Placebo in sports nutrition: a proof-of-principle study involving caffeine supplementation

Original article

Running head: Placebo in sports nutrition

Bryan Saunders[1], Luana Farias de Oliveira[1], Rafael Pires da Silva[1], Vitor de Salles Painelli[1], Livia Souza Gonçalves[1], Guilherme Yamaguchi[1], Thiago Mutti[1], Erika Maciel[3], Hamilton Roschel[1, 2], Guilherme Giannini Artioli[1, 2], Bruno Gualano[1, 2]

[1]Applied Physiology & Nutrition Research Group, University of São Paulo, Brazil.

[2] Rheumatology Division, School of Medicine, University of São Paulo, Brazil.

Correspondence:

Bruno Gualano

Av. Mello de Moraes 65

Butanta, 05508-030

Sao Paulo, SP, Brazil.

Phone: +55 11 3091-3096

Fax: +55 11 3813-5921

E-mail: gualano@usp.br

Abstract

We investigated the effects of supplement identification on exercise performance with caffeine supplementation. Forty-two trained cyclists (age 37±8 y, body mass [BM] 74.3±8.4 kg, height 1.76±0.06 m, maximum oxygen uptake 50.0±6.8 $ml \cdot kg^{-1} \cdot min^{-1}$) performed a ~30 min cycling time-trial 1 h following either 6 $mg \cdot kg^{-1}$BM caffeine (CAF) or placebo (PLA) supplementation and one control (CON) session without supplementation. Participants identified which supplement they believed they had ingested ("caffeine", "placebo", "don't know") pre- and post-exercise. Subsequently, participants were allocated to subgroups for analysis according to their identifications. Overall and subgroup analyses were performed using mixed-model and magnitude based inference analyses. Caffeine improved performance vs. PLA and CON ($P \leq 0.001$). Correct pre- and post-exercise identification of caffeine in CAF improved exercise performance (+4.8 and +6.5%) vs. CON, with slightly greater relative increases than the overall effect of caffeine (+4.1%). Performance was not different between PLA and CON within subgroups (all $P > 0.05$), although there was a tendency towards improved performance when participants believed they had ingested caffeine post-exercise ($P=0.06$; 87% *likely beneficial*). Participants who correctly identified placebo in PLA showed *possible harmful effects* on performance compared to CON. Supplement identification appeared to influence exercise outcome and may be a source of bias in sports nutrition.

Key words: Placebo effect; nocebo effect; expectancy; exercise performance; caffeine supplementation; supplement identification; cycling time-trial

Introduction

Contemporary investigations into the effects of nutritional interventions on exercise generally employ double-blind and placebo controlled study designs to ensure there is no bias from the prior knowledge of which substance has been ingested and that comparisons can be made against an appropriate control. The placebo effect, namely a positive outcome brought about purely from the belief that one has received a positive intervention (Clark et al., 2000), can mask the true effect of an intervention. The nocebo effect is directly opposite to this in that a negative outcome occurs following the administration of an intervention (Benedetti et al., 2007; (Lundby et al,, 2012; Pollo et al., 2012).

Caffeine-based investigations can be difficult to blind due to the associated side-effects at high doses (*i.e.*, >2-3 $mg \cdot kg^{-1}BM$), namely tachycardia and agitation (Graham & Spriet, 1995), and common knowledge thereof. Once an individual believes that they have ingested a performance enhancing substance, several behaviours may be modified that can contribute to exercise performance (Beedie et al., 2006). This may lead to many of the participants beginning exercise with a greater expectancy due to the occurrence of physiological side effects making it difficult to separate the true effect of caffeine from its associated placebo effect. However, most studies do not control whether blinding of the intervention was successful; determination of an individual's belief of what they have ingested prior to exercise may lead to further investigation into the effects of preconceptions (placebo effect) on exercise.

In addition to preconceptions, it would be reasonable to suggest that any behavioural processes that might have been modified prior to exercise might also change throughout exercise on the basis of new information (Beedie et al., 2006). This might relate to an individual's perceived effort throughout exercise, which may or may not be influenced by the intervention itself. An individual who believed they had ingested placebo prior to exercise but then changes opinion due to a good start may influence their pacing accordingly throughout the test. Conversely, someone who expects to improve performance due to preconceived opinion of ingesting the active substance, but subsequently struggles to perform, might suffer a reduction in performance due to a further lack of motivation. Therefore, it would also be of interest to determine the individual's perception of what was ingested following exercise to determine whether the initial opinion has been modified throughout the protocol.

Therefore, to advance the knowledge on the influence of the placebo effect in sports nutrition, we investigated the effect of supplement identification following caffeine ingestion on exercise performance. We hypothesised that caffeine supplementation would improve exercise performance regardless of proper identification, and that improvements would be greatest in those who correctly guessed they had taken caffeine, while participants ingesting placebo but guessed they had ingested caffeine would also improve their exercise performance.

Materials and Methods

Participants

Forty-two trained male cyclists (Table 1) volunteered and gave their written informed consent to participate in this study. The exclusion criteria included the use of beta-alanine and creatine in the past 6 months, the presence of any musculoskeletal disorder, or the current or past use of anabolic steroids or other illicit performance-enhancing drugs. Habitual caffeine consumption (Table 1) was assessed prior to inclusion in the study via a Food Frequency Questionnaire adapted from two previously developed and validated questionnaires (Bühler et al., 2014 and Fred Hutchinson Cancer Research Center, 2004). Although these data were not used to exclude any participant *per se*, any participant ingesting caffeine as a dietary supplement was not included in the study since these individuals may or may not have been more susceptible to correct supplement identification due to experience. The study was approved by the University of São Paulo's Ethics Review Committee as part of a larger thematic project, the remaining data of which is presented elsewhere.

Experimental Design

All participants attended the laboratory on six separate occasions following a minimum 6-h fasting period. All trials were performed at the same time of day for each participant (between 08:00 and 20:00) to ensure results were not affected by circadian variation (Reilly and Brooks, 1986). All tests were performed on a cycle ergometer (Lode Excalibur, Germany) and separated by a minimum of 72 h. The first session comprised of an incremental cycling test to exhaustion to determine VO_{2max} and maximal cycling output (W_{max}). In the remaining five sessions, participants performed a simulated time trial, namely two familiarisation sessions and three main

trials (caffeine - CAF, placebo - PLA, and control - CON). Twenty-four hours prior to the main trials, participants were required to refrain from alcohol, caffeine and any unaccustomed strenuous exercise. Food intake was monitored during the 24-h period prior to the main trials using a food diary. Food diaries were analysed by a nutritionist immediately prior to the experimental sessions to ensure that participants had not consumed any caffeine containing foods while energy and macronutrient intake was analysed at a later time by the same nutritionist using specific software (Avanutri online, Avanutri, Rio de Janeiro, Brazil).

Main trials were performed in a double-blind, randomised, counterbalance and cross-over manner. For the CAF and PLA trials, participants ingested a capsule containing either 6 $mg \cdot kg^{-1}$BM of caffeine or dextrose alongside 500 mL of water. Participants were then required to remain seated for 1 h prior to the commencement of the main exercise protocol. During the CON trial, participants followed the same procedures although they did not consume any capsule prior to exercise. Participants were allowed access to their phones or own reading material throughout this waiting period. Blinding occurred via an outside researcher who prepared each participant's supplements in identical looking opaque capsules. Participants were randomly assigned to each experimental condition using a Latin Square model (Mason et al., 2003).

In each supplementation trial, participants were required to respond to a standardised question immediately prior to exercise (*i.e.* 1 h post-supplement ingestion) and again immediately following completion of the exercise. The question related to their belief of which supplement they had taken and was given with the option of choosing one of

three possible answers (*i.e.* “Which supplement do you think you have ingested?” a) Caffeine b) Placebo c) Don’t know). They were also asked to state the reason they had chosen their answer (Supplementary Tables 1 and 2). Based upon each participant’s answer, subgroups were composed according to the supplement trial (*i.e.* CAF or PLA), supplement identification (*i.e.* “correct; “don’t know”; “wrong”), and the moment in which the question was answered (*i.e.* Pre-exercise identification; Post-exercise identification).

Experimental Procedures

Incremental cycling capacity test

Each participant performed a graded cycle capacity test to exhaustion on a cycle ergometer (Lode Excalibur, Germany) to determine individual VO_{2max} and W_{max}. Individual set up of the cycle ergometer (saddle and handlebar height and length) was determined prior to the maximal test, recorded electronically and maintained for all subsequent trials. Participants were required to perform four submaximal 4-min stages starting at 75 W; this was increased by 50 W each stage until 225 W. Thereafter, workload was increased by 30 W every minute until volitional exhaustion. Ventilatory and gas exchange measurements were recorded using a portable breath-by-breath system (K4 b^2, Cosmed, Italy) which has previously been validated (McLaughlin et al., 2001); the highest value averaged over a 30-s period during the test was defined as VO_{2max}. The last completed stage plus the fraction of time spent in the final non-completed stage multiplied by 30 W was defined as a participant’s W_{max}.

Cycling Time-Trial (TT)

The cycling TT was performed on a cycle ergometer (Lode Excalibur, Germany). Participants were required to perform a 5-min cycling warm up performed at 125 W followed immediately by the TT. Participants performed the TT in which they were required to complete a predetermined amount of work equivalent to 25 min at 85% of their individual W_{max} in the fastest possible time; this was based on the protocol of Jeukendrup et al. (2008).

The formula for total amount of work to be performed was as follows:

$$\text{Total amount of work} = 0.85 \times W_{max} \times 1500 \text{ s}$$

The average amount of work to be completed for all participants was 420.3 ± 68.6 kJ. The cycle ergometer was set in linear mode, meaning work load was cadence-dependent according to the formula:

$$W = \alpha \times (\text{rev} \cdot \text{min}^{-1})^2$$

The α value was based on each participant's W_{max} so that they were working at 85% W_{max} when cycling at a cadence of 95 rev·min^{-1}. Participants were instructed to complete the exercise in the fastest possible time. No motivation or specific information was given to the participants during the test although they were informed when they had completed 25%, 50%, 75% and 90% of the exercise. Mean power output (MPO, W) was recorded as the outcome measure for the TT. In order to determine the reliability of the test, we conducted a further test-retest study on 50 participants who completed the TT on two occasions. There was no significant difference in MPO between tests (227.2 ± 35.4 and 224.5 ± 34.7W) with a coefficient of variation of 3.0 ± 2.3%.

Statistical Analysis

Exercise data (MPO) was compared by mixed model analysis in order to determine the effect of supplementation on exercise. To ensure there was no learning effect, the effect of trial order was determined with trial considered a fixed factor and participants a random factor. For the overall analysis, supplementation was assumed as a fixed factor and participants as a random factor. To investigate the effect of expectation on exercise, further sub-analyses were performed according to pre- and post-exercise responses to the questionnaire. Participants were grouped according to their supplement identification ("correct"; "don't know"; "wrong") in CAF and PLA and subsequent exercise data within these subgroups was compared to CON. Analyses of these data were performed in an identical manner to the overall data, assuming supplementation as a fixed factor and participant as a random factor. Tukey post-hoc tests were performed whenever a significant F-value was obtained and the significance level was previously set at $P \leq 0.05$. All these analyses were conducted using SAS software (SAS® version 9.3, Cary, NC, USA) and are presented as mean ± 1SD unless otherwise stated. Magnitude based inferences (MBI; Batterham and Hopkins, 2006) were used to determine the practical significance of caffeine on TT performance using a spreadsheet to establish the likelihood of a meaningful effect on exercise capacity. The smallest worthwhile improvement in MPO was calculated using half the CV of the test (Hopkins, 2004; Paton and Hopkins, 2006). Qualitative descriptors were assigned to the positive percentile scores as follows: <1%, *almost certainly not*; 1-5%, *very unlikely*; 5-25%, *unlikely*; 25-75%, *possibly*; 75-95%, *likely*; 95-99%, *very likely*; >99%, *almost certainly* (Hopkins, 2002). Additionally, the estimated means and SDs from CAF and PLA, separated according to supplement identification, were used to calculate Cohen's *d* (Cohen, 1988) effect sizes and confidence intervals (CI) to plot between-trial comparisons. It is important to note that

direct comparisons could not be made between the overall effects vs. the within subgroup effects (e.g., “Overall CAF vs. CON” versus any sub-group within CAF) since this would result in analysis of duplicate data (considering some of the participants’ data in overall CAF and PLA are also included within their specific sub-groups). Therefore, these comparisons and subsequent interpretation were based upon MBIs, percentage and absolute changes, and individual responses.

Results

Questionnaires

Pre-exercise identification

In CAF, seventeen participants correctly identified caffeine, while twelve incorrectly identified placebo with a further thirteen choosing "don't know". Seventeen participants correctly identified placebo in PLA, eight believed they had ingested caffeine, and the remaining seventeen chose "don't know".

Post-exercise identification

Twenty participants correctly identified the supplement following exercise in CAF, while fourteen were incorrect and a further eight chose "don't know". Eighteen participants correctly assumed that they had taken placebo in PLA, while eleven believed they had ingested caffeine, and thirteen were unsure as to what they had ingested choosing "don't know".

A total of thirteen and fourteen participants changed their supplement identification in CAF and PLA from pre- to post-exercise. Six participants correctly identified caffeine post-exercise having previously been incorrect ("placebo", N = 3) or choosing "don't know" (N = 3). Three participants who had correctly identified caffeine changed their mind to placebo (N = 1) or "don't know" (N = 2) following exercise, while four participants who chose "don't know" prior to exercise incorrectly guessed that they had ingested placebo. Six participants changed their previously unsure ("don't know", N = 5) and incorrect ("caffeine", N = 1) opinions to correctly identify placebo in PLA. Five participants changed their opinion to "don't know" (N = 2) and caffeine (N = 3) having correctly identified placebo prior to exercise. Two participants who chose

246 "don't know" pre-exercise, incorrectly identified caffeine at post-exercise and one
247 participant changed his pre-exercise identification of "caffeine" to "don't know" at
248 post-exercise.
249
250 *Exercise results*
251 *Overall*
252 There was no effect of trial order on MPO (P = 0.58). There was an overall effect of
253 supplement on MPO (P = 0.0002) with *post hoc* analyses revealing an improved
254 performance in CAF vs. PLA (+3.0 ± 5.8%, 234.2 ± 36.7 vs. 228.0 ± 37.6 W, P =
255 0.007; 91% *likely beneficial*) and vs. CON (+4.1 ± 6.2%, 234.2 ± 36.7 vs. 225.7 ±
256 38.4 W, P = 0.0002; 99% *very likely beneficial*), but no difference between PLA and
257 CON (P = 0.50; 24% *unlikely beneficial*). Twenty-three participants improved above
258 the variation of the test in CAF and twelve in PLA.
259
260 *Pre-exercise identification*
261 Correct supplement identification in CAF resulted in improved MPO ($P \leq 0.001$;
262 100% *almost certainly beneficial*) compared to CON (Table 2). Similarly, incorrect
263 identification in CAF resulted in improved performance compared to CON (P =
264 0.003; 99% *very likely beneficial*), but there was no difference for participants who
265 chose "don't know" (P = 0.95; 16% *unlikely beneficial*) (Table 2). Effect sizes and
266 CIs are presented in Figure 1. Eleven of the seventeen participants who correctly
267 identified caffeine improved above the variation of the test, while four of thirteen
268 were improved having chosen "don't know" and eight of twelve having incorrectly
269 identified placebo (Figure 2).
270

There were no statistical differences in MPO between PLA and CON within supplement identification subgroups (all $P > 0.05$; Table 2), although magnitude based inferences suggested correct identification of "placebo" in PLA led to *possibly harmful* effects on performance. Effect sizes and CIs are presented in Figure 1. Four participants who correctly identified "placebo" showed performance reductions above the variation of the test. Twelve participants improved above the variation of the test in PLA; three who correctly identified placebo, five who chose "don't know" and four who believed they ingested caffeine (Figure 2).

Post-exercise identification

Participants who correctly identified caffeine in CAF improved MPO compared to CON ($P \leq 0.001$; 100% *almost certainly beneficial;* Table 2) Participants who incorrectly identified placebo in CAF also improved performance compared to CON ($P = 0.03$; 90% *likely beneficial;* Table 2), but there was no difference in performance in those who did not identify any supplement ($P > 0.05$; 58% *likely trivial*; Table 2). Effect sizes and CIs are presented in Figure 1. Fifteen of the twenty participants who correctly identified caffeine improved above the variation of the test, while seven of fourteen improved despite incorrectly identifying placebo. Of the eight who chose "don't know", only one improved performance (Figure 2).

Performance was not statistically different between PLA and CON for participants who chose "don't know" ($P > 0.05$; Table 2). There was a tendency towards improved MPO (+3.7 ± 6.3%, $P = 0.06$; 87% *likely beneficial*) in those who incorrectly believed they had ingested caffeine in PLA (Table 2), while MBIs suggested a *possibly harmful* effect of correct identification of placebo (-1.6 ± 4.9%) and only a 1% chance

296 of being positive. Effect sizes and CIs are presented in Figure 1. Six participants
297 improved above the variation having incorrectly identified caffeine, while five
298 improved having chosen “don’t know”. Only one participant improved having
299 correctly identified placebo while six worsened performance (Figure 2).
300
301 *Food intake*
302 Absolute and relative carbohydrate, protein, and fat intake in the 24 h prior to the
303 main trials were not significantly different (all $P > 0.05$). Similarly, total caloric
304 intake was not different prior to any trial ($P = 0.93$).

Discussion

This study showed that correct identification of caffeine, particularly post-exercise, improved cycling performance with greater relative improvements than the overall effect of caffeine. Furthermore, there was an apparent improved performance in PLA for participants who believed they had ingested caffeine, although this was based upon post-exercise supplement identification only, while correct identification of placebo, both pre- and post-exercise, may possibly have led to performance impairments.

This study employed trained cyclists, the majority of whom were competing at national and international level. Although none took caffeine as a supplement, all participants were aware of the substance and its purported ergogenic effect. Thus, it is reasonable to suggest that any individual who identified the supplement ingested as caffeine will have had the belief that their performance would improve accordingly. Indeed, correct identification of caffeine ingestion resulted in an improved performance with greater relative improvements than the overall effect of caffeine (Pre-exercise: +4.8% and Post-exercise: +6.5% vs. Overall: +4.1%; Figure 2). The questionnaire allowed an uncertainty regarding which supplement had been ingested ("don't know"). Thus, analysing participants who chose this response would theoretically allow determination of the "true effect" of caffeine since the individual would not be biased by opinion. Surprisingly, however, performance was unaffected with caffeine when participants were unsure as to what they had ingested, but was improved when they incorrectly identified placebo (Table 2). We can only speculate as to the reason for these unexpected findings; perhaps the physiological mechanisms by which caffeine improves performance were a greater stimulus in participants

believing they had ingested a placebo substance, or there may have been an increased motivation in these participants. Nonetheless, this was not directly measured here though future investigation should consider this.

Interestingly, post-exercise identification of caffeine in PLA showed a tendency towards improved performance despite participants having ingested no active substance. Increases in this subgroup were *likely beneficial*, above the variation of the test (+3.7 vs. +3.0%) and very close to the overall beneficial effect of caffeine shown in the current study (~4.0%). Beedie et al. (2006) previously investigated the effects of expectation on performance; participants were informed that they had ingested either 4.5 or 9.0 $mg \cdot kg^{-1}$BM prior to exercise although caffeine was not administered on any occasion. Despite this, the authors showed a *likely beneficial* 2.2% in 10 km TT performance when participants believed they had ingested caffeine, which is similar to the performance increase of ~3.5% according to post-exercise caffeine identification in PLA in the current study. Taken together, these results support the notion that the belief that one has ingested an active supplement can strongly influence the outcome of an exercise task (Clark et al., 2000). Furthermore, it seems reasonable to speculate that expectation, which is highly variable among individuals, is a factor that can potentially account for some of the variability in responses to certain interventions in sports nutrition.

Indeed, it is apparent that correct identification of placebo may have impeded performance with *possibly harmful* effects and a total of four (pre-exercise identification) and six (post-exercise identification) participants worsening performance beyond the variation of the test. The nocebo effect is directly opposite to

the placebo effect in that a negative outcome occurs following the administration of an inert intervention (Benedetti et al., 2007). This phenomenon has been shown to reduce exercise performance (Lundby et al,, 2012; Pollo et al., 2012) and increase ratings of perceived exertion (Bottoms et al., 2014), but it has been rarely addressed scientifically, particularly in sports nutrition. Interestingly, based upon our findings, it appears that correct identification of placebo by some athletes expecting to receive a potential ergogenic aid may result in the nocebo effect, possibly by frustrating their expectations. However, the opposite appeared true in individuals who believed they had ingested placebo when taking caffeine. While it remains unclear as to why and how active and non-active substances can differently modulate expectations and performance, this study provides some evidence to suggest that the nocebo effect may play a role in performance outcomes and should be accounted for within any experimental investigation or clinical intervention in sports nutrition.

Correct (+4.8%) and incorrect (+7.3%) pre-exercise supplement identification in CAF resulted in performance improvements above the overall effect (+4.1%). Post-exercise, incorrect placebo identification fell below this overall improvement (+3.3%) while correct identification of caffeine improved further (+6.5%). These changes are due to a number of participants changing their opinion from pre- to post-exercise, likely due to stimuli relating to the exercise (Beedie et al., 2006). The majority of the stated reasons for believing caffeine had been ingested prior to exercise were due to the sensation of caffeine associated side effects, specifically tachycardia, alertness and trembling. Additionally, a number of participants' reasons for identifying caffeine post-exercise appear to be due to stimuli felt throughout the exercise test, namely "feeling better" or "less tired". This was particular true with respect to the eleven

individuals who changed their opinion to caffeine, six of whom (four in CAF; two in PLA) improved their performance above the variation of the test. Thus, it could be suggested that post-exercise supplement identification may be the most accurate measurement relating to perception since it incorporates both conceptions prior to (*i.e.*, side-effects) and during (*i.e.*, side-effects and performance effects) the exercise. However, the main limitation of this study is that we did not determine why participants changed their opinion. Furthermore, it cannot be fully elucidated whether any participant's change in supplement identification resulted from their performance or whether it shaped the performance itself. Nonetheless, these data support the notion that preconceptions may be further modified by factors intrinsic to exercise (Beedie et al., 2006), and thus should be taken into account. Future research should include pre- and post-exercise questionnaires including the opportunity to discuss why opinions were modified.

The results of this study highlight the necessity in assessing a participants' perception of what they have ingested in order to distinguish the true effect of a supplement from its placebo effect. Importantly, simply including a placebo group may not be sufficient to effectively blind an experiment; active nutrients and drugs, such as caffeine, beta-alanine, sodium bicarbonate and creatine, may cause side effects or changes in performance, which are clues leading subjects to identify the treatment. To avoid bias in the analysis of results, it would be prudent to test the efficacy of the blinding procedure by asking participants to identify the supplement ingested. Comprehensive assessment of data according to perceptions of the supplement ingested could allow for more definitive conclusions on the actual effects of active nutrients in sports nutrition. In contrast to the undesirable effect of preconception in

research, any such bias may prove positive in a real world setting. It would be reasonable to suggest that an athlete may benefit solely from the belief that he has ingested an active supplement, a notion previously suggested to have some scientific basis (de la Fuente-Fernandez et al., 2002; Yang et al., 2002).

Perspective

Correct identification of caffeine, particularly after exercise, appeared to improve cycling performance to a greater extent than the overall effect of caffeine. Furthermore, participants who believed they had ingested caffeine while ingesting placebo also appeared to improve their performance while correct identification of placebo may lead to possible impairments in performance for some individuals. Altogether, these results suggest that an individual's perception of whether they have ingested an active supplement contributes greatly to their exercise performance, although the mechanisms by which this influences performance remain to be fully elucidated. Scientists must be encouraged to systematically test whether their blinding procedure was effective when interpreting data as this is likely a source of bias in sports nutrition.

Acknowledgements

Bryan Saunders, Vitor de Salles Painelli, Guilherme Giannini Artioli and Bruno Gualano have been financially supported by Fundação de Amparo à Pesquisa do Estado de Sao Paulo (FAPESP grants number: 2011/19513-2, 2013/04806-0, 2014/11948-8 and 2013/14746-4). Bryan Saunders (150513/2015-1), Bruno Gualano and Hamiton Roschel have been financially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

We wish to thank the Laboratório de Determinantes Energéticos de Desempenho Esportivo (LADESP) for access to the cycle ergometer used in this study and the volunteers for their efforts.

Conflict of interest

The authors declare that they do not have conflict of interests.

References

1. Batterham AM, Hopkins WG. Making meaningful inferences about magnitudes. *Int J Sports Physiol Perf* 2006: 1(1): 50-57.
2. Beedie CJ, Foad AJ. The placebo effect in sports performance. *Sports Med* 2009: 39(4): 313-329.
3. Beedie CJ, Stuart EM, Coleman DA, Foad AJ. Placebo effects of caffeine on cycling performance. *Med Sci Sports Exerc* 2006: 38(12): 2159-2164.
4. Benedetti F, Amanzio M, Vighetti S, Asteggiano G. The biochemical and neuroendocrine bases of the hyperalgesic nocebo effect. *J Neurosci* 2006: 26(46): 12014-1202
5. Bottoms L, Buscombe R, Nicholettos A. The placebo and nocebo effects on peak minute power during incremental arm crank ergometry. *Eur J Sport Sci* 2014: 14(4): 362-367.
6. Bühler ES, Dirk WLS, Schlegel KG, Winkler S. Development of a tool to assess the caffeine intake among teenagers and young adults. Ernaehrungs Umschau 2014: 61(4): 58–63.
7. Clark VR, Hopkins WG, Hawley JA, Burke LM. Placebo effect of carbohydrate feeding during a 4-km cycling time trial. *Med Sci Sports Exerc* 2000: 32:1642-1647.
8. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988. p. 20.
9. de la Fuente-Fernandez R, Phillips AG, Zamburlini M, Sossi V, Calne DB, Ruth TJ., Stoessl AJ. Dopamine release in human ventral striatum and expectation of reward. *Behav Brain Res* 2002: 136(2): 359-63.

462 10. Fred Hutchinson Cancer Research Center Web site [Internet]. Seattle (WA):
463 Specific Food Questionnaire: Caffeine Questionnaire. Available from:
464 http://www.fredhutch.org.
465 11. Graham TE, Spriet LL. Metabolic, catecholamine, and exercise performance
466 responses to various doses of caffeine. *J Appl Physiol* 1995: 78(3): 867-874.
467 12. Hopkins WG. Probabilities of clinical or practical significance. Sportscience
468 2002: 6, 431 Available from: http://www.sportsci.org/jour/0201/wghprob.htm.
469 13. Hopkins WG. How to interpret changes in an athletic performance test.
470 Sportscience 2004, 8(1), pp.1-7.
471 14. Jeukendrup AE, Hopkins S, Aragón-Vargas LF, Hulston C. No effect of
472 carbohydrate feeding on 16 km cycling time trial performance. *Eur J Appl*
473 *Physiol* 2008: 104(5): 831-837.
474 15. Lundby C, Millet GP, Calbet JA, Bärtsch P, Subudhi AW. Does 'altitude
475 training'increase exercise performance in elite athletes? *Brit J Sport Med* 2012:
476 14: bjsports-2012.
477 16. Mason RL, Gunst RF, Hess JL. Statistical design and analysis of experiments:
478 with applications to engineering and science. 2nd ed. John Wiley & Sons, Inc,
479 Hoboken, New Jersey; 2003, p. 328-31.
480 17. McLaughlin JE, King GA, Howley ET, Bassett DR Jr, Ainsworth BE.
481 Validation of the COSMED K4 b2 Portable Metabolic System. *Int J Sports*
482 *Med* 2001: 22: 280-284.
483 18. Paton CD, Hopkins WG. Variation in performance of elite cyclists from race
484 to race. Eur J Sport Sci 2006: 6(1): 25-31.
485 19. Pollo A1, Carlino E, Vase L, Benedetti F. Preventing motor training through
486 nocebo suggestions. *Eur J Appl Physiol* 2012: 112(11): 3893-903.

20. Reilly T, Brooks GA. Exercise and the circadian variation in body temperature measures. *Int J Sports Med* 1986: 7(6): 358-362.

21. Spriet LL. Exercise and sport performance with low doses of caffeine. *Sports Med* 2014: 44(2): 175-184.

22. Yang EV, Bane CM, MacCallum RC, Kiecolt-Glaser, JK, Malarkey, WB, Glaser, R. Stress-related modulation of matrix metalloproteinase expression. *J Neuroimmunol* 2002: 133 (1-2): 144-50.

Authorship information

Significant manuscript writer – Bryan Saunders

Significant manuscript reviewer/reviser – Bruno Gualano, Hamilton Roschel, Guilherme Giannini Artioli

Concept and design – Bruno Gualano, Bryan Saunders, Hamilton Roschel, Guilherme Giannini Artioli

Data acquisition – Bryan Saunders, Luana Farias de Oliveira, Rafael Pires da Silva, Vitor de Salles Painelli, Livia Souza Gonçalves, Guilherme Yamaguchi, Thiago Mutti, Erika Maciel

Data analysis and interpretation – Bryan Saunders, Luana Farias de Oliveira, Rafael Pires da Silva, Livia Souza Gonçalves, Guilherme Yamaguchi, Thiago Mutti, Erika Maciel

Statistical expertise – Bryan Saunders, Hamilton Roschel, Vitor de Salles Painelli

Figures

Figure 1. Effect sizes compared to CON in CAF and PLA separated into subgroups based upon supplement identification pre- and post-exercise. Panel A displays CAF vs. CON pre-exercise. Panel B displays CAF vs. CON post-exercise. Panel C displays PLA vs. CON pre-exercise. Panel D displays PLA vs. CON post-exercise.

Figure 2. Individual percentage change from CON in CAF (Panel A) and PLA (Panel B) organised according to supplement identification subgroups pre- and post-exercise. The grey dotted line represents the natural variation of the test (±3.0%) while the black dotted line represents the mean overall improvement with caffeine (+4.1%). The number of participants who improved above, were within, or worsened beyond the natural variation of the test in each subgroup is displayed below each graph.

522 **Table 1.**[1]
523

Characteristic		Mean (SD)	Range	
			Minimum	Maximum
Age (y)		37 (8)	18	55
Height (cm)		1.76 (0.06)	1.60	1.89
Body mass (kg)		74.3 (8.4)	58.9	93.0
Experience (y)		12 (11)	1	40
Weekly training	Duration (h)	11 (5)	4	25
	Distance (km)	272 (119)	50	500
VO_{2max}	Absolute ($L \cdot min^{-1}$)	3.7 (0.5)	2.8	4.8
	Relative ($ml \cdot kg \cdot min^{-1}$)	50.0 (6.8)	33.6	64.5
HR_{max} ($beats \cdot min^{-1}$)		182 (11)	158	201
W_{max}	Absolute (W)	329.7 (53.8)	181.4	439.0
Habitual caffeine intake ($mg \cdot day^{-1}$)		192 (156)	1.77	583.0

524

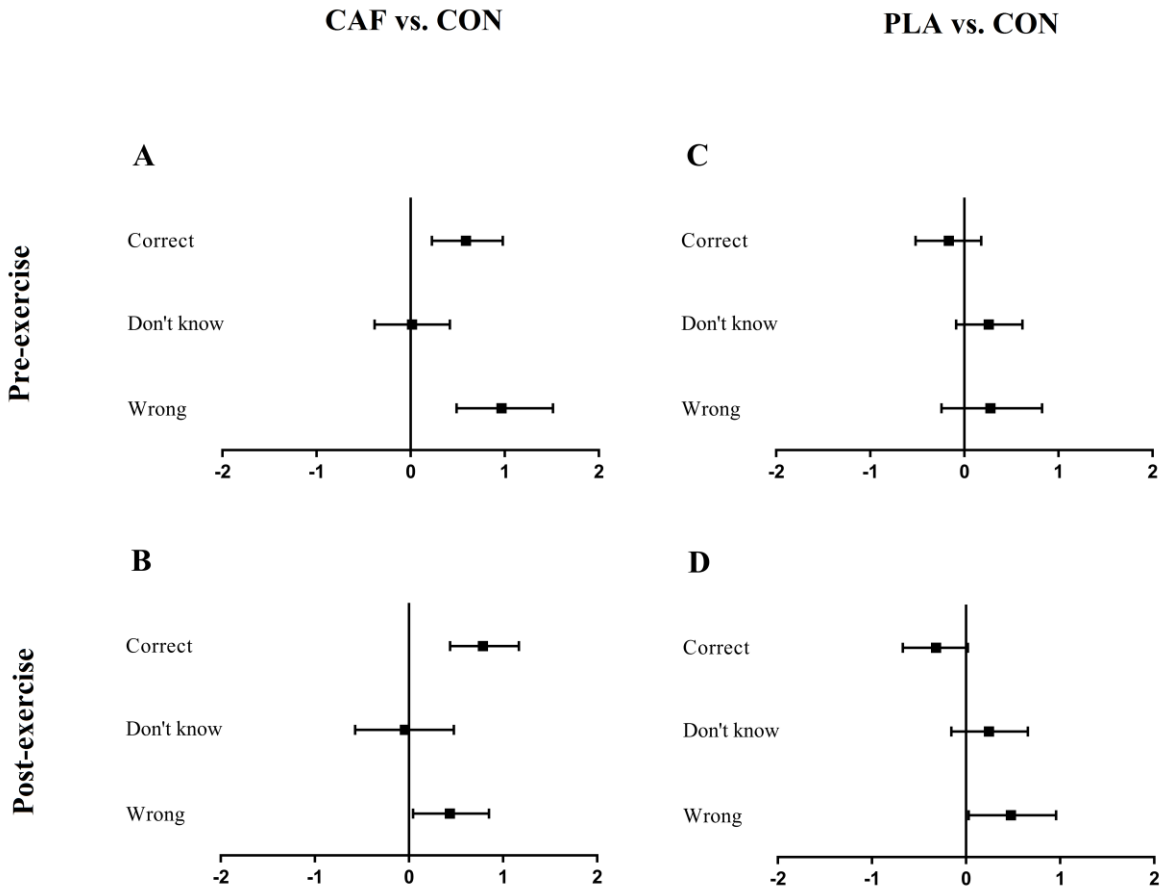
[1] Table 1. Participant characteristics

Table 2.[2]

		Pre-exercise identification			Post-exercise identification		
		Correct	**Don't know**	**Wrong**	**Correct**	**Don't know**	**Wrong**
	N	17	13	12	20	8	14
CAF	**MPO (W)**	233.9 ± 41.1*	230.9 ± 32.6	238.2 ± 37.1^	236.4 ± 37.2*	234.1 ± 35.8	231.1 ± 39.0#
CON		223.7 ± 40.7	230.8 ± 36.0	223.2 ± 40.3	223.2 ± 39.7	236.0 ± 43.0	223.6 ± 35.6
CAF vs. CON	**% difference**	+4.8 ± 4.7	+0.4 ± 5.0	+7.3 ± 7.5	+6.5 ± 6.6	-0.3 ± 4.2	+3.3 ± 5.2
MBI	**% chance of being beneficial/trivial/harmful**	99/1/0	13/76/11	99/1/0	100/0/0	7/65/28	87/13/0
	N	17	17	8	18	13	11
PLA	**MPO (W)**	240.0 ± 40.8	226.5 ± 30.0	205.6 ± 38.3	230.5 ± 43.4	223.6 ± 36.9	229.0 ± 30.1$
CON		242.0 ± 36.9	221.4 ± 35.8	200.5 ± 34.4	233.8 ± 39.9	218.3 ± 42.7	221.3 ± 30.6
PLA vs. CON	**% difference**	-1.0 ± 5.0	+3.0 ± 8.0	+2.4 ± 5.1	-1.6 ± 4.9	+3.2 ± 7.6	+3.7 ± 6.3
MBI	**% chance of being beneficial/trivial/harmful**	3/72/25	62/37/1	61/37/2	0/60/40	62/36/2	84/16/0

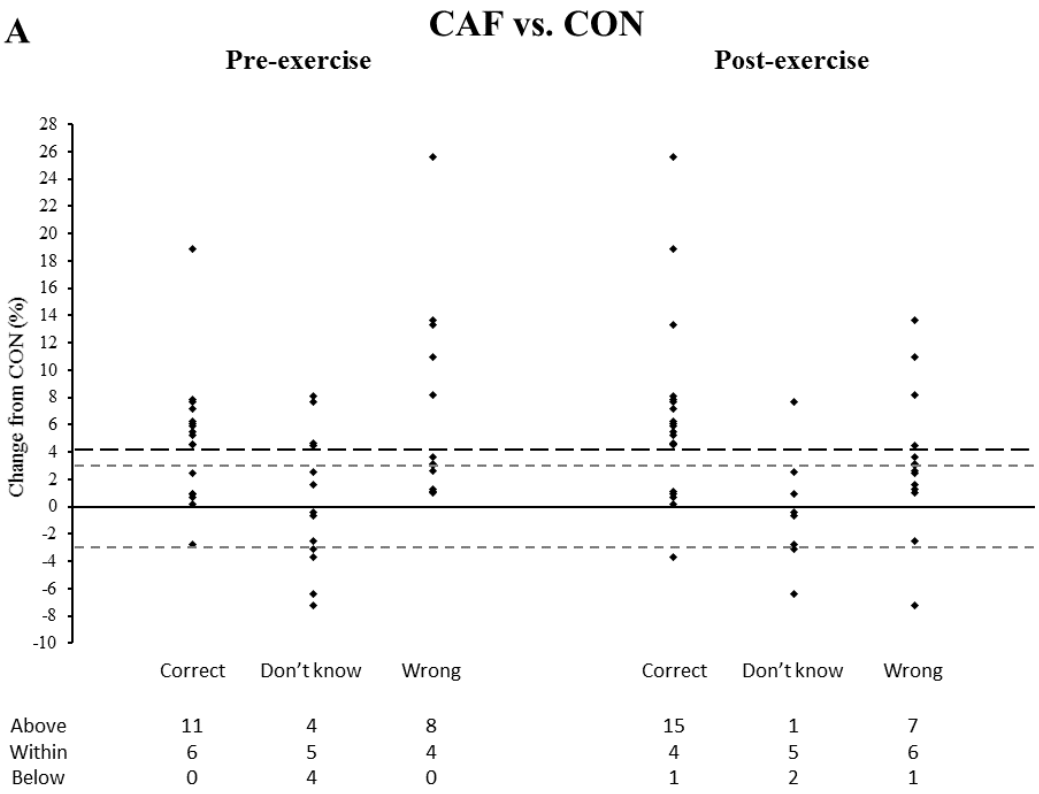
[2] Table 2. MPO in CAF, PLA and CON, and % absolute difference from CON in CAF and PLA, when categorising individuals into their pre- and post-exercise supplement identification responses. *P ⩽ 0.001 from CON. ^P ⩽ 0.01 from CON. #P ⩽ 0.05 from CON. $P ⩽ 0.06 from CON. MPO = Mean power output; CAF = Caffeine trial; PLA = Placebo trial; CON = Control trial; MBI = Magnitude based inferences.

527



528

529 **Figure 1.**
530


A
CAF vs. CON
Pre-exercise
Post-exercise
Change from CON (%)
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0
-2
-4
-6
-8
-10
Correct
Don't know
Wrong
Correct
Don't know
Wrong
Above 11 4 8 15 1 7
Within 6 5 4 4 5 6
Below 0 4 0 1 2 1

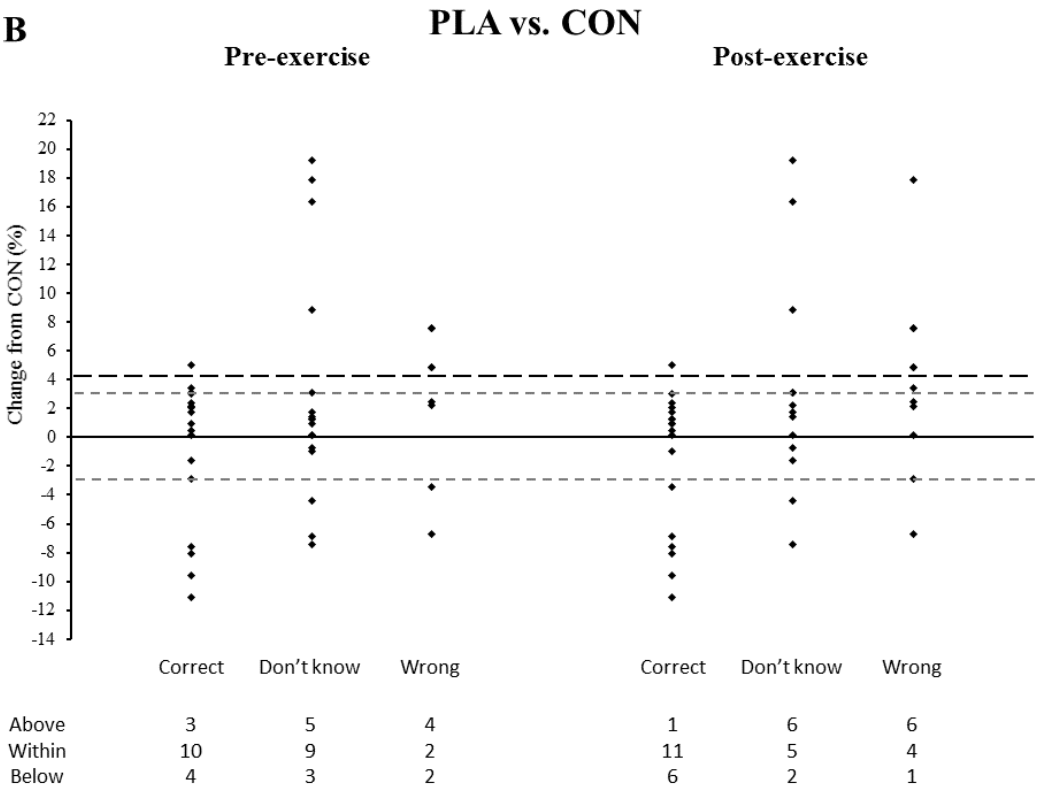

B
PLA vs. CON
Pre-exercise
Post-exercise
Change from CON (%)
22
20
18
16
14
12
10
8
6
4
2
0
-2
-4
-6
-8
-10
-12
-14
Correct
Don't know
Wrong
Correct
Don't know
Wrong
Above 3 5 4 1 6 6
Within 10 9 2 11 5 4
Below 4 3 2 6 2 1

Figure 2.

534