Effects of Aquatic Therapy on Adults with Multiple Sclerosis

Ellen Broach and John Dattilo

Research on the effects of aquatic therapy (AT) on people with multiple sclerosis (MS) is limited. To address the need for research on this topic a replication and extension of Broach, Groff, Dattilo, Yaffe, and Gast (1998) was conducted to examine effects of AT on participants' gross motor activity and fatigue. This single subject, multiple-probe design across participants was used with four adults with relapsing-remitting MS (ages 30–53). This study incorporated two gross motor dependent measures used by Broach et al.: walking up and down stairs and rotations on a bicycle. The study also included measurements of rotations on an upper ergometer and fatigue. Fatigue was examined via a mental and physical fatigue scale. This study supported previous findings that participation in AT by individuals with MS resulted in improvements in gross motor behaviors, and on walking stairs and bicycling. Mixed results prevented conclusions regarding the effect of AT on mental fatigue.

KEY WORDS: *Aquatic Therapy, Aquatic Exercise, Multiple Sclerosis, Therapeutic Recreation, Strength, Fatigue*

Multiple sclerosis (MS) is a degenerative neurological disorder characterized by the demyelinization of the central nervous system (CNS) pathways (Paty & Ebers, 1998). The term "multiple" is used because the disease usually affects more than one area of the nervous system (Schapiro, 1987). It is the demyelinization of the nerves that brings about MS symptoms, including muscular weakness, spasticity, hypersensitivity to heat, decreased

Ellen Broach is an Assistant Professor in the Department of Recreation and Sports Management at Georgia Southern University and John Dattilo is a Professor in the Department of Recreation and Leisure Studies at the University of Georgia. *The authors would like to thank the study participants and the Shepherd Center in Atlanta, GA for their financial assistance and support in the organization of this project.*

visual acuity, pain, emotional problems, cognitive dysfunction, and fatigue (National Multiple Sclerosis Society, 1992). Fatigue is a common complaint of individuals with MS and is a major factor in their psychological and social well-being (Deluca, Johnson, & Natelson, 1993; Sandroni, Walker, & Starr, 1992). The most common type of fatigue reported by individuals with MS results from poor conductivity of demyelinated or remyelinated fibers (Holtakers, 1993), which may be exacerbated by an increase in body temperature. Deluca et al. (1993) explained that individuals with MS often state that they could manage the problems and symptoms that MS brings if they were not so fatigued and had more energy. Despite the threat of fatigue from physical activity, exercise appears to improve both cardiorespiratory fitness and skeletal muscle function (Ponichtera-Mulcare, 1993). Therefore, Holtakers recommended exercise at submaximal levels for individuals with MS to avoid fatiguing while improving fitness, coordination, circulation, as well as reducing stress and spasticity.

MS is characterized by a variable pattern of progression typically including exacerbations of acute symptoms. Most individuals with MS have one or several exacerbations that pass into remission and do not reappear for years, if at all (Schapiro, 1987). The recovery from an exacerbation may be complete or partial (Rudick, 1990). Rudick added that the rise in the body core temperature from physical exercise causes inflammation of the nerves that may temporarily intensify the symptoms associated with MS. While the National MS Society recommended physical activity to cope with symptoms of MS, some individuals may experience adverse effects from certain land-based physical activity programs (Holtakers, 1993; Schapiro, Petajan, Kosich, Molk, & Fenney, 1988). Therefore, functional interventions performed in the water can be a useful alternative to certain land-based treatment because participants can typically participate with lower body temperature, less muscle stress, less energy expenditure, less guarded actions associated with fear of falling, and improved performance.

Aquatic therapy is an intervention that can be used to facilitate physical and psychological improvement while promoting acquisition of skills such as swimming and water exercise associated with leisure participation (Broach & Dattilo, 1996). Broach and Dattilo presented a review of the literature concerning effects of AT using prescribed swimming and exercise to improve health. The authors concluded that, while most of the studies reviewed were specific to aquatic exercise and not prescribed interventions, AT appears to have numerous psychological and physical implications including improved endurance, pulmonary function strength, self-concept, and mood, and decreased perceptions of pain. They further stated that research supports the belief that participation in an appropriate AT intervention can provide individuals with a realistic solution for improving functional ability while engaging in enjoyable activity.

Despite the potential benefits of AT, there is limited research examining the effects of water activity on individuals with MS. The research that has been conducted lacks rigor with only three studies examining physical effects of aquatic exercise on MS (see Broach et al., 1998; Gehlsen, Grigsby, & Winant, 1984; Gehlsen et al., 1986).

Gehlsen, Grigsby, and Winant (1984) conducted pre, mid, and post training measurements of isokinetic peak torque, work, power, and fatigue on a group of 10 individuals with MS. The 10 week training program (3 × week; 1 hour sessions) consisted of swimming and or calisthenics at an intensity of approximately 60–75% of each patient's maximal heart rate. Results showed significant improvement in isokinetic peak torque during knee extensions. Significant effects were not seen in knee flexion. Peak torque changes occurred over the first five weeks, between pre and mid trial measurements, with a slight loss during the second week of training. Measurements of fatigability showed improvement for lower extremities. Upper extremity measurements of force, work, and power showed significant improvement at all measurement tensions between pre and mid training, as well as mid and post training. While Gehlsen and colleagues found improved muscle performance, there were mixed results, a small sample size, and no control group.

In a subsequent nonexperimental study, Gehlsen and colleagues (1986) examined the effects of aquatic exercise on the gait of people with MS. No improvements in gait were found for 11 volunteers with MS who participated in a 10-week aquatic exercise program consisting of swimming and shallow water exercise (3 × week; 1 hour sessions). The authors concluded that aquatic exercise appeared to have no effect on gait parameters, yet participants reported other benefits such as lower perceptions of fatigue and increased ability to work harder. The possibility of decreased perceptions of fatigue and ability to work harder support the need to conduct additional research on effects of AT.

To address this lack of information pertaining to the effects of an AT program on individuals with MS, a single-subject multiple baseline study was conducted across three participants by Broach et al. (1998). Dependent measures were selected based on their connection to activities of daily living (vacuuming, riding a bicycle, climbing stairs, putting cans on a shelf, and buttoning a shirt). Results indicated improved gross motor upper-extremity (vacuuming) and lower extremity functioning (climbing stairs and bicycling). While two primary symptoms associated with MS include decreased gross motor ability and increased fatigue (Schapiro, 1995), there has been little examination of the effects of AT on these symptoms. Therefore, systematic replication and extension of the Broach et al. study across other participants was conducted to enhance understanding of effects of AT on gross motor ability and fatigue. This study replicated and extended the AT intervention and incorporated two of the previously used gross motor dependent measures: (a) walking up and down stairs and (b) rotations on a bicycle. The Broach et al. study was extended by including the following primary dependent measures not in the initial study: (a) pedaling an upper extremity ergometer and (b) perceived fatigue. The Rehabilitation Center research review committee suggested using an upper extremity ergometer as a reliable measure of upper extremity gross motor ability.

It is helpful when implementing an intervention to examine acceptability of the intervention from a societal perspective (Geller, 1991) and the subsequent likelihood of the application of the intervention (Baer & Swartz, 1991). Geller stated that the worth of an intervention is called its social validity. Although indicators of social validity were not intended to answer the primary questions of the study, suggestions by Wolf (1978) were followed to assess perceptions of family, friends, and relevant professionals regarding the importance of study goals, the appropriateness of procedures implemented, and importance of intervention effects.

Method

Participants

Participants were selected from a pool of volunteers in a large city in the southeastern United States. They were recruited by posting flyers in the MS center at the physical rehabilitation facility where the study was conducted and by a written announcement in the newsletter of the local chapter of the MS society. Participants included four individuals with relapsing-remitting MS who were unemployed due to complications related to MS. Participants were required to be at least 18 years old, obtain a doctor's order for AT, have no prior AT, and have no medical conditions precluding participation as indicated by a health history questionnaire. Additional criteria for participation were associated with their ability to complete actions used for data collection, such as sitting independently on the stationary bicycle and initiating at least one rotation, walking up and down 4–6 inch steps holding onto a railing for support, and initiating one rotation of an upper extremity ergometer.

Dorothy. Dorothy was a 53-year-old woman who lived alone and used a cane for assistance in ambulating. Regarding changes in her lifestyle that were associated with complications from MS she stated:

> I worked for 39 years but I can't work anymore because of my memory . . . I used to run, go to festivals where we would be out all day long, go shopping with my daughter, and entertain, but I cannot do that anymore. I am accepting of my disability and am happy that I have things that I can still enjoy.

Dorothy noted that fatigue limited shoulder range of motion, pain, and difficulty with walking decreased her physical activity. Dorothy emphasized that she hoped the AT program would be an enjoyable experience that would help improve her physical ability and endurance.

Teresa. Teresa was a 52-year-old woman who used a walker to ambulate. She stated that she had to resign from her previous employment because she could not perform some of the duties of a medical secretary due to forgetfulness and deteriorating fine motor skills reducing her ability to write. Concerning her current leisure activities she stated:

> Typically I spend the day straightening my house, walk to the mailbox, watch TV, and play with my dog. My son does much of the mopping that I use to enjoy . . . I just can't anymore. Fatigue, lack of money, the heat, and my leg hurting and dragging keep me from doing more.

Teresa stated that she hoped the AT would improve her walking, and provide her with an enjoyable activity that she could do independently.

Anita. Anita, diagnosed with MS for three years, was a 30-year-old woman who resided in an apartment. She was ambulatory but had to quit her job in human resources as a result "of the inconsistency of my dependability because of the MS." Anita reported that due to fatigue and dizziness associated with MS, she had progressively stopped many recreation activities. "I try to get 12 hours of sleep so I do not have to take a nap during the day." She stated that fatigue and money were major barriers to leisure activities. Anita reported that:

> If I want to clean my house I have to take it room by room on different days, it just physically wears me out. Can you believe I use to enjoy weekly 20 mile day hikes and working out? I also do not have the money I use to. I just have to get used to it. I hope my endurance will pick up.

Margo. Margo was a married, 43-year-old ambulatory woman who resided at home with her husband and four children. She stated that balance, dizziness, and fatigue were affected by MS and that her symptoms had increased dramatically in the past two years. Margo had to quit working as a speech pathologist because of fatigue and balance problems. However, she stated that she stays quite busy taking care of the household chores and finances, and assuring the children get to and from school and after school activities. Margo stated that she ceased her favorite activities of cross-stitching and shopping because of MS. Regarding the AT program, she stated: "I am at least interested in maintenance. I would also like to loose some weight and get a little stronger."

Setting and Personnel

The study was conducted in an accessible pool equipped with multiple methods of entry. The pool had a 30' long, 1.5' deep sitting area. The water temperature was regulated between 85° and 87° throughout the intervention. An aquatic therapist who was a certified Therapeutic Recreation Specialist, MS Aquatic Exercise instructor, and Aquatic Exercise Association Instructor directed the intervention.

Dependent Measures and Data Collection

The primary method of data collection included observations of walking up and down stairs, pedaling an upper-ergometer, bicycling, and response to perceived fatigue that was collected in a time series manner. Secondary information was obtained via social validity questionnaires. Secondary measures were used to provide additional information but were not considered primary measures because data were not collected at prescribed intervals throughout the study.

Walking up and down stairs. Participants stood facing a flight of eight, 8″ steps, used the handrail for balance, walked up eight steps, turned around, and walked down the steps when the observer stated "go" on a cue of "ready, set, go." Participants were directed to complete this sequence as often as possible within 3-min. An event recording system was used to count each step (Tawney & Gast, 1984).

Pedaling an upper extremity ergometer. Participants were positioned standing at the Tru Kinetics Aero 841 upper ergometer with the right hand in the lowest position of the rotation. The ergometer base was set at the same distance for each participant throughout the study. Participants were directed to begin pedaling when the screen began to count down from 3-min and to stop when 3-min were completed. When the monitor began counting down, participants began to perform as many rotations as possible in 3-min. When the 3-min transpired, the distance in miles was recorded through automation.

Pedaling a stationary bicycle. Participants pedaled a Randel Windtrainer stationary bicycle for as many rotations as possible for 3-min. The degree of resistance was set on 20 for all participants at the beginning of baseline and was maintained throughout the study. Measurement always began with the left pedal at the bottom position. The seat height was set on the first day of data collection and remained the same throughout the study. The staff stated: "When the screen shows start, I want you to complete as many rotations as you are able to in 3 minutes. When three minutes have passed and I say 'stop', please stop pedaling." The distance cited on the automated recording system on the Randel Windtrainer was recorded in miles, as well as the average rpm.

Observer accuracy. Observers included the first author, an assistant, and two student volunteers. To establish observer accuracy, the primary observer trained volunteer observers in documentation, use of stopwatches, and observation procedures using a task analysis prior to the study. While the observers were trained in timing the stair measure, the bicycle and upper cycle had automated timing. Observers were trained to perform the procedures on volunteers. When 90% agreement with the primary observer was achieved with (a) completing the task analysis procedures for dependent measures, (b) timing of stair measures, and (c) tallying number of stairs walked, the volunteers were deemed ready.

Observer agreement. To assess inter-rater reliability, 80% of the data collection sessions were observed by two observers (Tawney & Gast, 1984). Percent agreement for each session was calculated. If an agreement score of less than 80% was obtained, the discrepancies would have been examined and the observers would have been retrained in data collection procedures. However, observer agreement on data collection procedures was at least 90% on the dependent measures throughout the study.

Fatigue. Perceptions of fatigue were monitored using the Fatigue Questionnaire (Chalder et al., 1993), a self-report instrument designed to ascertain perceptions of physical and mental fatigue by assessing symptom severity, detecting epidemiological fatigue cases, and estimating changes in perceived fatigue. Participants completed the Fatigue Questionnaire when other dependent measures were collected to evaluate changes in perceived fatigue.

Chalder et al. (1993) analyzed data from 374 individuals with chronic fatigue syndrome. A principal component analysis was used to find a combination of variables that adequately explained variation. The principal component analysis was in agreement with the pattern in the first two eigen vectors that supported the two constructs of mental and physical fatigue, and the notion of a two factor solution. Internal consistency was tested by calculating Cronbach's alpha (range .88–.90) and by randomly dividing the items of the scale into two subscales and correlating them both with part one (r = .8613) and part two (r = .8466). Results showed an acceptable level of internal consistency. Since specificity/ sensitivity values of 58 and 74, respectively, on the CIS-R were very high (with few false positives), validity was supported.

Social Validity of Intervention

The social validity of projected outcomes, procedures, and effects were assessed. Projected outcome measures were critiqued by a physician, the rehabilitation center research review committee, and by the local MS society director. All reviewers agreed that projected outcomes were socially significant and relevant. Appropriateness of study procedures was assessed every fifth session of the intervention using five open-ended questions that reflected participants' opinions regarding program importance, satisfaction with the program, and suggestions for changes. To address social importance of effects, a questionnaire was administered on the first day of the study and on the last day of the intervention to participants, an adult family member, and the AT instructor. The questionnaire asked respondents to rate (from not at all to very much) the value of AT as well as how they thought AT will/has affect(ed) the participant's (a) physical ability, (b) fatigue level, and (c) enjoyment.

Experimental Design and Procedures

A single subject, multiple probe design across 4 participants was used to assess effects of AT. The multiple probe design is similar to the multiple baseline design in that the intervention is systematically introduced to only one participant at a time. Entry into the intervention is staggered to demonstrate any effect of the intervention while helping control for effects to internal validity (Tawney & Gast, 1984). In multiple probe design participants' baseline data are not collected on a continuous basis; rather, probe trails are observed intermittently. Continuous data collection in baseline occurs only for one randomly selected participant at a time prior to entering the intervention. In general when employing this design, the investigator observes the dependent measures of each participant in a baseline condition until the criteria of an appropriate level and trend is established. Level refers to the magnitude of data indicated by the ordinate value while trend refers to the direction (accelerating, decelerating, zero-celerating). Participants were involved in the AT program for at least 8 weeks as they completed the three study phases of baseline, AT intervention, and follow-up.

Baseline. Baseline consisted of continuous data collection trials and probe trials of appropriate participants. During continuous data collection trials, participants came to the rehabilitation center gymnasium Monday, Wednesday, and Saturday at 9:00 a.m. for testing on the stairs, upper ergometer, bicycle, and perceived fatigue measures. Participants were offered the opportunity for free swim in the pool after observations. All measures were completed in the same order to prevent multitest interference with carryover effects of sequencing. In this multiple-probe design, all four participants received a probe trial of the above dependent measures on the first day of baseline. Then, the first randomly selected participant, Dorothy, received continuous probes until she achieved a stable level and a non-accelerating trend of behavior on the dependent measures. As the intervention was introduced to Dorothy, all participants were again probed on the dependent measures. Then, continuous measures were taken in baseline on the second randomly selected participant, Teresa, for five sessions and until a stable level and/or non-accelerating trend was achieved on the dependent measures. As Teresa entered the AT condition, the dependent measures were probed again for all participants and taken continuously for the third participant (Anita). The sequence of probing and the staggered entry into the AT condition continued until all participants were introduced to the intervention.

Intervention. During intervention, participants attended the MS AT program and were continuously measured on each primary dependent measure prior to each intervention. Based on the previous findings of Broach et al. (1998), the AT protocol was modified by adding two AT strategies (aqua steps and proprioceptive neuromuscular facilitation patterns). AT was conducted in a small group at regularly scheduled times (3 × week). The protocol followed a four phase format: 10 minutes of walking warm-ups, 10 minutes of step exercises, 18 minutes of upper body exercises, and 7 minutes of lower body exercises.

Intervention procedural reliability involved observers timing the duration of each exercise sequence using a stopwatch. Duration of occurrence using the gross method of agreement was calculated by comparing the number of seconds recorded during each event of the intervention including warm-up, upper extremity, lower extremity, and step exercises. In addition, procedural reliability was assessed for the AT exercises comparing the list of potential exercises in each exercise sequence to the actual performance of exercises. At least an 80% agreement was maintained throughout the study.

Follow-up. All participants were re-tested on all dependent measures 2, 4, and 8 weeks after the intervention ended. Identical procedures employed during baseline-probe were utilized during follow-up.

Results

Walking Up and Down Stairs

Dorothy. As seen in Figure 1, during baseline Dorothy showed a stable, slightly decelerating trend in the number of stairs she walked (M = 142). During intervention, after an initial latency period that coincided with absences and reported back pain, she demonstrated a variable, yet accelerating trend in number of stairs walked for the last 66% of intervention resulting in a mean increase of 14 steps. During follow-up, the average number of steps Dorothy walked during the intervention was maintained with a mean score equal to the intervention mean score.

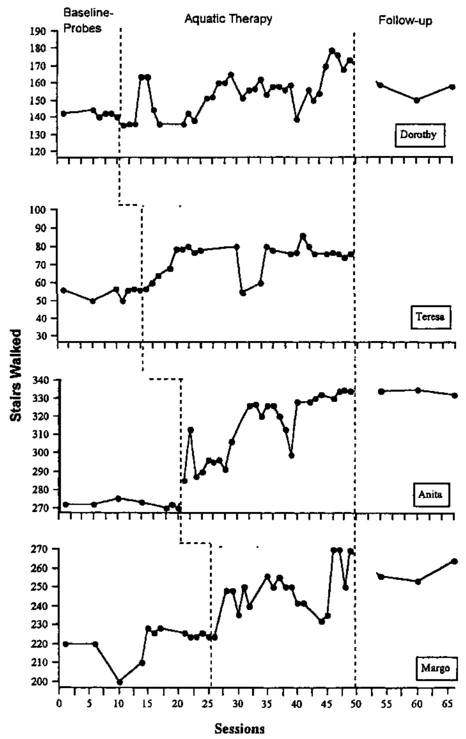

FIGURE 1. FREQUENCY OF STAIRS WALKED ACROSS PARTICIPANTS.

Teresa. The number of stairs that Teresa walked during baseline-probe was slightly variable, with a zero-celerating trend that was stable for the last 3 sessions (M = 55). During intervention, Teresa showed an accelerating trend in stairs walked and then maintained the improved level throughout the intervention except for sessions 29–32, which coincided with an illness and missed sessions (a 19 step improvement over baseline-probes mean). Teresa did not participate in lower extremity follow-up measures due to severe pain and weakness caused by bone spurs in her ankles.

Anita. Baseline data for Anita were stable and zero-celerating (*M* = 272). During intervention, data for stair walking showed an immediate level change and a variable, accelerating trend resulting in a mean score 58 steps higher than the baseline-probe mean. Anita's lower scores during session 23–26 and 39 of the intervention coincided with reported migraine headaches. During follow-up, Anita's behavior of walking stairs showed a stable zero-celerating trend and achieved a higher mean score than her intervention mean score (*M* = 334).

Margo. Margo demonstrated an initial variable and slightly accelerating trend; however, her behavior stabilized with a zero-celerating trend for the last eight measures of baseline (*M* = 224). During intervention, Margo's stairs walked were variable, yet accelerating resulting in a 26 point mean improvement over baseline and only a 5% overlap of scores with baseline. Margo's lower scores and absences during sessions 40–45 coincided with reported knee pain and a fall. While Margo's follow-up mean of 258 was higher than the intervention mean, it was slightly lower than three of the four last sessions of the intervention.

Pedaling an Upper Extremity Ergometer

Dorothy. As seen in Figure 2, during baseline Dorothy demonstrated a stable, zero-celerating trend in miles pedaled (*M* = .15). Upon entering intervention, Dorothy's behavior in pedaling showed a 7-session latency, followed by a variable, accelerating trend (.06 mile mean improvement over baseline). Although Dorothy's first two scores during follow-up were slightly lower than her last four scores of intervention, she demonstrated a slightly accelerating trend and maintained a

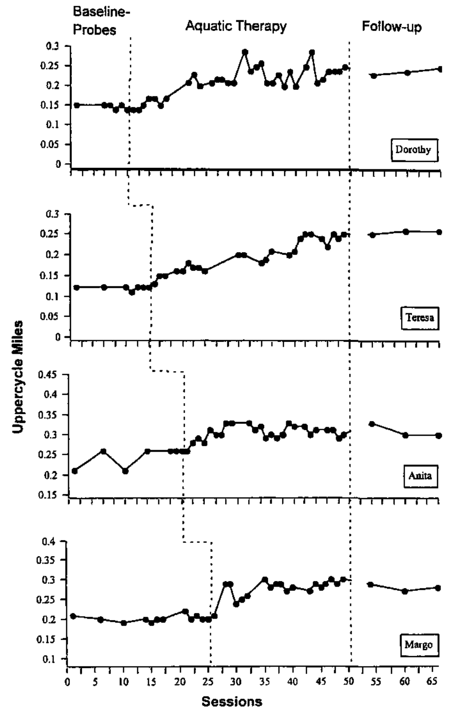

FIGURE 2. UPPER ERGOMETER MILES ACROSS PARTICIPANTS.

mean score of .24 which was higher than her intervention mean score.

Teresa. Teresa's data were stable, zero-celerating during baseline (*M* = .12). Once the intervention began, Teresa's uppercycle distance changed to a stable, accelerating trend (.08 mile mean improvement over baseline). Follow-up scores were stable and zero-celerating with a mean of .31, which was above the intervention mean. Follow-up indicated that effects were maintained.

Anita. During baseline, Anita's distance pedaled was variable for the first three sessions, but stable with a zero-celerating trend for the last four sessions (*M* = .25). Anita's intervention data for pedaling an uppercycle were initially accelerating then became zero-celerating for most of the intervention resulting in an improved mean score from baseline-probes (.06 mile mean improvement over baseline). Despite a slightly decelerating trend during follow-up and an initial score above the last 9 scores of intervention, a mean score was achieved equal to the intervention mean. Although these data indicated maintenance of effects achieved during intervention for Anita, data indicated a slightly decelerating trend.

Margo. Margo's uppercycle distance was stable and zero-celerating during baseline (M = .2). Although upper ergometer distance was initially variable and accelerating upon initiation of the intervention, Margo's behavior stabilized and became zero-celerating for the remaining 70% of intervention. Therefore, the mean score for pedaling the uppercycle during the intervention was higher than baseline (M = .28). During Margo's follow-up, a zero-celerating trend and a maintenance of the intervention mean was observed.

Pedaling a Stationary Bicycle

Dorothy. As seen in Figure 3, Dorothy's behavior in miles pedaled was stable and zero-celerating during baseline (M = .93). During intervention, she bicycled fewer miles for the first session and then showed a stable, accelerating trend and improved mean score (.28 mean mile improvement over baseline-probe). During follow-up, Dorothy's behavior was stable and slightly accelerating while maintaining a mean level above the intervention mean (M = 1.3).

Teresa. Distance bicycled by Teresa was stable, zero-celerating during baseline (M = .6). Upon initiation of the AT, distance bicycled showed an immediate level change. A stable and accelerating trend was observed during intervention (.37 mean mile improvement over baseline). No follow-up scores on this measure were obtained because of severe pain associated with bone spurs in Teresa's ankles.

Anita. Anita demonstrated a stable and zero-celerating trend during baseline (M = 1.29 miles). During intervention, after an immediate level change, on the average, Anita pedaled slightly longer distances that were

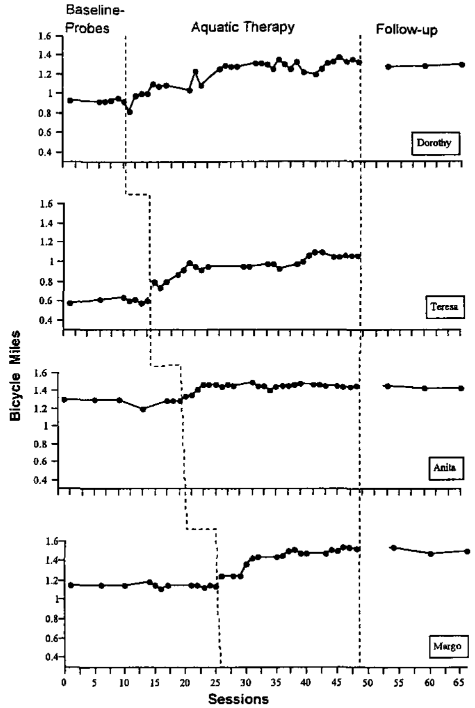

FIGURE 3. NUMBER OF MILES BICYCLED ACROSS PARTICIPANTS.

zero-celerating for most of the intervention (.15 mile mean improvement over baseline). Anita's behavior during follow-up was stable and zero-celerating with a mean score equal to her intervention mean score, indicating that she maintained her bicycling scores achieved during intervention.

Margo. During baseline, Margo's bicycling was stable with a zero-celerating trend (M = 1.15 miles). During intervention, she initially increased distance she bicycled during the first 51% of the intervention resulting in an accelerating trend. After the initial 51% of the intervention, she maintained longer distance and showed a zero-celerating trend for the remaining intervention (.30-mile mean improvement over baseline). During follow-up, Margo's behavior in pedaling a bicycle was zero-celerating with a mean indicating maintenance (M = 1.51).

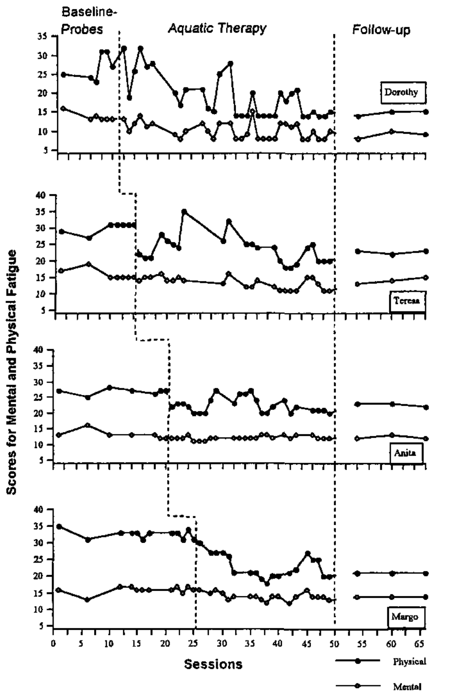

FIGURE 4. FATIGUE QUESTIONNAIRE SCORES ACROSS PARTICIPANTS.

Physical Fatigue

Physical fatigue scores had a potential range of 7 to 35, with the lower score representing less perception of physical fatigue.

Dorothy. As seen in Figure 4, baseline reports of Dorothy's physical fatigue accelerated. (M = 27). While her behaviors remained unstable during intervention, after the initial 6 sessions her physical fatigue was lower or equal to baseline for over 90% of sessions. Dorothy's overall mean score of 19 for physical fatigue during intervention was 8 points below baseline mean with a 21% overlap of baseline and intervention scores. Physical fatigue mean for the final 50% of intervention was 11 points below the baseline mean of 14. Dorothy's follow-up mean score of 15 for physical fatigue was below the intervention mean.

Teresa. Baseline probe scores for physical fatigue were stable and zero-celerating for 5 sessions prior to intervention (M = 30). Upon entering intervention, Teresa's scores dropped immediately and continued in a therapeutic direction with a variable decelerating trend during intervention. There was only 13% overlap of scores between conditions (6 point mean improvement over baseline). Teresa's follow-up scores for physical fatigue were zero-celerating resulting in a mean score below intervention mean (M = 23).

Anita. Physical fatigue reported during baseline was stable with a zero-celerating trend (M = 26). During intervention, Anita showed a variable, decelerating trend in her reported physical fatigue with a low 16% overlap of scores between conditions (3 point mean improvement compared to baseline). While Anita's follow-up physical fatigue scores were higher than her last 3 intervention scores, scores were decelerating, and she maintained a mean equal to intervention mean.

Margo. Margo demonstrated a slightly variable, zero-celerating trend during baseline with a mean of 32. Margo's decelerating scores during intervention resulted in a 6-point intervention mean improvement over baseline with only a 5% overlap of scores between conditions. Margo's follow-up scores were zero-celerating with a 5-point improved mean over intervention mean.

Mental Fatigue

Mental fatigue scores had a potential range of 4 to 20, with the lowest possible score of 4 representing less perception of physical fatigue.

Dorothy. As seen in Figure 4, Dorothy's reported mental fatigue was stable, zero-celerating for 5 sessions prior to intervention (M = 14). During the intervention, her mental fatigue exhibited a variable, decelerating trend (4 point mean improvement over baseline). Dorothy's follow-up scores for mental fatigue were zero-celerating with a mean one point below intervention mean.

Teresa. Mental fatigue scores were stable and zero-celerating for 5 measures prior to intervention (M = 16). During intervention her scores were slightly variable with a decelerating trend after the first 4 scores. There was a 30% overlap of scores between baseline and intervention and a 3 point mean improvement from baseline to intervention. Follow-up scores showed an accelerating trend with a mean above the intervention mean but below the baseline-probes mean, and therefore, maintenance of the intervention did not occur (M = 14).

Anita. Anita's mental fatigue scores were zero-celerating during baseline (M = 13). During the intervention reported fatigue continued a zero-celerating trend with 88% overlap of scores between baseline and intervention (1 point mean improvement). Anita's follow-up scores were zero-celerating with a mean score equal to intervention mean. Although Anita's intervention mean score was 1 point lower than her baseline mean, the high percent overlap of scores suggested that the AT intervention did not effect Anita's mental fatigue.

Margo. Margo's mental fatigue scores showed a stable zero-celerating trend during baseline and a stable zero-celerating trend during intervention with a 1 point mean improvement over baseline mean of 16. Margo's 68% overlap of scores between conditions was due to a low score from the second baseline score. Her percent overlap, without considering the 2nd session, is 26%. Margo's follow-up mental fatigue exhibited a zero-celerating trend that resulted in a mean score one point below intervention mean.

Social Validity Questionnaires

To assess the social validity of goals, participants, family members, and AT instructor responded to questions regarding the expected (pre-intervention score) and perceived outcome of each goal (physical ability and fatigue) and overall value of AT. Respondents agreed that before the intervention they expected it to be effective and after the intervention was over that they perceived the intervention to have been effective. Specifically, participants and family members reported that AT had positive effect on various physical behaviors and fatigue. To examine the social validity of the AT procedures, a questionnaire was administered every 5 sessions (a total 9 times) determined if the intervention was meeting participant expectations. The responses to the open ended questions indicated that the participants were satisfied with the AT procedures. In addition participants stated they felt that the most important outcomes of AT included increased freedom of movement, relaxation, energy, social interaction, and the feeling of being "normal" while participating in activity in the water. Many suggestions for the AT program were related to environmental factors (e.g., water temperature, time of the class, and accessibility concerns) rather than the specific AT.

Discussion

The study examined effects of participation in AT on gross motor behavior and fatigue. Visual analysis of data suggests that all participants increased their motor ability in riding a stationary bicycle, walking up and down stairs, riding an uppercycle, and decreased their physical fatigue. The data suggest some improvements in mental fatigue of some participants. However, because results were not replicated across all participants, effects of AT intervention were not established for mental fatigue.

Although research examining effects of exercise on gross motor behavior of people with MS is limited, results of this study support and extend the work of Broach et al. (1998) who found increases in lower extremity behaviors (including stairs walked and rotations pedaled on a stationary bicycle). In addition, the 4-week follow-up scores reported by Broach et al. indicated maintenance of the intervention for participants, which was similar to results of this study using an 8-week follow-up.

Relative to upper extremity behavior, Broach et al. (1998) did not examine participants' behavior of pedaling an upper ergometer. However, Broach et al. did find improvements in participants' behavior using a vacuum cleaner, which was representative of an upper extremity gross motor behavior. To improve reliability of measurements, a calibrated upper ergometer was used in this study. Improvements in upper extremity behavior associated with participating in AT for this study coincided to the findings reported by Broach et al.

Participants' reported that after they participated in AT physical fatigue was reduced while measures of mental fatigue did not indicate positive changes across conditions. Improved perceptions of fatigue correspond to findings of Broach et al. (1998) who used the Fatigue Questionnaire (Chalder et al., 1993) as a secondary measure. The three participants in the Broach et al. study reported improved levels of physical fatigue after the intervention when compared to pre-interventions reports. Gehlsen et al. (1986) additionally examined fatigue levels associated with aquatic exercise. Gehlsen et al. found that participants reported that they felt less fatigued and that they had the ability to work harder in various tasks after participating in the aquatic exercise program. Gehlsen, Grigsby, and Winant (1984) found that participants' lower extremity muscular fatigue calculated in arm pulls on a swim bench and via an isokenetic dynamometer for lower extremities significantly decreased. The results of this study support the reported findings for physical fatigue. However, while Broach et al. found small improvements in participants' mental fatigue, no changes in mental fatigue were reflected in this study.

While cool water decreases fatiguing effects from the rise in core temperature, physical properties of the water that include hydrostatic pressure and buoyancy result in increased support and reduction of gravity effects (Campion, 1990). However, while the water environment is potentially less fatiguing because outcomes can be achieved with less effort, caution must be taken to assure that participants do not increase fatigue levels from excessive exercise. The greater ease of movement could cause muscle fatigue unless the participant and therapist monitor activity level (Skinner & Thompson, 1989). Considering the cautions of Skinner and Thompson, Shapiro (1995) advised that exercising muscles not effected by MS may help relieve fatigue. Results of this investigation and the aforementioned research support Shapiro's contention that exercise may decrease effects of physical fatigue. AT in this study also appeared to offset the typical difficulties of MS that relate to fatigue such as the ability to do daily tasks such as walking up and down stairs. Further research is needed to determine if outcomes of this study can be replicated.

The social validity questionnaires indicated that participants and family members perceived AT to have a positive effect on various physical behaviors and fatigue. In addition, participants stated they felt that the most important outcome of AT included increased freedom of movement, relaxation, energy, and social interactions. Freedom of movement and relaxation occurs because of the support offered by the water, the consequential reduction of the effects of gravity and physiological responses to immersion (Becker & Cole, 1997; Campion, 1990). Increased energy reported by participants could be an indicator of decreased physical fatigue that was one goal of this program. The participants' reports of increased social interaction were consistent with findings from previous research (Weiss & Jamieson, 1989) that examined perceptions of social enjoyment. The researchers reported that 90 of 100 participants who participated in aquatic exercise from 8-weeks to 5-years perceived improved emotional states in part because they viewed the aquatic exercise as a support group, made new friends, and found it easy to converse with others. The AT program used in this study required participants to engage in small groups which resulted in social contact during the program. Since social interaction may promote activity adherence (Poldichak, 1991), it may be beneficial for future research to examine social interactions which occur in conjunction with AT.

A single subject, multiple probe design across participants was selected for this study. This single subject design appeared to be appropriate for this study because data were collected every time a participant attended a class. This design enabled the observation of overall trends in behavior despite periodic fluctuations in functional ability as is typical with MS. However, this presents a limitation related to generalizations. Since the study included only four participants, generalizations are limited. Since results associated with bicycling and walking stairs for this study replicated findings by Broach et al. (1998), the ability to generalize results associated with these dependent measures is higher than those associated with the uppercycle and perceptions of physical fatigue. Because AT is generally conducted one-on-one to meet specific treatment goals, large experimental studies are difficult to conduct. Therefore, it is recommended that additional replications using single subject design be conducted to further generalize findings.

Conclusions

Based on the overall findings, this AT program for individuals with MS appears to have merit. Results of this study indicated that people with MS should be encouraged to participate in AT interventions similar to the one in this study with the intent of facilitating improvements in gross motor behavior and physical fatigue. Similar to other individuals with chronic disease, people with MS should have medical approval and supervision by a qualified professional trained in exercise and the AT techniques for individuals with MS before initiating an AT program. When using AT with individuals with MS it is important to assure water temperature does not enhance fatigue symptoms. In addition, it is helpful to use verbal cueing to increase the likelihood participants will evaluate their fatigue and pace of participation. While AT is not a substitute for other treatment or therapies, it is one method that can help individuals manage MS. Results of this study provide support for use of AT with people with MS. Further research is needed to explore the impact of AT on physical behaviors, activity adherence factors such as perceptions of social support and enjoyment, fatigue, and other quality of life indicators of people with MS.

References

Baer, D. M., & Schwartz, I. S. (1991). If reliance on epidemiology were to become epidemic, we would need to assess its social validity. *Journal of Applied Behavior Analysis, 24,* 231–234.

Becker, B. E., & Cole, A. J. (1997). *Comprehensive aquatic therapy.* Newton, MA: Butterworth-Heinemann.

Broach, E., & Dattilo, J. (1996). Aquatic therapy: A viable therapeutic recreation intervention. *Therapeutic Recreation Journal, (30),* 213–239.

Broach, E., Groff, D., Dattilo, J., Yaffe, R., & Gast, D. (1998). The effects of aquatic therapy on adults with multiple sclerosis. *Annual in Therapeutic Recreation, 7,* 1–20.

Campion, M. R. (1990). *Adult hydrotherapy.* Oxford: Heinemann Medical Books.

Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., Wessely, S., Wright, D., & Wallace, E. P. (1993). Development of a fatigue scale. *Journal of Psychosomatic Research, 37*(2), 147–153.

Deluca, J., Johnson, S., & Natelson, B. H. (1993). Information processing efficiency in chronic fatigue syndrome and multiple sclerosis. *Archives of Neurology, 50,* 301–304.

Gehlsen, G., Beekman, K., Assmann, N., Winant, D., Seidle, M., & Carter, A. (1986). Gait characteristics in multiple sclerosis: Progressive changes and effects of exercise on parameters. *Archives of Physical Medical Rehabilitation, 67,* 536–539.

Gehlsen, G., Grigsby, S. A., & Winant, D. M. (1984). Effects of an aquatic fitness program on the muscular strength and endurance of patients with multiple sclerosis. *Physical Therapy, 64*(5), 653–657.

Geller, E. S. (1991). Where's the validity in social validity? *Journal of Applied Behavior Analysis, 24,* 179–184.

Holtakers, T. R. (1993). Multiple sclerosis and exercise. *Achieving a balance: A conference on positive lifestyles for persons of varying abilities* (pp. 16–17). Normandale, MN: Vinland National Center.

National Multiple Sclerosis Society. (1992). *What everyone should know about multiple sclerosis.* South Deerfield, MA: Author.

Paty, D. W., & Ebers, G. C. (1998). *Multiple sclerosis.* Philadelphia: F. A. Davis.

Poldichak, W. (1991). Establishing the fun in leisure. *Leisure Sciences, 13,* 123–136.

Ponichtera-Mulcare, J. A. (1993). Exercise and multiple sclerosis. *Medicine and science in sport medicine, 25,* 451–465.

Rudick, R. A. (1990). Helping patients live with multiple sclerosis. *Postgraduate Medicine, 88,* 197–207.

Sandroni, P., Walker, C., & Starr, A. (1992). Fatigue in patients with multiple sclerosis. *Archives in Neurology, 49,* 517–524.

Schapiro, R. (1987). *Symptom management in multiple sclerosis.* New York: Demos.

Schapiro, R. (1995). Exercise and multiple sclerosis. *MS Management,* 2(2), 35–38.

Schapiro, R., Petajan, J., Kosich, D., Molk, B., & Fenney, J. (1988). Role of cardiovascular fitness in multiple sclerosis: a pilot study. *Neuro Rehab, 2,* 43–49.

Skinner, A. T., & Thomson, A. M. (1989). *Duffield's exercise in water* (3rd ed.). London: Bailleire Tindall.

Tawney, J., & Gast, D. (1984). *Single-subject research in special education.* Columbus, OH: Merrill.

Weiss, C. R., & Jamieson, N. B. (1989). Women, subjective depression, and water exercise. *Health Care for Women International, 10,* 75–88.

Wolf, M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis found its heart. *Journal of Applied Behavior Analysis, 11,* 203–214.