Adult Performance on Three Tests of Equilibrium

Anne G. Fisher, Shirley E. Wietlisbach, Julia L. Wilbarger

Key Words: development • equilibrium • sensory integration

The purpose of this investigation was twofold: (a) to determine preliminary norms for young adult males and young adult females on three clinical tests of equilibrium, Tilt Board Tip (TBT), Tilt Board Reach (TBR), and Flat Board Reach (FBR), and (b) to determine sex differences in equilibrium performance. The three equilibrium tests were administered to a sample of 25 men and 29 women between the ages of 20 and 30 years. The results revealed a significant sex difference on TBT, with men obtaining significantly higher scores than women. FBR approached significance, and no significant difference was found on the TBR. These results support earlier findings of sex differences in performance on a variety of equilibrium measures and suggest that different normative criteria should be used to evaluate the performance of males and females on some tests of equilibrium. Possible reasons for males obtaining higher scores on TBT and for lack of differences between male and female performance on FBR and TBR are discussed.

Recently, occupational therapists have begun evaluating and treating an increasing number of young adults with sensory integrative dysfunction and/or learning disabilities. A major complaint and presenting problem of these clients is a history of difficulty with motor coordination. Inadequate postural responses, both a symptom and a hypothesized cause of such motor coordination problems, indicate that an assessment of equilibrium should be included in the occupational therapy evaluation of these clients. However, relatively few objective equilibrium tools are suitable for detecting equilibrium deficits in young adults, and none have been standardized. Furthermore, few studies have been done that describe the qualitative behaviors thought to constitute a typical equilibrium response in the young adult.

Literature Review

Occupational therapists commonly use two types of equilibrium assessments. The first, subjective clinical evaluation of equilibrium, typically involves making judgments about the quality of responses of various parts of the body when a person’s center of gravity is either actively or passively displaced by pushing or tilting. A variety of problems are associated with using subjective measures of equilibrium. The literature is inconsistent with regard to describing (a) the normal equilibrium response and (b) the extent to which a particular body part moves in an optimal equilibrium response (Fiorentino, 1973; Fisher, 1984; Gilfoyle, Grady, & Moore, 1981; Martin, 1965; Radmark, 1944; Weisz, 1938). The general lack of agreement and clarity in the description of a typical equilibrium response has resulted in an absence of objective grading criteria from which to develop valid clinical assessments of equilibrium. Therapists must, therefore, rely on subjective impressions alone. Ayres (1972) stated that a subjective clinical evaluation depends on the skill level and experience of the individual therapist. In addition, clinical observations not based on objective grading criteria can detect only gross deficits in equilibrium and, thus, may lack the sensitivity necessary to detect subtle differences in performance (Fisher).

The second type of assessment, objective evaluation, deals with the quantitative aspects of equilibrium which typically have been evaluated by measuring the length of time a person can stand on one foot with eyes open or closed, or the distance a person can walk on a balance beam (cf. Ayres, 1980; Bruininks, 1978). Objective evaluations do not always yield clinically useful information. While a person may perform within normal standards for his or her age, the quality of the response may still suggest dysfunction.

With these needs in mind, the first author in a
previous study with Bundy (1982) selected and modified four clinical tests of equilibrium designed to evaluate equilibrium reactions in both a qualitative and quantitative manner. The first author (1984) further defined and developed objective grading criteria to score both the qualitative (behavioral score) and quantitative (angle score) aspect of three of the clinical evaluation tools, Tilt Board Tip (TBT), Flat Board Reach (FBR), and Tilt Board Reach (TBR). The first author demonstrated that angle scores on all three tests were reliable and valid measures of equilibrium. Individual angle scores correlated significantly with age, and when the three angle scores were summed together, the total angle score demonstrated higher correlations with age than did existing equilibrium tests. The angle scores also significantly discriminated between normal boys and boys with sensory integrative (vestibular) dysfunction. The reliability of scoring was demonstrated to be high (inter- and intrarater reliabilities for angle scores, r = .98).

Several prior investigations, using these three clinical tests, found angle scores to be sensitive to differences in equilibrium performance between boys and girls (Fisher & Bundy, 1982; Fisher, 1984; Izraelevitz, Fisher, & Bundy, 1985). The first author and colleagues consistently found a significant sex difference on TBT. Subjects ranged from 3 to 13 years of age; boys obtained higher scores in all three studies. The first author found a significant age/sex interaction on FBR; boys obtained lower scores than did girls, and scores equalized in adolescence. Izraelevitz et al., in their study of 3- to 6-year-olds, found that the age/sex interaction on FBR only approached significance whereas the first author and Bundy found no significant sex differences or age/sex interaction on FBR. Finally, no significant sex differences or age/sex interactions were found on TBR in any of the studies.

Other researchers also support the finding of such age- and sex-related differences in performance on equilibrium tests (Espenschade, Dable, & Schoendube, 1955; Fregly & Graybiel, 1968). Williams (1983) found that “when balance performance is looked at across a wide range of ages and tasks, there is little or no differences between boys and girls in balance performance” (p. 269). However, when the nature of the equilibrium task is considered, significant differences in performance between male and female subjects have been found on a variety of equilibrium tasks. These differences have been reviewed by the first author (1984) and by Williams. The evidence supporting sex differences and age/sex interaction on tests of equilibrium suggests that different criteria should be used for assessing males and females at different ages.

Investigations of equilibrium across many ages appear to indicate that the rate of developmental change varies at certain ages, depends on the test being administered, and varies for both males and females. However, equilibrium research, in general, has offered little insight into the development of equilibrium across the entire life span. The developmental sequence of equilibrium reactions has generally been described as increasing in childhood and slowing down or plateauing in adolescence (Fisher, 1984; Williams, 1983). Izraelevitz et al. (1985) investigated the change in equilibrium reactions in normal 3-, 4-, and 5-year-old preschoolers. They found that TBT, TBR, and FBR angle scores increased significantly with age. These findings suggest that these three tests measure aspects of equilibrium that develop gradually between the ages of 3 and 6 years. However, the tests that demonstrated the greatest change with age and/or were the best predictors of age were not consistent between the sexes. The first author (1984), in a study of 3- to 13-year-olds, found that each of the three angle scores increased with age in younger children and appeared to plateau before 13 years of age. The age at which the scores plateaued differed for each test and for boys and girls.

In summary, the clinical usefulness of the angle scores for the three tests of equilibrium (TBT, FBR, and TBR) is well established with children. However, the first author and colleagues (Fisher & Bundy, 1982; Fisher, 1984; Izraelevitz et al., 1985) did not assess children older than 13 years. Given the potential clinical usefulness of these tests and the need for equilibrium assessments suitable for young adults, we decided to investigate the equilibrium response of normal young adults using these three tests of equilibrium. The resulting preliminary normative data were expected to provide the basis for evaluating young adults with suspected equilibrium deficits. Specifically, this study addressed two questions:

1. What is the normal range of angle scores for young adult males and young adult females on the three clinical tests of equilibrium (TBT, FBR, and TBR)?

2. Do the mean angle scores of young adult males and young adult females of the three tests differ significantly?

Method

Subjects

The sample consisted of 25 men and 29 women. The mean age of the men was 25.2 years (SD = 3.4 years), and the mean age of the women was 25.5 years (SD = 3.1 years). A two-tailed t test revealed that there was no significant difference between the mean age of men and women (t[51] = -0.33, p = .745). The subjects were students or friends of students living in the
Tests

The subjects were videotaped while they were performing the three clinical tests of equilibrium developed by the first author and Bundy (1982; Fisher, 1984).

Tilt Board Tip (TBT). The subjects were asked to stand on a tilt board 34 cm wide, 29 cm deep, and 17 cm high. Subjects aligned their feet along an 11 cm wide stripe painted across the center of the board. They were instructed to face the camera and try to remain on the tilt board as long as they could while it was tipped. An attempt was made to tip the board at the rate of 10°-15°/sec, but the actual rate of tipping was somewhat faster (approximately 35°/sec; range = 20°-40°/sec). The increased speed of tilt appeared to be related to the increased effort required to tip an adult person.

Tilt Board Reach (TBR). Standing on a larger but lower tilt board, 44.5 cm square and 12.7 cm high, the subjects reached with their preferred (ipsilateral) hand toward a target positioned at shoulder height and at a distance to their side that would elicit the maximal reaching response. Subjects were instructed to begin with their feet aligned along an 11 cm wide line in the center of the board and to reach as far as possible toward the target without moving the foot ipsilaterally to the hand used for reaching, but they could move the contralateral foot if they wanted to do so.

Flat Board Reach (FBR). The test procedure was the same as for TBR, with the exception that the subjects stood on a stable platform, 25.5 cm square and 6.0 cm high.

Test Procedure

The subjects were tested individually in a well-lit, distraction-free room with a minimum of furniture and a background free of diagonal lines. Before testing, subjects were oriented to the testing procedure and asked to sign an informed consent form. They were then asked to remove their shoes and socks. The three tests of equilibrium were administered to each subject. Test order was systematically varied across subjects to control for the effects of order and fatigue. For each test, the subject was given one or two practice trials to clarify any confusion regarding the expected performance (e.g., subjects were not supposed to lunge at the target). The next two or three trials were videotaped. An attempt was made to videotape responses that were typical of the best performance. Testing took approximately 15 min.

Scoring

The criteria developed by the first author (Fisher, 1984) were used to do the scoring, with the exception that minor modifications were made for videotaping (a 35 mm camera with a motor-driven shutter was used in previous studies). A time increment counter was used with the video equipment to record elapsed time, in milliseconds, directly on the videotape. This allowed individual frames of the videotape to be identified for scoring. For TBR and FBR, this frame was the last frame before the subject (a) no longer maintained his or her center of gravity over his or her base or of the support or (b) showed a marked change in posture indicating that he or she no longer maintained control of his or her center of gravity over his or her base or of the support. For TBT, the scored frame was the frame just prior to the point at which (a) the subject's center of gravity was no longer over the base of support, (b) the subject's feet began to slide, or (c) the subject began to step off the board.

Three angle scores were obtained. A goniometer was used to measure the angles in degrees on a 19 in. video display. The angle score for TBR and FBR was the angle of deviation of the trunk from the vertical position. This angle was measured by placing one arm of the goniometer along an imaginary line from the subject's chin to his or her crotch and the other arm of the goniometer along the vertical axis. The angle score for TBT was found by measuring the change in angle of the tilt board relative to the horizontal starting position. All of the videotaped responses were scored and the highest angle score for each test was chosen for analysis. For further details of the scoring procedure see the studies by the first author (1984) and by Izraelevitz et al. (1985).

Results

The means, standard deviations, and ranges of the three angle scores and the total angle score for young adult males and young adult females are shown in Table 1. Because the sample was not randomly selected and because a significant correlation was found between FBR and TBR ($r = .66, p < .0001$), the Bartlett Test of Sphericity was performed to determine whether there was a significant lack of homogeneity of variance between test scores (SPSS Inc., 1983). The results of the Bartlett test were significant ($p < .0001$). Therefore, the hypothesis of homogeneity was rejected and the data analyzed with a multivariate analysis of variance (MANOVA), which is a more powerful test under such conditions. Correlations between TBT and either TBR or FBR were not significant ($r < .20$).
Since the Wilks's criterion was used, the MANOVA did not generate a mean square within or between groups. Therefore, t tests were used to test for differences in mean angle scores between sexes. Because the total number of t tests performed, two-tailed t tests were used to decrease the probability of a Type I error. As shown in Table 2, the two-tailed t tests indicated that the only significant difference between groups was for TBT; FBR approached significance. A t test used to test the hypothesis that there would be significant sex differences for the total angle scores indicated that the difference was not significant.

As the probability existed that each test measured different sex-related differences in equilibrium, a stepwise discriminant analysis was used to determine which combination of angle scores could best discriminate between the equilibrium performance of young men and young women. The level of significance required to enter the discriminant equation was set at p < .05. At Step 1, TBT angle score entered the equation (Wilks's lambda = .8796, equivalent F[1,50] = 6.84, p = .012). At Step 2, FBR angle score approached but did not reach significance (one-tailed p = .06). Therefore, only TBT angle score was entered into the equation and was able to correctly classify 64.8% of the cases; chance, alone, would have predicted 50%. As shown in Table 3, the TBT angle score was a better predictor of male than of female performance.

Results of the two-way MANOVA with repeated measures on angle scores revealed a significant Sex X Angle interaction (Wilks's lambda = .81126, approximate multivariate F[2,104] = 5.82, p = .005). Since the Wilks's criterion was used, the MANOVA did not generate a mean square within or between groups. Therefore, t tests were used to test for differences in mean angle scores between sexes. Because of the total number of t tests performed, two-tailed t tests were used to decrease the probability of a Type I error. As shown in Table 2, the two-tailed t tests indicated that the only significant difference between groups was for TBT; FBR approached significance. A t test used to test the hypothesis that there would be significant sex differences for the total angle scores indicated that the difference was not significant.

As the probability existed that each test measured different sex-related differences in equilibrium, a stepwise discriminant analysis was used to determine which combination of angle scores could best discriminate between the equilibrium performance of young men and young women. The level of significance required to enter the discriminant equation was set at p < .05. At Step 1, TBT angle score entered the equation (Wilks's lambda = .8796, equivalent F[1,50] = 6.84, p = .012). At Step 2, FBR angle score approached but did not reach significance (one-tailed p = .06). Therefore, only TBT angle score was entered into the equation and was able to correctly classify 64.8% of the cases; chance, alone, would have predicted 50%. As shown in Table 3, the TBT angle score was a better predictor of male than of female performance.

### Table 1
Means and Standard Deviations of Angle Scores for Men and Women

<table>
<thead>
<tr>
<th>Test</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>M</td>
</tr>
<tr>
<td>TBT</td>
<td>25</td>
<td>37.8</td>
</tr>
<tr>
<td>TBR</td>
<td>25</td>
<td>54.2</td>
</tr>
<tr>
<td>FBR</td>
<td>25</td>
<td>53.5</td>
</tr>
<tr>
<td>Total</td>
<td>75</td>
<td>51.5</td>
</tr>
</tbody>
</table>

Note. TBT = Tilt Board Tip. TBR = Tilt Board Reach. FBR = Flat Board Reach.

### Table 2
Results of Two-Tailed t Tests for Angle Scores Between Men and Women

<table>
<thead>
<tr>
<th>Test</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBT</td>
<td>52</td>
<td>2.66</td>
<td>.010</td>
</tr>
<tr>
<td>TBR</td>
<td>51</td>
<td>-0.86</td>
<td>.393</td>
</tr>
<tr>
<td>FBR</td>
<td>52</td>
<td>-1.91</td>
<td>.061</td>
</tr>
<tr>
<td>Total</td>
<td>51</td>
<td>0.07</td>
<td>.941</td>
</tr>
</tbody>
</table>

Note. TBT = Tilt Board Tip. TBR = Tilt Board Reach. FBR = Flat Board Reach.

### Table 3
Discriminant Analysis Classification Results for Tilt Board Tip (TBT)

<table>
<thead>
<tr>
<th></th>
<th>Predicted Group Membership</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
</tr>
<tr>
<td>Actual Group Membership</td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>18</td>
</tr>
<tr>
<td>Women</td>
<td>12</td>
</tr>
</tbody>
</table>

Discussion

Consistent with the three previous studies, the results of this study found that males obtained higher angle scores on TBT than did females, and no sex differences were found on TBR. The results for FBR have varied across studies. Because males consistently obtained higher angle scores on TBT, it appears that TBT measures some aspect of equilibrium that is sensitive to performance differences between males and females throughout the period between early childhood and young adulthood. This finding that the pattern of gender similarities (i.e., TBR angle score) and differences (i.e., TBT angle score) has been so consistent throughout the period of development between 3 and 30 years of age has led to speculation about underlying factors that might influence equilibrium abilities. Thomas and French (1985), in their meta-analysis of gender differences in motor performance, suggested biology, environment, and the interaction between them, as potential sources of influence. In their discussion of biological factors, they stated that prior to puberty "gender sameness rather than difference is a more appropriate descriptor of biological characteristics" (p. 260).

According to Thomas and French (1985), a child's performance may be influenced to a large extent by his or her perception of society's expectations for sex-appropriate behavior. This environmental influence is compounded when physical educators, parents, and perhaps therapists, have different expectations for males and females. By treating perceived differences as natural or biological, they may reinforce or even expand environmentally induced gender differences in motor performance (Thomas & French).

The lack of significant sex differences in performance between males and females on TBR and FBR across investigations suggests that there are no biological differences or environmental influences that affect performance on these two tests. In contrast, the consistent finding of male "superiority" in TBT angle scores suggests an environmental factor that exists for TBT but not for TBR or FBR.

Gender differences prior to puberty in balance... are environmentally induced. This conclusion is based on the small effect sizes, usually less than 0.50 of a standard deviation unit, as well as documented observations that treatment, expectations, and practice opportunities differ by gender. If
One possibility is that FBR and TBR are less familiar to members of both sexes and that performance is therefore not influenced by such environmental factors as practice or expectation. TBT, in contrast, may be more familiar, and/or be influenced by such environmentally induced factors as practice and or risk taking. For example, boys may have more practice than girls with skateboarding, surfing, and skiing—activities that involve balance skills similar to those tested by TBT. The finding that only the TBT angle score entered the discriminant function and that it was a better predictor of male than of female performance (see Table 3) seems to support this conclusion.

However, these hypotheses of (a) overall biological sameness and (b) environmental influence on TBT only, if valid, must also explain why these similarities and differences persist into young adulthood and be shown to apply equally to the quantitative and quality of response led to the conclusion that adult males would have been greater than those of overall, appeared to be more confident in their approach to the task whereas on FBR, females appeared to use a more cautious strategy. Therefore, performance on TBT because the extent of the difference between males and females was no greater in young adult males and females performed differently and be shown to apply equally to the quantitative and quality of the subjects’ performance on each task.

According to Thomas and French (1985), the performance of males and females will diverge after puberty (with male performance surpassing female performance), but only if the greater size and strength of males result in a biological advantage. Such a biological advantage would appear not to be relevant to performance on FBR or TBR, or the angle scores of young adult males would have been greater than those of young adult females. Likewise, such a postpuberty biological advantage would appear not to affect performance on TBT because the extent of the difference between males and females was no greater in young adulthood than it was in early childhood (Fisher & Bundy, 1982; Fishet, 1984; Izraelevitz et al., 1985). Thus, the hypothesis of biological sameness, even postpuberty, appears to be supported for angle scores.

The hypothesis that environmental differences affected only the quality of performance on TBT may not be supported by further research. The authors’ subjective observations of the subjects’ overall pattern and quality of response led to the conclusion that young adult males and females performed differently on each of the three tests. These differences were most marked for TBT and for FBR. On TBT, males, overall, appeared to be more confident in their approach to the task whereas on FBR, females appeared to use a more cautious strategy. Therefore, further research is needed on the qualitative differences in performance between males and females of all ages. Such research should give consideration to the task characteristics of each test, as well as to the effect of prior experience with such diverse activities as ballet, diving, ice hockey, and soccer. TBT requires a rapid response to an imposed perturbation. Experience with postural adjustments, perhaps obtained through participation in competitive games, may give males a slight advantage. FBR and TBR in contrast, require predominantly the ability to control self-imposed, slow perturbation of the body’s center of gravity. Therefore, performance of these two tests may be less influenced by prior experience because of fewer factors of demand. Alternatively, both sexes may have equivalent experience with activities involving controlled perturbations.

Finally, it should be noted that mean differences do not imply individual sex differences. Quantitatively and qualitatively some females performed similarly to the highest scoring males. Participation in competitive games is not uniquely male, nor do all males exhibit this behavior.

Summary and Conclusions

The results of this investigation of three clinical measures of equilibrium support the results of previous investigations. Specifically, it was found that males consistently obtained higher angle scores on TBT and that there were no sex differences for TBR or FBR. This further supports the view that age and sex differences must be taken into account when criteria for the evaluation of equilibrium performance are established. However, to improve the validity of the preliminary norms established in this study, it is recommended that a future study include a larger and more heterogeneous sample of young adults. Furthermore, to determine the nature of differences between males and females, the behavioral aspects of equilibrium performance need to be investigated. Consideration should be given to testing the hypothesis that environmental factors may account for differences in performance on TBT as well as influence the pattern of response on all three tests and that environmental factors may be less sex- than experience-related.

Acknowledgments

The second and third authors’ contributions to this investigation were completed in partial fulfillment of the requirements of the Master of Science in Occupational Therapy at Sargent College of Allied Health Professions, Boston University, and were supported in part by the Dudley Allen Sargent Research Fund.

References


