**Research Proposal**

**Background:**

To date the process of prescribing training has relied upon the experience and intuition of those involved (i.e. coaches and athletes), as the necessary research in this area is lacking18. Over the past four decades, the scientific basis for prescribing training programs has advanced little beyond Banister and colleagues’ seminal work (Borresen & Lambert, 2009; Calvert et al., 1976). This is in marked contrast to the advances made in our understanding of the adaptions that result from training (Mann et al., 2013). However, this could change with the capability to measure individuals’ training and racing accurately and in detail in the field. The resulting large volumes of field measurements allows sports analytics to contribute to our understanding of effective training program prescription (Passfield et al., 2017). Furthermore, detailed monitoring of training and performance in the field provides an opportunity to reverse the usual scientific paradigm for research on this topic. Specifically, instead of conducting experiments to compare the effects of specific (laboratory-based) training regimens, we can measure study participants’ training, and track their resulting changes in performance. It may then be possible to determine which aspects of their monitored training is most effective, given sufficient data. With this scientific paradigm the method of enquiry consists of identifying which training led to the observed changes in performance, rather than trying to evaluate how performance changes in response to a restricted laboratory-based training protocol. Here the bigger the data, the better the insight, as effective training may be identified more clearly when the number of participants involved and the diversity of their training is greater. Exploring a wide range of training regimes with large numbers of participants is not a viable option for laboratory-based research, but in a field study it becomes quite plausible. Thus, this study aims to investigate the relationship between training and resultant changes in performance using a prospective study design in a large group of cyclists with a wide range of training regimes.

**Experimental Design:**

Participants will be recruited via the user database of the Golden Cheetah (GC) online training analysis software and be enrolled in the study for a period of 12 weeks. As usual for these existing GC user, they will be required to conduct their normal cycling training, and submit their training data to the Golden Cheetah software platform. Once per 4 week period, participants will be required to undertake a self-administered cycling test in order for their performance level to be tracked over the course of the season. In addition, participants will be required to conduct a standardised warm-up of 10 min at power outputs of 100W, and 10 mins at 150W - and to provide perceptual ratings following each training session/race that they undertake.

**Methods:**

*Collection of cyclist training data*

Cyclists will need to have their own power meters and that it is in good working order to be able to take part in the study. This will be made clear to participants before they consent to participate in the study. Participants will also be informed of the need to conduct the zero offset of their power meters prior to each training session and race in accordance with the manufacturer’s instructions. Participants will also be informed of the need to measure power values at a frequency of 1Hz. Participants will also be asked to wear a heart rate monitor during each training session/race. Power output and heart rate data will be collected from each participant for a 12 week period. Following each training or racing session, participants will be required to upload their data files to GC, with them being stored on a secure online file store (Google, San Francisco, USA). Data will be later analysed using Microsoft Excel and R.

In addition to uploading data files to the GC platform, participants will also be asked to provide perceptual measures for each training session/race. These will also be recorded using the GC platform. Following each training session/race, participants will be asked to provide a session-RPE score on a scale of 1-10 (Foster et al., 2001). Finally, if there was no data recorded on that day, they will be asked to record the reason (1=planned rest day, 2= did not train due to illness, 3= did not train due to injury.

*Self-administered exercise test*

Participants’ road bicycles will be equipped with a power meter and a magnet for direct cadence measurement. Participants will be requested to zero-offset their power meter device prior to all tests according to the manufacturer’s instructions. Participants will be required to refrain from heavy exercise in the 24 h prior to the self-administered tests and from food intake in the 3 h prior to tests. Prior to each of the test, cyclists will be required to warm-up for 10 min at power outputs of 100W and 10 mins at 150W. During the standardised warm-up, participants will be asked to record a single rating of perceived exertion for each work rate (100 & 150W) using the Borg CR-10 scale (Borg, 1998).

After a warm up, the participant was instructed to cycle as fast as possible for exercise durations in the order of 12, 7 and 3 min. The cyclist was instructed to continue low- intensity exercise for 30 min between each effort (i.e. recovery intensity) which has been found to be adequate for determining a valid CP. During each effort the cyclist will be free to alter their own gear ratio and cadence. The cyclist was also able to see their time, power output and cadence throughout the tests. The same road bike was used for all tests. Participants will be required to either complete the efforts on a turbo trainer, or as part of a training ride. If completing efforts as part of a training ride, participants will be request to complete the same route for each test over the course of the 12 week period. During the test participants will also be requested to measure their heart rate. Following the tests, participants will be required to upload their data file to GC in order for the results to be analysed.

**Data analysis:**

Linear mixed models will be used to compare differences in mean total distance, time and mean power output overall, between training and racing days and the percentage of exercise intensity spent across each month of the 12 week observation period. Where necessary, models will be fit with random intercept and slope to account for variable rates of change between each athlete and selected as the parsimonious model when minimising the AIC value. Training distribution and density profiles will be used to provide a quantification of each participant’s training load. A multi-resolution elastic net will be used to link the training distribution profiles with resulting changes in performance in the self-administered cycling tests over the 12 week observation period.