr>
<td><strong>Negative affect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Activity Scale for the Elderly</td>
<td>.138</td>
<td>.019</td>
<td>1.49</td>
<td>.225</td>
</tr>
<tr>
<td>steps/day*</td>
<td>.351</td>
<td>.123</td>
<td>11.1</td>
<td>.001</td>
</tr>
<tr>
<td>moderate and vigorous physical activity* (min/day)</td>
<td>.247</td>
<td>.061</td>
<td>4.94</td>
<td>.029</td>
</tr>
<tr>
<td>accelerometer counts/day*</td>
<td>.290</td>
<td>.085</td>
<td>7.05</td>
<td>.010</td>
</tr>
<tr>
<td><strong>Satisfaction with life</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Activity Scale for the Elderly</td>
<td>.055</td>
<td>.003</td>
<td>0.24</td>
<td>.625</td>
</tr>
<tr>
<td>steps/day*</td>
<td>.426</td>
<td>.181</td>
<td>17.5</td>
<td>.000</td>
</tr>
<tr>
<td>moderate and vigorous physical activity (min/day)</td>
<td>.220</td>
<td>.409</td>
<td>3.93</td>
<td>.051</td>
</tr>
<tr>
<td>accelerometer counts/day*</td>
<td>.299</td>
<td>.089</td>
<td>7.54</td>
<td>.007</td>
</tr>
</tbody>
</table>

*Significant at \(p < .05\).
.022). No other significant differences were observed between correlation coefficients.

### Discussion

These results suggest that depending on measurement technique, higher volume of PA is related to a favorable mental health profile (i.e., lower NAF, higher PAF, and SWL) in older adults. To date, two other studies have examined PA and mental health using objective PA monitoring. Consistent with the current results, Fox et al. (2007) reported significant relationships between total PA measured by either accelerometer or SWL ($r = .17$) and lower anxiety ($r = .16$) and depression ($r = .18$). In contrast, in the current study we found slightly higher correlations, as well as a significant relationship between ACC MV and NAF that was not detected by Fox and colleagues. Differences in the actual amount of both total PA and moderate-to-vigorous-intensity PA of the participants in our study versus that of Fox et al.
might explain the discrepancy, although these data were not presented in Fox et al.'s article. Our results are also consistent with those of Yoshiuchi et al. (2007), who reported an inverse relationship between both steps per day and moderate-intensity PA and depression in older adults. In general, the results of the current study, combined with those of previous researchers, are consistent with the existing PA and mental health literature (Arent et al., 2000; Mummery et al., 2004) and suggest that accumulating a greater amount of PA is positively related to mental health, although the PA is not required to be structured exercise per se. Future work should examine how a change in PA affects mental health to establish cause and effect.

A major result of the current study was that although PA-measurement techniques showed moderate to strong correlations with each other, objective techniques (i.e., steps/day, ACC CTS, and ACC MV) were more sensitive in detecting the small but significant relationship between PA and PAF, NAF, and SWL relative to the subjective (i.e., PASE) PA-measurement technique. This suggests that the ability of these objective monitors to predict mental health might be superior to that of selected subjective measures.

There has been a lack of focused older adult studies comparing objective and subjective PA measures. Borrowing from the general adult empirical literature, Strath et al. (2004) found that self-report questionnaires were limited in validity for detecting ubiquitous, light- to moderate-intensity PA. Consistent with their conclusions, the current study found that both steps per day and ACC CTS significantly predicted PA–mental health relationships, whereas the subjective measure employed did not. Therefore, the current results combined with other research regarding the accuracy of subjective versus objective monitoring indicate that further investigation is needed into the predictive validity of a larger variety of subjective PA-measurement instruments for characterizing relationships between PA and mental health.

A strength of the current study was the use of objective PA monitors, which more accurately estimate PA than do subjective recall questionnaires. It should be noted that the current study only examined one questionnaire for comparison with the objective devices. There is a plethora of PA-recall measures published in the literature (Kriska & Casperson, 1997), and it is likely that they vary somewhat as to the accuracy with which each measures PA. A notable volume of research, however, suggests that in general, self-report questionnaires provide a less accurate indication of PA behavior than do objective measures (Ainsworth et al., 2000; Strath et al., 2004; Wareham & Rennie, 1998). This relationship materialized in the current study when we compared different measurement techniques in their ability to predict mental health outcomes and provides a clear example of how subjective instruments, such as the one used in the current study, might be problematic for this area of research. Furthermore, with several different subjective measures of PA currently published in studies of PA and mental health, making comparisons between studies or conclusions about the body of research as a whole is problematic given the nuances of each questionnaire. The use of objective monitors provides the capability to better standardize PA measurement, allowing for easier comparison and condensation of results from multiple studies of PA and mental health. In addition, the fact that intensity, frequency, and duration of PA can be examined objectively with accelerometers free from psychological bias.
such as memory opens the door to new research possibilities and potentially new inquiries into the underlying mechanisms linking PA to mental health.

Although the current results are critical of subjective PA measures, this is not to say that subjective measures do not have validity or advantages, particularly for providing general relationships in large-scale, population-based studies. Furthermore, the use of objective monitors does harbor some limitations. For one, accurate PA monitoring requires compliance from participants, including remembering to wear the device and to place it in the correct location. The issue of compliance is certainly an obstacle for research with these devices and can be partially overcome with techniques such as careful instruction and participant follow-up. In addition, although accelerometers offer more detailed information about PA than pedometers do and might be more accurate than subjective PA measures, the accuracy with which they currently predict PA intensity in older adults is still under examination. Older adults are known to have lower metabolic rates and oxygen consumption than younger adults (Harris & Benedict, 1918), and the cut points generated by the current analyses of accelerometer data might not be appropriate. Finally, these data were collected on a relatively healthy, predominantly White sample of older adults. The nature of the relationship between PA and mental health in other ethnic groups or participants with health impairments cannot be concluded from these results. Future research must address these concerns.

Despite these limitations, the results of the current study support previous reports that PA is related to mental health in older adults and extend the literature to show that objective instruments might be preferable for accuracy of PA measurement and relationship detection. In the future, data from prospective studies using objective techniques might be key for establishing clear links between amount of PA and specific mental health benefits among older adults. Dissemination of this research can then assist practitioners in prescribing optimal amounts of PA for improvement and maintenance of psychological well-being in the older adult population.

Acknowledgments

This work was partially supported by funding received from a University of Wisconsin–Milwaukee Graduate School Research Award, the Center for Age and Community Fellowship Award, the Wisconsin Department of Health and Family Services, and a Career Development Award from the National Institute on Aging (K01AG025962). The authors would like to thank Mark A. Parmenter and Janet E. Laatsch for data-collection assistance.

References


Yoshiuchi, K., Nakahara, R., Kumano, H., Kuboki, T., Togo, F., Watanabe, E., et al. (2006). Yearlong physical activity and depressive symptoms in older Japanese adults: Cross-sectional data from the Nakanojo study. *The American Journal of Geriatric Psychiatry, 14*(7), 621–624. Author: Read proofs carefully. This is your ONLY opportunity to make changes. NO further alterations will be allowed after this point.