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 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 athletes from sprint events incorporating starts from blocks. This study analysed all available World and European Championship reaction time data from 1999 to 2014 to examine the effect of rule changes on competition reaction times at major championships. The exponentially modified Gaussian distribution was used to model the reaction times and make comparisons relative to sex, ruling periods and competition rounds. Revised reaction time thresholds of 115 ms and 119 ms were calculated for men and women respectively, indicating that the current 100 ms rule results in some false starts not being detected in competitive athletics. The conclusions of this study, based on the IAAF approved starting technology, propose a revision of the false start detection threshold is carried out and that men and women 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). Accurate identification of the precise RT latency period after the start signal for sprint starts at major championships is vital 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). Before 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 for two false starts in a race. 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 limit used to detect a false start is measured as the time taken by an athlete to produce a predefined force or acceleration threshold on the starting blocks. Lipps, Galecki et al. (2011) stated 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) suggested that valid RTs of 85 ms are possible and Brown, Kenwell et al. (2008) present similar results observing 21% of recorded RTs being below the 100 ms threshold. While such RTs are possible, the recording systems used in both studies measured the time to the first change in force on the blocks produced by the athlete rather than the time taken to produce a predefined force. In contrast, Komi, Ishikawa et al. (2009) used IAAF approved blocks together with additional sensor technologies and recommended that the IAAF should abandon the 100 ms threshold and the measurement technology currently employed.

Notwithstanding the results of the previous studies, an effective review of the 100 ms threshold requires analysis of data recorded using IAAF approved starting technologies. All of the previous studies have been based on restricted sample sizes and many do not include data from athletes of international major championship status. Appropriate statistical analysis of a large sample RTs of men and women at major international championships could be used to derive an appropriate temporal threshold for sprint start RTs in competitive athletics.

Previous studies have provided 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), but IAAF rule 161.2 continues to govern competitions for men and women alike. Given the physiological differences in strength in men and women, it is possible that a sex difference in RT's could exist and that the rules for men and women 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 men and three women, 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 for men and women 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 RTs 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 RT threshold for men and women 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.

Methods

Data Acquisition

RT data was collated from the IAAF results (http://www.iaaf.org/results) and the European Athletics websites (http://www.european-athletics.org/results/) and includes complete Championship event data for: the 60 m and 60 m hurdles indoor championships; 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 RT has less influence on the performance and therefore would not provide valid information on the quickest possible RT of an elite athlete. The data contained race RTs for 1303 and 1007 men and women, respectively, with a total of 4560 and 3999 RT records for performance of men and women respectively. While the available data contained 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. Based on 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 current 100 ms threshold.

Exploratory Analysis

Descriptive statistics for the centre and variability of RTs relative to sex and ruling periods were calculated. Box-and-whiskers plots provide visualisation of the RTs of men and women 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 RTs (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 $0 \le \tau \le \text{range}(RT)$. 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 were 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 EMGD 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 RT data came from an underlying EMGD. The null hypothesis assumed that the data fitted the distribution well, while the alternative assumed that the distribution was 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 study 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 $\Pr(RT \ge T) = 1 - \Pr(RT < T)$ 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 men and women, an EMGD was estimated to fit the sex 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 for men and women. The median and its confidence band are reported as a positive skew was observed in the RTs. 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. Box-and-whisker plots for the RTs by ruling period are presented in Figure 1. The left panel represents the men while the women 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.

Sex	Ruling Period	N	Median	95% Median CI	No. False Starts
Men	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
Women	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 RT data (ms).

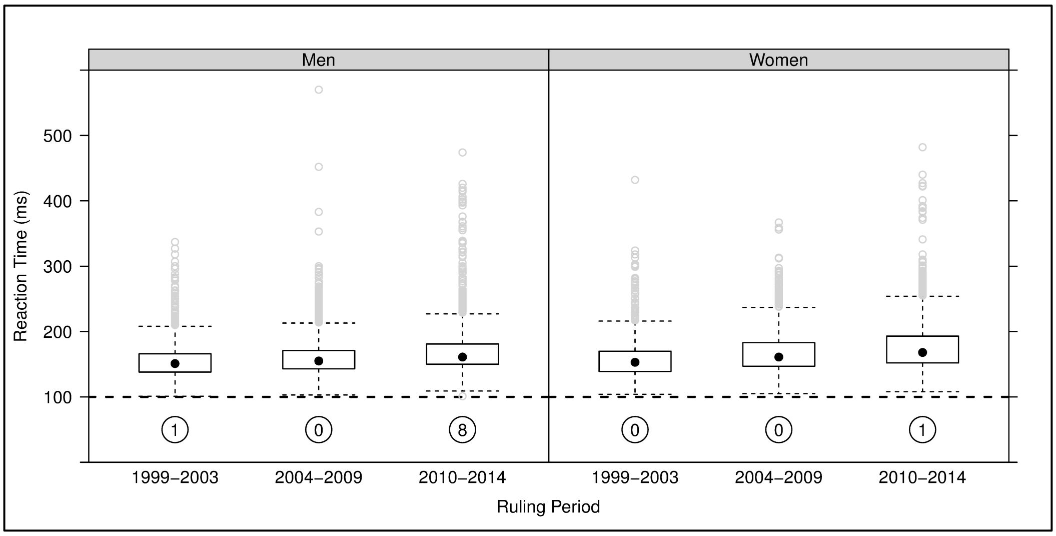

Figure 1: RTs across ruling periods.
Box-and-whisker plots of valid RTs of sprinters across World and European Championships 1999-2014. The numbers circled are the number of disqualifications recorded as false starts.

The results of fitting the EMGD for men and women 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, since races of greater importance may encourage athletes to take risks and anticipate the start signal. The parameters column presented in Table 2 provides 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, 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
Men	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
Women	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 EMGDs.

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 men and women.

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 ruling periods the respective probabilities are 0.95, 0.98 and 1.00 for men and 0.97, 0.99 and 1.00 for women. 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 men and women respectively.

The revision of the false start disqualification limit calculated from the historical data relevant to sex resulted in both thresholds exceeding the current 100 ms limit. Men's event data proposed a revised limit of 115 ms while the women's event data suggested a limit of 119 ms.

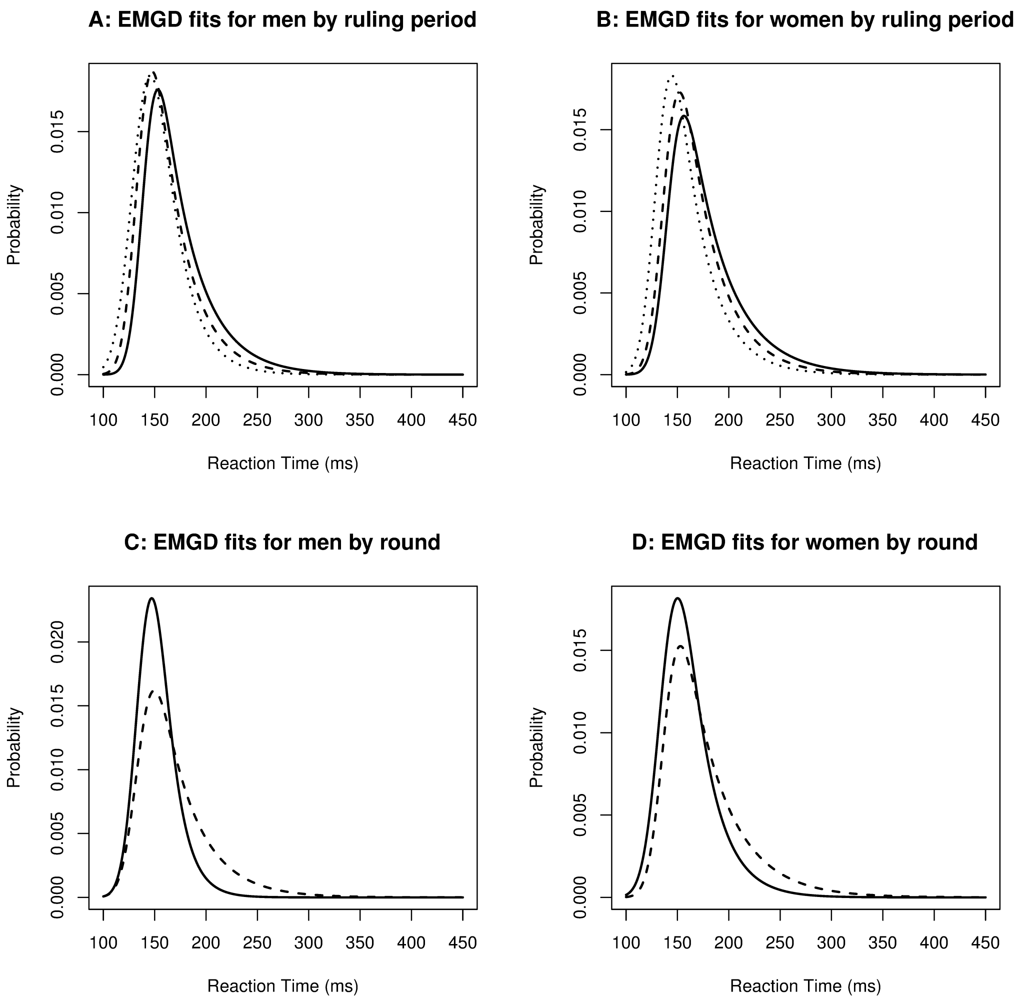

Figure 2: EMGD fits to RT data.
A and B present the estimated distributions for men and women respectively for the three ruling periods. The dotted line in each plot relates to the ruling period from 1999-2003, the dashed line represents the introduction of the group warning in January 2004, while the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

Discussion

The 95% confidence interval for the median RT 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, and at typical velocity of 10 m/s this difference represents a 30 cm advantage. If the 100 ms threshold is subject to not detecting valid false starts an anticipated start could alter the results of an event.

Ruling Periods

The median and 95% confidence band for each ruling period shown in Table 1 indicates shifts in the distributions of the RTs of athletes in periods after a rule change of approximately 5 ms for men and 7.5 ms for women. 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 for ruling periods, the EMGD of men and women has 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 densities for women being shifted further to the right. The probability of observing a RT greater than 120 ms increases for men and women 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, irrespective of sex. This suggests that the minimum observed RT of athletes recorded using the current IAAF approved technology and rules is greater than 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 suggests that athletes are more likely to anticipate the start signal when the race is more important. 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. A sex difference is evident by the comparison of the plots with the RT distribution for women 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 women, this may be due to the fact that men 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 sex was observed across both ruling periods and competition rounds. This prominent sex difference supports the case for independent false start detection limits for men and women as suggested by Lipps, Galecki et al. (2011). While this study has continuously reported a sex difference in RTs of athletes, this sex difference may be a consequence of the currently employed starting blocks. The removal of a force based threshold would establish if a true sex difference exists in the psychological RTs of athletes.

RT's Revision

Revised RT thresholds were established for men and women independently 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 men be increased to 115 ms, while women should be increased to 119 ms, due to the evident sex difference reported throughout this study and in Lipps, Galecki et al. (2011), under the constraints of the currently employed starting blocks.

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) suggested 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 initial response of an elite athlete after a start signal.

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, under the IAAF approved starting blocks. 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 men and women be independently revised to account for the strength differences between the sexes. Revised RT thresholds, estimated from the historical data, for both men and women 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.