Does Polarized Training Improve Performance in Recreational Runners?

Iker Muñoz, Stephen Seiler, Javier Bautista, Javier España, Eneko Larumbe, and Jonathan Esteve-Lanao

Purpose: To quantify the impact of training-intensity distribution on 10K performance in recreational athletes.

Methods: 30 endurance runners were randomly assigned to a training program emphasizing low-intensity, sub-ventilatory-threshold (VT), polarized endurance-training distribution (PET) or a moderately high-intensity (between-thresholds) endurance-training program (BThET). Before the study, the subjects performed a maximal exercise test to determine VT and respiratory-compensation threshold (RCT), which allowed training to be controlled based on heart rate during each training session over the 10-wk intervention period. Subjects performed a 10-km race on the same course before and after the intervention period. Training was quantified based on the cumulative time spent in 3 intensity zones: zone 1 (low intensity, <VT), zone 2 (moderate intensity, between VT and RCT), and zone 3 (high intensity, >RCT). The contribution of total training time in each zone was controlled to have more low-intensity training in PET (±77/3/20), whereas for BThET the distribution was higher in zone 2 and lower in zone 1 (±46/35/19).

Results: Both groups significantly improved their 10K time (39min18s ± 4min54s vs 37min19s ± 4min42s, \( P < .0001 \) for PET; 39min24s ± 3min54s vs 38min0s ± 4min24s, \( P < .001 \) for BThET). Improvements were 5.0% vs 3.6%, ~41 s difference at post-training-intervention. This difference was not significant. However, a subset analysis comparing the 12 runners who actually performed the most PET (n = 6) and BThET (n = 16) distributions showed greater improvement in PET by 1.29 standardized Cohen effect-size units (90% CI 0.31–2.27, \( P = .038 \)).

Conclusions: Polarized training can stimulate greater training effects than between-thresholds training in recreational runners.

Keywords: training zones, running performance, optimal distribution, training volume, training periodization

In recent years, training-intensity distribution has received attention as a potential determinant of endurance-training impact.\(^1,3\) Two consistent characteristics of elite endurance athletes’ training are large total training volume and high percentage of volume performed at intensity below the first lactate or ventilatory threshold. This pattern of emphasizing training below but also above the lactate threshold range has been termed polarized endurance training (PET).\(^2\) Paradoxically, it differs markedly from American College of Sports Medicine guidelines aimed at sedentary and low-active populations, which emphasize training at an intensity approximating the traditional lactate threshold.\(^5\)

There are substantial descriptive data supporting the “polarized” distribution in well-trained and highly trained endurance athletes from a variety of sports.\(^4,6,7\) Experimental and correlational data from well-trained subjects suggest that overemphasizing training at threshold-range intensities (moderately high-intensity [between-thresholds] endurance training; BThET) may be ineffective, or even counterproductive,\(^7,8\) particularly with respect to inducing positive adaptations in the blood lactate–power profile. Most recently, quasi-experimental data from national-team Chinese speed skaters\(^9\) and a case study of an elite 1500-m runner\(^10\) demonstrated improved performance when athletes altered their intensity distribution from a highly threshold-oriented focus to a more polarized intensity distribution. For highly trained athletes training 10 to 25 h/wk, this training distribution may allow maximal adaptive signaling while minimizing autonomic and hormonal stress responses\(^11\) and reducing the risk of overtraining.\(^12,13\)

For recreational athletes performing a much smaller total training volume (ie, 3–5 h/wk) it is unknown what intensity distribution is optimal or if intensity distribution is critical at all. An argument can be made for recreational athletes to perform more training at or above the lactate threshold. This intensity might maximize adaptive signaling given the limited total stimulus. Overtraining is not likely given the greater recovery time when training is
much less frequent. At the same time, a polarized training approach may also be beneficial for recreational athletes by emphasizing the avoidance of a monotone intensity distribution and keeping “easy training easy and hard training hard.” Foster et al\textsuperscript{14} previously suggested that among athletes, it is common for low-intensity sessions to be performed harder than prescribed and higher-intensity sessions to be performed at a lower than prescribed intensity. This is also a likely scenario for time-limited recreational athletes.

The goal of this study was therefore to compare the performance impact of 2 models of training-intensity distribution in a group of recreational endurance athletes training only about 4 h/wk. We hypothesized that a PET model would elicit greater performance improvements than a BThET model.

Methods

Experimental Approach to the Problem

A 2-group pretest–posttest quasi-experimental design was used. Different training programs assigned to each group were considered the independent variable. Competitive performance on two 10-km races was assessed before and after the experimental period and served as the dependent variable. One group of athletes performed a relatively higher percentage of their total training volume in zone 1, below their ventilatory threshold (VT). The second group trained relatively more in zone 2, between VT and respiratory-compensation threshold (RCT), while training less in zone 1. Both training programs were equal in volume in zone 3 (ie, intensities ≥85% $V_{O_{2\text{max}}}$), time (10 weeks), and load. To ensure that total training loads (ie, volume × intensity) were similar in both study groups, we used a modified version of the training-impulse (TRIMP) approach.\textsuperscript{15}

Subjects

Thirty-two recreational Spaniard runners (mean competition experience ≥5.5 yr) were initially recruited to participate in this study. They were regularly participating in 10K and half-marathon events. All subjects lived and trained in the area around Madrid, Spain (~600-m altitude).

Only the data of the subjects who met the following conditions were included: (1) completion of at least 98% of all the planned training sessions, (2) complete heart-rate (HR) recordings of each training session over the total training period, (3) performing regular training session under the supervision of one of the authors (J E-L), (4) showing no signs or symptoms of overtraining over the entire training period,\textsuperscript{7,13} and (5) performing the two 10-km races before and after the intervention training period.

Runners were randomly assigned to 2 different training groups (each $n = 16$) for a 10-week period, following a training program with increased contribution of zone 1 and increased contribution of zone 2 (BThET group), relative to the normal training pattern observed in this population.\textsuperscript{7} The institutional research ethics committee (European University of Madrid) approved the study, and the subjects provided informed consent before participation.

Training and Periodization

The training plan of 1 of the groups (PET) was designed to achieve a total percentage distribution in zones 1, 2, and 3 of ~75/5/20 based on HR distribution. The other group (BThET) followed a training plan designed to achieve a total percentage distribution in zones 1, 2, and 3 of ~45/35/20. The 2 training programs were designed to reach a similar score in the 2 groups for both total TRIMP accumulated over the 10-week macrocycle (~3500 TRIMPs) and mean TRIMP accumulated each week (mean of ~350 TRIMPs/wk; Figure 1).

Daily training loads were based on time goals rather than distance, with the intent of controlling the relative time in each zone for each athlete. All runners shared the same coach (J E-L) and training locations. They trained over different surfaces, alternating dirt and grass tracks, urban roads, and synthetic track. The tracks where these athletes trained also had several rolling paths, so time in zone, rather than distance, was set as the main variable during these workouts.

In both PET and BThET, the whole macrocycle was divided into a common preparatory 8-week period, followed by a specific 6-week period and a final competitive 4-week period.

Before initiation of the training intervention, all the recruited runners performed the same initial 8-week program of training, with 100% zone 1 training during weeks 1 to 3, progressing from 130 to 190 TRIMPs. From week 4 to week 7, they progressed from 88/5/7% distribution in week 4 to 54/27/19% in week 6 and from 269 to 347 TRIMPs. In the last preparatory week, TRIMP was reduced to 276, with a 78/14/8% distribution, allowing a good recovery. Pretraining physiological testing was performed at the end of this 8-week baseline period.

The intervention was conducted during the specific and competitive period (total 10 wk). The specific period had two 3-week mesocycles following a 2:1 load structure (ie, 2 weeks of high load followed by an “easy” week). Each 4-week mesocycle had a 3:1 load structure (preparatory and competitive periods).

Running distance averaged ~50 km/wk in both groups over the study, increasing through the specific period to reach a maximum of ~70 km/wk in the 13th week and finally decreasing to a mean of 30 km/wk before the posttraining 10K race (week 18). The weekly scheduled program included 2 “hard” sessions per week (interval or repetition workouts at high intensities) and 1 or 2 strength-training sessions per week. The remaining sessions were composed of continuous training (to be performed in zone 1 for the PET group or zone 2 for the BThET group). Both groups averaged the same training frequency (5–6 d/wk, according to their performance and time availability).
Strength Training During the Study Period
Strength training was identical for all subjects and was not related to the experimental design. During the initial 8-week period, the runners performed 2 d/wk isometric and own-body-weight plus medicine-ball circuit training and core and proprioceptive exercises.

During the specific period (weeks 9–14), short to long running intervals with weighted belts (3–5% of body weight) were performed 1 d/wk. Low-intensity plyometrics of 150 to 300 total jumps per session were also performed 1 d/wk.

During the competition period (weeks 15–18), subjects performed strength training only once a week.

Baseline Laboratory Testing and Performance Test
Subjects reported to the laboratory (~600-m altitude) at the beginning of the training period to perform a physiological (ramp) test on a treadmill (Technogym Run Race 1400 HC, Gambettola, Italy) for VT and RCT determination. After a general warm-up, starting at 8 km/h, running velocity was increased by 0.5 km/h every 30 seconds until volitional exhaustion. During the tests, gas-exchange data were collected continuously using an automated breath-by-breath system (Vmax 29C, Sensormedics, Yorba Linda, CA, USA). The following variables were measured: oxygen uptake (VO₂), pulmonary ventilation, ventilatory equivalents for oxygen and carbon dioxide, and end-tidal partial pressure of oxygen and carbon dioxide.

VO₂max was recorded as the highest VO₂ value obtained for any continuous 1-minute period during the tests, and typical determination criteria were used. The VT and RCT were determined using the criteria previously described. HR (beats/min) was continuously monitored during the tests (Accurex Plus, Polar Electro Oy, Finland).

At the end of the 8-week preparatory period, all runners performed the same 10-km race (pretraining 10K) held in Alcalá de Henares, Madrid. This competition was used (1) to determine initial performance level and to ensure baseline performance levels in both groups before the start of the study and (2) to compare the magnitude of changes in performance in both groups over the training period. It was preceded by the typical precompetitive rest period (ie, 3–4 d of “easy” training). At the end of the competitive period (week 18) subjects performed another 10K race all together (posttraining 10K) in Aranjuez, Madrid. It was not possible to perform physiological testing at the end of the 10-week intervention period.

Figure 1 — Training load and intensity distribution between groups. Abbreviations: Z, zone; PET, polarized endurance-training distribution; BThET, between-thresholds endurance-training.
Quantification of Exercise Load in Training

HR was measured continuously for all subjects during each training session over the macrocycle. The following variables were quantified: total time spent in each intensity zone (zone 1, HR below the VT; zone 2, HR between VT and RCT; zone 3, HR above RCT) and total load (TRIMP score) as explained hereafter. Previous research on trained endurance athletes has shown that HR values at VT and RCT determined during laboratory testing remain stable over the season despite significant improvements in the workload eliciting both thresholds.16 Thus, a single test performed during the preparatory period (as used here) appears valid for training monitoring based solely on target HR values at VT and RCT.16

We estimated total exercise load (ie, intensity × volume) using an approach to calculating the TRIMP based on Foster et al.15 This method, which has been used to estimate total exercise load in 3-week professional cycling races17,18 and regular training of well-trained endurance runners,7,19 uses HR data during exercise to integrate both total volume and intensity relative to 3 intensity zones. The score for each zone is computed by multiplying the accumulated duration in this zone by an intensity-weighted multiplier (eg, 1 min in zone 1 is given a score of 1, 1 min in zone 2 is given a score of 2, and 1 min in zone 3 is given a score of 3). Total TRIMP load is then obtained by summing the 3 zone scores.

Statistical Analyses

To ensure that the fitness and competition levels of both groups were similar at baseline, mean values of all the variables indicative of fitness levels (VO2max, VT, and RCT, etc) and performance (pretraining 10K) were compared between groups using an independent-samples t test. To ensure that total training load (volume × intensity) and distribution in intensity zones were similar and different, respectively, in the 2 groups during the training period, mean values of total TRIMP score; total time spent in zones 1, 2, and 3; and percentage of time spent in zones 1, 2, and 3 (over the 10-wk intervention period) were also compared. A 2 × 2 mixed ANOVA was performed (comparing the 2 groups as between-subjects factor and pre–post measures as within-subject factor) to assess the differences in posttraining 10K. Cohen d was used as a measure of effect size, using the reference values of small (d = 0.2), medium (d = 0.5) and large (d = 0.8) for interpreting them as suggested by Cohen.20 Eta-squared (η2) was also used as a measure of effect size, considering small effect sizes ≥0.01, moderate around 0.06, and large ≥0.14.20 The Bonferroni method was used for all ANOVA pairwise comparisons. In addition, qualitative analysis of confidence intervals of differences was performed using the spreadsheet from Hopkins.21 A threshold of 1.3% of improvement (equivalent to 30 s) was set for these inferences.

Descriptive data are reported as mean ± SD, and the level of significance was set at P ≤ .05 for all statistical analyses. A subset analysis was conducted with the athletes whose training-intensity distribution was most highly PET oriented and most highly BThET oriented.

Results

Baseline Laboratory Tests and Pretraining 10K

Two subjects (1 from each intervention group) were excluded from analysis due to incomplete training-data recordings, leaving 15 subjects in each group. There were no significant differences in age, weight, height, or body-mass index between the 2 groups. Furthermore, VO2max, VT and RCT; zone 3, HR above RCT) and total load (TRIMP score) as explained hereafter. Previous research on trained endurance athletes has shown that HR values at VT and RCT determined during laboratory testing remain stable over the season despite significant improvements in the workload eliciting both thresholds.16 Thus, a single test performed during the preparatory period (as used here) appears valid for training monitoring based solely on target HR values at VT and RCT.16

No significant differences were found between groups in total TRIMP score or mean weekly TRIMP score, although the BThET group trained descriptively more TRIMPs than the PET group (Table 2).

As prescribed, significant differences were found between groups for total time in zone 1 (F1,28 = 26.87,
Polarized Training in Recreational Runners

Significant performance improvements from the pretraining 10K to the postraining 10K were observed in both PET (39min18s ± 4min54s vs 37min19s ± 4min42s, 5.0% ± 3.3%, P < .0001) and BThET (39min24s ± 3min54s vs 38min0s ± 4min24s, 3.5 ± 3.0%, P < .001; Figure 2).

These differences were not significant between groups in the postraining 10K (F1,28 = 1.1516, P = .226, η2 = .051). Expressed in standardized Cohen effect-size units,20 the difference in improvement between groups was 0.43 (90% CI –0.17 to 1.04, P = .23). Figure 3 represents individual improvements per group.

The qualitative analysis of these differences21 showed that PET has a 43% probability of being linked to a beneficial or substantially positive effect over BThET and a 66% probability of being linked to a negligible or trivial effect.

Due to variation in the actual execution of the training plans, we also identified the subsets of athletes in each group whose training-intensity distribution was most highly zone 1 oriented and most highly zone 2 oriented. Six subjects from PET were identified whose mean intensity distribution was 78% zone 1, 11% zone 2, and 11% zone 3 during the intervention period. Their mean 10K improvement was 7.0% ± 3.6%. Six subjects from BThET were identified with an intensity distribution of 32% zone 1, 53% zone 2, and 16% zone 3. Their mean 10K improvement was 3.5% ± 3.0%.

Table 2 Results of Training Load Over the 10-Week Intervention Period, Mean ± SD

<table>
<thead>
<tr>
<th>Group</th>
<th>PET (n = 15)</th>
<th>BThET (n = 15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total running time (h)</td>
<td>39.1 ± 7.9</td>
<td>36.3 ± 8.1</td>
</tr>
<tr>
<td>Total time in zone 1 (h)</td>
<td>28.5 ± 6.3*</td>
<td>16.7 ± 6.2</td>
</tr>
<tr>
<td>Total time in zone 2 (h)</td>
<td>5.3 ± 2.7*</td>
<td>13.9 ± 8.8</td>
</tr>
<tr>
<td>Total time in zone 3 (h)</td>
<td>5.3 ± 1.7</td>
<td>5.6 ± 1.6</td>
</tr>
<tr>
<td>Total time in zone 1 (%)</td>
<td>72.9 ± 5.6*</td>
<td>46.8 ± 15.2</td>
</tr>
<tr>
<td>Total time in zone 2 (%)</td>
<td>13.5 ± 5.6*</td>
<td>37.3 ± 16.1</td>
</tr>
<tr>
<td>Total time in zone 3 (%)</td>
<td>13.6 ± 4.3</td>
<td>15.8 ± 4.1</td>
</tr>
<tr>
<td>Total TRIMPs</td>
<td>3299 ± 670</td>
<td>3691 ± 982</td>
</tr>
<tr>
<td>Mean TRIMPs/wk</td>
<td>330 ± 67</td>
<td>370 ± 98</td>
</tr>
</tbody>
</table>

Abbreviations: PET, polarized endurance-training distribution; BThET, between-thresholds endurance-training; TRIMP, training impulse.

*P < .05.

Figure 2 — Performance time changes (pre-training-intervention 10K vs post-training-intervention 10K in PET and BThET). Abbreviations: PET, polarized endurance-training distribution; BThET, between-thresholds endurance-training.
was 1.6% ± 4%. Expressed in standardized Cohen effect-size units, the difference in improvement between the subgroups of runners training most in zone 1 and those training most in zone 2 was 1.29 (90% CI 0.31–2.27, \( P = .038 \); Figure 4). These 2 subgroups did not differ in training experience (8 ± 3 vs 6 ± 5 y), preintervention 10K time (42 ± 6 vs 41 ± 3 min), or total training time during the intervention (37 ± 9 vs 37 ± 11 h).

**Figure 3** — Percentage improvement in performance (10K PRE vs 10K POST in PET and BThET). Abbreviations: PET, polarized endurance-training distribution; BThET, between-thresholds endurance-training; PRE, pre-training intervention; POST, post-training intervention.

**Figure 4** — Performance improvement (pre-training-intervention 10K vs post-training-intervention 10K in PET and BThET, subset analysis with extreme distribution cases). Abbreviations: PET, polarized endurance-training distribution; BThET, between-thresholds endurance-training.
Discussion

The key finding of the current study was that both between-thresholds-emphasis training and training with greater emphasis on a polarized intensity distribution over 10 weeks resulted in significant performance improvements in a 10K performance test. Mean improvements in the 2 groups were 3.5% in BThET and 5.0% in PET, or 84 and 119 seconds, respectively. This improvement was similar to other studies about performance in 10K runners.22,23 Given the high standard deviation, there were no significant differences between groups. The Hopkins' qualitative analysis is consistent with the conclusion that there is not enough evidence in the overall findings to support one approach over the other.

However, the magnitude of effect size, percentage of improvement, and eta-squared showed a superior effect in zone 1 for the PET group.

Many runners (from both groups) were not strict about prescribed intensity distribution. When those who trained clearly according to the specific intensity prescriptions were compared, the PET group improved significantly more than the BThET group. On the other hand, despite outcome differences, PET has been suggested to be easier since less distance might be covered and therefore this is less probability of injury,24 and this might cause greater adherence.25 However, according to our data (not shown), estimated running distance during the intervention period was not different between groups (~480 total km/group). This was probably because, although not significant, BThET average training load was higher (mean ~370 TRIMPs/wk vs ~330 in PET). Taken all together, the 5% vs 3.5% difference in overall improvement (about ~35 s difference during an ~40-min race) could be meaningful from a performance point of view.

With a statistical power of .80 and an α of .05, the sample size required for repeated-measures within-between interaction should be 40 subjects in each group (total sample 80) to reach a significant difference in performance times. The high standard deviation found in the current study suggests that individual variation to training is substantial, at least for this performance level. Further research is needed to determine what extent it is basically a question of sample size or different responses in less-trained athletes. Moreover, better control of workloads is needed in future studies with equated total loads and accurate daily distance recordings. Other studies in the field on endurance training have shown no differences when comparing different treatments (like continuous vs interval training) with low-experienced runners with equated loads.26

A recent study27 observed superior performance effects of polarized training in a group of cyclists with similar level to the current study’s runners. This study was well controlled and even more strongly emphasized the difference in intensity distribution between the 2 groups by eliminating all training above the RCT (zone 3) in their between-thresholds group. While there is strong agreement that successful endurance athletes characteristically perform about 80% of their training sessions at subthreshold intensity, the distribution of training intensity in the remaining 20% remains a topic of debate. However, how intensity and accumulated duration interact to drive physiological adaptation has only recently been investigated.28

Another limitation of these studies is that habitual off-training-workout daily physical activity was not considered. As shown by Hautala et al29 with low-fit subjects, a high amount of habitual light-intensity physical activity might be associated with a superior training response. Finally, we did not consider the quantification of strength training, which was the same for both groups, but it may have some particular impact in conjunction with aerobic-training-intensity distribution as has been discussed elsewhere.30

The current study is the third one to test the polarized training approach with an experimental design. Previous data were collected with an initial sample of 20 highly trained runners, only 12 of whom met the inclusion criteria.8 Recent cited research was conducted with 12 cyclists in a crossover fashion.27 Ensuring accurate and complete daily recordings from many subjects, plus the challenges of supervising a large sample for the same competition goals, makes it difficult to increase the sample size from this level.

In case of spending larger amounts of time to train, polarized distribution may prevent overtraining or diminishing returns in performance.11 This may be the independent training effect of subthreshold-intensity training on the lactate profile.8 In addition, it may improve the quality of the higher-intensity sessions by preventing fatigue and staleness.1-4,11 However, in this study we do not have physiological data at posttest to evaluate these possible explanations. We just found that the difference between those who trained with the most and the least polarized distribution showed a big effect size. Further research is needed to link the growing descriptive and experimental support for a polarized training-intensity distribution to underlying physiological stress-adaptation mechanisms.36

Practical Application and Conclusions

A greater focus on between-thresholds training is not associated with greater performance improvement in runners despite a low weekly training volume. A polarized training approach tended to stimulate greater performance enhancement, but this difference was not statistically significant due to substantial variation in both training responses and adherence to the specific training prescriptions. These 2 sources of variation are also important in a practical setting.

References
