Supplementary methods

(a) Rider selection

The 45 riders that did not finish the race were excluded a-priori, leaving 153 riders. To increase standardisation of the photos as much as possible, I excluded riders wearing caps (N = 6), sunglasses on their forehead (N = 7) or beards (N = 2) on their photo, as well as those that were not photographed from the front (i.e. where only one ear was visible) (N = 13), those with unusual haircuts (N = 2) or unusual facial expressions (N = 2), and those where the lighting was obviously different (N = 8). For two riders no photo was available. Finally, I excluded another 7 riders whose height and weight were not available on the Tour de France website. This left a pool of 104 riders, from which 80 were selected at random.

(b) Data collection

Surveys were created using the FluidSurveys website (http://www.fluidsurveys.com), using maximum privacy settings (e.g. IP address and referring URL were not recorded). Both versions of the survey were online alternatingly from July 24 until August 10 2012. The link was distributed to all members of the Institute of Evolutionary Biology and Environmental Studies of the University of Zurich, and was posted on the EvolDir (http://life.biology.mcmaster.ca/~brian/evoldir.html) and Oikos-listan (http://wallace.teorekol.lu.se/mailman/listinfo/oikos-listan) mailing lists. Recipients were asked to forward the link to friends and colleagues, and to post it on the various social media sites.

Participants had the opportunity to leave information on their sex, nationality, age, highest level of education ('No secondary education', 'Secondary education', 'Post-secondary or trade school', 'Bachelor or equivalent', 'Master or equivalent', 'Doctoral or equivalent', 'None of the above'), relationship status ('Single and searching for a partner', 'Single and not searching for a partner', 'In a stable relationship for less than 6 months', 'In a stable relationship for more than 6 months', 'Difficult to say...'), and their sexual orientation (using the Kinsey scale, which goes from 0 (exclusively heterosexual) to 6 (exclusively homosexual)). Additionally, women were asked whether they used a hormonal contraceptive, and if they did not, for the average length of their cycle, and how many days had passed since the start of their last period. The last questions included the additional options 'I don't know', 'In menopause' and 'Pregnant'.

(c) Inferring female fertile phase

When possible, I inferred whether non-pill-using women (who were not pregnant or menopausal) were in the high conception probability (i.e. late-follicular) phase of their menstrual cycle. Here I assumed that this covers a 6-day period, ending on the day of ovulation [1], and that ovulation takes place in the middle of the cycle [2]. Women who reported a number of days since the start of their last menstrual period exceeding the typical length of their cycle were excluded, as were women with menstrual cycles shorter than 21 days and longer than 35 days.

Various rules exist to determine the average timing of ovulation within the menstrual cycle, as well as the length and timing of the high conception probability phase relative to ovulation [e.g. 2, 3]. Also, inferences based on average cycle length are relatively imprecise as they ignore within- and among-individual variation in cycle length and the timing of ovulation, which can be substantial [1, 4]. Finally, any self-reported data inherently contains errors. However, although they are thus inherently conservative, tests for an effect of female reproductive period on the relationship between performance and attractiveness, masculinity and likeability were insensitive to variations on the criteria outlined above.

As all pill-users are women, and only non-pill using women are potentially in the fertile part of their cycle, I created a new 'sex/fertility' variable which combines these variables (levels: 'man', 'pill-using woman', 'non-pill-using woman in non-fertile part of cycle', 'non-pill-using woman in fertile part of cycle'). In case of a significant 'sex/fertility' effect, post-hoc tests comparing specific groups were used to find out what was driving this effect (sex, pill-use and/or timing in cycle).

(d) Variation in facial expression

Although portraits were standardised in terms of lighting, distance and background, there was substantial variation in facial expression. To be able to account for this, all portraits were additionally categorised as 'not smiling', 'smiling a little' and 'smiling'. Although to some degree this classification is subjective, independent classifications by four additional persons (two men and two women) were strongly correlated with the classification used here (Spearman $\rho = 0.77$-0.84). Furthermore, results were not influenced by whose classification was used.

(e) Rider height and weight

Given the positive correlation between height and weight ($r = 0.64$, $P < 0.001$) (figure S1A), I performed a principal component analysis (PCA) and used the two principal components in all analyses. Principal component (PC) 1 explains 82% of all variation and loads positively on both height and weight. I will refer to PC1 as 'size', with riders with high values of PC1 being heavier and taller. PC2 loads positively on weight and negatively on height. Thereby PC2 is a measure of size-corrected weight very similar to the body mass index (BMI), which is calculated as weight divided by height squared. Indeed, the correlation between PC2 and BMI is effectively 1 (figure S1B). Here I refer to PC2 as 'relative weight', with riders with high values of PC2 being relatively heavy for their height.

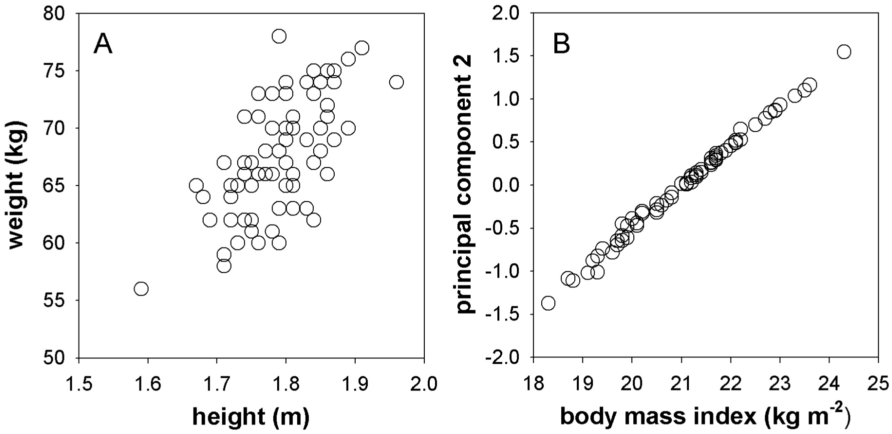

Figure S1. Rider height and weight. (A) The relationship between height and weight and (B) the relationship between body mass index and principal component 2, based on height and weight.

(f) Quantifying performance

Professional cycling, and especially a three-week stage race like the Tour de France, involves a substantial amount of tactics and teamwork. To obtain a measure of performance that is least influenced by this, and that gets as close as possible to being an objective measure of a rider's physical performance, I took for each rider the time it took them to complete the 6.4 km long prologue (a short individual time trial held before the first stage), the first individual time trial (ITT1; stage 9: 41.5 km), the second individual time trial (ITT2; stage 19: 53.5 km), and finally the complete race (subtracting the times for the prologue and the two time trials). These four measures of performance were significantly positively correlated ($0.26 < r < 0.76$, all $P \leq 0.001$) (figure S2).

Subsequently, I performed a principal component analysis (on the correlation matrix) and extracted the first principal component (PC1), which explained 61.2% of the variation (eigenvector of PC1: prologue: 0.44, ITT1: 0.58, ITT2: 0.47, total time: 0.50). To facilitate interpretation (i.e. to ensure that faster riders have higher values), PC1 was multiplied with -1. From here on I will refer to PC1 as 'performance'.

Note that, depending on their role in the team (helper or leader), riders vary in their motivation to finish in as little time as possible. Although this role is likely to be correlated with their abilities, this will introduce some error. Also, I only measure one aspect of performance, emphasising endurance capacity and the ability to perform consistently over a relatively long period of time, and thereby this measure of performance is biased against sprinters and pure climbers.

In addition to the estimation of the relationship between performance and attractiveness, I used a series of linear models to test which rider-specific variables (age, size (i.e. PC1, see above), relative weight (PC2) and facial expression) correlate with performance. Initial models included both linear and quadratic terms of all covariates. Backward elimination of non-significant terms was performed, starting with the least significant quadratic terms. Having arrived at the final model, all eliminated variables were reintroduced one-by-one, and for the non-significant terms these are the parameter estimates, F- and P-values that are presented. Although riders were randomised both within and between surveys, I additionally tested for an inadvertent difference in performance between riders from the two surveys, and whether performance was non-random with respect to the location of a rider within each survey.

A reasonable number of the 80 riders included in this study (n=32) also participated in the 2013 Tour de France. To test whether the above definition of performance, as well as its relationship with attractiveness, is repeatable across years, I calculated performance during the 2013 Tour in a similar manner as I did for 2012, using a rider's time for the first and second time trial, as well as their total time minus their time for the two time trials. Note that, unlike in 2012, in 2013 there was no prologue.

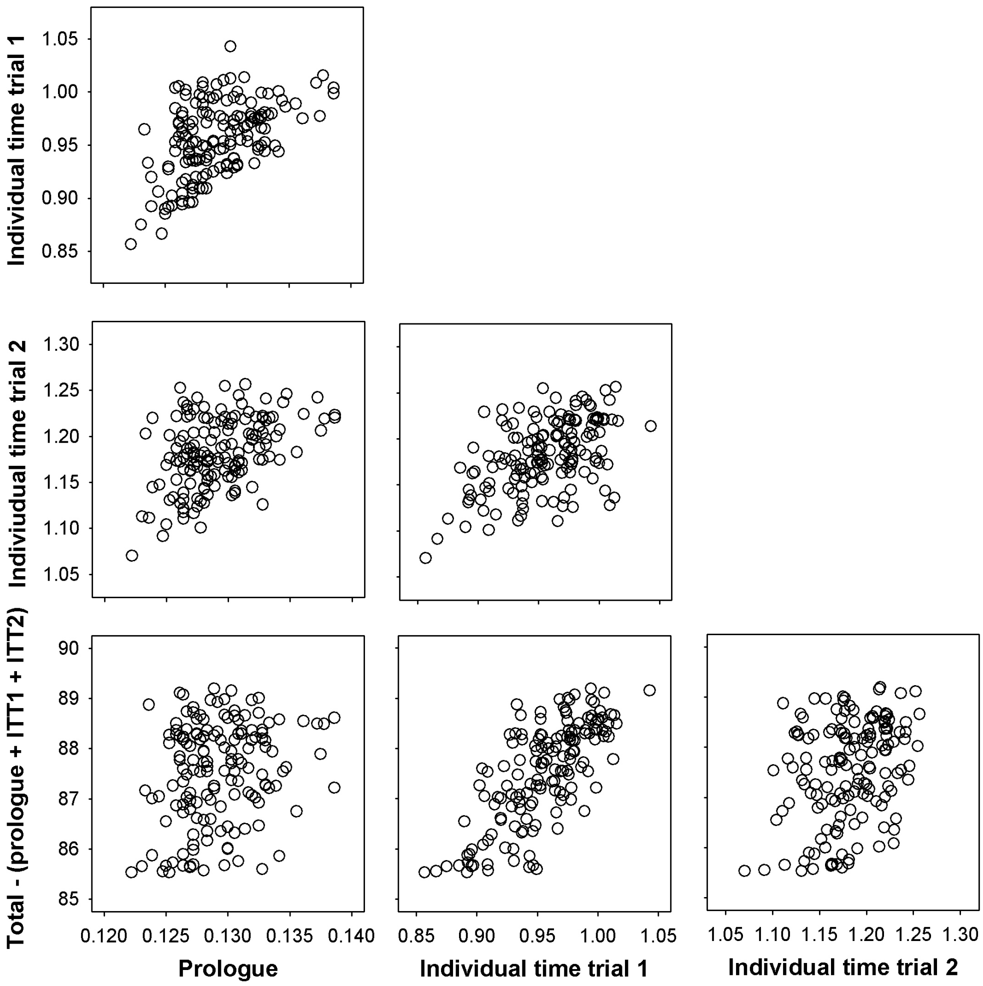

Figure S2. Measuring performance. The correlations among the time in hours taken by each rider to complete the prologue, the first individual time trial (ITT1), the second individual time trial (ITT2) and the total race (minus prologue, ITT1 and ITT2).

References

1. Wilcox A. J., Weinberg C. R., Baird D. D. 1995 Timing of sexual intercourse in relation to ovulation - effects on the probability of conception, survival of the pregnancy, and sex of the baby. *N. Engl. J. Med.* **333**, 1517-1521. (doi:10.1056/nejm199512073332301).

2. Lamprecht V. M., GrummerStrawn L. 1996 Development of new formulas to identify the fertile time of the menstrual cycle. *Contraception* **54**, 339-343. (doi:10.1016/s0010-7824(96)00202-8).

3. Arevalo M., Sinai I., Jennings V. 1999 A fixed formula to define the fertile window of the menstrual cycle as the basis of a simple method of natural family planning. *Contraception* **60**, 357-360. (doi:10.1016/s0010-7824(99)00106-7).

4. Fehring R. J., Schneider M., Raviele K. 2006 Variability in the phases of the menstrual cycle. *J. Obstet. Gynecol. Neonatal Nurs.* **35**, 376-384. (doi:10.1111/j.1552-6909.2006.00051.x).