Psychology and Aging
2005, Vol. 20, No. 2, 272–284

0882-7974/05/$12.00 DOI: 10.1037/0882-7974.20.2.272

Physical Activity and Psychological Well-Being in Advanced Age: A Meta-Analysis of Intervention Studies

Yael Netz
Zinman College of Physical Education and Sport Sciences, Wingate Institute

Meng-Jia Wu
Michigan State University

Betsy Jane Becker and Gershon Tenenbaum
Florida State University

A meta-analysis examined data from 36 studies linking physical activity to well-being in older adults without clinical disorders. The weighted mean-change effect size for treatment groups ($d^{C}. = 0.24$) was almost 3 times the mean for control groups ($d^{C}. = 0.09$). Aerobic training was most beneficial ($d^{C}. = 0.29$), and moderate intensity activity was the most beneficial activity level ($d^{C}. = 0.34$). Longer exercise duration was less beneficial for several types of well-being, though findings are inconclusive. Physical activity had the strongest effects on self-efficacy ($d^{C}. = 0.38$), and improvements in cardiovascular status, strength, and functional capacity were linked to well-being improvement overall. Social–cognitive theory is used to explain the effect of physical activity on well-being.

Keywords: psychological well-being, physical activity, meta-analysis, intervention studies

Accumulating evidence supports the popular belief that physical activity is associated with psychological health. The components of psychological health, however, are not yet clearly determined. Careful review of the literature presents a myriad of operational definitions of the terms *psychological health* and *psychological well-being* (Brown, 1992; Gauvin & Spence, 1996; McAuley & Rudolph, 1995). It is generally agreed that psychological well-being is a multifaceted phenomenon (Gauvin & Spence, 1996), particularly in the aging population (McAuley & Rudolph, 1995; Rejesky & Mihalko, 2001; Spirduso & Cronin, 2001; Stewart & King, 1991).

Our meta-analysis examined the effects of organized physical activity on the well-being of older adults without clinical disorders. On the basis of a conceptual framework proposed for evaluating well-being in older age (McAuley & Rudolph, 1995; Spirduso & Cronin, 2001; Stewart & King, 1991), we considered four general components: (a) emotional well-being (i.e., state and trait anxiety, stress, tension, state and trait depression, anger, confusion, energy, vigor, fatigue, positive affect, negative affect, and optimism), (b) self-perceptions (i.e., self-efficacy, self-worth, self-esteem, self-concept, body image, perceived physical fitness, sense of mastery, and locus of control), (c) bodily well-being (i.e., pain and perception of physical symptoms), and (d) global perceptions such as life-satisfaction and overall well-being.

Yael Netz, School of Exercise and Sport Sciences, Zinman College of Physical Education and Sport Sciences, Wingate Institute, Netanya, Israel; Meng-Jia Wu, Department of Counseling, Educational Psychology, and Special Education, College of Education, Michigan State University; Betsy Jane Becker and Gershon Tenenbaum, Department of Educational Psychology and Learning Systems, College of Education, Florida State University.

Correspondence concerning this article should be addressed to Yael Netz, School of Exercise and Sport Sciences, Zinman College of Physical Education and Sport Sciences, Wingate Institute 42902, Israel. E-mail: neyael@wincol.ac.il

The lack of experimental evidence for a causal link between physical activity and improved psychological well-being has been noted by most qualitative reviews studying the aging population (Brown, 1992; McAuley & Rudolph, 1995; O'Connor, Aenchbacher, & Dishman, 1993; Spirduso & Cronin, 2001). The authors of those reviews argued that although habitual exercise may improve psychological well-being, a strong sense of well-being may also be necessary to comply with a habitual and intensive exercise program. Therefore, in the current review only comparative or pretest–posttest treatment studies were included, whereas correlational studies, which examine relations rather than treatment effects, were excluded.

Most previous meta-analyses did not focus on older adults. Furthermore, rather than focusing on well-being as a multifaceted phenomenon, they focused on selected components of well-being, mostly on negative emotions and quite often in clinical populations. Because clinical populations are not only different from populations without clinical disorders but also differ on each particular clinical status (Martinsen & Stephens, 1994), we did not include clinical populations in our meta-analysis that focused on well-being in a broader view.

Two meta-analyses concentrated on anxiety in multiaged populations. Petruzzello, Landers, Hatfield, Kubitz, and Salazar (1991) indicated that acute and chronic physical activities reduced state anxiety, but their reported treatment-versus-comparison-group standardized-mean-difference effect size ($\bar{d}$) was rather small ($\bar{d} = 0.24$). According to the same report, chronic physical activity has also been related to small reductions in trait anxiety (with a mean effect of 0.34), but this result was computed across both comparison and single-group studies. McDonald and Hodgdon's (1991) meta-analytic results were comparable to those of Petruzzello et al. (1991), showing effect sizes of 0.28 for state anxiety and 0.25 for trait anxiety. However, McDonald and Hodgdon looked only at studies investigating aerobic fitness training.

Other meta-analyses have concentrated on depression in multiaged populations. According to North, McCullagh, and Tran (1990), the effect of physical activity on depression is reportedly moderate ($\bar{d} = 0.53$), but clinically depressed participants showed large decreases in depression as a result of physical activity, whereas persons from the general population showed small or no decreases due to lower floor and higher ceiling effects for the clinical sample. However, it is difficult to judge the stability of the between-groups comparisons reported by North et al., because they used problematic as well as conservative tests of differences among effects. Hedges and Olkin (1985, p. 148) discussed the problems with using standard analysis of variance *F* tests and associated post hoc contrasts in meta-analysis. Alternative tests (that do not assume homogeneity of variance) may have revealed differences that North et al. did not detect. Regardless of this weakness, this finding is consistent with the relatively large effect size ($\bar{d} = 0.72$) of exercise on clinically depressed individuals reported by Craft and Landers (1998).

A meta-analysis of the effects of exercise on mood in older adults (Arent, Landers, & Etnier, 2000) indicated an overall effect size of 0.24 for treatment versus comparison effects, and a mean change effect size of 0.38 standard deviations for single-group pretest–posttest studies. (Below we analyze change measures and denote them as d^C values to distinguish them from treatment-control effects). However, Arent et al. (2000) did not include other components of well-being such as self-perception or life-satisfaction. Some evidence suggests that exercise efficacy or mastery cognitions play a moderating role in enhancing affective responses (Jerome et al., 2002; Marquez, Jerome, McAuley, Snook, & Canaklisova, 2002; McAuley, Talbot, & Martinez, 1999). Exercise or physical self-efficacy changes are greatest in those with low self-efficacy at the initiation of an exercise program (McAuley, 1994), in part because of greater ceiling effects. For older adults, whose self-efficacy may be deteriorating along with their functional abilities, physical activity may provide a mastery experience that leads to changes in affect. Thus, mastery–efficacy measures are included in McAuley and Rudolph's (1995) narrative review of physical activity, aging, and psychological well-being. Surprisingly, however, meta-analyses (McDonald & Hodgdon, 1991; Moritz, Feltz, Fahrbach, & Mack, 2000) and narrative reviews (Fox, 2000) on self-perception components such as self-efficacy, self-concept, or self-esteem have only considered younger populations.

Besides examining the overall effect of physical activity on psychological well-being, the current review also examines variables that potentially moderate this effect. Exercise dose is likely to be the most substantial moderating variable. Exercise dose has several components, including duration of each exercise session, number of sessions per week (or other time frame), number of weeks of participation, as well as exercise intensity. Meta-analyses on the general population usually find that longer duration of exercise is needed to alter affect. Petruzzello et al. (1991) suggested that exercise sessions lasting 21 to 30 min resulted in larger decreases in state anxiety ($\bar{d} = 0.41$) than shorter bouts of exercise ($\bar{d} = 0.04$), with effects of even longer sessions falling between these two values. Similarly, training regimens lasting 10 to 12 weeks resulted in larger effect sizes for trait anxiety ($\bar{d} = 0.50$) than were found for shorter periods of exercise (for which $\bar{d}$ ranged from 0.14 to 0.17); regimens lasting more than 15 weeks resulted in an effect size of 0.90, similar to North et al.'s (1990) findings of reduced depression after 17 weeks of exercising ($\bar{d} \geq 0.97$). In contrast, Arent et al.'s (2000) findings on experimental versus comparison effects in older adults indicated that exercise lasting 1 to 6 weeks resulted in a larger effect size ($\bar{d} = 0.48$) than exercise lasting over 12 weeks ($\bar{d} = 0.19$).

Intensity of exercise is another moderator of exercise effects. Although Raglin and Morgan (1985) contended that in order to alter affect, exercise intensity needs to be at least 60% of maximal aerobic power, Dishman (1986) proposed a minimum of at least 70% of maximal oxygen uptake or adjusted maximum heart rate and a duration of at least 20 min. However, narrative reviews (Dunn, Trivedi, & O'Neal, 2001; Ekkekakis & Petruzzello, 1999) and meta-analytic reviews (Craft & Landers, 1998) on the general population failed to reveal a consistent pattern of dose–response effects between exercise and affect. Arent et al. (2000), on the other hand, found that low-intensity exercise was more beneficial ($\bar{d} = 0.58$) for older adults than medium- ($\bar{d} = 0.26$) or high-intensity ($\bar{d} = 0.29$) exercise in experimental versus comparison effects.

One controversial issue related to the dose–response relationship is the dependence of psychological changes on physical changes. Dunn et al. (2001), in their narrative review, and Craft and Landers (1998), in their meta-analytic review, both concerning effects in the general population, were not able to provide evidence of such dependence, whereas Arent et al. (2000) were able to show some indirect support for this dependence in their meta-analysis on older adults. Yet Spirduso and Cronin (2001) in their qualitative review on older adults proposed that physical activity might enhance quality of life in older adults without improving cardiorespiratory status. They postulated that the act of exercising might be beneficial in and of itself.

Exercise mode may also be an important moderator of exercise effects on psychological well-being. Petruzzello et al. (1991) found that aerobic exercise showed an effect size of 0.26, higher than that for nonaerobic exercise (−0.05). Effect sizes for aerobic (0.36) and nonaerobic (−0.16) exercise on trait anxiety were not significantly different although they appear quite discrepant (Petruzzello et al., 1991).[1] In the North et al. (1990) meta-analysis, both aerobic and resistance training were associated with improvements in depression. Dunn et al. (2001), in their narrative review, reported that although most studies examined aerobic exercise, the few examining resistance training showed significant effects in reducing depressive symptoms. Moreover, Arent et al.'s (2000) meta-analysis on older adults revealed a larger gain effect size on mood for resistance training ($\bar{d}^C = 0.80$) than for cardiovascular ($\bar{d}^C = 0.26$) or mixed (cardiovascular plus resistance; $\bar{d}^C = 0.37$) training.[2]

Although McAuley and Rudolph's (1995) qualitative review on exercise, aging, and well-being recommended examining age as a moderator, age was not examined in Arent et al.'s (2000) meta-analysis on older adults. Other meta-analyses performed on mul-

[1] Petruzzello et al. (1991) also used standard analysis of variance *F* tests to examine differences among different categories of studies. Their results, like those of North et al. (1990), may have differed had they used more appropriate weighted analyses.

[2] It is hard to judge the tests of means reported by Arent et al. (2000) because they did not report the standard errors for their means, instead reporting the standard deviations of the averaged effect sizes.

tiaged populations did not include studies on participants older than 65 years. Thus, age was included as a moderator in the current meta-analysis. The current meta-analytic review examines the effects of all of these moderating variables in addition to the overall effect of various physical activities on a range of psychological well-being outcomes in selected advanced age categories.

Method

Selection of Studies

Studies were located through computer searches of three databases: MedLine, PsycINFO, and SPORTDiscus. Data were collected from studies reporting the effect of exercise or physical activity interventions on aspects of well-being of older adults. Key terms used in these searches focused on global and specific psychological terms such as *well-being*, *psychological health*, *anxiety*, *depression*, *body image*, *self-concept*, *self-efficacy*, *life satisfaction*, *self-perception*, and *happiness*. Terms used to locate older participants were *middle age*, *old age*, *elderly*, and *older adults*. Key terms for exercise were *physical activity*, *exercise*, *physical fitness*, *sport*, *endurance training*, and *strength training*. Cross-checking of references and scans of journals in gerontology, psychology, and exercise science ensured an extensive literature search. All English-language studies done before January 2004 that could be located were included. The literature search initially yielded over 250 studies.

Initially, studies that looked at clinical populations or participants with a mean age less than 55 were excluded. This age inclusion criterion was based on previous studies (Colcombe & Kramer, 2003; Martin & Sinden, 2001). However, any definition of old is arbitrary. For example, McAuley and Rudolph (1995) included studies with participants over the age of 45 in their narrative review on physical activity, aging, and well-being, ending up with an average age of 56.7 years. Three studies with a mean age of 54 were therefore included in our review, making our age inclusion criterion a mean age of 54 years and older. Correlational studies were also excluded, leaving 51 studies. Fifteen studies that did not provide sufficient information for computing effects were excluded. The remaining 36 studies were included in the current analysis, and all studies provided psychological status measures for each group in the study before and after intervention.

Coding of Studies

Samples within studies were coded for a number of characteristics reflecting potential moderating variables for the exercise–psychological well-being relationship. These characteristics represented (a) the study design, (b) the participants, (c) the physical-fitness improvement of participants, (d) the exercise activity, and (e) the psychological well-being measure used.

Design characteristics. Samples were classified as either treatment groups or control groups. All treatment groups received some type of light or moderate to intensive exercise intervention. Control groups included participants who did not receive any exercise intervention, who attended social activities (with no exercise component), or who experienced very light exercise and were labeled as control groups by the authors of the studies. Also we coded (a) whether participants were assigned to treatment and control groups using randomization and (b) whether participants were matched on certain characteristics such as gender across groups. Some studies used both methods of assigning participants to groups. As described below, separate effects were computed for each treatment and control sample in a study, and when possible for subgroups within the treatment and control samples.

Participant characteristics. Where possible, separate effect sizes were calculated for men and women and for participants of different ages. Group gender mix was coded (all female, all male, or mixed). Samples with mean participant ages between 54 and 64 years were classified as late middle age, those with average ages between 65 and 74 years were called young–old, and those with participants aged 75 or over (on average) were classified as old–old. Also whether participants were sedentary prior to intervention and whether they showed improvement in physical status was coded (this is discussed separately below).

Physical-fitness improvement of participants. Separate effect sizes were calculated for subsamples of participants who showed improvement (improved) and those showing none (not improved) on four dimensions of physical fitness when this information was reported by the authors of the primary studies. Samples for which no information was available on fitness were classified into a third no-information group. Three of the dimensions refer to the main categories of physical fitness (American College of Sports Medicine, 1998): *cardiovascular fitness,* assessed by heart rate reserve, oxygen consumption, and so forth; *strength,* assessed by tests assessing various groups of major muscles such as knee extensors and flexors, shoulder extensors and flexors, hand grip, and so forth; and *flexibility,* determined by assessing range of motion of main joints. The fourth dimension (a combination of these three) is used quite often in the aging population and includes functioning capacity assessed by tests such as stair climbing, lifting and carrying, or sit-to-stand ability.

Exercise characteristics. A number of potential moderating characteristics of exercise were identified from the available literature and previous meta-analyses (Arent et al., 2000; North et al., 1990; Petruzzello et al., 1991). Studies were coded for exercise type (aerobic, calisthenics, and resistance training), duration of exercise (weeks of exercise intervention), frequency of exercise (sessions per week), length of exercise session (minutes per session), and exercise intensity (hard, moderate, light, very light, and not applicable). The intensity of aerobic exercise and calisthenics were coded using the American College of Sports Medicine (2000) tables. In cases where only a verbal description regarding the intensity of exercise rather than specific information such as maximum oxygen consumption, maximal heart rate, or heart rate reserve was provided, two experts in physiology of exercise made a subjective assessment of the exercise intensity referring to the American College of Sports Medicine (2000) definitions. Intensity of resistance training was assessed in terms of repetitions and workload using standard tables (Bompa, 1999). Two experts in resistive training made a subjective assessment in cases where only verbal information was provided.

Psychological well-being characteristics. On the basis of previous work (Spirduso & Cronin, 2001; Stewart & King, 1991), we first coded the psychological well-being measures into four broad categories: measures of emotions, self-perceptions, bodily well-being, and global well-being. Because initial analyses showed considerable variation within the four broad categories, outcomes were further classified into 11 categories. Emotions were further classified into anger, anxiety (including stress and tension), confusion, depression, and positive affect; self-perceptions were broken down into view of self (including self-worth, self-esteem, self-concept, body image, perceived physical fitness, sense of mastery, and locus of control) and exercise or physical self-efficacy; bodily well-being measures were further classified into physical symptoms and energy; and global well-being was separated into overall well-being and life-satisfaction. We refer to this classification as the *measure type* variable.

Computation of Effect Sizes

Effect sizes were computed as standardized mean-change measures (Becker, 1988), representing the magnitude of the difference between the pretest and the posttest for each outcome. Standardized mean-change measures were computed separately for all available control and treatment groups for each study. If results were available for more specific subgroups (e.g., male and female treatment and control groups), effects were computed for the most specific groups for which data were available.

We present the formulas for effects from a simple study with one treatment and one control group. The mean-change measures for the treatment group (g^{trt}) and control group (g^{ctrl}) are

$$g^{\mathrm{trt}} = \frac{(\bar{Y}^{\mathrm{trt}} - \bar{X}^{\mathrm{trt}})}{S_X^{\mathrm{trt}}} \text{ and } g^{\mathrm{ctrl}} = \frac{(\bar{Y}^{\mathrm{ctrl}} - \bar{X}^{\mathrm{ctrl}})}{S_X^{\mathrm{ctrl}}},$$

where $\bar{Y}^{\mathrm{trt}}$ denotes the posttest mean of the treatment group, $\bar{X}^{\mathrm{trt}}$ denotes the pretest mean of the treatment group, and S_X^{trt} denotes the standard deviation of the treatment group on the pretest. $\bar{Y}^{\mathrm{ctrl}}$, $\bar{X}^{\mathrm{ctrl}}$, and $S_X{}^{\mathrm{ctrl}}$ are parallel statistics for the control group. A simple interpretation of the value of g is that it represents the change from before exercise (pretest) to postexercise in pretest standard deviation units. A g value of 0.10 for a treatment group would indicate that the participants' mean levels of well-being had improved because of exercise about one-tenth of a standard deviation, relative to their initial level.

All gs were corrected for small-sample bias. The unbiased value, denoted d_i^C for the ith sample, is obtained by correcting each g_i via $d_i^C = \{1 - [3/(4n_i - 5)]\}\ g_i$, where n_i is the size of the sample (Becker, 1988). Finally, signs of some effects were reversed from the value reported in the primary study, so that positive d_i^C values indicated improved psychological well-being. If a study reported lower postexercise scores on a depression measure (an improved level of depression), and the d_i^C computed on the basis of the reported means was -0.25, the effect would be reversed for analysis to $d_i^C = 0.25$, reflecting improved well-being.

Analyses

Weighted analyses developed by Hedges and Olkin (1985) were used. Below we use the symbol $d.$ to refer to the weighted mean change. Under fixed-effects models, each effect size was weighted by the inverse of its variance (v), so that $v(d_i^C) = 2(1 - r_i)/n_i + (d_i^C)^2/(2n_i)$, where n_i is the sample size, d_i^C is the unbiased effect size, and r_i is the pretest–posttest correlation for the ith sample (i.e., the correlation between the scores X and Y for sample i). Unfortunately, it is very rare to find reported values of r_i when the primary research studies do not investigate relations between measures, and indeed none of the studies in this synthesis had reported r_i. A conservative value of $r_i = .50$ was used for all variances. Specifically, for fixed-effects tests, we weighted by the inverse variance $w_i = 1/v(d_i^C) = 2n_i/[(d_i^C)^2 + 2]$.

In many cases, initial analyses suggested that fixed-effects or mixed models (including models with predictor variables such as exercise type, etc.) did not fully explain the observed variation between effects. In those cases, random or mixed-effects models were adopted that incorporated additional between-effects uncertainty into the effect variances. The estimate of between-samples variation was a simple method-of-moments estimator: $S_\delta^2 = [\Sigma(d_i^C - \bar{d}^C)^2/(k - 1)] - \bar{v}$, where $\bar{v}$ is the within-study sampling variance $v(d_{ij}^C)$ averaged across the k effects in the analysis. For categorical models, S_δ^2 was computed within each moderator category. The revised weights were computed as $w_i{}^* = 1/[v(d_i^C) + S_\delta^2] = 2\,n_i/[2n_i\,S_\delta^2 + (d_i^C)^2 + 2]$. In addition, regression analyses based on both fixed- and mixed-model assumptions were used. Mixed-model regression analyses followed the procedures outlined by Raudenbush (1994).

Dependence

A fundamental assumption of most standard analyses used in meta-analysis is the independence of effects in the analysis. When multiple outcomes have been measured for the same individuals, or when the same outcome is measured at several time points for the same individuals, effect sizes computed for those multiple outcomes or time points will not be independent. A number of ways exist to deal with dependence (Becker, 2000). The primary approach used in this synthesis was to separate effects into groups that included only (or primarily) independent effects. Grouping the effects according to measure type eliminated nearly all dependence because few studies used multiple measures of the same outcome type.

Coding Reliability

The study features mentioned above were independently coded by two coders. Coding reliabilities, computed as the percentage of agreement of the coded variables before resolution, ranged from 86.1% for computation of effect sizes to 100.0% for gender, exercise type, duration, and frequency of exercise. Discrepancies in effect computation arose mainly over decisions to reverse the signs of the d^Cs (to have positive d^Cs represent improvement) and decisions about whether to include or exclude specific outcomes. Reliability of effect-size computation rose to 93.6% when discrepancies due to differences in the signs of the d^Cs were disregarded. All discrepancies were resolved before analyses began.

Results

Description of Effects

The 81 independent samples from 36 studies in this synthesis provided 406 standardized mean-change effect sizes (d^Cs). Twenty-two studies had both treatment and control groups, and 14 studies had only treatment groups (eight of these had just one group). Of the 28 multigroup studies, 10 used both matching and randomization to assign participants to groups. Eleven randomized but did not match, two used only matching, and five used neither approach. The d^C values ranged from a minimum of -2.01 to a maximum of 1.72. Table 1 shows counts of the numbers of studies and samples. Table 2 gives descriptive statistics for several features of the 81 samples in our synthesis.

Publication Bias

Meta-analysis results may not represent the population of interest due to publication bias, which occurs when studies with statistically significant results are more likely to be published than studies with statistically nonsignificant results. To assess the presence publication bias, we created a funnel plot, which is shown in Figure 1.

In the funnel plot, sample sizes and estimated effects for all samples are plotted in a two-dimensional space, with effects on the horizontal axis and sample size on the vertical axis. Because effects based on small samples are expected to show more variability than effects from large samples, the plot should look like an inverted funnel if publication bias is not a problem (see, e.g., Egger, Smith, Schneider, & Minder, 1997). When publication bias has occurred, sections of the funnel may be missing or the plot may

Table 1
Numbers of Studies and Samples

Feature	Number
Studies (k = 36 studies)	
Design	
Studies with both control and treatment groups	22
Studies with treatment groups only	14
Gender	
Studies reporting results without separating men and women	22
Studies reporting results separately for men and women	7
Studies with women only	7
Studies with men only	0
Samples (k = 81 samples)	
Mean age	
Late middle age (54–64 years)	22
Young old (65–74 years)	50
Old old (> 74 years)	9

become very asymmetrical. For instance, if only positive significant results are published, the funnel may appear to have been cut in half because only large and positive effects will be available. No part of the funnel appears to be missing in Figure 1, and the distribution of the effect sizes does not show any strong pattern suggestive of bias.

However, there appear to be more positive effects than negative ones. A statistical test of symmetry in the plot (Egger et al., 1997) indicates significant asymmetry (intercept = 1.35, $SE = 0.310$, $z = 4.35$, $p < .001$). Although this test provides some evidence of potential publication bias in the data, there are other sources of asymmetry in funnel plots. Among those is true heterogeneity in effects that may result from different intensities of the intervention (here, exercise) that relate to sample sizes. Indeed sample size is negatively correlated with both the number of sessions per week ($r = -0.31$, $p < .0001$) and minutes per session ($r = -0.12$, $p = .0019$). Thus, further analyses investigate the roles of these variables in producing change in well-being.

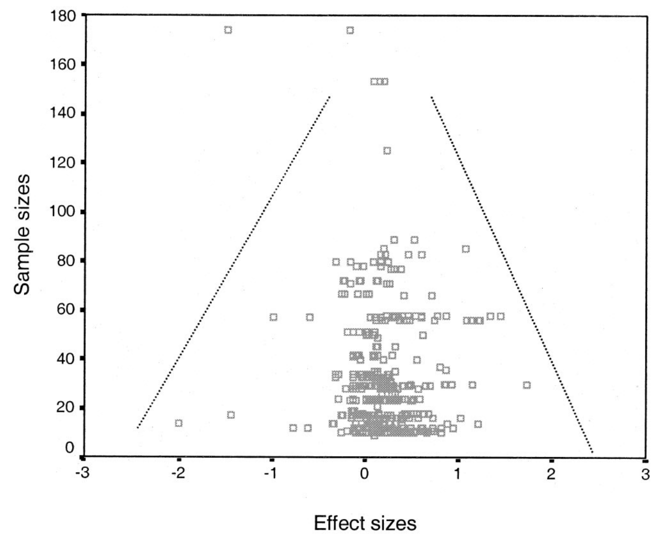

Figure 1. Funnel plot of all 406 effects from treatment and control samples.

Overall Analysis

Figure 2 shows a frequency histogram of all 406 effects. The effects are positively skewed, with a median of 0.16 and unweighted mean of $\bar{d}^C = 0.20$ ($SD = 0.36$). Over three fourths of the effects (79%) are positive.

The overall weighted mean effect under the fixed-effects model is $d^C. = 0.15$ with a standard error of 0.007 (the notation $d^C.$ distinguishes a weighted mean from an unweighted mean $\bar{d}^C$). However, the homogeneity test is significant, $Q(405) = 2{,}018.10$, $p < .001$, suggesting that all effects do not arise from a single population. Many potential moderator variables may account for differences among the effects. In addition, a more appropriate estimate of the average effect across all samples and measures is based on a random-effects model. This model suggests a mean effect of $d^C. = 0.19$ standard deviation units ($SE = 0.017$), which differs significantly from zero.

Treatment- and Control-Group Comparisons

All participants in treatment groups took part in exercise, whereas those in control groups experienced little to no exercise. To test whether exercise led to greater changes in well-being from pretest to posttest, we analyzed the effects separately for treatment and control groups. Of the 406 effects, 132 were from control groups and 274 were from treatment groups. The weighted mean change for treatment groups under the random-effects model was 0.24 ($SE = 0.023$, 95% confidence interval: 0.20 to 0.28). The treatment-group mean was almost three times the size of the weighted mean for control groups, $d^C. = 0.09$ ($SE = 0.024$, 95% confidence interval: 0.04 to 0.14). This difference was significant, $Q_B(1) = 66.12$, $p < .0001$, where B refers to between-groups variation, indicating that, on average, participants in treatment groups showed significantly greater improvement in their psychological well-being than those in control groups.[3] However, unexplained variation remained in both groups, $Q_W(404) = 1{,}952.03$, $p < .0005$, where W refers to within-group variation, under the fixed-effects model.

Table 2
Study Features for 81 Samples

Study feature	*M*	*SD*	Range
Sample size (number per group)	38.5	26.4	9–174
Age of participants	66.4	7.5	40–101
Number of effects	5.2	3.3	1–14
Exercise dose			
Duration (number of weeks of exercise)	19.6	17.5	1–52
Frequency (number of sessions per week)	2.7	1.1	0.5[a]–5.0
Length (minutes per exercise session for treatment groups)	61.0	49.7	30–210

[a] A frequency of 0.5 sessions per week indicates the exercise was conducted once during each 2-week period.

Categorical Analyses for Treatment and Control Samples

Moderators were next explored separately for treatment and control samples. Effects were categorized into smaller subgroups according to gender, age, exercise type, exercise intensity, and type of measure outcomes. Analyses of mean changes based on the random-effects model are shown in Table 3. Homogeneity tests (Q statistics) were computed under the fixed-effects model.

Age. For treatment samples, the means for the three age subgroups differed, $Q_B(2) = 135.73$, $p < .0001$. The late middle-age subgroup (average age less than 65 years) had the largest mean change ($d^C. = 0.33$), and the old–old samples (average age above 74 years) had the smallest mean ($d^C. = 0.11$). In control groups, the largest mean change was in the young–old ($d^C. = 0.10$), and, as in the treatment groups, the old–old samples had the smallest mean ($d^C. = 0.02$).

Physical-fitness improvement status. To explore the relationship between change in well-being and physical improvement on four dimensions (cardiovascular, strength, flexibility, and func-

[3] Analyses of 182 differences between paired treatment and control effects from the 22 studies that had both types of samples showed similar results. In these studies, the treatment groups outscored controls by 0.12 standard deviations ($SE = 0.016$).

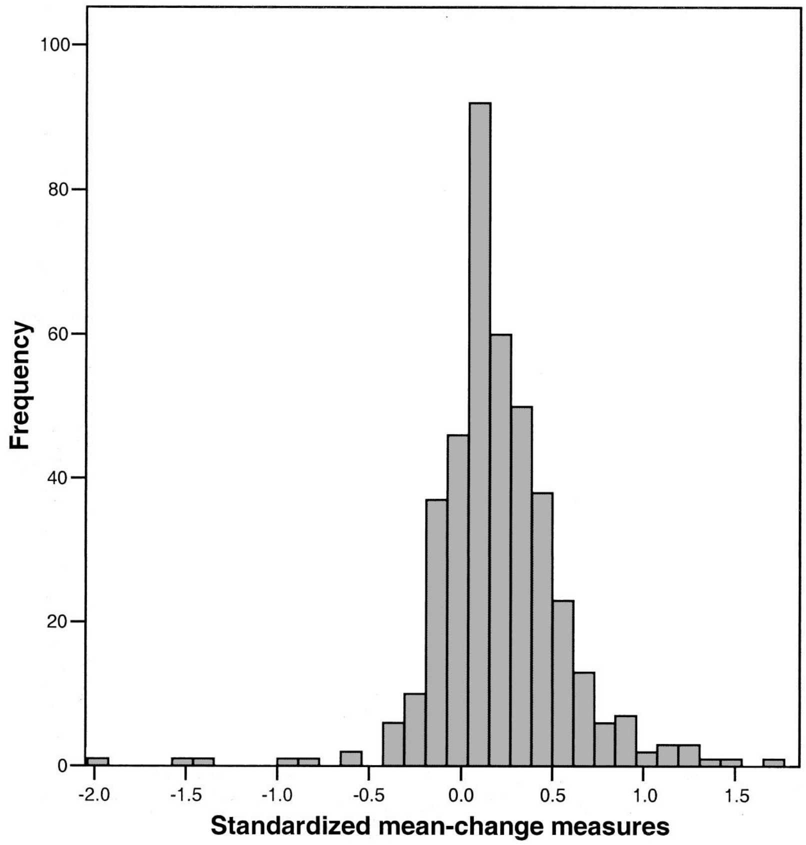

Figure 2. Histogram of change measures.

tional capacity), we analyzed only the samples containing information on improvement (either improved or not improved). In treatment groups, the mean effects for improved groups were significantly larger than the means for not improved samples on all physiological dimensions except flexibility ($z = 0.76$).

No control samples showed improved strength and functional capacity, so comparisons could not be made for these two fitness dimensions. Control participants showing no cardiovascular improvement showed significantly more change in well-being than improved control samples ($z = -2.67$). No difference in change in well-being was found for control participants according to improvement in flexibility.

Prior exercise levels. Many studies had screened participants for prior levels of exercise; 104 treatment effects and 30 control effects came from participants who were required to have been sedentary. Fixed-effects tests showed larger effects for participants who were sedentary (see Table 3 for Q_B values).

Exercise type. Three basic types of exercises—aerobic, resistive, and calisthenics—were found. Some samples experienced more than one type of exercise, and they were separated into combined-exercise groups, such as aerobic–resistive. With one exception, the mean changes for the various exercise types in treatment groups were significantly larger than zero; only the subgroup receiving both aerobic and calisthenic exercise had a nonsignificant mean change of 0.20 ($SE = 0.122$). Aerobic exercise improved psychological well-being the most ($d^C. = 0.29$, $SE = 0.031$), followed closely by resistive exercise ($d^C. = 0.23$, $SE = 0.045$). The effects of these two exercise types do not differ significantly. Resistive exercise combined with aerobic exercise showed the smallest mean change ($d^C. = 0.09$, $SE = 0.037$), but nearly half of these studies measured life satisfaction rather than some more specific outcome (see below for more information on measure types). In the control groups, the mean for participants who had very light calisthenics ($d^C. = 0.11$, $SE = 0.043$) was similar to the mean for those who had no exercise ($d^C. = 0.08$, $SE = 0.029$), and both control groups showed mean changes significantly greater than zero.

Exercise intensity. The effects were also categorized into subgroups based on exercise intensity. In treatment groups, moderate exercise benefited older adults' psychological well-being the most ($d^C. = 0.34$, $SE = 0.041$), whereas light-intensity exercise benefited the least ($d^C. = 0.14$, $SE = 0.018$). Significant variation still remained in all the subgroups (all Q_W values were significant). Because all calisthenic exercise conducted in control groups was of very light intensity, and intensity was not applicable to the no-exercise control-group condition, the results for control groups were identical to the results using exercise type as the moderator.

Measure type. To investigate whether changes in well-being in the treatment and control groups related to the different aspects of psychological well-being measured across studies, and also to reduce the impact of data dependence due to multiple outcomes, we categorized treatment and control effects into 11 categories according to detailed descriptions of the well-being measures. The last section of Table 3 shows the analyses of the 11 measure-type

Table 3
Homogeneity Tests Under Fixed-Effects Models and Estimated Mean Changes in Psychological Well-Being for Treatment and Control Groups Under Random-Effects Models With Moderators

	Treatment groups					Control groups				
Moderator	Number of effects (k)	Weighted mean d^C	SE	95% CI	Q and z tests[a,b]	Number of effects (k)	Weighted mean d^C	SE	95% CI	Q and z tests[a,b]
Overall	274	0.24	0.023	0.20–0.29	Q_W = 1,647.06*	132	0.09	0.024	0.04–0.14	Q_W = 304.97*
Age					Q_B = 135.73*					Q_B = 7.69*
Late middle age	102	0.33	0.034	0.26–0.39	Q_W = 544.74*	19	0.06	0.036	−0.01–0.13	Q_W = 22.93
Young-old	159	0.20	0.031	0.13–0.26	Q_W = 911.15*	103	0.10	0.030	0.05–0.16	Q_W = 246.18*
Old-old	13	0.11	0.082	−0.05–0.27	Q_W = 55.43*	10	0.02	0.063	−0.11–0.14	Q_W = 28.18*
Physical-fitness improvement status										
Cardiovascular					Q_B = 40.87*					Q_B = 11.43*
Improved	84	0.32	0.042	0.23–0.40	Q_W = 475.79*	28	0.01	0.063	−0.11–0.13	Q_W = 54.49*
Not improved	99	0.20	0.039	0.12–0.27	Q_W = 344.95*	52	0.20	0.033	0.14–0.26	Q_W = 91.50*
z					z = 2.09*					z = −2.67*
Strength					Q_B = 2.80					
Improved	44	0.20	0.036	0.13–0.28	Q_W = 88.09*	0				
Not improved	22	0.09	0.037	0.01–0.16	Q_W = 26.55	55	0.02	0.039	−0.05–0.10	Q_W = 106.04*
z					z = 2.13*					
Flexibility					Q_B = 7.39*					Q_B = 9.03*
Improved	19	0.18	0.140	−0.09–0.45	Q_W = 73.03*	11	−0.14	0.140	−0.41–0.13	Q_W = 28.85*
Not improved	17	0.07	0.037	−0.00–0.14	Q_W = 0.28	21	0.11	0.032	0.05–0.17	Q_W = 20.35
z					z = 0.76*					z = −1.74
Functional capacity					Q_B = 13.57**					
Improved	3	0.32	0.053	0.21–0.42	Q_W = 0.78					
Not improved	25	0.09	0.032	0.03–0.15	Q_W = 26.86*	22	0.02	0.078	−0.13–0.17	Q_W = 44.55*
					z = 3.71*					
Prior exercise level					Q_B = 48.36*					Q_B = 20.26*
Sedentary	104	0.35	0.047	0.26–0.44	Q_W = 1,059.58*	30	0.14	0.050	0.05–0.24	Q_W = 58.85*
Not sedentary	170	0.17	0.021	0.13–0.22	Q_W = 539.11*	102	0.08	0.028	0.02–0.13	Q_W = 225.87*
Exercise type					Q_B = 38.95*					Q_B = 14.28*
Aerobic-calisthenics	22	0.20	0.122	−0.04–0.44	Q_W = 75.31*					
Aerobic-resistive	22	0.09	0.037	0.02–0.16	Q_W = 26.55					
Aerobic-calisthenics-resistive	4	0.17	0.086	0.00–0.34	Q_W = 1.45					
Aerobic	151	0.29	0.031	0.23–0.35	Q_W = 1,153.12*					
Resistive	34	0.23	0.045	0.14–0.32	Q_W = 72.40*					
Calisthenics	41	0.15	0.052	0.05–0.26	Q_W = 279.27*	47	0.11	0.043	0.03–0.20	Q_W = 94.27*
None						85	0.08	0.029	0.02–0.13	Q_W = 196.42*
Exercise intensity					Q_B = 65.84*					Q_B = 14.28*
Hard	93	0.26	0.026	0.18–0.34	Q_W = 280.10*					
Moderate	106	0.34	0.041	0.26–0.42	Q_W = 894.92*					
Light	42	0.14	0.018	0.01–0.27	Q_W = 194.46*					
Very light	33	0.19	0.060	0.07–0.30	Q_W = 211.73*	47	0.11	0.043	0.03–0.20	Q_W = 94.27*
Not applicable						85	0.08	0.029	0.02–0.13	Q_W = 196.42*
Measure type					Q_B = 110.22*					Q_B = 38.74*
Anger	13	0.29	0.076	0.14–0.44	Q_W = 25.45*	6	0.17	0.057	0.06–0.28	Q_W = 1.92
Anxiety	51	0.23	0.048	0.14–0.33	Q_W = 281.79*	18	−0.01	0.068	−0.14–0.12	Q_W = 31.23*
Confusion	12	0.15	0.065	0.02–0.28	Q_W = 15.49	5	0.15	0.071	0.01–0.29	Q_W = 2.72
Depression	32	0.29	0.066	0.16–0.42	Q_W = 125.54*	18	0.19	0.078	0.04–0.35	Q_W = 60.39*
Energy	22	0.18	0.126	−0.07–0.48	Q_W = 73.25*	13	0.06	0.090	−0.11–0.24	Q_W = 29.58*

Table 3 (*continued*)

Moderator	Treatment groups: Number of effects (k)	Weighted mean d^C	*SE*	95% CI	Q and z tests[a,b]	Control groups: Number of effects (k)	Weighted mean d^C	*SE*	95% CI	Q and z tests[a,b]
Measure type (*continued*)										
Overall well-being	11	0.37	0.113	0.15–0.59	$Q_W = 37.25$*	4	0.07	0.056	−0.04–0.18	$Q_W = 3.77$
Life satisfaction	48	0.08	0.024	0.03–0.13	$Q_W = 78.23$*	34	0.09	0.031	0.03–0.15	$Q_W = 57.50$*
Physical symptoms	3	0.63	0.139	0.36–0.90	$Q_W = 5.52$	2	0.35	0.126	0.11–0.60	$Q_W = 1.43$
Positive affect	9	0.28	0.098	0.09–0.47	$Q_W = 30.01$*	5	0.20	0.068	0.07–0.33	$Q_W = 7.38$
Self-efficacy	47	0.38	0.073	0.24–0.52	$Q_W = 827.22$*	14	−0.00	0.121	−0.24–0.24	$Q_w = 60.01$*
View of self	26	0.16	0.024	0.11–0.21	$Q_W = 37.09$	13	−0.02	0.044	−0.10–0.07	$Q_W = 10.32$
Gender					$Q_B = 21.57$*					$Q_B = 6.38$*
Female only	97	0.24	0.041	0.16–0.32	$Q_W = 411.79$*	55	0.12	0.031	0.06–0.18	$Q_w = 75.42$*
Male only	48	0.24	0.058	0.12–0.35	$Q_W = 359.22$*	20	0.03	0.095	−0.16–0.22	$Q_W = 47.40$*
Both	129	0.25	0.030	0.19–0.30	$Q_W = 854.49$*	57	0.08	0.033	0.01–0.14	$Q_W = 175.76$*
Randomization					$Q_B = 12.17$*					$Q_B = 0.03$
Used	169	0.17	0.024	0.12–0.21	$Q_W = 825.95$*	117	0.10	0.027	0.05–0.15	$Q_W = 290.52$*
Not used	52	0.25	0.066	0.13–0.38	$Q_W = 268.39$*	15	0.07	0.033	0.00–0.13	$Q_W = 14.41$
Matching					$Q_B = 0.68$					$Q_B = 9.99$*
Used	78	0.15	0.046	0.06–0.24	$Q_W = 551.03$*	56	0.12	0.04	0.04–0.19	$Q_W = 120.82$*
Not used	143	0.20	0.027	0.15–0.26	$Q_W = 554.79$*	76	0.04	0.016	0.01–0.07	$Q_W = 174.15$*

Note. W = within; B = between.
[a] Two-group random-effects z tests of mean differences between levels of the moderator variable are shown only for physical improvement status because the other moderators have more than two levels.
[b] Q statistics were computed based on the fixed-effects model.
* $p < .05$.

groups. Mean effects differed significantly among the measure types for both treatment samples, $Q_B(10) = 110.22$, $p < .001$, and control samples, $Q_B(10) = 38.74$, $p < .001$. Grouping effects by measure type created three homogeneous subgroups of effects from treatment samples (for the outcomes of confusion, physical symptoms, and view of self) and six homogeneous sets of effects from control samples (anger, confusion, physical symptoms, positive affect, overall well-being, and view of self). Homogeneity was indicated by nonsignificant Q_W values.

Differences between the treatment- and control-group means under the random-effects model were significant for only four measure types: anxiety ($z = 2.88$, $p < .02$), overall well-being ($z = 2.38$, $p < .01$), self-efficacy ($z = 2.69$, $p < .01$), and view of self ($z = 3.59$, $p < .01$). The mean effects in Table 3 show that exercise had the largest impact on improving physical symptoms in treatment samples ($d^C_. = 0.63$, $SE = 0.139$). Control groups also showed large changes for physical-symptoms measures ($d^C_. = 0.35$, $SE = 0.126$). However, only five effects represented physical symptoms. Exercise had the least impact on life satisfaction ($d^C_. = 0.08$, $SE = 0.023$ for treatment groups). The only nonsignificant impact of exercise for treatment samples was on energy (95% confidence interval: −0.07 to 0.48).

Gender. Gender differences were found for both treatment, $Q_B(2) = 21.57$, and control samples, $Q_B(2) = 6.38$. However, because the gender composition of the samples did not provide much explanatory power, and differences were not significant under a random-effects model, gender was not included in other analyses.

Design issues. Many studies had used randomization and/or matching to assign participants to treatment and control groups. Treatment and control groups from studies using randomization showed similar changes, whereas changes from samples that were not randomly assigned appear to differ (0.25 for treatment groups vs. 0.07 for controls). However, when paired differences were computed for only studies that had both treatment and control samples, no differences due to use of randomization were found, $Q_B(2) = 3.49$, $p = .06$. Similar results were found for the use of matching, $Q_B(2) = 0.24$, $p = .62$. Because the paired analyses are based on within-study comparisons, these design differences can be considered negligible.

Weighted Multiple Regression Analyses of Exercise Dose

To determine whether changes in well-being were predicted by exercise dose (weeks of exercise, sessions per week, and session length in minutes), we used weighted multiple regression analyses. Initially, participant age was included in the form of a dichotomous variable (contrasting younger and older participants). However, age was not significant in any of the models and was not useful in accounting for the variance among the effects. To achieve more parsimonious models, we excluded age from further analyses.

A regression model predicting effect size from weeks, number of weekly sessions, and minutes per session was run first for the sets of all treatment and control effects separately, then for each measure type. Mixed-effects models, which account for between-studies uncertainty, were used to explore the relationship between exercise-dose variables and the effects; results are shown in Table 4. The slopes for weeks of exercise were negative and significant for the full sets of treatment groups ($b_{weeks} = -0.006$, $p < .05$) and control groups ($b_{weeks} = -0.006$, $p < .05$), controlling for session frequency and length. This indicates that longer exercise programs showed either less positive change or actual reductions in levels of psychological well-being. The other dose variables (frequency and length) were positive but not significant in the treatment and control samples.

Regressions were next computed separately within each measure type for the treatment samples. The impact of exercise duration in weeks (controlling for sessions per week and session length) was inconsistent across measure types. For anxiety ($b_{weeks} = -0.009$, $p < .01$), depression ($b_{weeks} = -0.011$, $p < .01$), and self-efficacy ($b_{weeks} = -0.049$, $p < .05$), decreases in well-being are found for longer interventions. The remaining measure types showed b_{weeks} slopes that did not differ from zero. Inconsistent relations were also found for number of sessions per week, which was significant and positive for anxiety ($b_{sessions} = 0.167$, $p < .01$) and self-efficacy ($b_{sessions} = 0.493$, $p < .01$) after we controlled for weeks of exercise and session length. Finally, after controlling for weeks and days of exercise, we found that longer exercise sessions led to positive significant changes in anxiety ($b_{minutes} = 0.008$, $p < .05$). The remaining b_{weeks} slopes of measures types did not differ from zero. In general, the session length (in minutes) did not relate significantly to the outcome.

Discussion

The present meta-analysis of comparative data showed a small but significant effect ($d^C_. = 0.19$) of exercise on well-being in older adults without clinical disorders. However, treatment groups showed almost three times as much pretest–posttest change ($d^C_. = 0.24$) as did control groups ($d^C_. = 0.09$), suggesting a causal effect for physical activity on psychological well-being enhancement, in accordance with qualitative reviews studying the aging population (Brown, 1992; McAuley & Rudolph, 1995; O'Connor et al., 1993; Spirduso & Cronin, 2001).

Our meta-analysis supports the perception that psychological well-being is a multifaceted phenomenon, especially in older adults. Previous meta-analyses examining the effect of physical activity on psychological variables focused on emotions, specifically negative emotions such as anxiety and depression. On the basis of the conceptual framework proposed for evaluating well-being in older age (Stewart & King, 1991), we also examined constructs related to self-perception and global well-being. Indeed, our results indicate that those constructs, specifically self-efficacy, overall well-being, and view of self, were significantly affected by physical activity, showing the largest treatment-control differences.

These findings support the literature arguing that physical activity might provide a mastery experience for older adults whose physical self-efficacy may be deteriorating along with their functional abilities (McAuley & Katula, 1998; McAuley & Rudolph, 1995; McAuley, Shaffer, & Rudolph, 1995). The significant moderating effect of improvements in cardiovascular capacity and strength, and specifically in daily functioning (functional capacity), adds more support to that mastery experience. Moreover, self-efficacy and view of self have been shown to moderate people's affective reactions (Jerome et al., 2002; Marquez et al., 2002; McAuley et al., 1999) and reduce anxiety level possibly by dampening biological stress reactions such as cortisol concentrations (Bandura, 1991; Rudolph & McAuley, 1995). This may explain the positive alteration in anxiety found in the current meta-analysis

Table 4
Results of the Mixed-Model Multiple-Regression Analyses

Outcome	Number of effects (k)	Q_{Model}[a]	Q_{Error}[b]	b_{weeks} (SE)	$\|z\|$	$b_{sessions}$ (SE)	$\|z\|$	$b_{minutes}$ (SE)	$\|z\|$
Control	112	8.27*	107.32	−0.006 (0.0021)	2.86*	0.037 (0.0262)	1.41	0.0006 (0.0005)	1.20
Treatment	240	8.76*	224.20*	−0.006 (0.0020)	3.00*	0.036 (0.0249)	1.45	0.0005 (0.0006)	0.83
Anger	10	1.67	5.42	0.010 (0.0279)	0.36	−0.145 (0.1227)	1.18	0.0014 (0.0081)	0.17
Anxiety	46	10.62*	45.28	−0.009 (0.0032)	2.81*	0.167 (0.0589)	2.84*	0.0084 (0.0042)	2.00*
Confusion	9	1.50	4.68	0.029 (0.0277)	1.05	0.042 (0.1112)	0.38	0.0072 (0.0073)	0.99
Depression	29	8.47*	22.25	−0.011 (0.0039)	2.82*	0.049 (0.0668)	0.73	0.0014 (0.0014)	1.00
Energy	16	1.69	10.55	0.038 (0.0469)	0.81	−0.357 (0.2809)	1.27	−0.0078 (0.0120)	0.65
Overall well-being	11	0.94	6.80	−0.002 (0.0387)	0.05	0.097 (0.1580)	0.61	0.008 (0.0118)	0.68
Life satisfaction	44	4.43	57.07*	0.003 (0.0034)	0.88	−0.015 (0.0294)	0.51	−0.0013 (0.0009)	1.44
Physical symptoms	2								
Positive affect	9	3.91	5.15	−0.032 (0.0198)	1.62	0.405 (0.3446)	1.18	−0.0056 (0.0125)	0.45
Self-efficacy	46	11.47*	42.44	−0.049 (0.0225)	2.18*	0.493 (0.1513)	3.26*	0.0057 (0.0038)	1.50
View of self	20	8.78*	24.61	−0.007 (0.0080)	0.88	0.014 (0.0506)	0.28	0.0000 (0.0016)	0.00

Note. Results could not be obtained for physical symptoms effects because at least three studies are needed to estimate any regression model.
[a] $df = 3$. [b] $df = k - 4$.
* $p < .05$.

as a result of engaging in physical activity. On the other hand, no treatment-control differences were found on depression, anger, and confusion, possibly because of (a) the nonclinical disorders nature of the study population, which may have led to (b) a floor effect, that is, the absence of such negative mood states in this population. Most instruments assessing these feeling states are designed for use in clinical populations.

Another well-being component that did not show more improvement for treatment groups than controls is life satisfaction. This finding supports Rejeski and Mihalko's (2001) view that life satisfaction is not affected merely by improved daily functioning and efficacy. People may readily admit to compromised function yet are satisfied with their limitations and with life in general. Losses of physical competence are particularly difficult for individuals who have high internal standards like athletes and musicians (Spirduso, 1995), for whom physical achievement plays a large, if not dominating, role in life. Consequently, older adults who place lower importance on their physical function may express more satisfaction with their functional state than do their peers who have functional limitations but continue to place a relatively high value on their physical function (Rejeski & Mihalko, 2001). Thus, the absence of exercise effects on life satisfaction in a population without clinical disorders may result from a relatively high level of life satisfaction the older adults regularly experience and the lack of sufficient sensitivity to small changes in the instruments used to measure life satisfaction.

As for the moderating effect of mode of exercise, aerobic training, closely followed by resistive training, was most significant in affecting psychological well-being. Studies comparing aerobic training with resistive training produced conflicting results. Previous meta-analyses have indicated anxiety reduction following aerobic exercise but no improvements due to resistance training (Landers & Petruzzello, 1994; Petruzzello et al., 1991). Conversely, Arent et al. (2000) indicated the superiority of resistive training over aerobic training in altering mood in older adults. Dunn et al.'s (2001) narrative review identified more studies involving aerobic exercise, but the few that examined resistance training showed significant effects in reducing depressive symptoms. Perhaps, in older adults without clinical disorders, both aerobic and resistive training increase feelings of mastery and self-efficacy thus improving global well-being, whereas aerobic training is more dominant in alleviating symptoms of depression in the clinical population.

Yet another central issue in the physical activity–well-being relation is the moderating effect of exercise dose. Although Dishman (1986) contended that high intensity is needed to alter affect, and others (Dunn et al., 2001; Ekkekakis & Petruzzello, 1999; Rejeski, 1994) failed to reveal a consistent pattern of intensity effect on psychological factors, our findings indicate that moderate exercise benefited most the psychological well-being of older adults, followed by high-intensity exercise. Because moderate intensity has also been recommended for enhancing physical health (Pate et al., 1995), the implications are that high-intensity physical activity is not needed in order to promote physical or psychological health.

Other dose–response issues concern the duration of exercise, its frequency, and session length. The current results on exercise program duration are inconsistent with previous meta-analyses performed on younger adults, which indicated that longer periods of training are associated with the largest effect size on anxiety (Petruzzello et al., 1991) and depression (North et al., 1990). Furthermore, the current findings are also inconsistent with the conclusions of the narrative review on the aging population (McAuley & Rudolph, 1995), suggesting that programs lasting less than 10 weeks have a less consistent effect on psychological well-being than longer study protocols. In contrast, our results, consistent with those of Arent et al. (2000) on mood change, show that longer periods of physical activity are associated with less positive change in well-being. Perhaps after relatively little exercise activity, individuals can realize the increased weight they can lift or the increased distance they can walk, which may stimulate an increased sense of control and general well-being (Spirduso, 1995). As for frequency and length of sessions, the impact was not consistent across all well-being measures. However, improvements in self-efficacy and anxiety were positively associated with increased frequency of physical activity.

Our findings on the moderating effect of mean age of samples suggest a gradual decrease in the effect size of physical activity on well-being as one ages. These findings support the results of Ruuskanen and Ruoppila (1995) who reported that active older adults at the age of 60–75 had fewer depressive symptoms than their nonactive peers. Among the oldest old (76+ years), however, this activity effect was not prominent. It is possible that the self-efficacy–affect relationship demonstrated in young populations (Jerome et al., 2002; Marquez et al., 2002) weakens as a person gets older. If biochemical activity is involved in this relationship (Rudolph & McAuley, 1995), and certain intensity is needed for affect alteration, people may not reach that level of intensity as they get older. Also, as people age, their psychological well-being becomes more multifaceted (more complex) as more functional and social losses are involved. Thus, the ability of physical activity to affect well-being may decrease. As for gender, consistent with McAuley and Rudolph's (1995) conclusions, the current results suggest that the well-being of older men and women is not differentially affected by physical activity.

Quite a few mediating mechanisms have been proposed for explaining the physical activity–well-being relationship. Our findings of the large effect of physical activity on self-efficacy, view of self, and global well-being, along with the moderating effect of improved fitness in the causal link between physical activity and well-being (specifically fitness related to daily functioning), support, to a large extent, social–cognitive theory. Self-efficacy, the primary variable in this theory, concerns the individual's beliefs in his or her capabilities to successfully execute necessary courses of action to satisfy situational demands, and it is the most salient variable affecting well-being and psychological health (Bandura, 1991). For older adults, whose self-efficacy may be deteriorating along with their functional abilities, physical activity may provide a mastery experience that leads to increased self-efficacy (McAuley, Shaffer, & Rudolph, 1995), which in turn leads to improved psychological well-being (McAuley & Rudolph, 1995). Furthermore, this mastery experience may occur with moderate-intensity (Pate et al., 1995), short-duration physical activity. For example, significant gains in muscle strength and mass as well as an improvement in bone density were reported following only 2 weeks, and in another study, very old frail individuals increased their muscle strength by almost 180% in 8 weeks (Evans, 1999). On the other hand, within shorter time frames, high frequency of activity is recommended for increasing self-efficacy. This result is consistent with the recommendation to exercise every day, if possible, to promote physical fitness (Pate et al., 1995).

Although the effect sizes found in the present study were significant, their magnitudes are small. One explanation for this is the no-clinical-disorders nature of the population. Many well-being measures, specifically emotional well-being measures, are designed for use in clinical populations. The fact that no treatment-control differences were found on depression, confusion and, anger are consistent with this hypothesis. Another explanation is that effects may be reduced because of other methodological problems typical to experimental studies of older adults (O'Connor et al., 1993), such as participant selection, lack of adequate control groups, failure to control for extraneous variables, and so forth. However, it might well be the case that older adults without clinical disorders do not suffer depression, confusion, and anger to the extent that physical activity will result in a positive effect.

Finally, the studies published in the literature did not allow us to estimate effects pertaining to the minimum time, intensity, and mode of exercise needed to achieve a meaningful psychological effect for exercise engagement in older adults. Advanced research must target the environmental thresholds (i.e., exercise mode, duration, and intensity) and age category, which signify the beginning of psychological gains, and possibly psychological declines associated with various types of physical activities. More specifically, researchers must use methods of inquiry that explore the underlying mechanisms that link psychological benefits to improved functionality, such as resistive training, and the ones that link psychological benefits to well-being in the absence of physical benefit but gain in mental benefits (e.g., Feldenkrais, Yoga, or Pilates). A more comprehensive examination of moderating and mediating variables must be implemented to enable understanding and promotion of physical-activity programs in advanced age.

References

References marked with an asterisk indicate studies included in the meta-analysis.

American College of Sports Medicine. (1998). Exercise and physical activity for older adults. *Medicine and Science in Sports and Exercise, 30,* 992–1008.

American College of Sports Medicine. (2000). *ACSM's guidelines for exercise testing and prescription.* Baltimore: Lippincott Williams & Wilkins.

Arent, S. M., Landers, D. M., & Etnier, J. L. (2000). The effects of exercise on mood in older adults: A meta-analytic review. *Journal of Aging and Physical Activity, 8,* 407–430.

Bandura, A. (1991). Self-efficacy mechanism in physiological activation and health-promoting behavior. In J. Madden (Ed.), *Neurobiology of learning, emotion, and affect* (pp. 229–269). New York: Raven Press.

Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology, 41,* 257–278.

Becker, B. J. (2000). Multivariate meta-analysis. In H. E. A. Tinsley & S. Brown (Eds.), *Handbook of applied and multivariate statistics and mathematical modeling* (pp. 499–525). San Diego: Academic Press.

*Bennett, J., Carmack, M. A., & Gardner, V. J. (1982). The effect of a program of physical exercise on depression in older adults. *The Physical Educator, 39,* 21–24.

Bompa, T. O. (1999). *Periodization training for sports.* Champaign, IL: Human Kinetics.

Brown, D. R. (1992). Physical activity, ageing, and psychological well-being: An overview of the research. *Canadian Journal of Sport Sciences, 17,* 185–193.

*Brown, D. R., Wang, Y., Ward, A., Ebbeling, C. B., Fortlage, L., Puleo, E., et al. (1995). Chronic psychological effects of exercise and exercise plus cognitive strategies. *Medicine and Science in Sports and Exercise, 27,* 765–775.

*Chin-A-Paw, M. J., de-Jong, N., Schouten, E. G., Van-Staveren, W. A., & Kok, F. J. (2002). Physical exercise or micronutrient supplementation for the well being of the frail elderly? A randomized controlled trial. *British Journal of Sports Medicine, 36,* 126–131.

Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychological Science, 14,* 125–130.

*Cowper, P. A., Morey, M. S., Bearon, L. B., Sullivan, R. J., DiPasquale, R. C., Crowley, G. M., et al. (1991). The impact of supervised exercise on the psychological well-being and health status of older veterans. *The Journal of Applied Gerontology, 10,* 469–485.

Craft, L. L., & Landers, D. M. (1998). The effect of exercise on clinical depression and depression resulting from mental illness: A meta-analysis. *Journal of Sport and Exercise Psychology, 20,* 339–357.

*Damush, T. M., & Damush, J. G. (1999). The effect of strength training on strength and health-related quality of life in older adult women. *The Gerontologist, 39,* 705–710.

Dishman, R. K. (1986). Mental health. In V. Seefeldt (Ed.), *Physical activity and well-being* (pp. 303–341). Reston, VA: American Alliance for Health, Physical Education, Recreation & Dance.

Dunn, A. L., Trivedi, M. H., & O'Neal, H. A. (2001). Physical activity dose-response effects on outcomes of depression and anxiety. *Medicine & Science in Sports & Exercise, 33,* S587–S597.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple graphical test. *British Medical Journal, 315,* 629–634.

Ekkekakis, P., & Petruzzello, S. J. (1999). Acute aerobic exercise and affect: Current status, problems and prospects regarding dose-response. *Sports Medicine, 28,* 337–374.

*Emery, C. F., & Gatz, M. (1990). Psychological and cognitive effects of an exercise program for community-residing older adults. *The Gerontologist, 30,* 184–188.

Evans, W. J. (1999). Exercise training guidelines for the elderly. *Medicine & Science in Sports & Exercise, 31,* 12–17.

Fox, K. R. (2000). Self-esteem, self-perceptions and exercise. *International Journal of Sport Psychology, 31,* 228–240.

Gauvin, L., & Spence, J. C. (1996). Physical activity and psychological well-being: Knowledge base, current issues, and caveats. *Nutrition Reviews, 54,* S53–S67.

*Hassmen, P., & Koivula, N. (1997). Mood, physical working capacity and cognitive performance in the elderly as related to physical activity. *Aging: Clinical and Experimental Research, 9,* 136–142.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

*Hickey, T., Wolf, F. M., Robins, L. S., Wagner, M. B., & Harik, W. (1995). Physical activity training for functional mobility in older persons. *The Journal of Applied Gerontology, 14,* 357–371.

Jerome, G. J., Marquez, D. X., McAuley, E., Canaklisoval, S., Snook, E., & Vickers, M. (2002). Self-efficacy effects on feeling states in women. *International Journal of Behavioral Medicine, 9,* 139–154.

*Jette, A. M., Harris, B. A., Sleeper, L., Lachman, M. E., Heislein, D., Giorgetti, M., & Levenson, C. (1996). A home-based exercise program for nondisabled older adults. *Journal of the American Geriatrics Society, 44,* 644–649.

*Katula, J. A., Blissmer, B. J., & McAuley, E. (1999). Exercise intensity and self-efficacy effects on anxiety reduction in healthy, older adults. *Journal of Behavioral Medicine, 22,* 233–247.

*King, A. C., Pruitt, L. A., Phillips, W., Oka, R., Rodenburg, A., & Haskell, W. L. (2000). Comparative effects of two physical activity programs on measured and perceived physical functioning and other health-related quality of life outcomes in older adults. *Journals of Gerontology, Series A: Biological Sciences and Medical Sciences, 55,* M74–M83.

*King, A. C., Taylor, C. B., & Haskell, W. L. (1993). Effects of differing intensities and formats of 12 months of exercise training on psychological outcomes in older adults. *Health Psychology, 12,* 292–300.

Landers, D. M., & Petruzzello, S. J. (1994). Physical activity, fitness, and anxiety. In C. B. Bouchard, R. J. Shephard, & T. Stephens (Eds.), *Physical activity, fitness and health: International proceedings and consensus statement* (pp. 868–882). Champaign, IL: Human Kinetics.

Marquez, D. X., Jerome, G. J., McAuley, E., Snook, E. M., & Canaklisova, S. (2002). Self-efficacy manipulation and state anxiety responses to exercise in low active women. *Psychology and Health, 17,* 783–791.

Martin, K. A., & Sinden, A. R. (2001). Who will stay and who will go? A review of older adults' adherence to randomized controlled trials of exercise. *Journal of Aging and Physical Activity, 9,* 91–114.

Martinsen, E. W., & Stephens, T. (1994). Exercise and mental health in clinical and free-living populations. In R. Dishman (Ed.), *Advances in exercise adherence* (pp. 55–71). Champaign, IL: Human Kinetics.

McAuley, E. (1994). Physical activity and psychosocial outcomes. In C. B. Bouchard, R. J. Shephard, & T. Stephens (Eds.), *Physical activity, fitness and health: International proceedings and consensus statement* (pp. 1040–1055). Champaign, IL: Human Kinetics.

*McAuley, E., Bane, S. M., & Mihalko, S. L. (1995). Exercise in middle-aged adults: Self-efficacy and self-presentational outcomes. *Preventive Medicine, 24,* 319–328.

*McAuley, E., Blissmer, B., Marquez, D. X., Jerome, G. J., Kramer, A. F., & Katula, J. (2000). Social relations, physical activity, and well-being in older adults. *Preventive Medicine, 31,* 608–617.

*McAuley, E., Jerome, G. J., Marquez, D. X., Elavsky, S., & Blissmer, B. (2003). Exercise self-efficacy in older adults: Social, affective, and behavioral influences. *Annals of Behavioral Medicine, 25,* 1–7.

McAuley, E., & Katula, J. (1998). Physical activity interventions in the elderly: Influence on physical health and psychological function. In R. Shulz, G. Maddox, & M. P. Lawton (Eds.), *Annual review of gerontology and geriatrics* (Vol. 18, pp. 111–154). New York: Springer Publishing Company.

*McAuley, E., Katula, J., Mihalko, S. L., Blissmer, B., Duncan, T. E., Pena, M., & Dunn, E. (1999). Mode of physical activity and self-efficacy in older adults: A latent growth curve analysis. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 54,* P283–P292.

*McAuley, E., Marquez, D. X., Jerome, G. J., Blissmer, B., & Katula, J. (2002). Physical activity and physique anxiety in older adults: Fitness and efficacy influences. *Aging and Mental Health, 6,* 222–230.

*McAuley, E., Mihalko, S. L., & Bane, S. M. (1997). Exercise and self-esteem in middle-aged adults: Multidimensional relationships and physical fitness and self-efficacy influences. *Journal of Behavioral Medicine, 20,* 67–83.

McAuley, E., & Rudolph, D. (1995). Physical activity, aging, and psychological well-being. *Journal of Aging and Physical Activity, 3,* 67–96.

McAuley, E., Shaffer, S., & Rudolph, D. (1995). Self-efficacy and affective responses to exercise testing in older males. *International Journal of Aging and Human Development, 41,* 13–27.

McAuley, E., Talbot, H. M., & Martinez, S. (1999). Manipulating self-efficacy in the exercise environment in women: Influences on affective responses. *Health Psychology, 18,* 288–294.

McDonald, D. G., & Hodgdon, J. A. (1991). *Psychological effects of aerobic fitness training.* New York: Springer-Verlag.

*McNeil, J. K., LeBlanc, E. M., & Joyner, M. (1991). The effect of

exercise on depressive symptoms in the moderately depressed elderly. *Psychology and Aging, 6,* 487–488.
*Mihalko, S. L., & McAuley, E. (1996). Strength training effects on subjective well-being and physical function in the elderly. *Journal of Aging and Physical Activity, 4,* 56–68.
Moritz, S. E., Feltz, D. L., Fahrbach, K. R., & Mack, D. E. (2000). The relation of self-efficacy measures to sport performance: A meta-analytic review. *Research Quarterly for Exercise and Sport, 71,* 3–20.
*Netz, Y., Tenenbaum, G., & Sagiv, M. (1988). Pattern of psychological fitness as related to pattern of physical fitness among older adults. *Perceptual and Motor Skills, 67,* 647–655.
*Nieman, D. C., Warren, B. J., Dotson, R. G., Butterworth, D. E., & Henson, D. (1993). Physical activity, psychological well-being, and mood state in elderly women. *Journal of Aging and Physical Activity, 1,* 22–33.
North, T. C., McCullagh, P., & Tran, Z. V. (1990). Effect of exercise on depression. *Exercise and Sport Science Reviews, 18,* 379–415.
O'Connor, P. J., Aenchbacher, L. E., & Dishman, R. K. (1993). Physical activity and depression in the elderly. *Journal of Aging and Physical Activity, 1,* 34–58.
Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., et al. (1995). Physical activity and public health. *The Journal of the American Medical Association, 273,* 402–407.
*Perri, S., & Templer, D. I. (1985). The effects of an aerobic exercise program on psychological variables in older adults. *The International Journal of Aging and Human Development, 20,* 167–171.
*Perrig-Chiello, P., Perrig, W. J., Ehrsam, R., Staehelin, H. B., & Krings, F. (1998). The effects of resistance training on well-being and memory in elderly volunteers. *Age and Ageing, 27,* 469–475.
Petruzzello, S. J., Landers, D. M., Hatfield, B. D., Kubitz, K. A., & Salazar, W. (1991). A meta-analysis on the anxiety-reducing effects of acute and chronic exercise. *Sports Medicine, 11,* 143–182.
*Pierce, E. F., & Pate, D. W. (1994). Mood alterations in older adults following acute exercise. *Perceptual and Motor Skills, 79,* 191–194.
Raglin, J. S., & Morgan, W. P. (1985). Influence of vigorous exercise on mood state. *Behavior Therapy, 8,* 179–183.
Raudenbush, S. W. (1994). Random effects models. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 301–321). New York: Russell Sage Foundation.
Rejeski, W. J. (1994). Dose-response issues from a psychosocial perspective. In C. B. Bouchard, R. J. Shephard, & T. Stephens (Eds.), *Physical activity, fitness and health: International proceedings and consensus statement* (pp. 1040–1055). Champaign, IL: Human Kinetics.
Rejeski, W. J., & Mihalko, S. L. (2001). Physical activity and quality of life in older adults. *Journals of Gerontology, Series A: Biological Sciences and Medical Sciences, 56,* 23–35.
Rudolph, D. L., & McAuley, E. (1995). Self-efficacy and salivary cortisol responses to acute exercise in physically active and less active adults. *Journal of Sport and Exercise Psychology, 17,* 206–213.
Ruuskanen, J. M., & Ruoppila, I. (1995). Physical activity and psychological well-being among people aged 65 to 84 years. *Age and Ageing, 24,* 292–296.
*Shin, Y. (1999). The effects of a walking exercise program on physical function and emotional state of elderly Korean women. *Public Health Nursing, 16,* 146–154.
Spirduso, W. W. (1995). *Physical dimensions of aging.* Champaign, IL: Human Kinetics.
Spirduso, W. W., & Cronin, D. L. (2001). Exercise dose-response effects on quality of life and independent living in older adults. *Medicine & Science in Sports & Exercise, 33,* S598–S608.
Stewart, A. L., & King, A. C. (1991). Evaluating the efficacy of physical activity for influencing quality of life outcomes in older adults. *Annals of Behavioral Medicine, 13,* 108–116.
*Stoll, O., & Alfermann, D. (2002). Effects of physical exercise on resources evaluation, body self-concept and well-being among older adults. *Anxiety, Stress and Coping: An International Journal, 15,* 311–319.
*Swoap, R. A., Norvell, N., Graves, J. E., & Pollock, M. L. (1994). High versus moderate intensity aerobic exercise in older adults: Psychological and physiological effects. *Journal of Aging and Physical Activity, 2,* 293–303.
*Tlokinski, M. K. (1997). Physical activity and depression and hypochondria: The impact of a prevention-intervention programme on the elderly people. In R. Lidor & M. Bar-Eli (Eds.), *Innovations in sport psychology: Linking theory and practice. Proceedings of the International Society of Sport Psychology* (pp. 697–699). Netanya, Israel: Zinman College of Physical Education and Sport Sciences, Wingate Institute.
*Tsutsumi, T., Don, B. M., Zaichkowsky, L. D., Takenaka, K., & Oka, K. (1998). Comparison of high and moderate intensity of strength training on mood and anxiety in older adults. *Perceptual and Motor Skills, 87,* 1003–1011.
*Valliant, P. M., & Asu, M. E. (1985). Exercise and its effects on cognition and physiology in older adults. *Perceptual and Motor Skills, 61,* 1031–1038.
*Watanabe, E., Takeshima, N., Okada, A., & Inomata, K. (2000). Comparison of water- and land-based exercise in the reduction of state anxiety among older adults. *Perceptual and Motor Skills, 91,* 97–104.
*Watanabe, E., Takeshima, N., Okada, A., & Inomata, K. (2001). Effects of increasing expenditure of energy during exercise on psychological well-being in older adults. *Perceptual and Motor Skills, 92,* 288–298.
*Williams, K., Gill, D. L., Butki, B., & Kim, B. J. (2000). A home-based intervention to improve balance, gait and self-confidence in older women. *Activities, Adaptation and Aging, 24,* 57–69.
*Williams, P., & Lord, S. R. (1997). Effects of group exercise on cognitive functioning and mood in older women. *Australian and New Zealand Journal of Public Health, 21,* 45–52.
*Wininger, S. R. (2002). The anxiolytic effect of aqua aerobics in elderly women. *Perceptual and Motor Skills, 94,* 338–340.

Received July 31, 2003
Revision received December 28, 2004
Accepted February 1, 2005 ■