THE EFFECTS OF THE FALSE START DISQUALIFICATION RULES ON THE REACTION TIMES OF ELITE SPRINTERS

Corresponding Author:
Kevin Christopher Brosnan
Department of Mathematics and Statistics,
University of Limerick,
Limerick,
Ireland.
Email: kevin.c.brosnan@ul.ie

Co-Authors:
Dr. Kevin Hayes
Department of Mathematics and Statistics,
University of Limerick,
Limerick,
Ireland.
Email: kevin.hayes@ul.ie

Dr. Andrew J. Harrison
Department of Physical Education and Sports Science,
University of Limerick,
Limerick,
Ireland.
Email: drew.harrison@ul.ie

THE EFFECTS OF THE FALSE START DISQUALIFICATION RULES ON THE REACTION TIMES OF ELITE SPRINTERS

Abstract

The 100 ms ruling for false start disqualification at elite athletic competitions governed by the International Association of Athletics Federations has been in force since the early 1990's. Throughout this period, significant changes have been made to the rules governing the disqualification of an athlete from an event. This paper analyses all World and European Championship reaction time data from 1999 to 2014 to examine the effect of these important rule changes on competition reaction times. The exponentially modified Gaussian distribution is used to model the reaction times of athletes and a comparison relative to sex, ruling periods and competition rounds is provided. Revised reaction time thresholds were calculated for male and female athletes at 115 ms and 119 ms respectively, indicating that the 100 ms ruling is subject to true false starts not being detected in competitive athletics. The conclusions of this study propose a revision of the false start detection threshold is carried out and that male and female athletes are governed independently due to the substantial evidence of a sex difference in reaction times of elite athletes.

Keywords: exponentially modified Gaussian distribution; auditory performance; athletics

Introduction

The reaction times (RTs) of elite 100 m sprinters has been singled out as a possible benchmark against which to gauge the absolute limits of human auditory performance (Lipps, Galecki et al. 2011). Accurately identifying the precise threshold beyond which an elite sprinter can react to the starter's pistol, and detecting occasions in competition where a highly-trained athlete has reacted so fast that (s)he must have breached this lower limit (a false start), are necessary for the fair and impartial refereeing of competitive athletics.

The International Association of Athletics Federations (IAAF) rule 161.2 stipulates that a false start occurs when a sprinter registers a RT less than 100 ms (International Association of Athletics Federations 2015). Prior to January 2004, a RT of less than 100 ms (i.e. a false start) resulted in a warning on a competitor and the competitor was disqualified if a RT of less than 100 ms was recorded on two separate starts. From January 2004, a false start by any competitor placed all the athletes in the race on a first warning. Subsequently any sprinter registering a false start would be automatically disqualified regardless of whether (s)he was the original offender. To eradicate gamesmanship (Athletics Australia 2009) the IAAF made significant changes to the false start disqualification rule in January 2010. Under the revised “No False Start” rule no warnings are issued and a competitor who false starts is automatically disqualified and removed from the race.

The 100 ms detection limit used to detect a false start is measured as the time taken by an athlete to produce a predefined force on the starting blocks. Lipps, Galecki et al. (2011) state that the IAAF justification for this 100 ms threshold is based on a considerably dated study involving eight non-elite Finish male sprinters (Mero and Komi 1990). Pain and Hibbs (2007) suggest that RTs of 85 ms are possible and Brown, Kenwell et al. (2008) present similar results observing 21% of recorded reaction times being below the 100 ms threshold. Whilst such RTs are possible the recording system used in both studies measures the time to the first reaction of the athlete and not the time taken to produce a predefined force. In contrast Komi, Ishikawa et al. (2009) provide a study utilising the existing technology set down by the IAAF, recommending that the IAAF abandon the 100 ms threshold and the measurement technology currently employed. Notwithstanding the results of the aforementioned studies a complete review of the 100 ms threshold requires analysing data recorded under the conditions of the current measurement systems utilised by the IAAF. In addition each of the studies carried out to date rely on a restricted sample size and many do not include athletes of significant international status. A large sample study, incorporating male and female athletes, with athletes of significant international status can provide an appropriate threshold which should be used in competitive athletics.

Previous studies have implied that empirical evidence of a sex difference in the strength of elite athletes (Lipps, Galecki et al. 2011) and of the general population (Der and Deary 2006) exist, regardless IAAF rule 161.2 continues to govern both male and female competitions. Given the physiological differences in strength of men and women it is justified that a sex difference would exist and that the rulings for male and female events should be derived independently.

A revision of the false start detection threshold was suggested in Komi, Ishikawa et al. (2009), where seven Finnish national-level sprinters, four male and three female, were studied. Of interest here is the common authorship of Komi in establishing the original study (Mero and Komi 1990) and Komi, Ishikawa et al. (2009) who have suggested a review of the 100 ms threshold. This raises questions over the validity of the current threshold, but the 2009 study is weakened by a lack of data and the fact that the data are based on subjects of limited international status. A re-evaluation of the 100 ms threshold will only be effective if the data utilised was recorded at competitions in which international athletes performed. This study utilised the large data set of RTs available from international competitions allowing the revision to be governed by performances of athletes of substantial international status, while also providing RTs from male and female athletes in comparison with the study by Mero and Komi (1990).

The initial aim of this study was to characterise how, if at all, the IAAF rule changes has influenced the reaction times of elite sprinters in top-tier competitions. It focused on statistically modelling the historical RT data and validating that a difference across ruling periods, sex and round of competition exists. The subsequent aim was to establish whether the data collected across multiple top-tier competitions could be combined and used to revise the IAAF 100 ms reaction time threshold for male and female sprinters accordingly.

The remainder of this paper provides an in-depth analysis of fifteen years of RT data. The methods section provides an overview of the collected data followed by an introduction to the statistical methods utilised in the analysis. The results section provides tables and figures which correspond to the application of the methodology discussed in the methods section to the sprint starts RT data. A discussion of these results follows providing an insight into the RT data across ruling periods, sex and competition rounds. The paper ends with a concise conclusion of the findings of this study.

Methods

Data

The RT data collated from the IAAF results website (http://www.iaaf.org/results) and the European Athletics website (http://www.european-athletics.org/results/) includes complete Championship event data for the 60 m and 60 m hurdles indoor championships and the 100 m, 110 m hurdles and the 200 m outdoor championships over the period from 1999 to 2014 inclusive. Events over distances greater than 200 m were excluded since the reaction time has less influence on the performance and therefore would not provide valid information on the quickest possible reaction time of an elite athlete. The data collected contained race RTs for 1303 and 1007 unique male and female athletes, respectively, with a total of 4560 and 3999 RT records for male and female performances respectively. While the data contains additional information on the finish time, wind speed and name of the sprinter, the only variables of interest in this study were the RT of the sprinter, the year of the event and the round of the competition. Following an extensive literature search this study constitutes the largest exploration of RTs of elite sprinters to date and far exceeds the work of Mero and Komi (1990) which is the basis for the 100 ms threshold.

Exploratory Analysis

The RT data are positively skewed. Thus the median was utilised as a measure of centrality and confidence intervals around the median at the 2.5 and 97.5 percentiles were constructed. Box-and-whiskers plots provide visualisation of the RTs of males and females across ruling periods. The exploratory analysis also examined the number of false starts recorded in each ruling period.

Statistical Modelling

The exponentially modified Gaussian distribution (EMGD) has been widely used to model reaction times (Dawson 1988). The EMGD is a convolution of an exponential distribution, controlling the tail of the distribution, and a Gaussian distribution to represent the group of quickest reactions. Its probability density function is:

$$f(RT|\mu,\sigma,\tau) = \frac{1}{\tau} exp\left\{\frac{\mu}{\tau} + \frac{\sigma^2}{2\tau^2} - \frac{RT}{\tau}\right\} \Phi\left(\frac{RT - \mu - \frac{\sigma^2}{\tau}}{\sigma}\right),$$

where, μ and σ represent the mean and standard deviation of the Gaussian component and τ is the rate parameter of the exponential distribution, while $\Phi(t)$ is the cumulative standard normal distribution function evaluated at *t*. The parameters are constrained such that $\min(RT) \leq \mu \leq \max(RT)$, $0 \leq \sigma \leq \text{range}(RT)$ and $0 \leq \tau \leq \text{range}(RT)$. If the RTs follow an EMGD with the above parameters its measure of location and scale can be represented as $\mathbb{E}(X) = \mu + \tau$ and $\text{Var}(X) = \sigma^2 + \tau^2$, respectively.

Estimation:

The parameters for each of the EMGD's reported are estimated using the maximum likelihood approach provided in the retimes package (Massidda 2013) based in the R statistical computing environment (R Core Team 2015). The maximum likelihood estimates of the parameters for the EMG distribution are described in detail in Silver, Ritchie et al. (2009).

Goodness-of-fit Statistics

Pearson's Chi-Squared goodness-of-fit was used to evaluate if the observed data came from an underlying EMGD. The null hypothesis assumes that the data fits the distribution well, while the alternative assumes that the distribution is not a true reflection of the observed data, more detail of this goodness-of-fit test can be found in Greenwood and Nikulin (1996). For this paper a significance level of $\alpha = 0.05$ was used, with $r - p - 1$ degrees of freedom, where *r* is the number of quantiles at which the fit is compared, 10 is used in all calculations in this study, and p is the number of estimated parameters which is 3 for all cases when using the EMGD. The critical value is taken from the Chi-squared tables with 6 degrees of freedom at a significance level of 0.05 which corresponds to a critical value of 12.592.

Comparison across ruling periods and rounds

To compare differences across ruling periods and rounds a simple probabilistic approach was taken. The probability of a RT above a given threshold, T, being observed can be expressed as $\Pr(RT \geq T) = 1 - \Pr(RT < T)$ by the total law of probability. In turn the $\Pr(RT < T) = F(T)$ where $F(T)$ is the cumulative distribution function (CDF) of the EMGD evaluated at *T*. The CDF is formulated by integrating the probability density function shown above, resulting in the CDF for the EMGD being:

$$F(T) = -\exp\left(\frac{-T}{\tau} + \frac{\mu}{\tau} + \frac{\sigma^2}{2\tau^2}\right)\Phi\left(\frac{T - \mu - \frac{\sigma^2}{\tau}}{\sigma}\right) + \Phi\left(\frac{T - \mu}{\sigma}\right),$$

where *T* is the threshold of interest, μ, σ and τ are the parameters of the EMGD defined earlier and Φ is again the CDF of the standard normal distribution.

Revision of the False Start Detection Threshold

For both males and females, an EMGD was estimated to fit the gender specific RTs across all ruling periods. The CDF of the EMGD presented earlier was utilised to find the threshold (*T*) which 99% of the observed RTs lay above. The identified threshold is the authors' proposal of the true quickest possible RT by an elite athlete under the current technology set down by the IAAF.

Results

Table 1 provides descriptive statistics for the RTs of elite sprinters relative to rule change periods set down by the IAAF and by sex. The rows of the table represent the combination of ruling period and sex. The columns provide the number of observations, the median, a 95% confidence interval around the median and the number of false starts recorded in the period.

Sex	Ruling Period	N	Median	95% Median CI	No. False Starts
Male	1999-2003	1407	151.0	(116.0, 230.0)	1
	2004-2009	1545	155.0	(121.0, 250.0)	0
	2010-2014	1608	161.0	(128.0, 282.7)	8
Female	1999-2003	1160	153.0	(119.0, 244.0)	0
	2004-2009	1237	161.0	(126.2, 256.8)	0
	2010-2014	1602	168.0	(130.0, 278.0)	1
Total	1999-2014	8559	159.0	(122.0, 259.0)	10

Table 1: Descriptive statistics for reaction times of elite sprinters (ms).

Box-and-whisker plots for the RTs by ruling period are presented in Figure 1. The left panel represents the male participants while the females are presented in the right panel. The *x*-axis for both panels corresponds to the three ruling periods with the commonly scaled *y*-axis representing RTs in milliseconds. The horizontal dashed line at 100 ms denotes the current disqualification limit – while RTs are recorded below 100 ms they correspond to false starts which are not a true reflection of the shortest valid RTs in elite athletics. The circled numbers below the threshold line provide the number of disqualifications due to false starts relative to sex and ruling period.

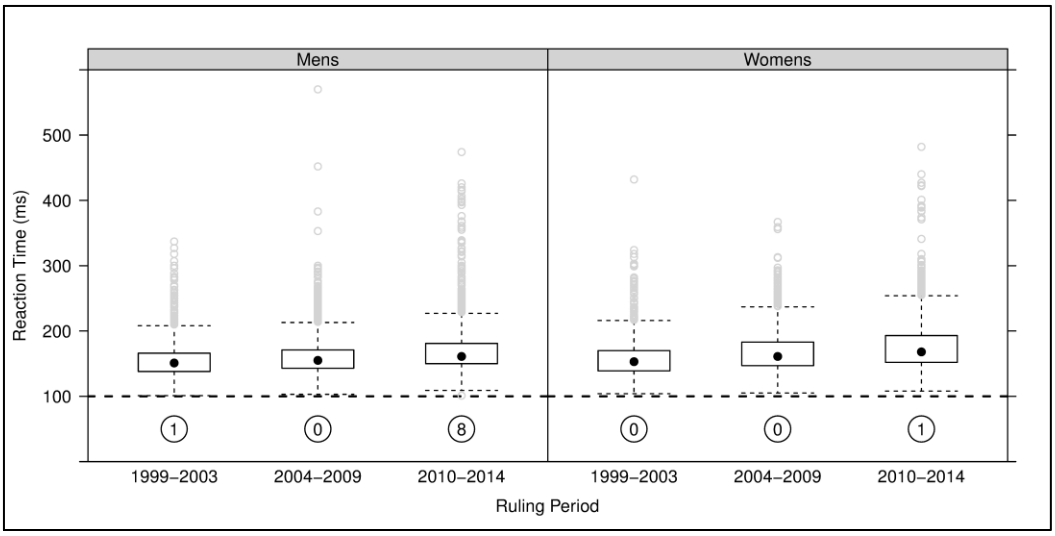

Figure 1: Reaction times across ruling periods.
Box-and-whisker plots of valid reaction times of sprinters across World and European Championships 1999-2014. The numbers circled are the number of disqualifications recorded as false starts.

The results of fitting the EMGD for males and females in each ruling period are provided in Table 2 and Figure 2. These also compare the distributional properties of RTs from first round events and finals, as races of significant value may encourage anticipation of the starter's pistol. The parameters column presented in Table 2 provide the maximum likelihood of the best estimates of the EMGD parameters given the data from the ruling period and sex of interest, the properties column shows the location and scale of the distribution, the form of which was discussed in the methods section earlier while the final two columns provide a measure of goodness-of-fit of the estimated EMGD to the observed data with the corresponding p-value for the result of the test.

		Parameters			**Properties**		**Goodness-of-fit**	
Sex	**Ruling Period**	μ	σ	τ	*E*(X)	*s*(X)	χ^2	p-value
Male	1999-2003	132.80	14.66	21.97	154.77	26.41	16.03	0.0136
	2004-2009	134.76	11.57	27.76	162.52	30.07	42.87	<0.0001
	2010-2014	139.95	10.92	32.33	172.28	34.12	60.25	<0.0001
	First Rounds	134.96	12.61	33.60	168.56	35.89	68.53	<0.0001
	Final Rounds	137.30	12.53	15.21	152.51	19.71	11.57	0.0723
Female	1999-2003	131.32	12.14	27.62	158.94	30.17	8.95	0.1765
	2004-2009	138.32	11.91	31.24	169.56	33.43	3.80	0.7039
	2010-2014	141.82	11.81	36.51	178.33	38.37	13.73	0.0328
	First Rounds	137.89	12.11	38.24	176.13	40.11	7.60	0.2688
	Final Rounds	136.90	14.04	24.15	161.05	27.93	7.97	0.2403

Table 2: Estimated parameters, distributional properties and goodness-of-fit statistics of the EMGD fits.

The estimated distributions for ruling periods, first round and finals by sex are shown in Figure 2. Panels A and B show the fit for each of the three ruling periods, the dotted line is the fit for the 1999-2003 ruling period, the dashed line is the estimated distribution for the 2004-2009 period, while the immediate disqualification ruling distribution is the solid line. Panel C and D provide a visualisation of the estimated EMGD fits for the first round heats, dashed line, and finals, solid line, for both males and females.

The probabilistic approach for identifying a difference across ruling periods and rounds is determined by calculating the probability of observing a RT greater than 120 ms. For male ruling periods the respective probabilities are 0.95, 0.98 and 1.00, female ruling period's probabilities are 0.97, 0.99 and 1.00. The probabilities of observing RTs above 120 ms in first rounds are 0.98 and 0.99 and in finals are 0.98 and 0.97 for males and females respectively.

The revision of the false start disqualification limit calculated from the historical data relevant to gender resulted in both thresholds exceeding the 100 ms limit. Male event data proposes a revised limit of 115 ms while the female event data is suggestive of a limit of 119 ms.

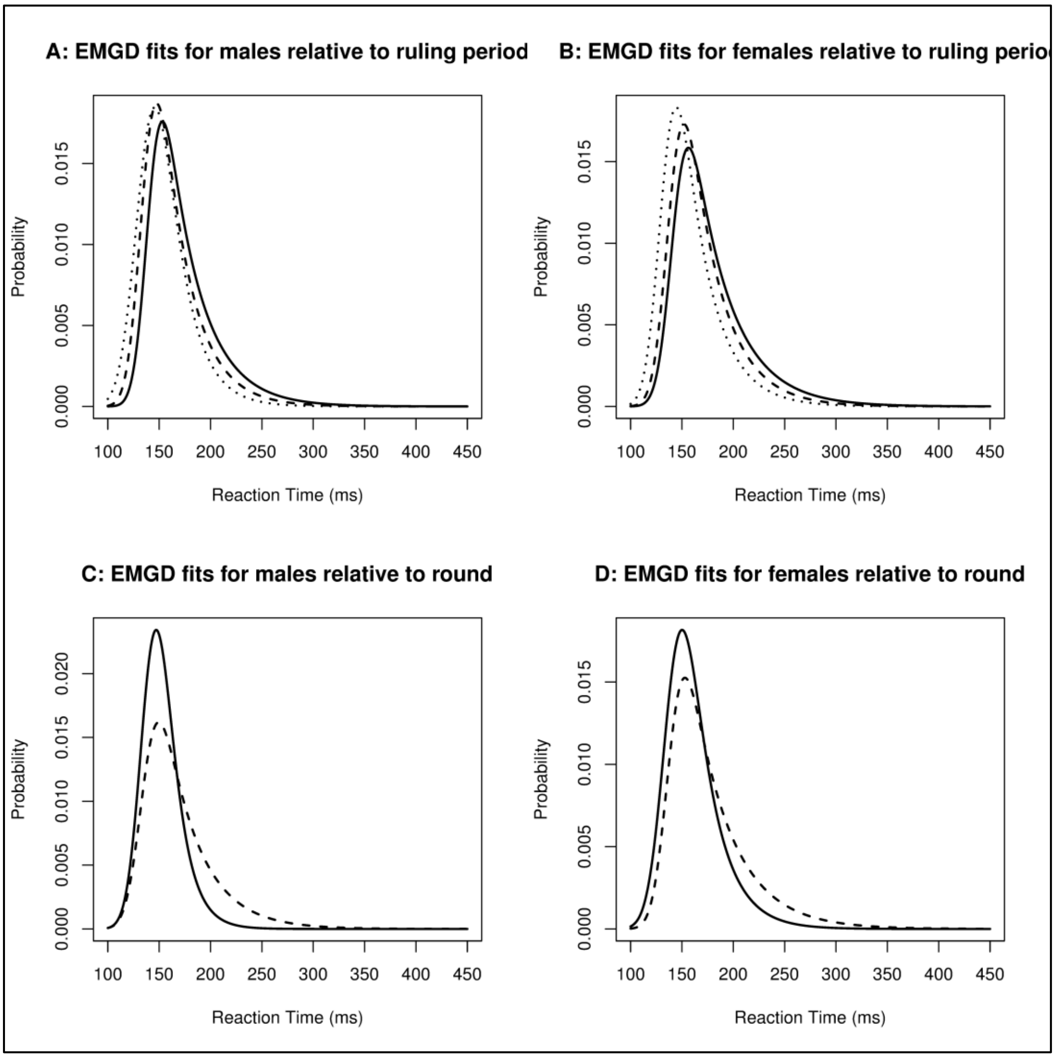

Figure 2: EMGD fits to reaction time data.
Panels A and B present the estimated distributions for males and females respectively for the three ruling periods. The dotted line in each plot relates to the ruling period from 1999-2003, the dashed line represents the introduction of the group warning in January 2004, while the solid line represents the current ruling of automatic disqualification which was introduced in January 2010. C and D present the estimated distributions for males and females respectively for first round heats and finals. The dashed line represents the first round reaction times of athletes between 1999 and 2014, while the solid line represents the reactions time in the finals of these same events.

Discussion

The 95% confidence interval for the median of the entire data shows a lower bound of 122 ms, 22 ms above the threshold set down by the IAAF. For a 100 m sprinter 22 ms can be a significant advantage. For example in the 2012 Olympic Games in London the winning margin between the first and second athletes in the women's 100 m event was only 30 ms, at 10 m/s a 30 ms difference represents a 30cm advantage. Thus an improved start regime could be the winning or losing of a 100 m sprint, particularly if the 100 ms threshold is indeed subject to not detecting valid false starts.

Ruling Periods

The median and 95% confidence band for each ruling period shown in Table 1 suggests shifts in the distributions of the RTs of athletes in periods after a rule change of approximately 5 ms for males and 7.5 ms for females. While the shift in RTs of sprinters after the rule change in 2004 is not particularly noticeable in Figure 1, a noticeable shift does appear after the rule change in January 2010. Figure 2 panels A and B, suggest that in the case of ruling periods, the EMGD for male and female athletes have an increasing trend over time. It is evident in both sexes that a shift over time has occurred with a significant shift in the distribution density after the automatic disqualification ruling in January 2010. The sex difference can also be identified as panels A and B of Figure 2 are on the same scale, with the females densities being shifted further to the right. The probability of observing a reaction time of greater than 120 ms increases for males and females relative to the ruling periods. The probability of observations greater than 120 ms following the introduction of the automatic disqualification rule in 2010 is 1, regardless of gender. This is suggestive that the true reaction time limit of athletes recorded using the current technology is in the region of 120 ms, some 20 ms above the enforced threshold.

Competition Rounds

A comparison of the RTs in first round heats and finals is shown in panels C and D of Figure 2. A reduction of the RTs in finals may suggest that athletes are more willing to engage in predicting the starter's pistol when a medal is on the line. A shift to the left indicating a decrease in RT is evident in both cases, thus suggesting that athletes are more likely to take risks in finals than in first round heats. The sex difference can be identified as the plots are on the same scale, with the females densities being shifted further to the right. The probabilistic approach for the difference between RTs greater than 120 ms in first rounds and finals only shows a relative difference for female athletes, this may be due to the fact that males are willing to take a risk at all levels whereas women require a significant incentive to take that risk, as discussed in Whitley, Nelson et al. (1999).

Sex Difference

The assertion by Lipps, Galecki et al. (2011) of a sex difference in RTs is evident throughout this study and can be further explained by the strength differences between men and women, allowing men to produce the required force to reach the false start threshold quicker. A difference in RT relevant to gender was seen to exist across both ruling periods and competition rounds. This prominent sex difference signifies a need for independent false start detection limits for male and female athletes as suggested by Lipps, Galecki et al. (2011).

RT’s Revision

Revised RT thresholds were established for males and females separately by estimating an EMGD fit for each of the sexes. The revised threshold was calculated as the RT for which 99% of the observed RTs lay above. As such it is suggested that the RT threshold for male athletes be increased to 115 ms, while females should be increased to 119 ms, due to the evident sex difference reported throughout this study and in Lipps, Galecki et al. (2011).

Technology

Throughout this paper there has been evidence that the 100 ms ruling as a disqualification limit for false starts in elite athletics is too short. This needs to be taken in the context of the current technology utilised to measure an athlete’s RT. The IAAF approved technology does not record the quickest RT, or the first movement of the athlete, but rather the addition of the time to produce a predefined force along with the RT. Pain and Hibbs (2007) suggest that RTs of 85 ms are possible; this paper does not contradict these findings but rather suggests that these reduced RTs are similar to the time until the very first reaction of an elite athlete after a starter's pistol. It is the authors’ belief that utilising high-speed cameras or laser technology to record RTs could result in RTs in the region of 85 ms.

Conclusions

From this study it is clear that the IAAF 100 ms false start threshold is inadequate for the fair and impartial refereeing of elite athletics. Similar to the work of Lipps, Galecki et al. (2011) a sex difference has also been identified from the historic data. It is suggested that the ruling for males and females be independently revised to account for the strength differences between the sexes.

Revised RT thresholds, estimated from the historical data, for both males and females are provided in the results section.

In conclusion, this study has examined a large dataset of RTs of international athletes, combined with an extensive literature review, it has shown that the current governance of false start disqualification is inadequate under the current measurement system utilised by the IAAF.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Supplementary Material

Supplement A: Reaction Time Data 1999-2014. (https://goo.gl/xYuQXH; .csv)

Supplement B: R Code for Reaction Time Analysis. (https://goo.gl/rBPCHn; .R)

References

Athletics Australia. (2009). "Implementation of IAAF "No False Start" Rule." from http://www.easternsuburbs.org.au/assets/console/customitem/attachments/New_Start_Rule_Guidelines_141109.pdf.
Brown, A. M., Z. R. Kenwell, B. K. V. Maraj and D. F. Collins (2008). ""Go" signal intensity influences the sprint start." Medicine and Science in Sports and Exercise **40**(6): 1142-1148.
Dawson, M. R. W. (1988). "Fitting the ex-Gaussian equation to reaction time distributions." Behavior Research Methods, Instruments, & Computers **20**(1): 54-57.
Der, G. and I. J. Deary (2006). "Age and sex differences in reaction time in adulthood: Results from the United Kingdom health and lifestyle survey." Psychology and Aging **21**(1): 62-73.
Greenwood, P. E. and M. S. Nikulin (1996). A guide to chi-squared testing, John Wiley & Sons.
International Association of Athletics Federations. (2015). "Competition Rules 2016-2017." from http://www.iaff.org/about-iaaf/documents/reules-regulations.
Komi, V., M. Ishikawa and S. Jukka (2009). "IAAF sprint start research project: Is the 100ms limit still valid." New studies in athletics **24**(1): 37-47.
Lipps, D. B., A. T. Galecki and J. A. Ashton-Miller (2011). "On the implications of a sex difference in the reaction times of sprinters at the Beijing Olympics." PLoS ONE **6**(10).
Massidda, D. (2013). "Retimes: Reaction Time Analysis." R package version 0.1-2.

Mero, A. and P. V. Komi (1990). "Reaction time and electromyographic activity during a sprint start." European Journal of Applied Physiology and Occupational Physiology **61**(1-2): 73-80.
Pain, M. T. G. and A. Hibbs (2007). "Sprint starts and the minimum auditory reaction time." Journal of Sports Sciences **25**(1): 79-86.
R Core Team (2015). "R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, 2012)." URL: http://www.R-project.org
Silver, J. D., M. E. Ritchie and G. K. Smyth (2009). "Microarray background correction: maximum likelihood estimation for the normal–exponential convolution." Biostatistics: kxn042.
Whitley, B. E., A. B. Nelson and C. J. Jones (1999). "Gender differences in cheating attitudes and classroom cheating behavior: A meta-analysis." Sex Roles **41**(9-10): 657-680.