Submission Type: Original Investigation

Title of the article

Relationships between internal and external match load indicators in soccer match officials

Full Names of the Authors:

Daniel Castillo\textsuperscript{1}, Matthew Weston\textsuperscript{2}, Shaun J. McLaren\textsuperscript{2}, Jesús Cámara\textsuperscript{1}, Javier Yanci\textsuperscript{1}

Institutional/Corporate Affiliations:

\textsuperscript{1}Faculty of Education and Sport, University of the Basque Country, UPV/EHU, Vitoria-Gasteiz, Spain, and \textsuperscript{2}Department of Sport & Exercise Sciences, School of Social Sciences, Business & Law, Teesside University, Middlesbrough, United Kingdom.

Running Head: Match loads in soccer referees.

Contact Details for the Corresponding Author:

Javier Yanci, PhD
Faculty of Education and Sport, University of the Basque Country, UPV/EHU, Lasarte 71, 01007, Vitoria-Gasteiz, Spain
E-mail address: javier.yanci@ehu.es

Manuscript word count: 4989
Abstract Word Count: 251
Number of Figures and Tables: 3
The aims of this study were to describe the internal and external match load (ML) of refereeing activity during official matches and also to investigate the relationship among the methods of ML quantification across a competitive soccer season. A further aim was to examine the usefulness of differential perceived exertion (dRPE) as a tool for monitoring internal ML in soccer referees. Twenty field referees (FR) and 43 assistant referees (AR) participated in this study. Data were collected from 30 competitive matches (FR = 20 observations, AR = 43 observations) and included measures of internal (Edwards’ heart rate derived training impulse [TRIMP_{EDW}]), external (total distance covered [TD], distance covered at high speeds [HSR] and player load [PL]) ML, differentiated ratings of perceived respiratory [sRPE_{res} ML] and leg muscle [sRPE_{mus} ML] exertion. Internal and external ML were all greater for FR when compared to AR (-19.7 to -72.5); with differences ranging from very likely very large to most likely extremely large. The relationships between internal ML and external ML indicators were, in most cases, unclear for FR (r < .35) and small to moderate for AR (r < .40). We found substantial differences between RPE_{res} and RPE_{mus} scores in both FR (.6 AU; ±90% confidence limits .4 AU) and AR (.4; ±.3 AU). These data demonstrate the multifaceted demands of soccer refereeing and thereby highlight the importance of monitoring both internal and external ML. Moreover, dRPE represent distinct dimensions of effort and may be useful in monitoring soccer referees ML during official matches.

**Key words:** perceived exertion, heart rate, training load, referee, GPS.
Quantifying the physical and physiological loads imposed by specific training drills and competition is important to understand the dose-response nature of the training process, with regards to optimizing the performance of athletes. An accurate and detailed understanding of competition demands can provide sport scientists and practitioners with an objective framework to prescribe the optimum training dose. Training loads (TL) and match loads (ML) may be expressed in terms of both external (physical demands, such as total distance covered, distance at certain velocities, accelerations, etc.) and internal (physiological demands, such as heart rate [HR] and ratings of perceived exertion [RPE]) components. Indeed, these ML indicators have been extensively analyzed using in both soccer players and in match officials.

As a result of recent developments in microsensor technology, some authors have suggested that player load (PL) - a vector magnitude representing the sum of accelerations recorded in the three principal axes of movement - could be a more suitable measure of external ML than locomotive demands alone, which neglect both energetically taxing changes in speed and the three-dimensional nature of movement and impacts typical to soccer players and officials. Likewise, while RPE represent a practical and valid measure of internal load, differential RPE (i.e. central ['respiratory': sRPE_res] and peripheral ['muscular': sRPE_mus] exertion) have gained recent attention within the team sport literature as measures which may improve the accuracy and sensitivity of internal load measurement by discriminating global perceived exertion into its specific physiological mediators. Furthermore, these subjective measures may be useful to sport scientists as they are inexpensive, accessible at all levels and are not prohibited by the rules of competition. While dRPE and PL have the potential to enhance the monitoring of internal and external loads during intermittent, stochastic activities such as team sport competition, there is no literature available to date which quantifies these measures in soccer referees during official matches. This information could provide unique and novel insights into the specific physical and physiological demands of match officials during competitive fixtures.

Knowledge of the relationships between internal and external ML permits for a better understanding of the dose-response nature of training and competition. Weston et al. observed a moderate association between HR and RPE in field referees (FR, r = .49), while Costa et al. observed small to moderate correlations between total distance covered and internal load measures (Edwards’ HR-derived training impulse [TRIMP_EDW], r = .22 and session-RPE [sRPE] TL, r = .38). Despite this, only a few studies have examined the internal-external ML relationships in assistant referees (AR). Given the recent development and use of novel measures of internal (i.e. sRPE_res and sRPE_mus) and external (i.e. PL) ML, the relationships between these variables and also traditional ML measures are of interest. While an examination of such may further advocate the criterion-related validity of dRPE and PL as useful monitoring tools in team sport players and match officials, this information is also likely to be useful to those responsible for the programming, monitoring and evaluation of TL in team sport match officials.

Therefore, the main purposes of this study were to describe internal and external match load of refereeing activity during official matches and to also investigate the relationship among the methods of match load quantification across a competitive
soccer season on match officials. A further aim was to examine the usefulness of dRPE as a tool for monitoring internal match loads in soccer referees.

**Methods**

**Participants**

Sixty-three soccer match officials who officiated in thirty soccer matches of the Spanish National Third Division across the 2014–15 competitive soccer season participated in this study. Match officials had at least ten years of officiating experience, with a minimum of six years at this particular level of competition. Of the 63 match officials, 20 were FR (age: 27.70 ± 6.20 yr, stature: 177.63 ± 6.74 cm, body mass: 74.07 ± 8.54 kg, BMI: 23.46 ± 2.18 kg·m$^{-2}$) and 43 were AR (age: 30.68 ± 9.60 yr; stature: 176.15 ± 5.62 cm; body mass: 75.05 ± 7.81 kg; BMI: 24.21 ± 2.51 kg·m$^{-2}$). All match officials trained at least three times a week and were involved in refereeing on average three times per month. This investigation was performed in accordance to the Declaration of Helsinki and was approved by the Ethics Committee of the University of the Basque Country (UPV/EHU).

**Design**

We used an observational design to examine the relationships between internal and external match load indicators in match officials. Data were collected from 30 competitive matches (FR = 20 observations, AR = 43 observations) and included measures of internal (TRIMP$_{EDW}$, sRPE$_{res}$ ML, sRPE$_{mus}$ ML) and external (total distance covered [TD], distance covered at high speeds [HSR] and PL) ML. Prior to the start of each match, the match officials performed a standardized 15 minutes warm-up including running, progressive sprints and stretching. However, this data was not included in the overall analysis.

**Internal Loads**

To quantify TRIMP$_{EDW}$, match officials’ HR was recorded continuously during the matches (Polar Team System™, Kempele, Finland) at 5 s intervals. HR during the 15 min half-time period was excluded from the analysis. Intensities of effort were subsequently calculated and expressed as percentages of each match official known maximal heart rate (HR$_{max}$) obtained during the match $^{26}$. The total time (min) spent in 5 arbitrary intensity zones was summated and multiplied by a specific weighing factor. These were: 1 for 50–60% HR$_{max}$, 2 for 60–70% HR$_{max}$, 3 for 70–80% HR$_{max}$, 4 for 80–90% HR$_{max}$ and 5 for 90–100% HR$_{max}$. The sum all 5 intensity zones represented TRIMP$_{EDW}^{29}$.

Using the CR10 scale, match officials provided differentiated ratings for their perceived respiratory (i.e. breathlessness; sRPE$_{res}$) and leg muscle (sRPE$_{mus}$) exertion $^{4}$. To calculate the RPE-derived ML, each score was multiplied by the match duration (min) as per Foster et al. $^{30}$. Match officials were fully habituated with the RPE procedures and scaling methods prior to this investigation.

**External Loads**
Referees’ match activities were monitored using microsensor units containing a 10 Hz global positioning system (GPS) and a 100 Hz triaxial accelerometer (MinimaxX v4.0, Catapult Innovations™, Melbourne, Australia). Microsensor units were harnessed in a tight-fit vest which was worn by the match officials throughout the games. The microsensor devices were activated 15 min prior to the start of each match, in accordance with the manufacturer’s recommendations. Data were downloaded post-match to a PC and analysed using a customized software package (Logan Plus v.4.4, Catapult Innovations™) [19]. We used TD (m) and HSR (> 13 km·h⁻¹) distance (m) recorded from the GPS within the microsensor units as our indicators of running-based external MLs [28]. Additionally, PL was computed as vector magnitude representing the sum of accelerations recorded in the anterior-posterior, mediolateral and vertical planes of movement, measured by the microsensor units’ 100 Hz tri-axial piezoelectric linear (Kionix: KXP94). The reliability and validity of these microsensor units for the measurement of TD, HSR and PL are reported elsewhere [31,32].

Data analysis

Results are presented as means ± standard deviations (SD). Prior to analyses, plots of the residuals versus the predicted values of all variables revealed no clear evidence of non-uniformity of error. To compare the differences in internal and external ML between FR and AR, a magnitude-based inference approach was used [33]. Data were log transformed and subsequently back transformed to represent the between-referee differences in ML’ as accurate percentages. Standardized thresholds of .2, .6, 1.2, 2.0 and 4.0 multiplied by the pooled between-referee SD were used to anchor small, moderate, large, very large and extremely large differences, respectively. Uncertainty in the estimates was then calculated based on the disposition of the 90% confidence limits (CL) for the respective mean difference in the relation to the standardized thresholds. The probability (percent chances) that the true between-referee differences in internal and external ML were the observed magnitude were then qualified via the following probabilistic terms: 25–75%, possibly; 75–95%, likely; 95–99.5%, very likely; >99.5%, most likely [33]. Inferences were classified as unclear if the 90% CL overlapped the thresholds for both substantially positive and negative thresholds by ≥5%. Between-subject correlations were calculated to examine the relationships between internal and external ML. For referees with repeated match samples, the mean value for each ML variable was used in replacement of the original data (n = 20, range = 2–4 matches). The following scale of magnitudes was used to interpret the correlation coefficients: <0.1, trivial; .1–0.3, small; .3–.5, moderate; .5–.7, large; .7–.9, very large; >.9, nearly perfect [33]. Confidence limits (90%) for the correlations were constructed using a bias corrected accelerated bootstrapping technique of 2000 samples with replacement from the original data (SPSS™ v.21, Armonk, NY: IBM Corp.). Magnitude-based inferences were subsequently applied to qualify the uncertainty in the correlation estimates, using the method previously described [33].

Results

The FR’ and AR’ internal and external MLs are presented in Table 1. Internal and external ML were all greater for FR when compared to AR, with differences ranging from very likely very large to most likely extremely large. Analysis of match sRPEₘus and sRPEₙ₃₃ scores revealed that for the FR, the difference between RPEₘus (7.1 ± 1.1 AU) and RPEₙ₃₃ (6.6 ± 1.1 AU) was likely small/possibly moderate (.6; ±90%
confidence limits .4 AU). For AR, the difference between RPE_mus (4.2 ± 1.5 AU) and 
RPE_res (3.8 ± 1.3 AU) was likely small (.4; ±.3 AU).

*** Table 1 approximately here ***

The relationships amongst internal and external MLs for FR and AR are presented in 
Tables 2 and 3, respectively. For FR, the relationships between internal and external 
load measures ranged from unclear to possibly moderate, while the relationships 
amongst internal and external load measures ranged from unclear to possibly very large 
(Table 2). For AR, the relationships between internal and external load measures ranged 
from unclear to likely moderate, while the relationships amongst internal and external 
load measures ranged from unclear to likely very large and likely large to very likely 
very large, respectively (Table 3).

*** Table 2 approximately here ***

*** Table 3 approximately here ***

Discussion

The aims of this study were to describe the match loads (ML) of soccer field and 
assistant referees across a competitive season of official matches and also to investigate 
the relationships between methods of internal and external ML quantification. A further 
aim was to examine the usefulness of differential ratings of perceived exertion (dRPE) 
as a tool for monitoring internal ML in soccer referees. The results of our study showed 
that, a) FR attain considerably higher internal and external MLs when compared with 
AR, b) the relationships between internal ML and external ML indicators were, in most 
cases, unclear for FR and small to moderate for AR, and c) dRPE represent distinct 
dimensions of effort in soccer referees during official matches.

Given the different roles undertaken by FR and AR during match play, and considering 
that assistant refereeing is limited to half of the length of the field, external ML 
performed by AR represents approximately half of the external ML performed by FR 34. 
Resultantly, AR also incur substantially lower internal ML when compared with FR 34. 
These notions are in agreement with our current data, which shows that internal and 
external ML were ~20–40% and ~50–70% lower, respectively, in AR when compared 
with FR. Others have reported total match distances of ~10,000 and ~5,000 m for FR 
and AR, respectively, across various levels of soccer competition 16,35. Likewise, 
Krustrup et al. 36 noted that both TD covered (FR, 10,270 ± 900 vs. AR, 6,760 ± 830 m) 
and distance covered above 18 km·h⁻¹ (FR, 1,920 ± 580 vs. AR, 970 ± 520 m) were 
more than double for FR when compared with AR. Regarding internal ML, the typical 
match intensity is greater for FR (85–90% HR_max) when compared with AR (77–79% 
HR_max) 12,35.

A unique aspect of the current study was the ability to quantify novel methods of 
internal and external ML indicators (i.e. dRPE and PL, respectively) in soccer referees 
during official matches. Differential RPE provide information on the perceived central 
(respiratory) and peripheral (leg muscle) internal ML 4,9,21,22, while PL represents the 
sum of external load incurred from multiplanar activities such as running (footfalls), 
acceleration/decelerations, changes of direction, and impacts to name a few 18,32. Our 
data again show that FR incurs greater PL and report greater dRPE when compared with
their AR counterparts. Taken together, these data support and add to the literature surrounding the demands of soccer match officials during competition. Knowledge of these different internal and external match responses could help inform the planning and progression of appropriate in-season training loads designed to prepare match officials for the physical and physiological requirements of competition.

Examination of the relationships between internal and external ML may help physical trainers of soccer referees know whether both ML methods are necessary to quantify match demands or use only one method is enough to quantify and organize the appropriate training doses, based on the desired training responses that are specific to match demands. The results of our investigation are in agreement with others, who have typically reported unclear/trivial through to moderate correlations between internal ML and intensity with external ML indicators in soccer referees. Costa et al. observed small and moderate associations between TD covered and both TRIMP_{EDW} (r = .22) and sRPE ML (r = .38) in Brazilian FR. Catteral et al. reported a trivial correlation (r = .15) between TD and mean %HR_{max} in professional FR, although Mallo et al. reported a moderate association (r = .50) between mean %HR_{max} and the time spent running at high speeds (>18 km·h^{-1}) in international FR. Likewise, moderate relationships (r = .31) have also been observed in international AR between mean %HR_{max} and the total number of high-intensity activities (>13 km·h^{-1}). It is likely that the associations between internal and external ML could be moderated by factors such as the individual fitness level of the referee and also acute physiological stress incurred as a result of physical (i.e., recent training, nutrition, etc.) and social (i.e., travel, sleep, etc.) factors. This may be one explanation for the typically low (unclear to moderate) correlations observed in our current investigation and within the work of others.

Due to associations between internal and external load measures were ranged from unclear to possibly moderate in our study, it seems that these constructs measure distinctly different match demands. We therefore recommend concurrent measures of match internal and external loads to help fully understand the true dose-response of referees’ during team-sports matches.

In line with the aims of our investigation, we chose to explore the associations between measures of internal and external ML only, rather than measures of internal intensity (i.e., sRPE, mean %HR_{max}, blood lactate concentration) and external ML. We feel that the latter may be conceptually unsound, given that measures of training and match load encompasses both the intensity and volume of the session. Consequently, the calculation of ML indicators (i.e., sRPE_{res} ML, sRPE_{mus} ML, TRIMP_{EDW}) provides a more robust index for investigation rather than intensity alone. Nonetheless, the work of others coincides with those results reported in our study, in which the relationships between internal and external ML indicators were typically more prominent in AR when compared with FR. The physical and physiological demands of a match are very different for FR and AR due to their disparate roles taken on the field. These findings may therefore be explained by the relatively short (one half of the field) and linear running patterns of AR in comparison with the stochastic and multi-directional movements of FR. The latter is likely to induce more variable match demands and associated internal responses, which could have mitigated the magnitude of the relationships between internal and external ML.

In our investigation, we chose not to pool our sample of match officials due to the very large / extremely large differences in internal and external ML between these two
groups. When concentrating on a more homogeneous subset of match officials (i.e. FR and AR), the strengths of relationships between internal and external ML are likely to be much lower than a pooled analysis which may result in spuriously high correlations that are only useful for confirming already obvious between-group differences. We acknowledge that our study involved a relatively small sample size, particularly for FR (n = 20), and our analysis of the relationships between internal and external ML was therefore restricted to a between-referee comparison. To determine if higher internal ML loads are associated with higher external ML, a within-subject design is the appropriate method as it permits the analysis of within-subject changes by removing between-subject differences. We therefore recommend future work in this area to utilize larger sample sizes and different competitive levels (i.e. elite referees) involving several repeated measures per referee, as well as examining the factors that may reasonably moderate the relationships between internal and external match loads, such as individual referee characteristics (e.g. physical fitness and acute physiological stress) and match-related contextual variables.

This is the first study in which dRPE have been collected on professional soccer referees to quantify internal ML. In our study, RPE_res and RPE_mus scores were in the range of 6-7 (‘very hard’). These ratings are typically lower than global RPE reported in elite soccer referees and may explained by differences in competition standard. A key finding of our investigation was the substantial differences observed between RPE_res and RPE_mus scores in both FR and AR. Match official perceived their leg muscle exertion to be greater than respiratory exertion - a finding consistent with soccer and Australian Football players. The results of our correlation analysis also suggest that there remains approximately 40% unexplained variance between RPE_res and RPE_mus during official competition. Taken together, these data indicate that while RPE_res and RPE_mus may not be mutually exclusive, dRPE do represent distinct internal constructs that are perceived differently by sub-elite soccer match officials. The very large correlation observed between RPE_res ML and RPE_mus ML is not surprising given that the augmentation of central and peripheral exertion during exercise is closely related, particularly during high-intensity intermittent activities. The substantial differences in the magnitudes of the relationships between RPE_res and RPE_mus with external ML indicate that these measures may each be influenced by dissimilar external loads. In agreement with others, we therefore believe our data supports the notion that dRPE represent a worthwhile addition to the monitoring of ML in soccer referees. Disassociations between RPE_res and RPE_mus may help assist in the monitoring and planning of training loads by informing individualized training or post-match recovery strategies, although such ideas warrant further investigation in both sub-elite and elite soccer match officials. Consequently, we encourage the collection of these measures in both future practice and research surrounding team-sport match officials.

Conclusions

Field referees attain considerably higher internal and external MLs when compared with AR during official competition, suggesting that the planning and progression of training activities should be different for these two groups. We found that the relationships between internal and external ML indicators were, in most cases, unclear for field referees and small to moderate for assistant referees, suggesting that these two factors are somewhat independent of one another in sub-elite referees. Finally, dRPE represent distinct dimensions of effort perception in soccer referees during official matches.
Practical Applications

Considering that FR covered almost twice total and high speed running (>13km·h\(^{-1}\)) distance, and registered higher internal loads (i.e. sRPE\(_{res}\) ML, sRPE\(_{mus}\) ML, TRIMP\(_{EDW}\)) than AR, we suggest that FR and AR should undertake different training regimes not only in relation to prescription training activities but also to overall training volume. Our data also highlights the importance monitoring both internal and external loads during matches and training to help manage workloads and prescribe appropriate training and recovery activities. Differential RPE could be a useful addition to the monitoring and programming of soccer referees’ training loads.

Acknowledgments

The authors would like to thank the match officials of the Comité Navarro de Árbitros de Fútbol (CNAF) for facilitating data collection and for the opportunity to carry out this investigation. This study was supported by the Basque Country Government for doctoral Research.

References


Legend of Tables and Figures

Table 1. Practical difference on internal and external match load (ML) between field (FR) and assistant referees (AR).

Table 2. Relationships (r; ±90% CL) between and amongst internal and external match loads for field referees (n = 20)

Table 3. Relationships (r; ±90% CL) between and amongst internal and external match loads for assistant referees (n = 43)