Cognitive-physical task interaction during self-paced cycling: A multiscale Granger Causality study

Chiara Avancini[a,b*], Daniele Marinazzo[c], Daniel Sanabria[a,b], Juan José Pérez-Díaz[d], José-Antonio Salas-Montoro[d] & Luis F. Ciria[a,b].

a. Mind Brain and Behavior Research Center, CIMCYC-UGR, University of Granada, Spain.

b. Department of Experimental Psychology, University of Granada, Spain

c. Department of Data Analysis, Faculty of Psychological and Educational Sciences, University of Ghent.

d. Department of Physical Education and Sport, Faculty of Sport Sciences, University of Granada, Spain.

*Corresponding author: Chiara Avancini (chiara.avancini@ugr.es, +34 691 078 206), Centro de Investigación Mente, Cerebro y Comportamiento (CIMCYC), Campus de Cartuja s/n, 18071 Granada, Spain

This manuscript is a preprint. All authors reviewed and approved the final version of this manuscript

Please cite as:
Avancini, C., Marinazzo, D., Sanabria, D., Pérez-Díaz, J.J., Salas-Montoro, J.A., & Ciria, L.F. (2023). Cognitive-physical task interaction during self-paced cycling: A multiscale Granger Causality study. *SportRxiv*

Abstract

Studying cognitive-physical interactions in self-paced high-intensity physical exercise presents the challenge of accounting for potential dual-task effects. In fact, self-pacing is thought to rely on top-down cognitive processing which makes it more susceptible to cognitive-physical interactions. Hence, even in paradigms where the experimental manipulation concerns the intensity of the exercise (i.e. high intensity versus low intensity) rather than its presence (i.e. exercise versus resting), performing the physical task might be more cognitively demanding in the higher intensity exercise condition. Here, we investigate the temporal dynamics of cognitive-physical interactions during dual-tasking by applying time-domain Granger Causality to data that combined indoor self-paced cycling and a cognitive task. Moreover, we investigate whether greater experience in self-pacing during cycling would reduce the need for exerting top-down control and therefore dual-task effects. We show that while cognitive and physical performance can interact in some individuals, better physical performance was not detrimental to cognitive performance in the expert cyclists group. We therefore propose that in self-paced physical exercise cognitive-physical interactions in expert cyclists are overall not confounded by dual-tasks interaction effects, although such interaction cannot be excluded for every single participant.

Keywords: Dual-task, cognitive load, top-down processing, stimulus-response conflict, physical exercise.

1 Introduction

Cognitive performance during physical exercise is typically studied using one of two methodologies: either comparing an exercise condition with a non-exercise resting condition (e.g., Audiffren et al., 2008) or comparing two (or more) exercise conditions at different intensities (e.g., Ciria et al., 2019). In the first case, participants' behavior is assessed in situations that differ in terms of physical and cognitive demands, as in the exercise conditions they perform a physical (e.g., walking on a treadmill or pedaling on an indoor bike) and cognitive (e.g., simple reaction times, RT) tasks at the same time. This is controlled, to some extent, in the second case, as participants are always subject to a dual-task situation. However, even in this case, performing the physical task might be more cognitively demanding in the higher intensity exercise conditions than in the lower intensity conditions, which might result in a potentially stronger interaction with the concurrent cognitive task. The study of these potential cognitive-physical interactions is particularly relevant in the exercise-cognition field, especially when examining the potential effect of acute physical exercise on cognitive processes (e.g., on memory or attention). Moreover, self-paced high-intensity exercise conditions, (e.g., a cycling time-trial where participants are instructed to perform their best for a given time) are highly susceptible to cognitive-physical interactions, since self-pacing is thought to rely on top-down cognitive processing (Holgado & Sanabria, 2021). This issue is addressed in the two experiments included in this report.

Cognitive-physical interaction effects have been studied mainly during walk, whereby participants' gait and cognitive performance is compared between single and dual-task conditions (see Al-Yahya et al., 2011, for a review). The potential cognitive-physical interaction has also been investigated through more intense exercise conditions, such as cycling (Brisswalter et al., 1995) or rowing (Duckworth et al., 2021). For instance, Brisswalter et al. (Brisswalter et al., 1995) reported a U-shape relationship between mean RT performance in a simple probe task and pedal rate, and a linear relationship between mean RT and VO_{2max}.

In cognitive-physical dual-task experiments, time series of physical and cognitive performance data are averaged within blocks of given lengths and analyzed using parametric or non-parametric statistical tools (analysis of variance, t-tests, Mann-Whitney test, linear regression etc.) (e.g., Brisswalter et al., 1995). Instead, here we use a novel approach based on Granger Causality (GC) (Granger, 1969), in order to exploit the information contained in the temporal fluctuations in these measures to quantify their interdependencies and respective modulations. GC analysis involves building an autoregressive model to predict the future values of the system under consideration, given its past. The prediction on the target given by the values of its own past is then compared with the one to which also the past values of the candidate driver are included. If the prediction improves (i.e., if the candidate driver adds relevant information on the future values of the target above and beyond the information provided by the past values of the driver alone) it is said that the driver has a Granger influence on (or Granger causes) the driver. GC is then a proxy for a dynamical influence. Here we applied a multiscale version of time-domain GC (Faes et al., 2017) to capture influences at different temporal scales. It is important to stress that GC is informative on effects/behaviors, as opposed to mechanisms (Barrett & Barnett, 2013).

The objective of the present study was twofold. First, we aimed at investigating temporal dynamics during dual-tasking involving self-paced cycling. This was addressed by applying GC to a dataset (Dataset 1) already published by our research group (Zandonai et al., 2021). Second, we studied the role of expertise in the potential interaction between physical and cognitive performance during a cycling self-paced time trial. If cyclists learn to efficiently self-pace through experience (Brick et al., 2016; Edwards & Polman, 2013; Holgado & Sanabria, 2021), they would reduce the need for exerting top-down control during cycling physical efforts. One would then expect a null or small interaction between the physical and cognitive tasks in expert cyclists, and a reliable interaction in non-expert endurance athletes, such as runners or swimmers.

Our specific pre-registered hypotheses https://doi.org/10.17605/OSF.IO/6QAR5 were the following[1]: 1) power output will influence heart rate (HR) in both experts and non-experts; 2) a significant bidirectional influence between RT and power output in the cycling task will be shown only in non-experts; 3) experts will outperform non-experts in the cycling task, resulting in a longer distance covered in the 30-minute session, higher watts/kg ratio developed during the experiment, higher HR, and higher ratings of perceived effort (RPE); 4) experts will outperform non-experts in the RT task, resulting in shorter overall RT, reduced congruency effect and conflict adaptation (see below for a description of the task).

2 Materials and Methods

2.1 Participants

[1] Note that the pre-registration included the recording of muscle oxygenation saturation and hemoglobin, but we could not finally do it due to technical issues.

2.1.1 Dataset 1

For Dataset 1, we re-analysed data from 23 healthy expert cyclists from a previous study (Zandonai et al., 2021). In the original experiment, 29 subjects completed an intense cycling exercise session under 3 different conditions: tramadol, paracetamol and placebo. Here, we analyzed the data from 23 participants (6 were discarded for technical issues) under the placebo condition to avoid any potential moderator effect of the drugs.

2.1.2 Dataset 2

For Dataset 2, we planned to collect data from 100 healthy athletes, 50 experienced cyclists and 50 non-cycling endurance athletes (i.e., runners and swimmers). Given the difficulty of estimating an effect size a priori, we aimed for a large sample size based on our previous experience recruiting this type of participants. In addition, we planned to monitor the Bayes Factor (BF) for between-group differences in GC parameters and the other dependent variables, and to stop the experiment whenever the BF reached moderate evidence to support (BF>6) or reject the null hypothesis (BF<1/6). Finally, due to the time and budget constraints, we recruited a total of 44 participants, composed of 21 expert cyclists (20 males, mean age 31.95 years, range 18-55 years) and 22 non-expert athletes (runners and swimmers; 17 males, mean age 25.63 years, range 18-55 years). Both expert cyclists and non-expert endurance athletes had at least 3 years of experience in their sport with a training routine of 4 or more days per week. We ensured that the non-experts did not include cycling in their training routine and had no previous cycling experience. Exclusion criteria were the presence of symptomatic cardiomyopathy, metabolic disorders, chronic obstructive pulmonary disease, epilepsy, therapy with b-blockers or medications that would alter cardiovascular function, hormonal therapy, smoking, or neurological disorders. Before taking part in the experiment, participants were informed about the experiment and provided written consent. They received a compensation of 10€ for their participation in the experiment. The experiment was approved by the local ethical committee (978/CEIH/2019) and was conducted following the Declaration of Helsinki.

2.2 Experimental design and procedure

2.2.1 Dataset 1

Dataset 1 corresponded to a self-paced high-intensity cycling session (indoor time-trial) lasting 20 min, in which participants were told to perform their best, avoiding premature extenuation, while responding, as fast and accurately as possible, to the Sustained Attention to Response Task (SART) (Robertson et al., 1997). More details about the procedure can be found in the original article (Zandonai et al., 2021).

2.2.2 Dataset 2

2.2.2.1 Design and procedure

Dataset 2 consisted of a between-participants design, with the main independent variable of Expertise (experts vs. non-experts). Participants performed a 30 min indoor high intensity self-paced cycling session and an auditory Simon task simultaneously. Participants were asked to maintain their coffee intake habit (i.e., to avoid it if not used to) and refrain from taking any other stimulants for 8 h before the experimental session, as well as avoid any intense physical exercise 24 h prior to the test (as in Dataset 1). When participants arrived at the laboratory, the cycle ergometer (SRM indoor trainer, SRM, Germany) was adjusted to their preferences. The experimenter adjusted the chest heart rate monitor (H10, Polar Electro, Kempele, Finland).

Power output was measured using the SRM indoor trainer and the Favero Assioma pedals (Favero Electronics SLR, Arcade, Italy). Auditory stimuli were presented through in-ear earphones (Hyperx, HP Inc., USA).

2.2.2.2 Cycling session

The session started with a 10-min warm-up at a power corresponding between 1.5 and 2.5 W·kg^{-1}. They were instructed to achieve the maximum mean power possible during the 30 min. Power and HR data were collected at a frequency of 1Hz. Perceived cognitive and physical effort was measured using a visual analogue scale.

2.2.2.3 Task

Participants were asked to perform an auditory version of the classic Simon task (Simon & Rudell, 1967). Recordings of the spoken words *"izquierda"* (left in Spanish) and *"derecha"* (right in Spanish) were presented to participants through the right or left earphone. Stimuli were considered congruent when the word meaning corresponded to the side from which they were played (e.g., listening to the word *"derecha"* through the right earphone) and incongruent when they did not (e.g., listening to the word *"derecha"* through the left earphone). Participants were asked to report by button pressing (with their thumb) the location depicted by the word meaning while ignoring its physical location (see Figure 1). For instance, if the word presented was *"izquierda"* the participant had to press a button with his left hand regardless of which headphone the word was played from. If the word presented was *"derecha"* the participant had to press a button with his right hand regardless of which headphone the word was played from. The response devices were placed on both sides of the bike's handlebar, so the participant was able to respond without moving the hands away from the bike. Speed and accuracy were stressed. Participants had a maximum of 750 ms to respond, after which responses were discarded. The participant's response and the following trial were separated by an Interstimulus interval (ITI) which was a random number between 800 and 1200 ms. If the participant did not respond, the ITI began after the response window of 750 ms ended. Testing was fixed in time duration, therefore the total number of trials depended on the participant's speed. During the warm-up, participants familiarized with the task through a 30 seconds practice block.

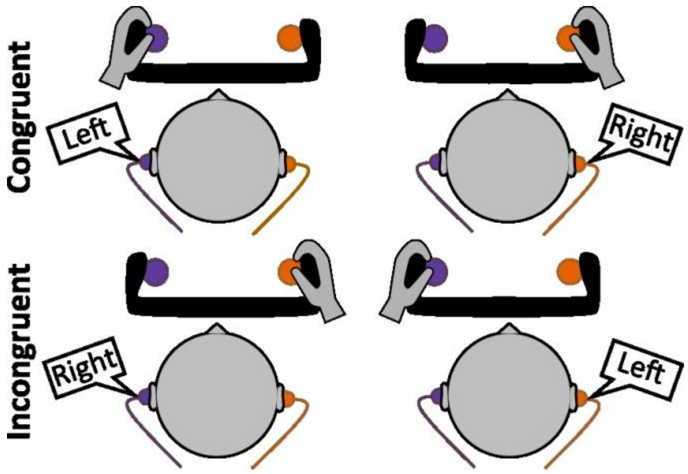


Figure 1: Auditory conflict task. In each trial, participants either heard the word *"left"* or *"right"* played from the left or right headphone. The trials were congruent when the meaning of the word matched the location from where it was played, or incongruent if otherwise. Presentation of congruent or incongruent trials were equally probable and randomized. Participants responded by button pressing. They were instructed to press either the left or right button according to the meaning of the word, ignoring the location from which it was played.

2.3 Preprocessing and statistical analysis

For Dataset 1, the preprocessing followed the procedure reported in Zandonai and colleagues (2021). For Dataset 2, behavioral data corresponding to incorrect responses and omissions were removed. To match the time series to RT, we selected the time series sampling points closest to the time points of the behavioral responses.

For both datasets, time series were detrended with the l_1 norm (Kim et al., 2009) and standardized prior to the analysis. We used time-domain GC (Granger, 1969), which establishes whether an autoregressive model of a target time series improves when another time series is included in the model, acting as a proxy for a dynamical influence. Among the several modifications to Granger's original conceptualization (see Shohaje & Fox, 2022 for a review), here we use a multiscale version of GC (Faes et al., 2017), allowing to assess Granger-causal influences broken down across several temporal time scales. The first scale contains all the temporal complexity of the time series (thus up to the Nyquist frequency). The second scale considers slower frequencies (up to half of the Nyquist frequency), the third one even slower (up to ⅓ of the Nyquist frequency) and so on. The approach used here, and described in detail in Faes and colleagues (2017), performs downsampling and averaging in a single step, allowing to mitigate problems arising by considering the two steps separately.

We downsampled the time series up to a factor 12, i.e. we used 12 scale values, in steps of 1. The order of the autoregressive model was chosen according to the Bayesian Information Criterion testing the values from 1 to 20.

In order to account for the simultaneous presence of short-term dynamics and long-range correlations, particularly prominent in the data under investigation, we complemented the model with a vector autoregressive fractionally integrated framework for Gaussian processes (Pinto et al., 2022).

Statistical significance was assessed for every participant by building a null distribution using iterative amplitude adjusted Fourier Transform surrogates preserving the spectrum (Schreiber & Schmitz, 1996), and checking whether the results fall outside the 95th percentile of the null distribution.

In Dataset 2, if significant effects were found in any of the participants, the BF (with the null hypothesis as denominator) for the GC parameters was calculated considering the independent variable “Expertise”. The BF was also calculated for the rest of between-group comparisons. To compare the conflict effect in the Simon task between groups, we computed the RT difference between incongruent and congruent trials for every participant. For the conflict adaptation effect, we first computed the congruency effect for previous congruent and previous incongruent trials, to then subtract the congruency effect of previous incongruent trials from that of previous congruent trials to obtain and index of the conflict adaptation effect.

Data and analytic code can be found on the OSF page of the project (https://doi.org/10.17605/OSF.IO/6QAR5).

3 Results

3.1 Granger causality

The GC analysis performed on Dataset 1 showed influence of power output on RT in only 4 participants (out of 23), influence of RT over power output in 5 participants, and influence of power output on HR in 11 participants (see Figures 2 and 3). In Dataset 2, RT to power influence was shown in 7 expert cyclists (out of 21), and 8 non-expert cyclists (out of 23). Larger GC values were obtained in all time scales, although BF analyses showed anecdotal evidence for the null in the case of time scales 1 to 5, and anecdotal evidence for the alternative hypothesis for time scales 6 to 12 (see Table 1 in the supplementary material). Power influence on RT was shown in 5 experts (out of 21) and 6 non-experts (out of 23). Again, larger values were obtained for non-experts than for experts in all time scales, albeit all between-group comparisons showed anecdotal evidence for the null (all BF_{10} <.72; see Table 1 in the supplementary material). As expected, there was an influence of power output and HR in 12 experts (out of 21) and 19 non-expert cyclists (out of 23), with larger values for non-experts than for experts in all time scales. Independent-samples BF t-tests showed anecdotal evidence for the null in all time scales (all BF_{10} <.78 see Table 1 in the supplementary material). Graphic representation of individual results are available in the supplementary material (supplementary figures 1 and 2 for the experts group of Dataset 1, supplementary figures 3 and 4 for the experts group of Dataset 2, and supplementary figures 5 and 6 for the non-experts group of Dataset 2).

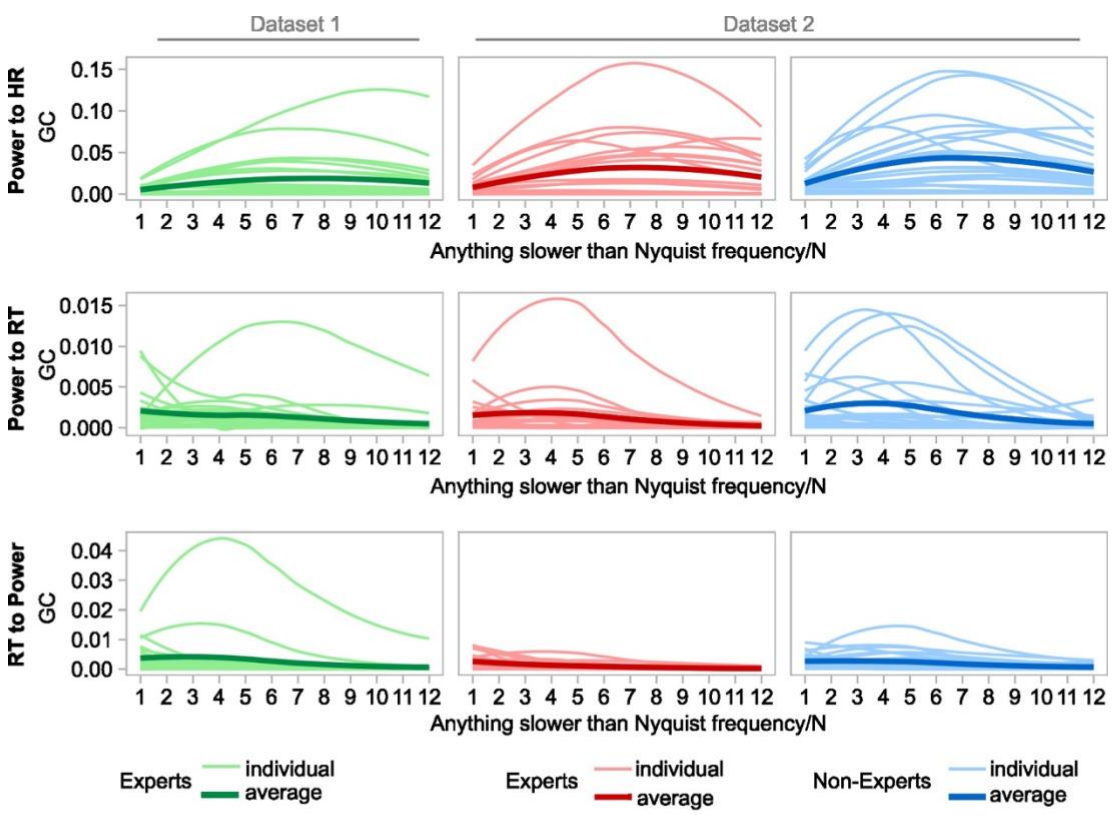
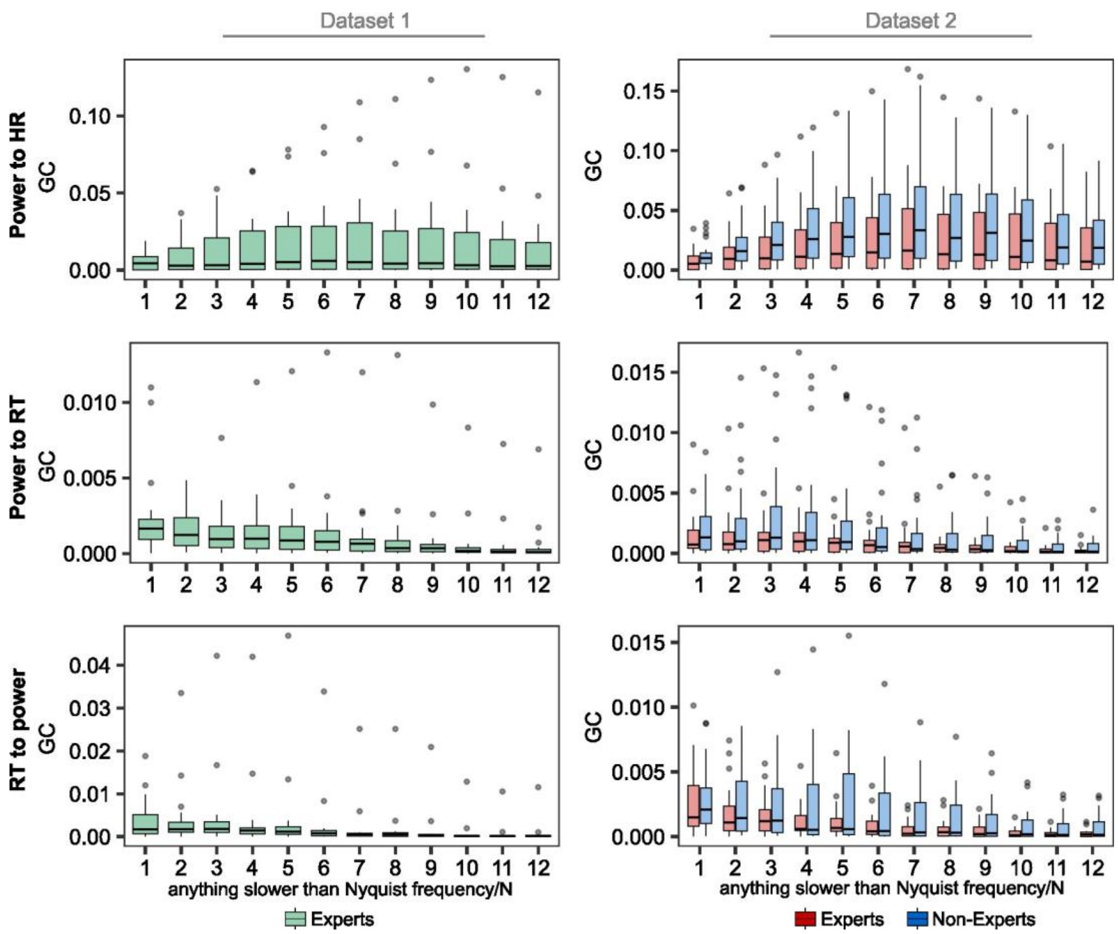


Figure 2. Individual GC estimates of power to HR (top row), power to RT (middle row), and RT to power (bottom row). Estimates are calculated over 12 time bins in the experts (Dataset 1), experts and non-experts (Dataset 2) groups. Thin lines represent individual participants' estimates and bold lines the group average.

3.2 Reaction times, heart rate and power output, perceived physical and cognitive effort

The analysis of the RT data showed anecdotal evidence for a null group difference in overall RT, BF_{10} = 0.45, anecdotal evidence for a larger congruency effect in non-experts than in experts, BF_{10} = 2.93, and anecdotal evidence for the null regarding the conflict adaptation effect, BF_{10} = 0.49. Strong evidence was shown for group differences in terms of overall power output, BF_{10} = 239189, relative power output (w/kg) BF_{10} = 145500000, and HR BF_{10} = 31.46. In terms of the perceived effort, the analysis showed anecdotal evidence for greater perceived physical



effort in the expert group than in the non-expert group, $BF_{10} = 1.844$, and anecdotal evidence for the null in the case of perceived cognitive effort, $BF_{10} = 0.43$ (see Figure 4).

Figure 3. Group GC estimates of power to HR (top row), power to RT (middle row), and RT to power (bottom row). Estimates are averaged over each of the 12 time bins in the experts and in the experts and non-experts (Dataset 2) groups. Box plots show the median (middle horizontal line), and 25th and 75th percentiles (bottom and top horizontal lines). The upper and lower whiskers indicate the 1.5 times the interquartile range above the 75th percentile and below the 25th percentile. Gray dots represent outlier values.

4 Discussion

People can perform two tasks at the same time but usually at the cost of shared resources and potential interaction effects (Pashler, 1994). This is what has been reported for the case of motor tasks, such as walking, and RT tasks (Al-Yahya et al., 2011), and could in turn explain at least part of the variance in studies investigating cognitive performance during physical exercise (e.g., cycling) that compare exercise condition(s) with a resting condition (Chang et al., 2014), or even in the case of two cycling conditions with different intensities (Ciria et al., 2019).

The results of Datasets 1 and 2 showed evidence of mutual interaction between power output and RT in some of the participants, with no evidence of group differences in GC indexes. In most cases when an interaction was present, its maximum was found at time scales slower than the one corresponding to the original sampling. For example, some individual and average curves peak at scale 7, corresponding to a frequency around 0.07 Hz, and to a period of about 14 seconds. In other words, the dynamical processes of the driver time series which are more informative in predicting the dynamical processes of the target time series are located in a temporal range centered at this frequency. This lack of strong cognitive-physical interactions in any of the samples tested here contrasts with the evidence for better physical performance of the expert cyclists, and anecdotal evidence for better performance (i.e., reduced congruency effect) in the cognitive task. The expected influence of power output on HR was detected in 11 out of 23 participants in Dataset 1 and 31 out of 44 participants in Dataset 2, which reflects the impact of workload on heart response (McCarthy & Wyatt, 2003).

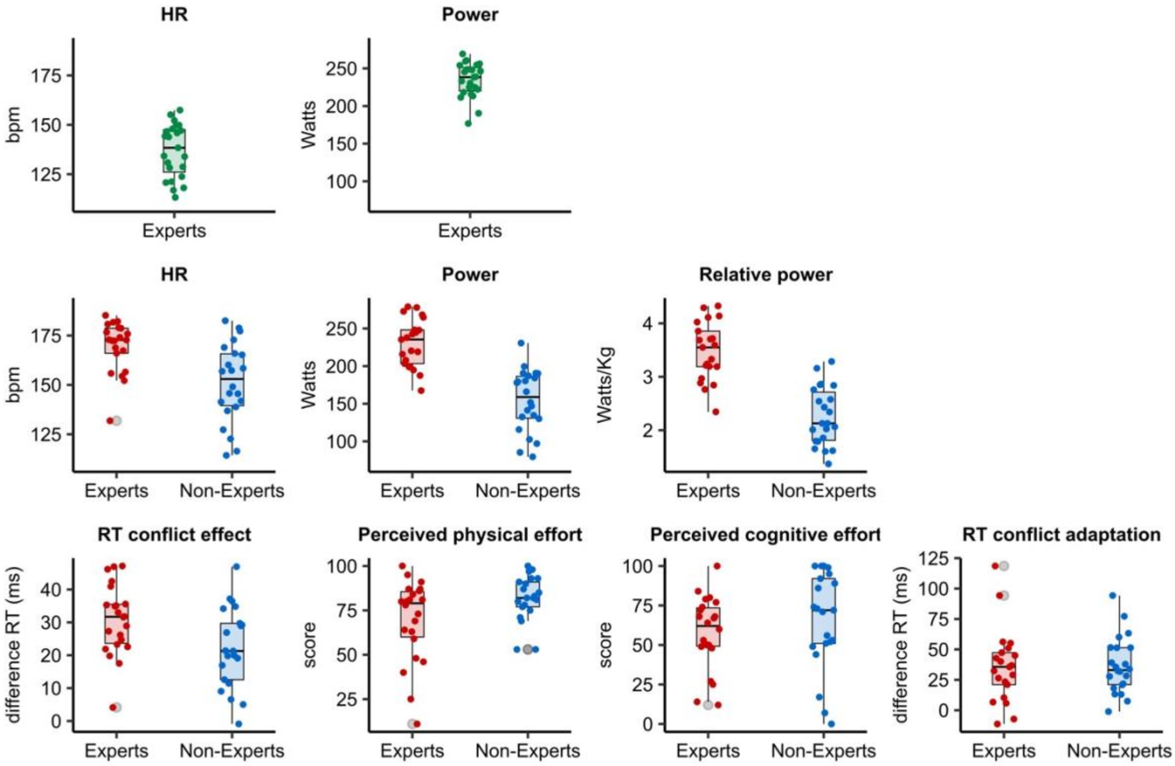


Figure 4. Group representation of the main variables measured in Dataset 1 and Dataset 2. Box plots depict the median (middle horizontal line), and 25th and 75th percentiles (bottom and top horizontal lines). The upper and lower whiskers indicate the 1.5 times the interquartile range above the 75th percentile and below the 25th percentile. Jittered dots are individual participants' means.

The GC results in both studies suggest that cognitive and physical performance can interact in some individuals, at least for the case of RT and self-paced high intense indoor cycling. Given that only a small portion of the sample showed that cognitive-physical interaction pattern, and that no clear group differences were reported, a potential ad-hoc explanation points to individual differences in self-pacing capacities/strategies, regardless of the particular expertise in the physical/motor cycling task. The lack of relationship between the RT and power output time series in the majority of the participants in Studies 1 and 2 could also be due to the use of a stationary indoor bike in a laboratory. Maintaining the desired cadence and effort pace seems much easier indoors than outdoors, where cyclists have to keep attending while riding in a changing environment and react rapidly to unexpected events that could compromise their safety (e.g., a pothole on the road, or a dog crossing the road). Hence, our results cannot be directly extrapolated to real cycling contexts, where the likelihood of physical and cognitive performance mutual influence might increase, and expertise could play a crucial role.

In our study, however, expertise in cycling did not seem to be important according to the results of the BF analysis, even if larger GC values were shown for non-experts than for experts for both the influence of RT to power output and power output to RT. Participants in the non-expert group had no prior experience in cycling, but were endurance athletes with experience in self-paced efforts, thereby explaining, at least partially, the lack of group differences in the GC indexes. In contrast, the analysis of central tendency measures, commonly used in this type of studies, showed strong evidence for group differences in terms of power output and HR, and anecdotal evidence for superior cognitive performance in experts. Together, these results could be taken as evidence of lack of dual-tasks interaction effects, as GC shows that better physical performance was not detrimental to cognitive performance in the expert cyclists group. In any case, the GC approach used here certainly provides more valuable information than those central tendency measures, at least for the purpose of looking at potential physical-cognitive performance interactions.

In conclusion, our study brings two major contributions. One is the evidence that dual-task effects are likely to be negligible in expert cyclists, suggesting that experiments using paradigms in which the experimental manipulation concerns exercise intensity may be robust to dual-task confounds, although interactions between power output and RT cannot be discarded, at least in some participants. However, our study did not allow us to neatly isolate the effects of expertise in such a relationship. We recruited non-cyclist athletes to control for fitness levels and address cycling-specific expertise. Nevertheless, endurance athletes regardless of discipline may still acquire important experience in self-paced physical exercise. As a result, our findings cannot be generalized to any other group than those considered in this study as, for example, non-athletes. Another important contribution concerns the use of GC as a way of determining the potential dynamical influence between time series of physical and cognitive performance data in research on cognitive performance during cycling. Our results highlight the importance of GC measures in carefully assessing individual cognitive-physical interaction beyond group effects.

Acknowledgments

This study was supported by a research project grant from the Spanish Ministry of Science and Innovation to Daniel Sanabria (PID2019-105635GB-I00); a postdoctoral fellowship by the Spanish Ministry for Science and Innovation awarded to Chiara Avancini (FJC2020-046310-

I); a postdoctoral fellowship from the Regional Government of Andalusia awarded to Luis F. Ciria (DOC_00225); a predoctoral fellowships by the Spanish Ministry of Universities awarded to Juan José Pérez-Díaz (FPU20/00611) and to José-Antonio Salas-Montoro (FPU17/04742).

Competing interests statement

The authors declare that they have no known competing interests.

Authors' contributions

Chiara Avancini: Methodology, Software, Experiment setup, Data curation, Writing, Reviewing, Editing, Figures; **Daniele Marinazzo:** Methodology, Analysis, Software, Writing, Reviewing, Editing; **Daniel Sanabria:** Conceptualization, Methodology, Analysis, Writing, Reviewing, Editing, Supervision; **Juan José Pérez-Díaz:** Data collection, Data curation; **José-Antonio Salas-Montoro:** Data collection, Reviewing; **Luis F. Ciria:** Conceptualization, Methodology, Writing, Reviewing, Editing, Supervision. All authors have read and approved the final version of the manuscript, and agree with the order of presentation of the authors.

References

Al-Yahya, E., Dawes, H., Smith, L., Dennis, A., Howells, K., & Cockburn, J. (2011). Cognitive motor interference while walking: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *35*(3), 715–728. https://doi.org/10.1016/j.neubiorev.2010.08.008

Audiffren, M., Tomporowski, P. D., & Zagrodnik, J. (2008). Acute aerobic exercise and information processing: Energizing motor processes during a choice reaction time task. *Acta Psychologica*, *129*(3), 410–419. https://doi.org/10.1016/j.actpsy.2008.09.006

Barrett, A., & Barnett, L. (2013). Granger causality is designed to measure effect, not mechanism. *Frontiers in Neuroinformatics*, *7*. https://www.frontiersin.org/articles/10.3389/fninf.2013.00006

Brick, N. E., MacIntyre, T. E., & Campbell, M. J. (2016). Thinking and Action: A Cognitive Perspective on Self-Regulation during Endurance Performance. *Frontiers in Physiology*, *7*. https://doi.org/10.3389/fphys.2016.00159

Brisswalter, J., Durand, M., Delignieres, D., & Legros, P. (1995). Optimal and non-optimal demand in a dual-task of pedaling and simple reaction time: Effects on energy expenditure and cognitive performance. *Journal of Human Movement Studies*, *29*, 15–34.

Chang, Y.-K., Chi, L., Etnier, J. L., Wang, C.-C., Chu, C.-H., & Zhou, C. (2014). Effect of acute aerobic exercise on cognitive performance: Role of cardiovascular fitness. *Psychology of Sport and Exercise*, *15*(5), 464–470. https://doi.org/10.1016/j.psychsport.2014.04.007

Ciria, L. F., Luque-Casado, A., Sanabria, D., Holgado, D., Ivanov, P. Ch., & Perakakis, P. (2019). Oscillatory brain activity during acute exercise: Tonic and transient neural response to an oddball task. *Psychophysiology*, *56*(5), e13326. https://doi.org/10.1111/psyp.13326

Ciria, L. F., Perakakis, P., Luque-Casado, A., & Sanabria, D. (2018). Physical exercise increases overall brain oscillatory activity but does not influence inhibitory control in young adults. *NeuroImage*, *181*, 203–210. https://doi.org/10.1016/j.neuroimage.2018.07.009

Duckworth, S. C., Higginbotham, C. S., Pederson, J. A., Rogers, R. R., Marshall, M. R., Williams, T. D., & Ballmann, C. G. (2021). Physical and Cognitive Performance During Upper-Extremity Versus Full-Body Exercise Under Dual Tasking Conditions. *Perceptual and Motor Skills*, *128*(1), 338–352. https://doi.org/10.1177/0031512520945088

Edwards, A. M., & Polman, R. C. J. (2013). Pacing and awareness: Brain regulation of physical activity. *Sports Medicine (Auckland, N.Z.)*, *43*(11), 1057–1064. https://doi.org/10.1007/s40279-013-0091-4

Faes, L., Nollo, G., Stramaglia, S., & Marinazzo, D. (2017). Multiscale Granger causality. *Physical Review E*, *96*(4), 042150. https://doi.org/10.1103/PhysRevE.96.042150

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, *37*(3), 424–438. https://doi.org/10.2307/1912791

Holgado, D., & Sanabria, D. (2021). Does self-paced exercise depend on executive processing? A narrative review of the current evidence. *International Review of Sport and Exercise Psychology*, *14*(1), 130–153. https://doi.org/10.1080/1750984X.2020.1774915

Kim, S.-J., Koh, K., Boyd, S., & Gorinevsky, D. (2009). L1 Trend Filtering. *SIAM Review*, *51*(2), 339–360. https://doi.org/10.1137/070690274

McCarthy, J. P., & Wyatt, F. B. (2003). Prediction equation: Power output from heart rate for cyclists. *International Sports Journal*, *7*(1), 56–61.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*(2), 220–244. https://doi.org/10.1037/0033-2909.116.2.220

Pinto, H., Pernice, R., Silva, M. E., Javorka, M., Faes, L., & Rocha, A. P. (2022). Multiscale partial information decomposition of dynamic processes with short and long-range correlations: Theory and application to cardiovascular control. *Physiological Measurement*, *43*(8), 085004. https://doi.org/10.1088/1361-6579/ac826c

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). `Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, *35*(6), 747–758. https://doi.org/10.1016/S0028-3932(97)00015-8

Schreiber, T., & Schmitz, A. (1996). Improved Surrogate Data for Nonlinearity Tests. *Physical Review Letters*, *77*(4), 635–638. https://doi.org/10.1103/PhysRevLett.77.635

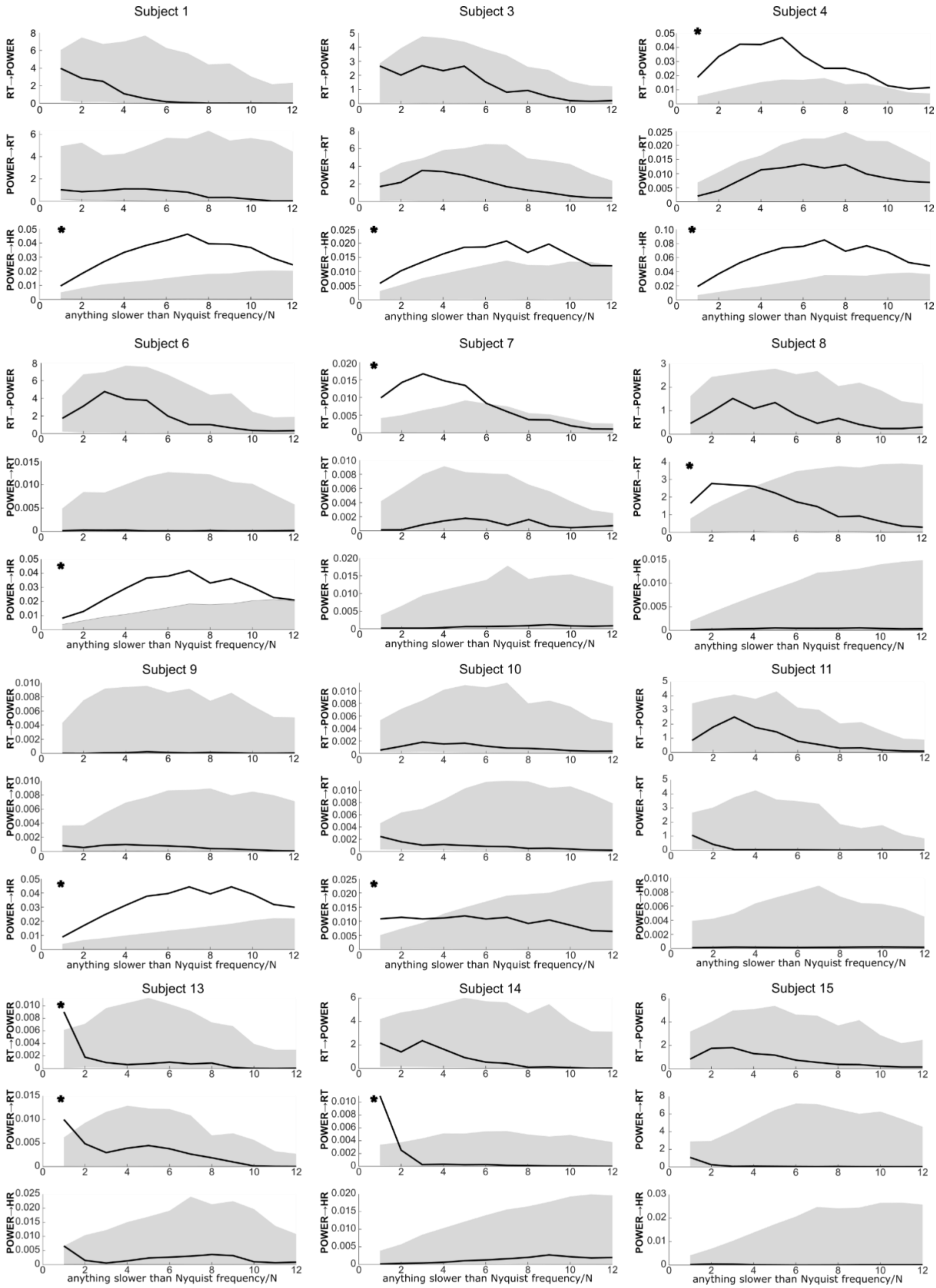
Shojaie, A., & Fox, E. B. (2022). Granger causality: A review and recent advances. *Annual Review of Statistics and Its Application*, *9*, 289-319.

Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology, 51*(3), 300–304.

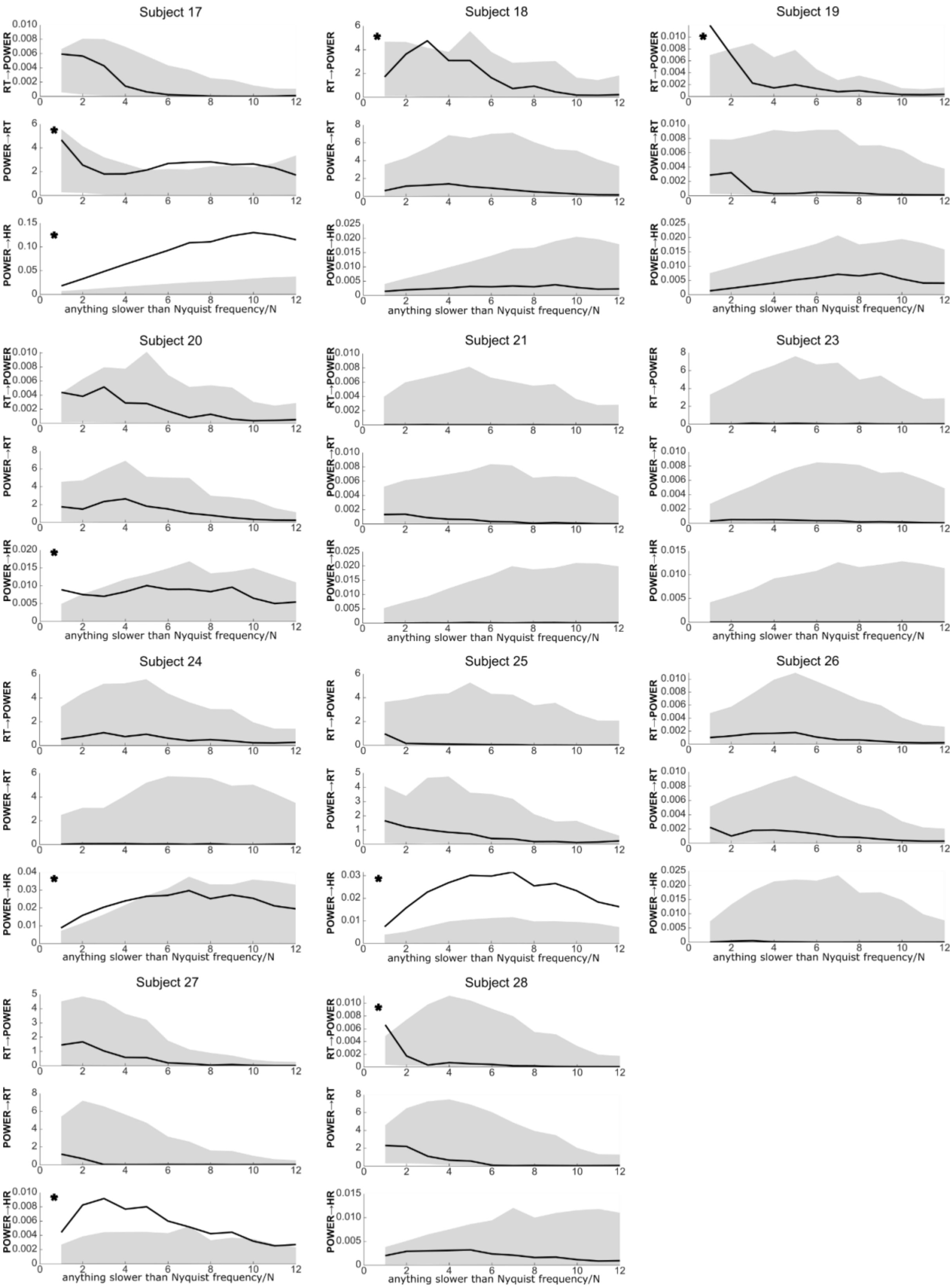
https://doi.org/10.1037/h0020586

Zandonai, T., Holgado, D., Ciria, L. F., Zabala, M., Hopker, J., Bekinschtein, T., & Sanabria, D. (2021). Novel evidence on the effect of tramadol on self-paced high-intensity cycling. *Journal of Sports Sciences*, *39*(13), 1452–1460. https://doi.org/10.1080/02640414.2021.1877440

Supplementary material

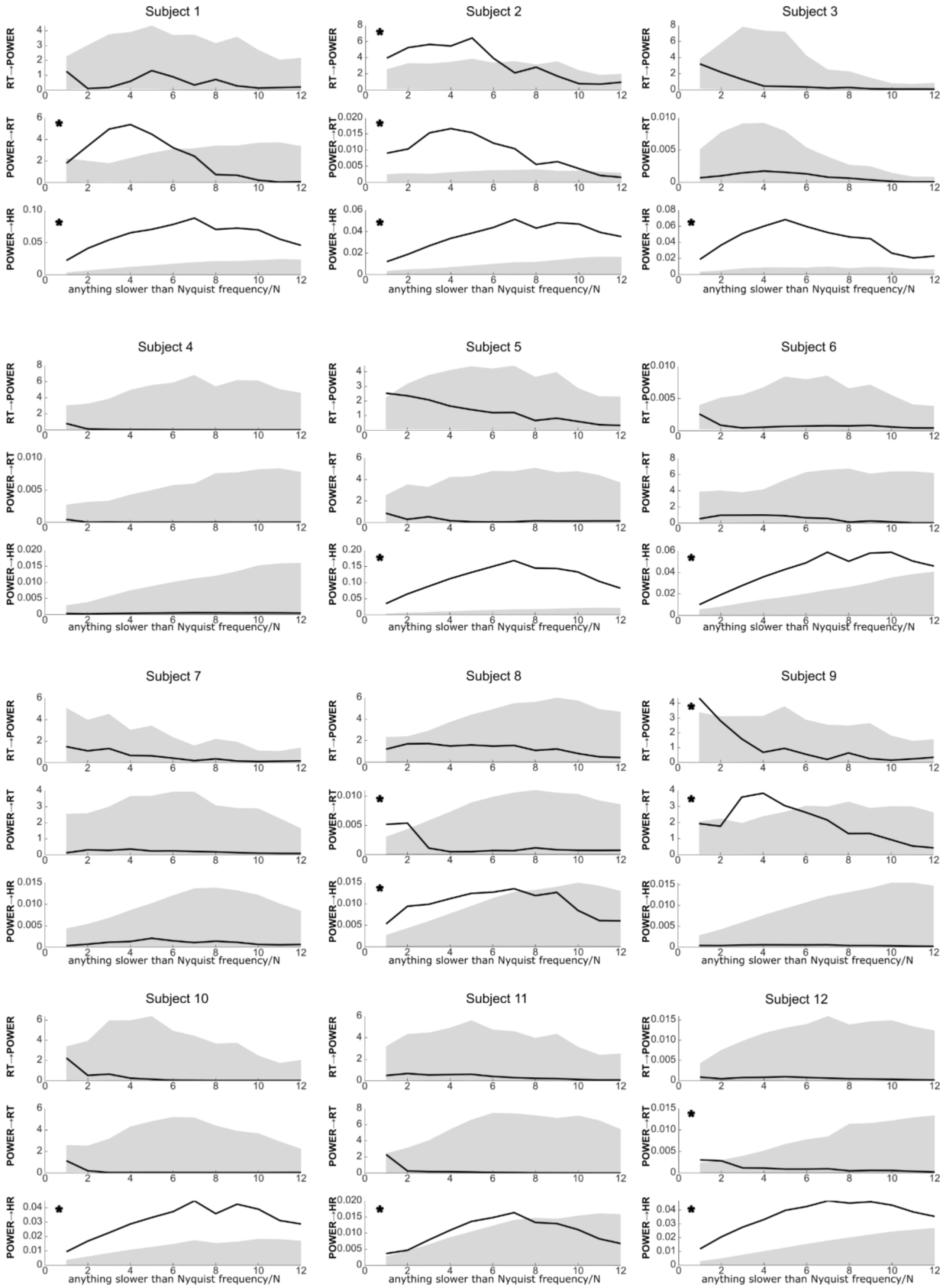

Subject 1
Subject 3
Subject 4
Subject 6
Subject 7
Subject 8
Subject 9
Subject 10
Subject 11
Subject 13
Subject 14
Subject 15
RT→POWER
POWER→RT
POWER→HR
anything slower than Nyquist frequency/N

Supplementary figure 1. Individual GC estimates of the experts group from Dataset 1 (subject 1 to subject 15). Estimates of power to HR, power to RT, and RT to power are calculated over 12 time bins. Bold lines represent individual participants' estimates and the gray shadowing the 5th to 95th percentiles of the surrogate null distribution. The asterisks denote results that fall outside the 95th percentile of the null distribution. Note that participants' numbers reflect the numbers assigned during data collection.

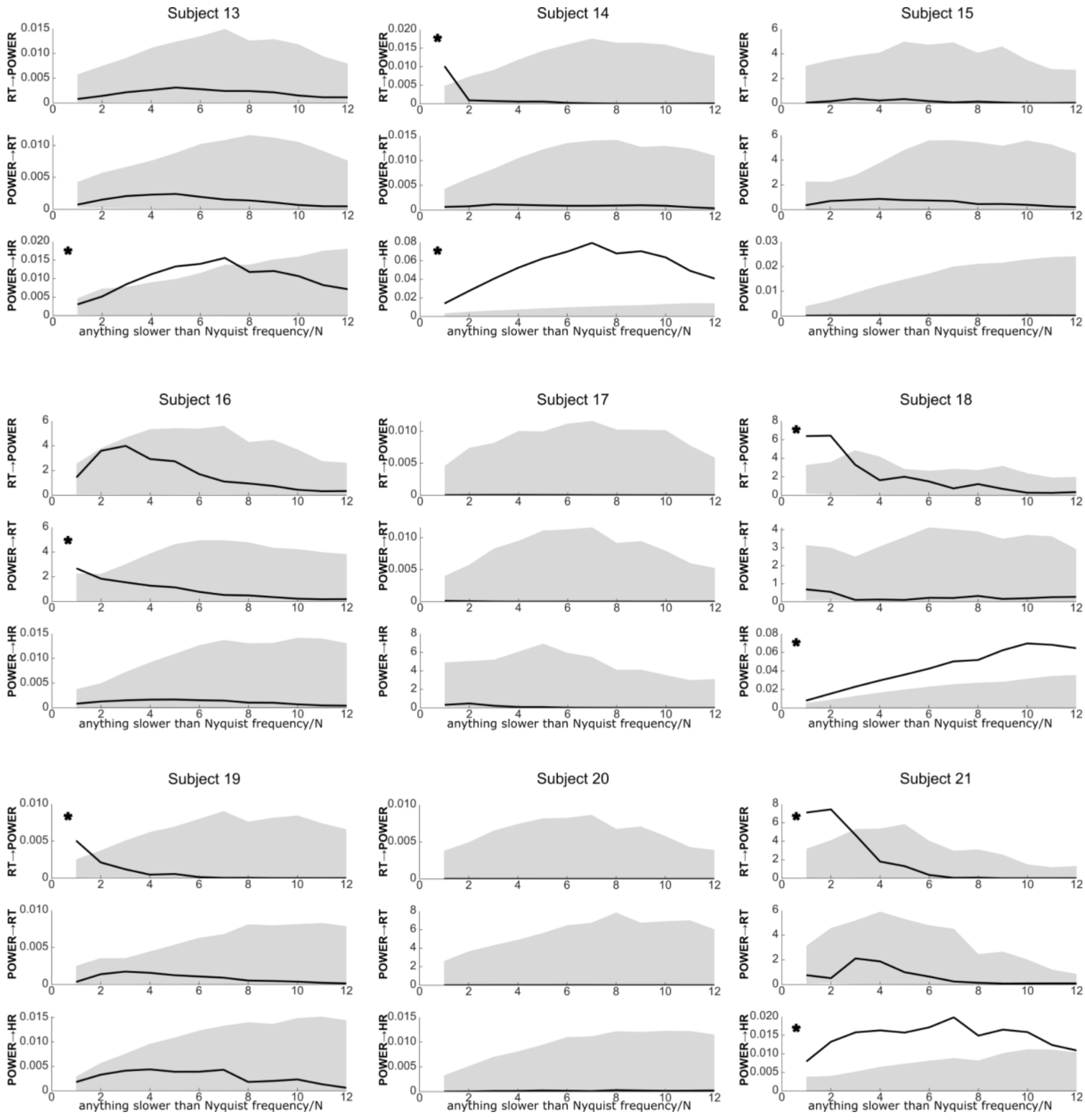


Supplementary figure 2. Individual GC estimates of the experts group from Dataset 1 (subject 17 to subject 28). Estimates of power to HR, power to RT, and RT to power are calculated over

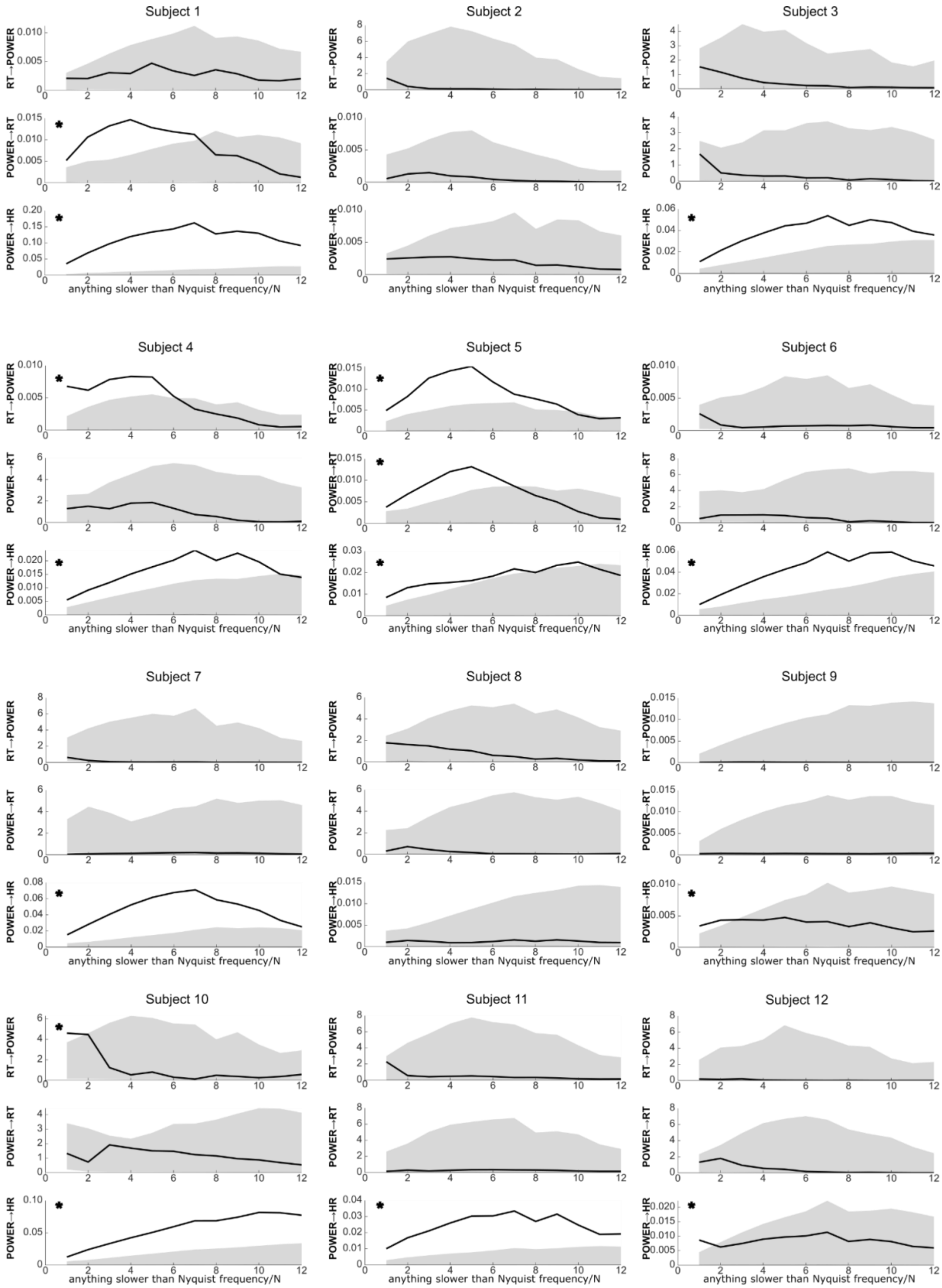
12 time bins. Bold lines represent individual participants' estimates and the gray shadowing the 5th to 95th percentiles of the surrogate null distribution. The asterisks denote results that fall outside the 99th percentile of the null distribution. Note that participants' numbers reflect the numbers assigned during data collection.


Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Subject 6
Subject 7
Subject 8
Subject 9
Subject 10
Subject 11
Subject 12
RT→POWER
POWER→RT
POWER→HR
anything slower than Nyquist frequency/N
*

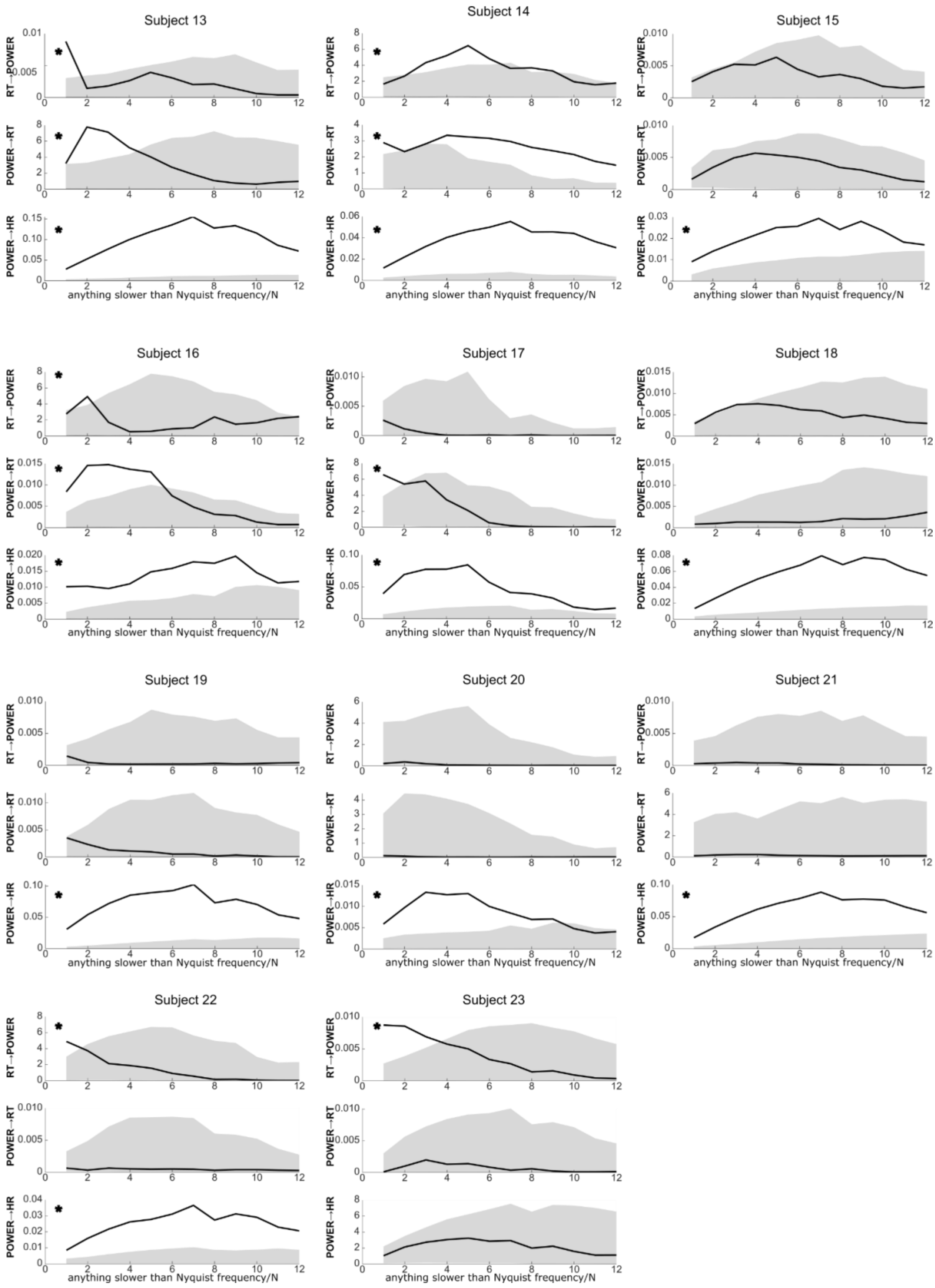
Supplementary figure 3. Individual GC estimates of the experts group from Dataset 2 (subject 1 to subject 12). Estimates of power to HR, power to RT, and RT to power are calculated over 12 time bins. Bold lines represent individual participants' estimates and the gray shadowing the 5th to 95th percentiles of the surrogate null distribution. The asterisks denote results that fall outside the 95th percentile of the null distribution.



Supplementary figure 4. Individual GC estimates of the experts group from Dataset 2 (subject 13 to subject 21). Estimates of power to HR, power to RT, and RT to power are calculated over 12 time bins. Bold lines represent individual participants' estimates and the gray shadowing the 5th to 95th percentiles of the surrogate null distribution. The asterisks denote results that fall outside the 95th percentile of the null distribution.


Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Subject 6
Subject 7
Subject 8
Subject 9
Subject 10
Subject 11
Subject 12
RT→POWER
POWER→RT
POWER→HR
*
anything slower than Nyquist frequency/N

Supplementary figure 5. Individual GC estimates of the non-experts group from Dataset 2 (subject 1 to subject 12). Estimates of power to HR, power to RT, and RT to power are calculated over 12 time bins. Bold lines represent individual participants' estimates and the gray shadowing the 5th to 95th percentiles of the surrogate null distribution. The asterisks denote results that fall outside the 95th percentile of the null distribution.


Subject 13
Subject 14
Subject 15
Subject 16
Subject 17
Subject 18
Subject 19
Subject 20
Subject 21
Subject 22
Subject 23
RT→POWER
POWER→RT
POWER→HR
anything slower than Nyquist frequency/N

Supplementary figure 6. Individual GC estimates of the non-experts group from Dataset 2 (subject 13 to subject 23). Estimates of power to HR, power to RT, and RT to power are calculated over 12 time bins. Bold lines represent individual participants' estimates and the gray shadowing the 5th to 95th percentiles of the surrogate null distribution. The asterisks denote results that fall outside the 95th percentile of the null distribution.

	RT to power		**power to RT**		**Power to HR**	
scale	**BF_{10}**	**error %**	**BF_{10}**	**error %**	**BF_{10}**	**error %**
1	0.299	0.006	0.348	0.006	0.782	0.006
2	0.419	0.006	0.515	0.006	0.625	0.006
3	0.582	0.006	0.464	0.006	0.566	0.006
4	0.914	0.007	0.410	0.006	0.521	0.006
5	0.916	0.007	0.436	0.006	0.499	0.006
6	1010	0.007	0.408	0.006	0.443	0.006
7	1.202	0.007	0.393	0.006	0.427	0.006
8	1.060	0.007	0.522	0.006	0.400	0.006
9	1.260	0.007	0.423	0.006	0.406	0.006
10	1.310	0.007	0.446	0.006	0.386	0.006
11	1.484	0.007	0.645	0.006	0.379	0.006
12	1.592	0.008	0.729	0.006	0.394	0.006

Table 1: Results of Bayesian independent-samples t-tests comparing GC parameters of the experts against the non-experts group of Dataset 2. BF are calculated with the null hypothesis as denominator, meaning that the larger the value, the more evidence is provided for the alternative hypothesis.