GENDER DIFFERENCES IN UNIVERSITY STUDENTS’ GYMNASTICS SPECIFIC SELF-EFFICACY AND PERFORMANCE

Dimitrios C. Milosis and Theophanis A. Siatras

Department of Physical Education and Sport Sciences, Aristotle University of Thessaloniki, Greece

Abstract
The aim of the study was to examine differences between male and female university students and their gymnastics specific self-efficacy and performance in gymnastics. In the study, 201 male and 160 female students and their 7 teachers participated voluntarily. The students attended the second year in the School of Physical Education and Sport Science (SPESS) and were obliged to take the course "Teaching of Gymnastics" for two semesters. Students’ self-efficacy was evaluated by suitable questionnaires at the beginning of the academic year, at the end of the first semester, and at the end of the academic year. Gymnastics skills’ difficulty was evaluated by teachers, and students’ performance was graded by their teachers at the end of the first semester, and at the end of the academic year. The main findings were: (a) both genders increased their self-efficacy during the academic year, while females had higher overall self-efficacy towards all gymnastics apparatus than males in the three measurements; (b) there were no gender differences in the average of the three measurements of self-efficacy toward common gymnastics skills; (c) there were no differences in gymnastics skills’ difficulty performed by males and females, and (d) females received higher grades than males. Considering the limitations of the study, the findings can be very useful for more effective organization and teaching of university gymnastics courses.

Keywords: Self-efficacy, Artistic gymnastics, Males, Females, Performance.

INTRODUCTION
Gymnastics is considered a very important sport for the motor development of children (Nilges, 1997) and for this reason it is included in the curricula of physical education at every level of education. Gymnastics is taught as a core compulsory applied course for two semesters in the second year to students in the School of Physical Education and Sport Science (SPESS). The syllabus of the course focuses on the development of the dominant movement patterns: locomotion, statics/balance, spring, landing, rotation/rolling, flight, and swing. The practical content of the course includes gymnastics skills at a basic level for each of the 4 apparatus for female and the 6 apparatus for male students respectively (MAG; WAG, Code of points, Fédération Internationale de Gymnastique, 2020a,b). Students must attend at least 70% of the three-hour weekly classes (two practical and one theoretical) to meet the requirements of the course in each semester. To successfully pass the exams, students must be graded with an average of at least 5 (out of 10) in practical (60%) and
theoretical (40%) examination of the course.

In the context of improving the quality of the course, some of the issues that concern teachers who teach gymnastics at the SPESS are poor performance of some students; non-participation in the practical part of the course for various reasons; failure to attend at least 70% of the three-hour weekly classes needed to fulfill the requirements to pass the course; dropout and abstention from the course exams at the end of semester. Specifically, our analysis of the data from the previous academic year included the present investigation showed that about 24% of students, did not attend at all in the fall semester, or did not engage in the practical course of gymnastics for a variety of reasons (unknown). About 14% did not participate at all in the exam, and 12% got below the baseline. In the spring semester, the situation was even worse: 29% did not attend, 23% did not participate in the exam, and 8% got below baseline (total percentage greater than 50%).

Self-efficacy is the self-evaluation of one’s competence to successfully perform a specific task or to obtain specific performance outcomes in a particular situation (Bandura, 1997; Pajares, 1996). Self-efficacy expectations influence the task selection; expenditure of effort; persistence in the face of barriers; resilience, and ultimately behavior (Bandura, 1997; Bandura, & Schunk, 1981); personal goals and performance (Locke. & Latham, 1990), and goal setting and commitment; cognitive actions, and affective processes (Maddux, 1995). Self-efficacy has been well-researched as one of the most influential psychological concepts affecting motivation to engage to achieve results in sport performance (Moritz, Feltz, Fahrbach, & Mack, 2000; Vealey, Hayashi, Garner-Holman, & Giacobbi, 1998). Considering that the development of specific skills and performance in gymnastics requires strong motivation, courage, determination, personal goals and commitment, it can be assumed that the importance of perceived self-efficacy is crucial (Ede, Hwang, & Feltz, 2011). Furthermore, research results have demonstrated that students’ skill-specific self-efficacy in gymnastics significantly predicted their performance at the end of the two semester course (Milosis et al., 2018).

Self-efficacy is a multidimensional construct (Zimmerman, 2000), directly related to a specific activity domain (Bandura, 1997; Maddux, 1995), and can be specific to a task at a particular level of performance (Yeo, & Neal, 2006). Considering the situational specific nature of self-efficacy, researchers typically develop self-efficacy scales specific to the research environment in which they are to be applied (Bandura, 1997; Lane, Devonport, Milton, & Williams, 2003). Such measurements evaluate students’ judgments about their capabilities to have a clear activity or task in mind, and allow researchers to relate self-efficacy to achievement and to predict performance (Bandura, 1986; Pajares, 1996). Correspondingly, in sport and physical activity settings, it has been found that task-specific scales predict better specific tasks (McAuley, & Gill, 1983). Significant positive correlations between self-efficacy and subsequent performance measurements have been reported for gymnastics among other sports (Ede, Hwang, & Feltz, 2011). For example, Weiss, Wiese, and Klint (1989) have concluded that artistic gymnasts with higher expectations of final achievement before a competition tended to be more successful than gymnasts with low expectations of success.

However, the nature of the activity and the situation under consideration are determining factors for the level of specificity where self-efficacy is measured (Bandura, 1992). For example, a high-level basketball player may have higher perceptions of self-efficacy on the basketball court, but might lack self-
confidence if required to perform a routine on the trampoline (McAuley, & Gill, 1983). However, it is possible that even within a domain of action one may have high self-efficacy for one parameter of performance but not another. For example, it is doubtful whether the performance of the forward roll could predict the performance on the cartwheel due to the different degree of difficulty of the skills (LaForge-MacKenie, & Sullivan, 2014). Thus, it is possible that a female student is highly efficacious for a turn or a jump performed on the floor apparatus, but inefficacious for the same skill executed on the balance beam. Correspondingly, a male student is likely to have a high level of self-efficacy to perform a handstand on the floor, but a lower level to execute the handstand on the parallel bars. Also, a male student may have higher self-efficacy to perform a forward tucked salto on the floor compared to a female student, but lower self-efficacy to perform a handspring on the floor compared to the same female student.

Many attempts have been made to investigate the factors that may affect different performance and achievement of the two genders’ defining perceptions about the self as a key parameter (e.g., Eisenberg, Martin, & Fabes, 1996). Gender differences in psychological parameters have been reported by many researchers and perhaps the most researched topic in sport psychology has been in the area of self-efficacy/self-confidence (Lirgg, George, Chase, & Ferguson, 1996).

In general, most self-efficacy research has concluded that male athletes consistently evaluate higher their overall physical competence and are more positive than female athletes about their ability and their performance expectations in most traditional sport activities (Moritz et al., 2000; Rattanakoses et al., 2009; Woodman, & Hardy’s, 2003). Conversely, it was hypothesized that females would give lower performance estimates, perform more poorly, report a lower ability level, and attribute success or failure to external causes when compared to males in a competitive task (Corbin, & Nix, 1979; Woodman, & Hardy, 2003).

However, it has been suggested that gender identity and gender stereotypes may affect sports competence and self-efficacy. Gender identity is a complex, multidimensional construct which evaluates one’s self-appraisal of being male or female (Egan, & Perry, 2001). Gender stereotypes relate to children’s beliefs about behavioral differences between the genders (and the desirability of such differences), specifying that certain behaviors are more important for, or more common to, one gender than the other. According to Choi (2004), self-efficacy is related to gender-role orientation, masculinity or femininity. Masculinity operationalizes self-efficacy through competitiveness, independence, aggressiveness, and assertiveness. Self-efficacy is related to femininity when it is in the domain of submissiveness, dependence, and interpersonal relationships. Earlier studies provided support for this theory. For example, findings supported that if an activity is considered more masculine, females will have less self-efficacy in the task (Corbin, & Nix, 1979; Lirgg et al., 1996; Sanguinetti, Lee, & Nelson, 1985). On the other hand, females had higher self-perceptions of ability and expectancies for success than males on activities considered more feminine, for example dance, gymnastics, figure skating, cheerleading, etc. (Lirgg et al, 1996; Clifton, & Gill, 1994). In general, females were less confident than males on a perceived masculine task, and males were less confident than females on a perceived feminine task (Lirgg, 1996).

Gymnastics has been described as a sport that includes different combinations of aesthetic and acrobatic dimensions (Goirand, 1996). According to the literature, those dimensions are marked by gendered stereotypes, defining the
aesthetics of a movement as feminine and the acrobatic dimension as masculine (Kirk, 2010). It has still been argued that the more decisive the force for the execution of a skill is, the more masculine it can be considered (Whitson, 1994). In gymnastics, males perform skills and exercises in 6 “men’s” apparatus, and females in 4 “women’s” apparatus. There are skills performed by both genders (e.g., cartwheel, forward or backward handspring or salto), skills performed only by males (e.g., cross support on still rings) or only by females (e.g., jumps, leaps, turns on balance beam).

Following the guidance obtained from the prior literature review, the aim of the study was to verify whether there were substantial differences between male and female students’ gymnastics specific self-efficacy and performance in gymnastics skills in a sample of university students. As gender differences existed in most sports self-efficacy in general, it was hypothesized that gender differences exist in gymnastics self-efficacy as well. Exploring students’ gymnastics-specific self-efficacy is important for a better understanding of how to better organize and teach a gymnastics course to increase male and female students’ self-efficacy beliefs and eventually their performance in gymnastics skills and exercises.

METHODS

In the study, 361 students (201 males; $M = 20.16$ years old, $SD = 2.44$, and 160 females; $M = 19.95$ years old, $SD = 2.30$) who attended the second year in the School of Physical Education and Sport Science (SPCSS) and their 7 teachers (over 20 years of experience) participated voluntarily. Data collected from five students (2 males and 3 females) who were competitive gymnastics athletes in the past were excluded from the statistical analyses. According to the curriculum of the SPCSS, all students were obliged to attend the course "Teaching of Gymnastics" for two semesters. The teaching of fundamental gymnastics skills for male and female students (e.g., forward, and backward roll, cartwheel, headstand, handstand) is included in the Greek physical education curriculum for primary and secondary education. However, due to a lack of gymnastics facilities and suitable equipment in Greek schools, most Greek students have been taught gymnastics at a very low level or not at all. Consequently, the participants that were involved in the present study had very little or no experience and were considered as novices in gymnastics. Participants were informed that the purpose of the research was to improve the course “Didactics of Gymnastics”, that their answers in the questionnaire were confidential and would not be disclosed to third parties, and that they would not affect their grades in the two semesters. All participants were fully informed about the purpose of the study and the implementation of the protocols, according to the Ethical Committee of the University guidelines. Written informed consent was obtained from all participants.

For the evaluation of students’ self-efficacy, two reliable and valid skill-specific self-efficacy questionnaires individualized for male and female students were used. Particularly, males’ questionnaire consisted of 33 skills-specific items divided into 6 subscales, one for each apparatus (11 items for MFX, 3 for PH, 3 for SR, 3 for MV, 7 for PB, and 6 for HB). Accordingly, females’ questionnaire consisted of 41 skills-specific items divided into 4 subscales, one for each apparatus (3 items for FV, 9 for UB, 14 for BB, and 15 for FFX). The level of difficulty of the skills presented to the students was evaluated by 7 teachers (with over 20 years of experience each) who taught in the SPCSS gymnastics courses and who had themselves responded to a properly designed questionnaire. To investigate gender differences in self-efficacy, only the items for the skills that were common to male and female students
were used [9 skills on the FX, 3 on the V, and 4 on the HB and UB (HBUB) (Table 1)]. The vault skills were presented (and performed) on the “vaulting horse” (pommel horse, but without the handles placed sideways) for females, and on the “vaulting table” for male students respectively. The squat through and the straddle over vault skills were performed using springboard and the handspring using mini trampoline for both genders.

Furthermore, video elicitation based on criteria for the optimal performance was used as a tool to provide students an accurate movement of execution of each gymnastic skill that was to perform. The instruments used in the present study were developed for the evaluation of second year students’ gymnastics self-efficacy, based on the curriculum of the compulsory course “Teaching of Gymnastics” of SPESS (Milosis et al., 2018).

Table 1
Common gymnastics skills for male and female students.

<table>
<thead>
<tr>
<th>Floor exercise</th>
<th>Horizontal bar/Uneven bars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward handspring</td>
<td>Glide kip</td>
</tr>
<tr>
<td>Back standing scale</td>
<td>From hang pull over to support</td>
</tr>
<tr>
<td>Salto forward tucked</td>
<td>Back hip circle</td>
</tr>
<tr>
<td>Handstand forward roll</td>
<td>Underswing dismount</td>
</tr>
<tr>
<td>Cartwheel</td>
<td></td>
</tr>
</tbody>
</table>

The questionnaire evaluating the level of difficulty of the gymnastics skills was completed by the teachers before the beginning of the first semester. The self-efficacy questionnaire was completed by the students three times during the academic year: (a) in the first gymnastics lesson at the beginning of the first semester; (b) in the last lesson of the first semester before the examinations, and (c) in the last lesson of the second semester before the final examinations for the academic year. Before the process began, it was highlighted to the students that the skills would be presented in random order regardless of their difficulty. In addition, students were asked not to comment on the difficulty of the skills so that they would not affect their co-examined peers either positively or negatively. Then, corresponding video clips related to the technical execution of each skill were projected.

Each video clip was screened twice so that the students could get the best picture of the performance required based on the criteria that were set in the design of the study. Immediately afterwards, students were asked to answer one self-efficacy question regarding the presented gymnastics skill. For example, for the skill “squat through” on the “vaulting horse” for female students, the corresponding video clip related to the technical execution was projected and the following question was posed to female students: “How confident are you that after your participation in the gymnastics courses you will be able to perform the ‘squat through’ on the ‘vaulting horse’ at the end of the academic year in the way shown in the video?” Afterwards, students rated their confidence on a 10-point scale (1 = not at all confident, 10 = extremely confident). The above procedure was followed throughout the questionnaire for both male and female
Students and lasted 30-40 minutes for each measurement.

Students’ grades from their teachers (n = 7) for the first and the second semester based on their performance in the practical examination were collected. Specifically, the students performed all the skills on the six apparatus for males and four apparatus for females and were evaluated by their teachers and graded from 1 to 10 for each skill according to the quality of execution. Each skill was evaluated and assessed in 2-7 predefined criteria for quality execution depending on the magnitude of the deviation from perfect execution. The content validity of the criteria used were tested in a previous study (Milosis et al., 2018). The average score of all apparatus (ALL) for males and females was calculated and was used in statistical analyses (Table 4).

The normality of the data was examined using Kolmogorov-Smirnov Test. A p value less than 0.05 was considered significant. The data were analysed using the SPSS software (SPSS v. 23, SPSS Statistics, IBM Corp., NY USA). The internal consistency of students’ self-efficacy towards each gymnastics apparatus was assessed using the Cronbach’s alpha coefficient. Differences between the evaluation of male and female skills’ difficulty by their teachers were examined using multivariate analysis of variance (MANOVA). To examine gender effects on students’ self-efficacy and grades, repeated-measures ANOVAs were used.

RESULTS

Reliability, repeatability, content, concurrent, and predictive validity of the instruments used were supported by the results of a previous study (Milosis et al., 2018). The frequency distributions of self-efficacy sum scores come close to a normal distribution. For both males and females, skewness and kurtosis had values between −1.0 and +1.0. Kolmogorov-Smirnov test of normality was not significant p > .05. The alpha reliability coefficient was satisfactory for all variables (Tables 1, 2).

Differences in students’ self-efficacy during the academic year

Repeated-measures analyses of variance were conducted in order to examine changes in students’ self-efficacy during the academic year. More specifically, they tested within-subjects effects between the three measurements of male and female students’ self-efficacy towards each gymnastics apparatus as illustrated in the following. Males’ apparatus MFX, PH, SR, MV, PB, HB, and the mean score of all apparatus (MALL); Females’ apparatus FV, UB, BB, FFX, and the mean score of all apparatus (FALL).

Differences in males’ self-efficacy during the academic year

Separate repeated-measures ANOVAs were conducted with students’ self-efficacy towards each gymnastics apparatus as dependent variable and time of measurement as the within-subjects factor. Results revealed that there was a significant within-subjects effect for students’ self-efficacy towards each gymnastics apparatus (see Table 2): (a) MFX, F(2,400) = 30.25, p < .001, η² = .13, (b) PH, F(2,400) = 44.07, p < .001, η² = .18, (c) SR, F(2,400) = 95.74, p < .001, η² = .32, (d) MV, F(2,400) = 46.18, p < .001, η² = .19, (e) PB, F(2,400) = 28.87, p < .001, η² = .13, (f) HB, F(2,400) = 26.20, p < .001, η² = .12, and (g) MALL, F(2,400) = 92.32, p < .001, η² = .32. Pairwise comparisons revealed that except for the measurements of the initial and the first semester for the PB, males’ self-efficacy increased significantly (p < .005) measurement by measurement (Table 1, Figure 1).
Table 1
Means (M), standard deviations (SD), alpha reliabilities, significant differences (F), and effect size ($\eta^2$) of males (n = 201) self-efficacy, for the three measurements.

| Abbreviations: MFX, males’ floor exercise; PH, pommel horse; SR, still rings; MV, males’ vault; PB, parallel bars; HB, horizontal bar; MALL, the mean score on all males’ apparatus. Excluding the means of the initial and first semester for the males’ PB (means in boldface), all means differed significantly between the three measurements (p < .05). All F statistics were statistically significant (p < .001). |

<table>
<thead>
<tr>
<th>Initial Measure</th>
<th>End of First Semester</th>
<th>End of Second Semester</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>MFX</td>
<td>6.21</td>
<td>1.50</td>
</tr>
<tr>
<td>PH</td>
<td>5.22</td>
<td>1.64</td>
</tr>
<tr>
<td>SR</td>
<td>4.60</td>
<td>2.07</td>
</tr>
<tr>
<td>MV</td>
<td>5.23</td>
<td>1.98</td>
</tr>
<tr>
<td>PB</td>
<td><strong>5.78</strong></td>
<td>1.63</td>
</tr>
<tr>
<td>HB</td>
<td>5.20</td>
<td>1.75</td>
</tr>
<tr>
<td>MALL</td>
<td>5.39</td>
<td>1.47</td>
</tr>
</tbody>
</table>

Table 2
Means (M), standard deviations (SD), alpha reliabilities, significant differences (F), and effect size ($\eta^2$) of females (n = 160) self-efficacy, for the three measurements.

| Abbreviations: FV, females’ vault; UB, uneven bar; BB, balance beam, FFX, females’ floor exercise, FALL, the mean score on all females’ apparatus. Excluding the means of the first and second semester for the females’ FFX (means in boldface), all means differed significantly between the three measurements (p < .05). All F statistics were statistically significant (p < .001). |

<table>
<thead>
<tr>
<th>Initial Measure</th>
<th>End of First Semester</th>
<th>End of Second Semester</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>FV</td>
<td>4.78</td>
<td>1.86</td>
</tr>
<tr>
<td>UB</td>
<td>5.97</td>
<td>1.48</td>
</tr>
<tr>
<td>BB</td>
<td>6.44</td>
<td>1.84</td>
</tr>
<tr>
<td>FFX</td>
<td>6.68</td>
<td>1.78</td>
</tr>
<tr>
<td>FALL</td>
<td>5.98</td>
<td>1.48</td>
</tr>
</tbody>
</table>
Figure 1. Males’ vs. females’ self-efficacy towards each gymnastics apparatus and the mean score of all apparatus for the three measurements. Excluding the measurements for the initial and first semester for the PB (male) and for the first and the second semester for the FFX (female) self-efficacy increased significantly ($p < .05$) measurement by measurement.

Abbreviations: MFX, males’ floor exercise; PH, pommel horse; SR, still rings; MV, males’ vault; PB, parallel bars; HB, horizontal bar; MALL, the mean score of all males’ apparatus; TMALL, teachers’ evaluation of difficulty for of all males’ apparatus; FV, females’ vault; UB, uneven bar; BB, balance beam, FFX, females’ floor exercise, FALL, the mean score of all females’ apparatus; TFALL, teachers’ evaluation of difficulty for of all females’ apparatus.
Figure 2. Teachers’ evaluation of gymnastics skills difficulty.
Abbreviations: TMFALL, teacher evaluation for the total number of gymnastics skills on the six males’ and four females’ apparatus; TMFFX, teacher evaluation for the nine common gymnastics skills for males and females on the floor exercise; TMFV, teacher evaluation for the three common gymnastics skills on the vault; THBUB, teacher evaluation for the four common gymnastics skills on the horizontal bar and the uneven bars.
Higher scores represent lower difficulty.

Table 3
Descriptive statistics for contrast effects followed the within subjects interaction between group and time.

<table>
<thead>
<tr>
<th></th>
<th>Initial measure</th>
<th>End of first semester</th>
<th>End of second semester</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>ALL</td>
<td>5.39</td>
<td>1.47</td>
<td>5.98</td>
</tr>
<tr>
<td>FX</td>
<td>6.09</td>
<td>1.37</td>
<td>5.95</td>
</tr>
<tr>
<td>V</td>
<td>5.32</td>
<td>1.87</td>
<td>4.80</td>
</tr>
<tr>
<td>HBUB</td>
<td>4.59</td>
<td>1.79</td>
<td>5.19</td>
</tr>
<tr>
<td>Grades</td>
<td>6.58</td>
<td>1.80</td>
<td>7.41</td>
</tr>
</tbody>
</table>

Abbreviations: M, Means; SD, Standard deviations; \( F \), Significant differences; \( \eta^2 \), Effect size; ALL, the total number of gymnastics skills on the six males’ and four females’ apparatus; FX, nine common gymnastics skills for males and females on the floor exercise; V, the three common gymnastics skills on the vault; HBUB, four common gymnastics skills on the HB and UB.
Means in boldface were increased more compared to the corresponding mean of the opposite gender relatively with the means of the prior measurement.
*p < .05, **p < .01, ***p < .001
Figure 3. Males’ vs. females’ self-efficacy towards all gymnastics apparatus, floor exercise, vault, horizontal bar and uneven bars.
Figure 4. Males’ vs. females’ average self-efficacy of the three measurements (initial, first, second semester) for the common skills, the total skills on all apparatus (6 for male and 4 for female) and grades (first, second semester).

Abbreviations: AMFFX, average self-efficacy of the three measurements for male and female common floor exercise skills; AMFV, average self-efficacy of the three measurements for male and female common vault skills; AMFHBUB, average self-efficacy of the three measurements for male and female common horizontal and uneven bars skills; AMFC, average self-efficacy of the three measurements for all male and female common skills; AMFALL, average self-efficacy of the three measurements for male and female skills of all apparatus (6 for males, 4 for females); Grades, average grades of the two semesters.

Differences in females’ self-efficacy during the academic year

Separate repeated-measures ANOVAs were conducted with students’ self-efficacy towards each gymnastics apparatus as dependent variable and time of measurement as the within-subjects factor. Results revealed that there was a significant within-subjects effect for students’ self-efficacy towards each gymnastics apparatus (see Table 1): (a) FV, $F(2,318) = 96.30, p < .001, \eta^2 = .38$, (b) UB, $F(2,318) = 51.26, p < .001, \eta^2 = .24$, (c) BB, $F(2,318) = 82.58, p < .001, \eta^2 = .34$, (d) FFX, $F(2,318) = 18.56, p < .001, \eta^2 = .11$, (e) FALL, $F(2,318) = 144.13, p < .001, \eta^2 = .48$. Pairwise comparisons revealed that except for the measurements for the second and the third semester for the FFX, females’ self-efficacy increased significantly ($p < .05$) measurement by measurement (Table 2, Figure 1).

Teachers’ evaluation of gymnastics skills difficulty

Multivariate analysis of variance (MANOVA) was conducted in order to examine differences between the evaluation of male and female skills difficulty by their teachers. Students’ gender was the independent variable and teachers’ evaluation of the difficulty of gymnastics skills for: (a) the total number of skills of the six males’ and four females’ apparatus (TMFALL), (b) nine common skills of the floor exercise (TMFFX), (c) the three common skills of the vault (TMFV), and (d) four common skills of the horizontal bar and uneven bars (THBUB) were the dependent variables. Results revealed a significant main effect for gender, Wilks’ $\lambda = .152, F(4,9) = 12.56, p < .001, \eta^2 = .85$. Given the significance of the overall test, the univariate main effects were examined. Not significant univariate main effects for gender were obtained for all variables (a) TMFALL, $F(1,12) = .23, p$
Gender differences on students’ self-efficacy and grades

Repeated-measures ANOVAs were conducted in order to examine gender effects on students’ self-efficacy and grades. More specifically, they tested (a) within-subjects effects between the three measurements, (b) between-group differences from first up to the third measurement, and (c) possible differences on the mean of the three measurements between males and females. For the estimation of students’ self-efficacy, students’ responses were computed to: (a) the total number of skills of the six males’ and four females’ apparatus (ALL), (b) nine common skills of the FX, (c) the three common skills of the V, and (d) four common skills of the HB and UB (HBUB).

Differences in self-efficacy towards ALL

A repeated-measures ANOVA was conducted with students’ self-efficacy towards ALL as dependent variable, time of measurement as the within-subjects factor and students’ gender as the between-subjects factor; it revealed that there was a significant within-subjects effect for the total sample, $F(2,359) = 215.41$, $p < .001$, $\eta^2 = .38$, and no significant within-subjects interaction between group (gender) and time of measurement, $F(2,359) = 1.79$, $p = .171$, $\eta^2 = .01$. Inspection of between-subjects effects revealed that there were no significant differences between the two groups across all measurements, $F(1,359) = .000$, $p = .999$, $\eta^2 = .00$. Moreover, contrast effects following the within-subjects interaction between group and time revealed that from initial measurement to first semester measurement, the self-efficacy towards ALL of females increased more compared to males’, $F(1,359) = 6.41$, $p < .01$, $\eta^2 = .02$ (Table 3, Figure 3).

Differences in self-efficacy towards FX

A repeated-measures ANOVA was conducted with students’ self-efficacy towards FX as dependent variable, time of measurement as the within-subjects factor and students’ gender as the between-subjects factor, revealed that there was a significant within-subjects effect for the total sample, $F(2,718) = 77.38$, $p < .001$, $\eta^2 = .18$, and a significant within-subjects interaction between gender and time of measurement, $F(2,718) = 3.63$, $p < .05$, $\eta^2 = .01$. Inspection of between-subjects effects revealed that there were no significant differences between the two groups across all measurements, $F(1,359) = .000$, $p = .999$, $\eta^2 = .00$. Moreover, contrast effects following the within-subjects interaction between group and time revealed that from initial measurement to first semester measurement the self-efficacy towards FX of females increased more compared to males’, $F(1,359) = 4.37$, $p < .05$, $\eta^2 = .01$ (Table 3, Figure 3).

Differences in self-efficacy towards V

A repeated-measures ANOVA was conducted with students’ self-efficacy towards V as dependent variable, time of measurement as the within-subjects factor and students’ gender as the between-subjects factor; it revealed that there was a significant within-subjects effect for the total sample, $F(2,718) = 140.32$, $p < .001$, $\eta^2 = .28$, and a significant within-subjects interaction between gender and time of measurement, $F(2,718) = 8.00$, $p < .001$, $\eta^2 = .06$, and the second semester measurement respectively, $F(1,359) = 17.11$, $p < .001$, $\eta^2 = .05$. Moreover, contrast effects following the within-subjects interaction between group and time revealed that from initial measurement to first semester measurement, the self-efficacy towards ALL of females increased more compared to males’, $F(1,359) = 14.38$, $p < .001$, $\eta^2 = .04$, the first, $F(1,359) = 24.28$, $p$
Inspection of between-subjects effects revealed that there were no significant differences between the two groups across all measurements, $F(1,359) = 1.14, p = .286, \eta^2 = .00$. However, subsequent ANOVAs revealed that males’ self-efficacy towards $V$ in the initial measurement was higher compared to females’, $F(1,359) = 9.39, p < .01, \eta^2 = .03$. Moreover, contrast effects following the within subjects interaction between group and time revealed that from first semester measurement to second semester measurement the self-efficacy towards $V$ of females increased more compared to males’, $F(1,359) = 6.48, p < .01, \eta^2 = .02$ (Table 3, Figure 3).

### Differences in self-efficacy towards HBUB

A repeated-measures ANOVA was conducted with students’ self-efficacy towards HBUB as dependent variable, time of measurement as the within-subjects factor and students’ gender as the between-subjects factor; it revealed that there was a significant within-subjects effect for the total sample, $F(2,718) = 71.40, p < .001, \eta^2 = .17$, and a significant within-subjects interaction between gender and time of measurement, $F(2,359) = 4.04, p < .05, \eta^2 = .01$. Inspection of between-subjects effects revealed that there were significant differences between the two groups across all measurements, $F(1,718) = 3.81, p < .05, \eta^2 = .01$. Subsequent ANOVAs revealed that females’ self-efficacy towards HBUB in the initial measurement was higher compared to males, $F(1,359) = 9.39, p < .01, \eta^2 = .03$. Moreover, contrast effects following the within subjects interaction between group and time revealed that from initial measurement to first semester measurement the self-efficacy towards HBUB of males increased more compared to females’, $F(1,359) = 7.61, p < .01, \eta^2 = .02$ (Table 3, Figure 3).

### Differences in students’ grades

Frequencies analyses conducted revealed that for the first semester, 10.5% of the males and 1.9% of the females were graded below 5, while 20.9% of the males and 28.1% of the females did not participate in the examinations. Correspondingly for the second semester, 13% of the males and 0.6% of the females were graded below 5, while 37.8% of the males and 35.6% of the females did not participate in the examinations. A multivariate analysis of variance (MANOVA), conducted with students’ grades in gymnastics for the first and the second semester as the dependent variables and students’ gender as the independent variable, revealed that there was a significant main effect of gender on students’ grades, $F(2,201) = 6.21, p < .01, \eta^2 = .06$. Subsequent ANOVAs revealed significant differences between males and females in average grades for the first semester, $F(1,201) = 11.74, p < .001, \eta^2 = .05$, and the second semester, $F(1,201) = 9.54, p < .01, \eta^2 = .04$, with females outperforming males in both cases (Table 3, Figure 3).

### Differences in students’ average self-efficacy and grades

Multivariate analysis of variance (MANOVA) was conducted in order to examine differences for the average of the three measurements of males’ and females’ self-efficacy and the grades for the first and the second semester. Students’ gender was the independent variable and the average score of their grades (first, second semester) and their self-efficacy measurements (initial, first, second semester) for: (a) nine common skills of the floor exercise (AMFFX), (b) the three common skills of the vault (AMFV), and (c) four common skills of the horizontal bar and uneven bars (AHBUB), (d) the total number of the common skills (AMFC), (e) the total number of skills of the six males’ and four females’ apparatus (AMFALL), were the dependent variables. Results revealed a significant main effect for gender, Wilks’ $\lambda = .619, F(5,206) = 25.36, p < .001, \eta^2 = .38$. Given the
significance of the overall test, the univariate main effects were examined. Significant univariate main effects for gender were obtained for the variables (a) AMFALL, \( F(1,210) = 7.30, p < .01, \eta^2 = .03 \) (males 6.27 ± 1.39, females 6.78 ± 1.32, mean ± SD) and (b) Grades, \( F(1,210) = 12.18, p < .001, \eta^2 = .06 \) (6.38 ± 1.79, 7.22 ± 1.66) and not significant for the other variables (a) AMFFX, \( F(1,210) = .74, p = .787, \eta^2 = .00 \) (6.56 ± 1.28, 6.50 ± 1.47), (b) AMFV, \( F(1,210) = 2.28, p = .132, \eta^2 = .01 \) (6.19 ± 1.43, 5.89 ± 1.48), (c) AHBUB, \( F(1,210) = .74, p = .787, \eta^2 = .00 \) (5.55 ± 1.61, 5.69 ± 1.61), and (d) AMFC, \( F(1,210) = .14, p = .709, \eta^2 = .00 \) (5.86 ± 1.36, 5.91 ± 1.45) (Figure 4).

**DISCUSSION**

The aim of the present study was to examine gymnastics specific self-efficacy differences between male and female university students. Results revealed that except for the measurements of the initial and the first semester for the PB and for the first and the second semester for the FFX, students’ self-efficacy significantly increased, measurement by measurement, for both genders (Table 1 & 2, Figure 1). These findings are consistent with the self-efficacy theory, according to which mastery experiences provide the most influential source of efficacy information (Bandura, 1986). In the first measurement of the current study, participants had no previous experience with either of the tasks and consequently relied on other sources of efficacy information, as, for example, perceptions of the gender-appropriateness of the task and conception of ability. In line with research findings, during the academic year, as learning progressed, students’ self-efficacy increased (Bandura, 1997; Schunk, 1996).

According to Schunk (1996), students enter learning situations with varying degrees of self-efficacy for learning. As they engage in activities, students are affected by personal factors (e.g., goal setting, information processing) and situational influences (e.g., rewards, teacher feedback) that provide them with cues about how well they are learning (Schunk, 1996). Self-efficacy is enhanced when students perceive they are performing well or becoming more skillful (Bandura, 1997; Schunk, 1996) and focus on personal achievements (Maddux, 1995). Lack of success or slow progress do not necessarily diminish self-efficacy if learners believe they know how to perform better, such as by working harder, seeking help, or switching to a more effective strategy (Schunk, 1996). In Figure 1 it appears that females in general tend to have higher self-efficacy beliefs for all gymnastics skill on the 4 women’s gymnastics apparatus compared to males (for the 6 men’s gymnastics apparatus). It is also shown that males realized that it was easier to perform the skills on FX and more difficult on SR, while females perceived as easier the skills on the BB and more difficult on V.

Results revealed a significant main effect for gender between the evaluation of male and female skills difficulty by their teachers. Although teachers showed a tendency to evaluate as more difficult the gymnastics skills performed on all apparatus (6 for males and 4 for females) for males and the common skills performed on FX, V, and HB/UB for females, examining the univariate main effects did not find significant effects for gender for all variables. Looking at the curriculum, it seems that it includes skills with increasing level of difficulty, and common skills for both genders, as well as skills that could be characterized as more suitable for males (e.g., dynamic skills on SR) or females (e.g., turns, leaps, and jumps on FX and/or BB). However, the non-significant differences regarding difficulty of skills performed on all apparatus that favor females could be explained by the fact that the curriculum includes proportionately easier skills for females compared to males. The evaluation
of male and female skills difficulty by the teachers was in complete agreement with the students' grades based on their performance. Specifically, as shown by the results, females had higher grades compared to males for the first and the second semester. In contrast, the evaluation of the common skills as more difficult for females, by the teachers, is likely, since the common skills are few, with an objectively higher level of difficulty. Furthermore, the successful execution of many of them is based to a decisive degree on strength and speed.

In relation to these results, the examination of students' self-efficacy towards ALL, and common skills on FX, V, and HB/UB, across all measurements revealed the following results. Females had significantly higher overall self-efficacy towards ALL than males’ in the initial, the first and the second semester measurement, and for the average of the three measurements. Furthermore, from the initial measurement to the first semester measurement, their self-efficacy increased more compared to that of males. Weitlauf, Cervone, Ronald and Wright (2001) also reported a significant increase in self-efficacy in women attending a voluntary self-defense training program. These findings are contrary to the findings from previous research where males reported higher levels of confidence in a variety of physical skills (Moritz et al., 2000; Rattanakoses et al., 2009; Woodman, & Hardy’s, 2003).

The findings of the current study could be explained partially by the fact that, according to teachers’ evaluation, females’ skills on the 4 women’s apparatus included more skills, in general, and many of them were evaluated as much easier compared to the set of skills related to the 6 apparatus of males. Furthermore, although females performed on a fewer apparatus compared to males (4 vs. 6), they performed more skills, including “female” skills (e.g., choreography, dance, turns, jumps, leaps), and this likely resulted in higher self-efficacy. In support, it has been found that females are more confident than males on a feminine type of tasks (Clifton, & Gill, 1994; Lirgg et al., 1996) and on selected physical activities, such as cheerleading, gymnastics, and dance (Eder & Parker, 1987). As supported by recent research (Clifton, & Gill, 1994; Lirgg et al., 1996; Sanguinetti et al., 1985), females do not display a lack of confidence in all situations. It seems that individuals’ expectancies for success increase when participating in activities deemed gender appropriate.

Considering the common skills on FX, there were no significant differences between the two groups, with the exception that from the initial measurement to the first semester measurement, the self-efficacy towards FX of females increased more compared to that of males. However, in the initial measurement, females had not significantly lower self-efficacy than males. The skills “backward roll to handstand” and “salto forward tucked” may have led to this result because those are the skills that need great speed and strength for effective and quality execution.

Regarding the common skills on V, there were no significant differences between the two groups across all measurements. However, self-efficacy of males towards V in the initial measurement was higher compared to that of females. Moreover, from the first semester measurement to the second semester measurement, the self-efficacy of females towards V increased more compared to that of males. It is very important to mention that the vault skills were presented (and performed) on the “vaulting horse” (pommel horse, but without handles, placed sideways) for females (it is easier to perform the skills) and on the “vaulting table” for male students. The squat through and the straggle over vault skills were performed using springboard and the handspring using mini trampoline for both
genders. However, to perform well on V requires great speed and strength for effective push down onto the springboard (on foot), and shoulder and hands block onto the vaulting horse/table. This is why males are more likely to perform these skills more easily and efficiently compared to females. Researchers reported that females were less confident than males on this perceived masculine task (Clifton, & Gill, 1994; Lirgg, et al., 1996).

Finally, there were significant differences towards HBUB between the two groups across all measurements. Specifically, females’ self-efficacy towards HBUB in the initial measurement was higher compared to males. In contrast, males’ self-efficacy towards HBUB increased more from the initial measurement to the first semester measurement compared to that of females. However, there were no significant differences between genders regarding the average of the self-efficacy for the common skills on FX, V and HBUB.

The results of the present study provide support for the arguments that there is a tendency for most gender differences in self-efficacy beliefs to be small and with no clear direction, and that there is a continuum of gender differences as opposed to the simple yes or no answer (Pajares, 2005). As reported by Clifton and Gill (1994), since educational gymnastics includes skills that can be characterized as feminine or masculine, it is possible for male and female students to show similar patterns in terms of self-efficacy beliefs. Furthermore, as educational gymnastics involves students working within their own abilities, both male and female university students may overestimate or underestimate their self-efficacy depending on how they perceive their ability (Milosis et al., 2018).

The findings of the present study should be viewed in light of some limitations. The data for the evaluation of self-efficacy were collected solely from university male and female students and this must be considered before we generalize the results to include students of other levels of education or gymnasts. Furthermore, the current study concerns students’ self-efficacy in terms of specific gymnastics skills taught in the course “Didactics of Gymnastics”. Expanding the study to include students in other individual or team sports courses may lead to different results, and thus could be an interesting inquiry for future research. The present study did not examine students’ perceptions about gender appropriateness of the gymnastics skills included in the curriculum. Such an assessment would to some extent explain the differences found in self-efficacy between genders.

Additionally, it would be beneficial to investigate students’ pre-existing implicit beliefs about intelligence, or conception of ability. It has been suggested that pre-existing belief systems, such as implicit theories of intelligence, or conceptions of ability, also exert an influence on self-efficacy beliefs, especially in adults (Jourden, Bandura, & Banfield, 1991). Future studies could also investigate other personality and social factors that influence different types of confidence and sources of confidence, such as the structure of the teaching environment, the academic program, and extracurricular involvement. Next, a qualitative research based on student interviews would also be of interest.

Self-efficacy acts as a cognitive mediator of behavior, including task selection, persistence in the face of barriers, resilience, goal setting and commitment, and effort expended (Bandura, 1997; Bandura, & Schunk, 1981; Locke, & Latham, 1990; Maddux, 1995). It has been shown that it is a significant predictor of students’ gymnastics performance (Milosis et al., 2018). It has been argued that in order to promote female’s self-efficacy, females should be engaged in environments that provide social support (e.g., positive feedback from teammates, encouragement
from significant others, a sense of belonging and control) while allowing them to develop their skills. Females often use self-regulation skills in a learning setting, and they express more confidence in the ability to accomplish tasks.

On the other hand, demonstration of ability (i.e., showing an ability by outperforming others or winning) and satisfaction with accomplishments were significantly more important for the promotion of male athletes’ self-efficacy (Vealey et al., 1998). Thus, teachers may need to be intentional with their teaching style when interacting with female versus male students, realizing the potential differences in self-efficacy between the genders. Generally, students gain confidence when they engage in environments structured to be inclusive, supportive of autonomy, challenging, motivating, to provide clear performance evaluation and to encourage self-regulation of cognitions and behavior, skill building, and healthy norms (Dzewaltowski, Estabrooks, Gyurcsik, & Johnston, 2002). Furthermore, by offering student activities that are considered gender-appropriate, they are likely to increase their expectations of success and consequently increase their self-efficacy (Lirgg et al., 1996).

CONCLUSION

Considering the limitations of the present study, the results provide useful information regarding the changes in university students’ skill-specific self-efficacy in gymnastics as a consequence of learning during an academic year. These results support the hypothesis that during the academic year, as learning progressed, students’ self-efficacy increased because students gained mastery experience and could see progress in their performance. In addition, tracking of self-efficacy and performance over a series of sessions may provide useful information about individual students and may help to early identify students that are low in self-efficacy, guide teachers’ judgements, and reduce failure rates. Thus, it is critical for teachers to formulate and maintain positive expectations for all students, to develop an inclusive, autonomy-supportive, challenging, and task/mastery-oriented environment, and to encourage self-regulation of cognitions and behavior.

Furthermore, exploring male and female students’ gymnastics-specific self-efficacy differences is important for a better understanding of how to better organize and teach a gymnastics course to increase male and female students’ self-efficacy beliefs and eventually their performance in gymnastics skills and exercises. Gender differences in competency beliefs influence students’ activity choices and participation and should be considered when studying achievement. Therefore, it is crucial for teachers to convey to students that the gender appropriateness of skills is socially constructed, not based on biological factors, and that gymnastics skills are developed through deliberate practice. By doing so, students may be encouraged to attempt a variety of skills, which, in turn, may positively influence their beliefs about what they are truly capable of achieving. Teachers must cultivate student's beliefs in their capabilities, while at the same time create conditions for the desired success to be achieved since it is usually easier to weaken self-efficacy beliefs through negative appraisals than to strengthen such beliefs through positive encouragement.

REFERENCES


Bandura, A. (1992). On rectifying the comparative anatomy of perceived control: Comments on” cognates of personal
control.” *Applied and Preventative Psychology*, 1, 121-126. doi: 10.1016/S0962-1849(05)80153-2


of Sport and Exercise Psychology, 18, 426-343.


Contemporary Educational Psychology,
25, 82-91. doi: 10.1006/ceps.1999.1016

Corresponding author:

Dimitrios Milosis
Aristotle University of Thessaloniki,
Department of Physical Education and
Sports Science, Laboratory of Evaluation
of Human Biological Performance - New
facilities
Thermi, 57001 Thessaloniki, Greece.
Email: dmylosis@phed.auth.gr

Article received: 2.4.2022
Article accepted: 13.6.2022