

Mini-review

Caffeine and cognitive performance: Persistent methodological challenges in caffeine research



Jack E. James*

Reykjavik University, Iceland
National University of Ireland, Galway, Ireland

ARTICLE INFO

Article history:

Received 7 March 2014
Received in revised form 23 May 2014
Accepted 27 May 2014
Available online 2 June 2014

Keywords:

Caffeine
Cognitive performance
Caffeine-naïve
Placebo control
Methodological challenges

ABSTRACT

Human cognitive performance is widely perceived to be enhanced by caffeine at usual dietary doses. However, the evidence for and against this belief continues to be vigorously contested. Controversy has centred on caffeine withdrawal and withdrawal reversal as potential sources of experimental confounding. In response, some researchers have enlisted “caffeine-naïve” experimental participants (persons alleged to consume little or no caffeine) assuming that they are not subject to withdrawal. This mini-review examines relevant research to illustrate general methodological challenges that have been the cause of enduring confusion in caffeine research. At issue are the processes of caffeine withdrawal and withdrawal reversal, the definition of caffeine-naïve, the population representativeness of participants deemed to be caffeine-naïve, and confounding due to caffeine tolerance. Attention to these processes is necessary if premature conclusions are to be avoided, and if caffeine's complex effects and the mechanisms responsible for those effects are to be illuminated. Strategies are described for future caffeine research aimed at minimising confounding from withdrawal and withdrawal reversal.

© 2014 Elsevier Inc. All rights reserved.

Contents

1. Introduction	117
2. Confounding due to reversal of withdrawal effects	118
2.1. Caffeine-naïve participants: Studies of low- or non-consumers	119
3. When is a participant caffeine-naïve?	119
3.1. Confirming low-consumer status and abstinence	119
4. Are caffeine-naïve participants representative?	120
5. Even if confounding due to withdrawal reversal is avoided what about tolerance?	121
6. Future research and conclusions	121
References	122

1. Introduction

Human cognitive performance is widely perceived to be enhanced by caffeine at usual dietary doses. Yet, the evidence for and against this belief continues to be vigorously contested (e.g., Childs and de Wit, 2006; Haskell et al., 2005; James, 1994; James and Rogers, 2005; Rogers et al., 2013; Smith et al., 2006). Much controversy centres on

potential experimental confounding from caffeine withdrawal and withdrawal reversal (Einöther and Giesbrecht, 2013; James and Rogers, 2005). One approach to addressing those sources of confounding has been to enlist experimental participants who habitually consume little or no caffeine and therefore are not subject to caffeine withdrawal (Borota et al., 2014; Childs and de Wit, 2006; Haskell et al., 2005; Hewlett and Smith, 2006; Rogers et al., 2013; Smith et al., 2006, 2013). The research design employed in the most recent of these studies (Borota et al., 2014) is representative, and this mini-review focuses on that study, while also referring to related studies, to elucidate persistent methodological challenges that have contributed to the enduring confusion.

* Reykjavik University, Menntavegur 1, 101 Reykjavík, Iceland. Tel.: +354 599 6449; fax: +354 599 6201.
E-mail address: jack@ru.is.

In brief, Borota et al. concluded that memory consolidation but not recognition memory is enhanced at 24 h when 200 mg of caffeine (the approximate equivalent of 1–2 cups of coffee) is ingested during initial task exposure. However, the study overlooked specific behavioural and pharmacological processes associated with caffeine exposure that obscure the drug's effects on cognitive performance. It is important that these overlooked processes are examined if premature conclusions concerning caffeine enhancement are to be avoided. Although concerned with the study of caffeine and cognitive performance, the present review does not examine in detail specific cognitive processes, an area of controversy in its own right that has been recently reviewed (Rogers, 2014). Rather, the emphasis here is on experimental design and ways of controlling recurring confounding that has dogged the systematic study of caffeine and cognitive performance. Additionally, it may be noted that most of the main issues examined in the present review are equally relevant to studies of the effects of caffeine on mood (cf., James and Gregg, 2004a; James and Rogers, 2005).

The earliest systematic studies of the psychopharmacology of caffeine were conducted a century ago (Hollingworth, 1912a,b), and for most of the intervening period, it has been believed that caffeine enhances human cognitive performance. That belief, however, is contestable on theoretical (e.g., James, 1994) and empirical grounds (e.g., James and Rogers, 2005). The problem is that a large body of research purporting to show caffeine enhancement shares a common flaw arising from uncritical adoption of standard placebo-controlled drug-trial methodology (James, 1994; James and Rogers, 2005). It has been common practice in placebo-controlled studies of caffeine to emulate gold-standard placebo-controlled methodology used to investigate other drugs such as new pharmaceuticals. Typically, studies have measured cognitive performance in healthy volunteers before and after double-blind administration of caffeine and placebo. Compared to baseline and placebo, performance has often been reported to improve following ingestion of caffeine, leading to the conclusion that caffeine enhances performance. However, critical examination of this standard research design shows that when used to examine the effects of caffeine on cognitive performance the findings it has yielded are, at best, ambiguous.

Because of the importance of ensuring that all participants are equivalent in systemic levels of the drug being investigated, it is usual in placebo-controlled trials for participants to be drug free when randomised to drug or placebo groups. While this strategy works well for drugs that are not in general use by populations from which study participants are drawn, suitability of the strategy is less certain when, as with caffeine, daily consumption is the norm. The daily diet of most people includes caffeine consumed in separate portions throughout the day, with fewer portions consumed later in the day, followed by overnight abstinence (James, 1997). With the half-life of caffeine in healthy adults being approximately five hours (Pfeifer and Notari, 1988), typical overnight abstinence of 10–14 h results in substantial elimination of systemic caffeine by early morning (Lelo et al., 1986a). In fact, it is common in placebo-controlled studies of caffeine for researchers to make a methodological convenience out of naturally-occurring overnight abstinence by simply asking participants to forgo their usual morning caffeine beverage prior to testing. However, it is this step, intended to standardise procedures by ensuring participants are “equivalent” at time of caffeine administration, that has long been a cause of serious confounding.

2. Confounding due to reversal of withdrawal effects

Caffeine exerts pharmacological actions at diverse sites, both centrally and peripherally, due mostly to antagonism of endogenous adenosine, with A_1 and A_{2A} receptors appearing to be the primary targets (Ferré, 2008). Repeated consumption of caffeine generally leads to the development of physical dependence, evidenced by the appearance of behavioural, physiological, and subjective withdrawal effects

provoked by abrupt cessation of use (Juliano and Griffiths, 2004). Although incompletely understood, the mechanism responsible for caffeine dependence is thought to involve adenosine upregulation resulting in hypersensitivity during abstinence. This hypothesis is consistent with symptoms of caffeine withdrawal, which include headache, tiredness/fatigue, decreased energy, decreased well-being, difficulty concentrating, irritability (Juliano and Griffiths, 2004), and importantly for present purposes, decreased cognitive performance (e.g., James, 1998; Rogers et al., 2003, 2013; Yeomans et al., 2002). Symptoms may be felt within about 12–16 h, generally peak at around 24–48 h, and usually abate within 3–5 days, although occasionally may continue for up to a week (Griffiths et al., 1990; Hughes et al., 1993). Cessation of as little as 100 mg (approximately one cup of instant coffee) per day, and possibly considerably less, can produce symptoms of withdrawal (Griffiths et al., 1990; Lieberman et al., 1987; Smit and Rogers, 2000).

The facts concerning caffeine withdrawal are critical for understanding the results of placebo-controlled trials of the effects of caffeine administration on cognitive performance. Having avoided caffeine since the evening before, participants in most studies will have entered the early stages of caffeine withdrawal by the time they are tested in the laboratory (typically, at least 12–14 h since caffeine was last ingested). Thus, the crucial question, illustrated in Fig. 1, is: To what extent is enhanced performance (attributable to caffeine) an indication of a genuine *net* effect of the drug or merely the result of *reversal of withdrawal*? A third possibility is that improvements in cognition are a combination of net effects and withdrawal reversal.

Of several approaches for overcoming confounding due to reversal of caffeine withdrawal, “long-term” withdrawal designs have proved to be the most successful (James and Rogers, 2005). These incorporate core features of the traditional drug-challenge paradigm, including double blinding and placebo control, combined with periods of abstinence long enough (several days to one week is usually sufficient) to remove withdrawal effects (see Table 1). Extending the abstinence period substantially beyond the traditional period of overnight or 24 h removes confounding due to withdrawal effects prior to administration of caffeine or placebo challenge. Studies that have employed designs incorporating long-term withdrawal have yielded consistent evidence of caffeine having little or no net benefit for cognitive performance for adults (James, 1998; James et al., 2005; Judelson et al., 2005; Rogers et al., 2005) and children (Heatherley et al., 2006).

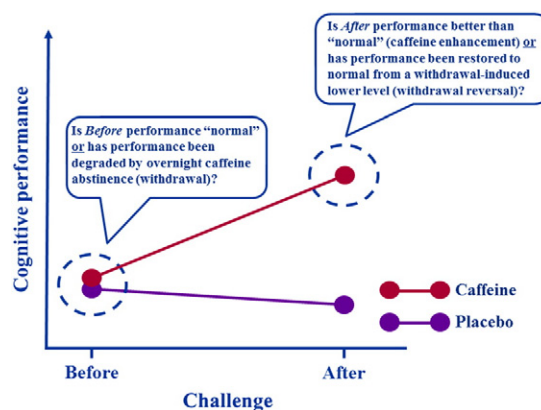


Fig. 1. Schematic representation of the results of a typical double-blind placebo-controlled experiment to test the effects of caffeine on cognitive performance by comparing performance before and after caffeine challenge. Note. This type of study design yields ambiguous results due to failure to control for withdrawal effects from overnight caffeine abstinence and withdrawal reversal when caffeine is administered. Specifically, improved performance after caffeine could be due to the drug producing either *net* benefit or *reversal of withdrawal* without net benefit (see text).

Table 1

Summary of a double-blind placebo-controlled crossover experimental protocol that incorporates alternating periods of “long-term” caffeine exposure and abstinence.

Week	Run-in days (Days 1–6)	“Challenge” (Day 7)	Effects revealed by challenge
1	Placebo	Placebo	Caffeine “wash out” simulates the “caffeine naïve” state, providing a suitable baseline for comparison with other conditions.
2	Placebo	Caffeine	Acute challenge reveals tolerance effects when compared with other conditions.
3	Caffeine	Placebo	Acute abstinence reveals withdrawal effects when compared with other conditions.
4	Caffeine	Caffeine	Simulates habitual consumption revealing the net effects of caffeine when compared with caffeine wash out, tolerance, and withdrawal.

Note. “Long-term” designs such as that summarised here have found that caffeine’s effects on cognitive performance are due primarily to reversal of negative withdrawal effects with little or no net benefit (James, 1998; James and Rogers, 2005; James et al., 2005).

2.1. Caffeine-naïve participants: Studies of low- or non-consumers

The foregoing discussion shows that success in determining the effects of caffeine on cognitive performance depends critically on being able to control against confounding due to withdrawal-reversal. One alternative experimental approach to that which requires regular consumers to be long-term abstinent is to employ the traditional drug-challenge protocol with “caffeine naïve” participants who habitually consume little or no caffeine. This has been a popular approach (e.g., Childs and de Wit, 2006; Hewlett and Smith, 2006; Haskell et al., 2005; Smith et al., 2006, 2013) and the one that Borota et al. (2014) employed in their study of caffeine and long-term memory. Since naïve participants are caffeine-free, they should not be subject to caffeine withdrawal and therefore reversal of withdrawal effects should not be a source of confounding. This strategy, however, creates various threats to experimental internal and external validity of which there are at least three that are not easily resolved. These relate to how “caffeine-naïve” is defined, the representativeness of participants deemed to be caffeine-naïve, and confounding due to caffeine tolerance.

3. When is a participant caffeine-naïve?

Considering the ubiquity of caffeine, it should be evident that caffeine-naïve participants are not necessarily easily found. Besides the naturally-caffeinated beverages of coffee and tea, caffeine is added to a wide range of beverages, including sodas, energy drinks, bottled water, and alcoholic drinks. Chocolate contains caffeine, and the drug is added to an increasing variety of foods, including confectionery, ice cream, chewing gum, yoghurt, breakfast cereal, cookies, flavoured milk, sunflower seeds, and beef jerky. In addition, caffeine is present in medications, including prescribed and over-the-counter compounds for weight loss, pain relief, colds and flu, and anti-somnogenic compounds. Miscellaneous products that contain caffeine include breath-freshener sprays and mints, skin lotions, cosmetics, soap, shampoo, caffeine aerosol inhaler, and illicit-drug compounds with which caffeine is frequently combined as a diluent (cutting agent). The near-universal consumption of caffeine is evidenced by its presence as a biologically significant contaminant of freshwater and marine systems in environments where there are no naturally-occurring sources of the drug (Rodríguez del Rey et al., 2012).

Who, then, can be said to be caffeine-naïve? Borota et al. (2014) merely labelled their participants “caffeine naïve”, with no definition offered to explain what that means. Studies with similar aims have adopted a variety of definitions. Haskell et al. (2005) defined their “habitual non-consumers” as drinking no tea and coffee, and reporting an intake of less than 50 mg/day “from other sources”. Childs and de Wit (2006) defined their “light, nondependent users” as reporting less than 300 mg per week. For Smith et al. (2006), non-consumers “were defined as individuals who reported that they did not ingest any drinks containing caffeine”, with no reference to other sources. In personal communication with the authors of the Borota et al. (2014) study, it was reported that participants’ self-reported regular intake was “less than 500 mg per week [and that] almost all subjects had intake less

than 70 mg per week” (Yassa, 2014). Although this is indeed a low level of exposure compared to population averages in the region of 300–400 mg per day, it is not zero exposure. In particular, average daily exposure for participants whose consumption was nearer to 500 mg per week is likely to have been sufficient to produce physical dependence (Lieberman et al., 1987).

Notably, Borota et al. monitored levels of salivary caffeine and paraxanthine, the main metabolite of caffeine in humans (Lelo et al., 1986a,b), during experimental sessions with participants, which led to some participants “who did not conform to protocol” being excluded. Specifically, participants were excluded if their salivary metabolites indicated that they had ingested caffeine before coming to the laboratory for either of two laboratory sessions conducted on separate days (Yassa, 2014). Additionally, it should be noted that measurement of salivary metabolites *in-session* provides a check on only one type of breach of protocol. Any participants who regularly consumed more than indicated by self-report need only to have abstained from caffeine for 24 h before attending laboratory sessions to escape detection. If that happened, those participants would be in a state of caffeine withdrawal and subject to the influence of withdrawal reversal.

This is not to suggest that participants wantonly deceive researchers. However, in light of caffeine’s ubiquity it cannot be assumed that participants are accurate in their perceptions of their actual level of exposure to caffeine from one day to the next. Illustrative of that point, participants in the Borota et al. (2014) study performed no better than chance when asked to say whether they had been administered caffeine or placebo in laboratory sessions, which is a finding that is consistent with results reported by others (e.g., James, 1998; James et al., 2005). Moreover, the fact that Borota et al. were forced to exclude some participants due to breaches of protocol revealed by salivary assays is evidence that self-reported intake cannot be relied upon as an entirely safe method of recruitment in experiments that require participants to be caffeine-naïve.

3.1. Confirming low-consumer status and abstinence

In studies involving ostensibly caffeine-naïve participants, the importance of objective verification of low-consumer status cannot be overemphasised. Similar importance attaches to the need for verification of adherence to protocol in studies that incorporate periods of caffeine abstinence (e.g., overnight), whether participants are caffeine naïve or habitual consumers. Above all, high priority should be given to objective verification of self-reported out-of-session exposure, which although potentially demanding of participants and research resources can be done successfully by analysing metabolites from as few as one blood or saliva sample taken in the late afternoon on successive days (James et al., 1988; Lelo et al., 1986b). In some of the aforementioned long-term experiments of caffeine and performance, saliva samples were collected daily over several weeks (James, 1998; James et al., 2005). Consequently, adherence was independently verifiable, allowing in- and out-of-session breaches of protocol to be detected. Conversely, since Borota et al. (2014) did not monitor out-of-session breaches, there is no way of knowing whether participants remained caffeine-naïve during the critical several days preceding laboratory sessions.

Thus, despite participants in the Borota et al. study ostensibly being caffeine-naïve, it is not possible to know whether some may have attended laboratory sessions in states of caffeine withdrawal. Therefore, withdrawal reversal cannot be ruled out as an explanation of the reported improvements in memory.

Of course, it follows that if attempts at objective verification are to be successful, then the intended veridicality method must be valid. This, however, has not always been the case. For example, in a study of the effects of ad libitum caffeine consumption on cognitive performance in healthy habitual caffeine consumers, Christopher et al. (2005) collected two saliva samples from each participant. One sample was taken “before work” after participants had consumed their morning caffeine, and the second was taken in the late-afternoon “after work” after participants had consumed “their normal amount of caffeine over the course of the day”. However, as pointed out previously (James and Rogers, 2005), the salivary results reported by Christopher et al. (2005) must be invalid (and unusable) because they are implausible. The average for all participants’ before-work salivary caffeine concentration was reported to be $>4 \mu\text{g/ml}$, whereas the estimated mean before-work self-reported caffeine intake was reported to be about 50–60 mg (less than 1 mg/kg of body weight or the approximate equivalent of one cup of tea). The problem is that salivary caffeine concentration could not possibly have been at the levels reported because those levels are far in excess of that which would be observed for persons with normal liver function.

A very approximate rule of thumb is that 1 mg/kg of caffeine results in caffeine plasma levels of about $1 \mu\text{g/ml}$ (James, 1991), and salivary caffeine concentration may be expected to be about 70% of plasma levels (Walther et al., 1983). As an empirical illustration, Childs and de Wit (2006) found that participants given 150 mg caffeine (the approximate equivalent of one strong cup of coffee) had a mean salivary caffeine concentration of $2.8 \mu\text{g/ml}$ when measured 90 min later. Making generous allowances for variation in method and timing, before-work salivary caffeine concentration in the Christopher et al. (2005) study would not be expected to exceed $1 \mu\text{g/ml}$ and would more likely have been in the region of about one-tenth of the levels that were reported. Moreover, mean after-work salivary caffeine concentration was marginally lower than the before-work level, suggesting that participants somehow had lower systemic levels of caffeine later compared to earlier in the day despite supposedly having consumed substantial amounts of caffeine throughout the day. The authors subsequently acknowledged that their salivary-assay methodology was “probably” flawed (Hewlett and Smith, 2006), but that was not accompanied by any tempering of the claim that caffeine was shown to have net benefits for cognitive performance (Hewlett and Smith, 2006; Smith et al., 2006, 2013).

Attempts to measure relevant periods of caffeine exposure and non-exposure using salivary caffeine assays have not always fared any better in studies of low- and non-consumers of caffeine. Smith et al. (2006) compared the effects of caffeine on cognitive performance following overnight abstinence in habitual consumers and non-consumers, and salivary caffeine was measured for the purpose of excluding participants who breached protocol. It is curious that when the same exclusion criterion (salivary caffeine $>2 \mu\text{g/ml}$) was applied to both groups, recent caffeine consumption was found to be twice as frequent (4 instances versus 2) among “non-consumers” than consumers. More noteworthy, however, is the choice of an exclusion criterion of $>2 \mu\text{g/ml}$, especially when applied to supposed overnight-abstinent non-consumers. Salivary caffeine concentration $>2 \mu\text{g/ml}$ is evidence of substantial caffeine ingestion; the approximate equivalent, for example, of a couple of standard caffeine beverages consumed within the previous couple of hours or larger amounts consumed more distally within the past 24 h.

Considering the relaxed exclusion criterion employed by Smith et al. (2006), it is possible that breach of protocol was rampant, with consumers failing to remain overnight abstinent and so-called non-consumers being abstinent neither long-term nor overnight. As such, the study fails as a test of differences in response to caffeine between consumers and non-consumers. Conversely, a recent study by Rogers

et al. (2013) is illuminating in part because it shared some features in common with the study by Smith et al. (2006). Specifically, both studies involved habitual consumers and low-consumers who were required to abstain overnight before being challenged with caffeine. The Rogers et al. (2013) study, however, used an exclusion criterion of $>0.2 \mu\text{g/ml}$, exactly 10-fold lower than that used by Smith et al. (2006). This difference in method could well explain the difference in findings from the two studies. Contrary to findings reported by Smith et al. (2006), Rogers et al. (2013) found that caffeine withdrawal negatively affected cognitive performance in habitual consumers and the effect was reversed following caffeine challenge. Additionally, caffeine had no effect on cognitive performance in low consumers.

In a study that compared the effects of repeated doses of caffeine in “high, low and non-consumers”, Hewlett and Smith (2006) reported mean levels of self-reported habitual caffeine consumption for the three groups, which included the surprisingly precise estimate of “0.0 mg/day” for the non-consumers. Although no attempt was made at objective verification, the authors purport to “show no effect of overnight caffeine withdrawal on mood and performance”. However, null results are always problematic to interpret, as failure to detect effects can so easily be due to methodological weaknesses. Such would appear to be the case with the Hewlett and Smith (2006) study, especially when their findings are considered in the context of other evidence concerning caffeine withdrawal.

Caffeine withdrawal effects have been rigorously investigated, and the evidence of their existence is extensive and unequivocal (Juliano and Griffiths, 2004). Thus, the most plausible explanations of Hewlett and Smith’s (2006) failure to observe any difference in withdrawal effects when comparing consumers and non-consumers is either that participants were not in fact overnight abstinent at the critical point of observation or reputed non-consumers were not what they were thought to be. Failure to objectively verify participants’ self-reports in that study means that the doubt cannot be resolved. However, considering the weight of evidence confirming the reality of caffeine withdrawal, weakness of methodology is the most plausible explanation of Hewlett and Smith’s (2006) failure to observe any such effects.

4. Are caffeine-naïve participants representative?

Considering the high population prevalence of daily caffeine consumption, caffeine-naïve persons, by definition, are a small self-selected minority. Consequently, the generalisability (i.e., external validity) of findings from caffeine-naïve participants to regular consumers is necessarily open to question. For example, negative reactions to caffeine are often reported by those who are infrequent consumers (Rogers et al., 1995), and this not only suggests reasons for their low use but may also reflect differences, including presence or absence of specific gene variants, that limit the potential to generalise between groups whose levels of caffeine consumption vary. Caffeine-induced anxiety may serve to illustrate the point. Given that caffeine’s effects are largely mediated by the adenosine receptor system, it is noteworthy that adenosine is also thought to be involved in the regulation of anxiety (Alsense et al., 2003). Variation in self-reported anxiety after caffeine administration has been found to be associated with two linked polymorphisms on the A_{2A} receptor gene (Alsense et al., 2003). Difference in genotype and phenotype between caffeine-naïve persons and regular consumers suggests that neither group is a suitable model of caffeine’s effects for the other. Accordingly, the assumption that results obtained from studies of habitual low- and non-consumers are generalisable to regular caffeine consumers is not merited.

Whereas Borota et al. (2014) included only caffeine naïve participants in their study in the mistaken belief that results would simply generalise to consumers, others have expressly examined differences between naïve participants (low- or non-consumers) and habitual consumers by including both groups in the same experiment. Understanding the differences has helped to elucidate the pattern of effects that

caffeine has on cognitive performance. In particular, results have been consistent with the predictions of withdrawal reversal in showing that performance is worse in caffeine withdrawn consumers than in naïve participants and that caffeine improves performance in consumers up to but not above that of naïve participants (Goldstein et al., 1969; Rogers et al., 1995, 2003, 2013). That is, caffeine withdrawal has negative effects on cognitive performance for consumers, whereas caffeine administration has little or no net benefit for either consumers or those who are habitual low- or non-consumers. The absence of comparison between caffeine-naïve participants and habitual consumers in the Borota et al. (2014) study means that effects on long-term memory for naïve participants (if that is what they were), while of potential interest in their own right, shed little if any light on effects for habitual consumers.

5. Even if confounding due to withdrawal reversal is avoided what about tolerance?

Use of naïve participants may avoid confounding due to withdrawal reversal, but simultaneously creates potential for confounding due to caffeine tolerance. Drug tolerance refers to the progressive reduction in responsiveness that sometimes accompanies repeated exposure to a drug, wherein the same drug dose has less effect following repeated use or an increased dose is required to produce effects previously experienced. Compared to caffeine withdrawal, which is pronounced and has been reliably characterised, caffeine tolerance in humans is more subtle and less well-delineated (e.g., James, 1997). In particular, there are differences in pattern between different measured outcomes. For example, whereas some indices of subjective experience appear to undergo complete tolerance (e.g., “mental alertness” and “jitteriness”), some objective indices of motor performance appear to be relatively free of tolerance (e.g., finger tapping speed) (Rogers et al., 2013). Because the time course of caffeine tolerance is broadly similar to that for withdrawal (i.e., it occurs within about 3–5 days), the long-term approach summarised in Table 1 protects against confounding due to tolerance (as well as withdrawal reversal), which is another reason for advocating long-term studies of caffeine.

The justifiable assumption that caffeine-naïve participants are non-tolerant to caffeine (i.e., if they have not experienced caffeine they should not have developed tolerance to its effects), could suggest that the study of naïve participants avoids confounding due to tolerance. In fact, using caffeine-naïve participants in this context can create more problems than it solves. To begin with, the fact that regular consumers are likely to have developed tolerance to certain caffeine effects (e.g., jitteriness) limits the extent to which findings can be generalised from non-tolerant naïve participants. Additionally, if the study protocol includes repeated administrations of caffeine, as is often the case, consumers and naïve participants may be expected to respond differently, again limiting generalisation of findings from one group to the other. That is, whereas consumers are likely to respond to repeated administrations of caffeine in ways that are relatively unaffected by tolerance (because tolerance has already developed due to their habitual consumption), naïve-participants are likely to respond in ways that are affected by tolerance induced by repeated administration of the drug. Thus, whenever participants in studies experience repeated administrations of caffeine, such as occurred in the study by Borota et al. (2014), the claim that findings are generalisable to regular caffeine consumers is unmerited. Again, study designs that incorporate periods of long-term caffeine abstinence and exposure (i.e., daily abstinence or consumption over successive days) offer the most viable option for teasing out similarities and differences in caffeine withdrawal and tolerance as experienced by consumers and non-consumers.

6. Future research and conclusions

Consumed by the majority of people worldwide, caffeine is the most widely and habitually ingested drug in history. In the absence of

substantial embellishments, the classic drug-challenge protocol, long considered the gold standard of drug research, is not a suitable design for revealing the effects of a drug that is both widely available and subject to development of withdrawal and tolerance. Nevertheless, studies of caffeine that incorporate standard drug-challenge design continue to be published, signifying need on the part of editors and reviewers to be attuned to the inherent ambiguity of results arising from such studies. The Borota et al. (2014) study is one such recent example.

Without seeking to foreclose on the ingenuity of future researchers to devise novel strategies for illuminating longstanding problems, the type of research design that currently holds most promise for elucidating caffeine's manifold subtle effects is one that extends the standard drug-challenge design by incorporating long-term manipulation of caffeine exposure and abstinence. The extended design summarised in Table 1 is suitable for use with both naïve participants and habitual consumers in studies that examine both groups separately or together. Moreover, whereas the classic drug-challenge protocol is favoured for revealing acute effects, the extended design advocated here elucidates both acute and long-term effects, enabling net caffeine effects to be isolated that are unconfounded by withdrawal and tolerance.

When the aim of research is to examine effects in persons who are caffeine naïve, there is a need for researchers to be rigorous in defining the population of interest. Indeed, it should be acknowledged that truly caffeine-naïve persons are rare in most populations, and that it is more reasonable to speak of low-consumption or a history of little exposure to caffeine. Although the level of exposure that defines “low” and “little” must inevitably be somewhat arbitrary, Rogers et al. (2013) have possibly progressed furthest in specifying empirically-sound parameters for recruiting study participants using self-reports of low exposure substantiated by objective measurement of salivary caffeine, with concentration of ≤ 0.2 $\mu\text{g}/\text{ml}$ indicating selection and >0.2 $\mu\text{g}/\text{ml}$ indicating exclusion.

Researchers should also be rigorous in specifying the reasons for choosing to examine low-consumers. Although low-consumers represent a population of interest in their own right, most of the interest they have attracted to date has been based on the false premise that findings are generalisable to habitual consumers. Again, the Borota et al. (2014) study is one such example. However, if generalisation to habitual consumers is the intended purpose of studying low-consumers, then study designs, as well as being long-term, should incorporate both low-consumers and habitual consumers to permit controlled comparisons of effects in both groups.

One source of discouragement from undertaking long-term caffeine studies relates to the challenges inherent in any long-term human research, including demands on participants as well as demands on research materials (e.g., daily monitoring of caffeine intake and other study parameters, and procurement and analysis of daily saliva samples) and research personnel needed to maintain participant observations over weeks rather than days. A relevant question concerns the exact meaning of “long-term”. The operational definition of long-term in the research design summarised in Table 1 is one week (or more precisely, 6 days), because that design is predicated on four consecutive one-week phases of exposure to and abstinence from caffeine. It is fortunate that due to relatively short half-life, caffeine withdrawal and tolerance substantially abate within one week, enabling key questions concerning a wide range of long-term effects, including psychological (e.g., performance and mood) and physiological (e.g., blood pressure; James and Gregg, 2004b) to be revealed within the time-frame of four consecutive weeks in repeated-measures designs in which participants serve as their own controls.

However, exposure to caffeine is essentially life-long for most people. Therefore, the operational definition of long-term that underpins the design summarised in Table 1, while useful for particular research aims should not be considered immutable for all research purposes. Because most pregnant women consume caffeine and caffeine crosses the placenta (Bonita et al., 2007; Brazier and Salle, 1981; Van't Hoff, 1982),

initial exposure for most people precedes birth (James, 1997). Exposure in the form of soft drinks typically continues throughout childhood, and although patterns of consumption may change (e.g. intake switches from soft drinks during childhood to coffee in adulthood) usage tends to consolidate during adolescence and early adulthood. Thereafter, consumption tends to stabilise, generally undergoing comparatively little change for the remainder of life. Lifelong exposure and unparalleled prevalence raise questions concerning possible implications of dietary caffeine for health. Thus, in addition to studies of the kind advocated here, ultimately there is need for studies that examine behavioural and pharmacological effects of caffeine not merely for a period of weeks, but also for months and possibly years.

Studies of caffeine-naïve persons are important for revealing potentially novel caffeine effects. However, findings from such studies are subject to formidable limitations. In particular, the study of caffeine-naïve persons alone cannot solve problems of confounding due to caffeine withdrawal and withdrawal reversal encountered in studies of habitual consumers. Major methodological challenges arise concerning the definition of “caffeine-naïve” (and the related concepts of “low-consumer” and “non-consumer”), the population representativeness of participants deemed to be caffeine-naïve, and confounding due to caffeine tolerance. By drawing attention to these key methodological challenges, the present mini-review aims to promote better controlled studies capable of delineating the complexities of caffeine's effects on cognitive performance and the mechanisms involved.

References

- Alsense K, Deckert J, Sand P, de Wit H. Association between A_{2A} receptor gene polymorphisms and caffeine-induced anxiety. *Neuropsychopharmacology* 2003;28:1694–702.
- Bonita JS, Mandarano M, Shuta D, Vinson J. Coffee and cardiovascular disease: *in vitro*, cellular, animal, and human studies. *Pharmacol Res* 2007;55:187–98.
- Borota D, et al. Post-study caffeine administration enhances memory consolidation in humans. *Nat Neurosci* 2014. <http://dx.doi.org/10.1038/nn.3623>.
- Brazier JL, Salle B. Conversion of theophylline to caffeine by the human fetus. *Semin Perinatol* 1981;5:315–20.
- Childs E, de Wit H. Subjective, behavioral, and physiological effects of acute caffeine in light, nondependent caffeine users. *Psychopharmacology (Berl)* 2006;185:514–23.
- Christopher G, Sutherland D, Smith A. Effects of caffeine in non-withdrawn volunteers. *Hum Psychopharmacol Clin Exp* 2005;20:47–53.
- Einöther SJ, Giesbrecht T. Caffeine as an attention enhancer: reviewing existing assumptions. *Psychopharmacology (Berl)* 2013;225:251–74.
- Ferré S. An update on the mechanisms of the psychostimulant effects of caffeine. *J Neurochem* 2008;105:1067–79.
- Goldstein A, Kaizer S, Whitby O. Psychotropic effects of caffeine in man. IV. Quantitative and qualitative differences associated with habituation to coffee. *Clin Pharmacol Ther* 1969;10:489–97.
- Griffiths RR, Evans SM, Heishman SJ, Preston KL, Sannerud CA, Wolf B, et al. Low-dose caffeine physical dependence in humans. *J Pharmacol Exp Ther* 1990;255:1123–32.
- Haskell CF, Kennedy DO, Wesnes KA, Scholey AB. Cognitive and mood improvements of caffeine in habitual consumers and habitual non-consumers of caffeine. *Psychopharmacology (Berl)* 2005;179:813–25.
- Heatherley SV, Hancock KMF, Rogers PJ. Psychostimulant and other effects of caffeine in 9- to 11-year-old children. *J Child Psychol Psychiatry* 2006;47:135–42.
- Hewlett PJ, Smith AP. Acute effects of caffeine in volunteers with different patterns of regular consumption. *Hum Psychopharmacol Clin Exp* 2006;21:167–80.
- Hollingworth HL. The influence of caffeine on mental and motor efficiency. *Arch Psychol* 1912a;22:1–166.
- Hollingworth HL. The influence of caffeine on the speed and quality of performance in typewriting. *Psychol Rev* 1912b;19:66–73.
- Hughes JR, Oliveto AH, Bickel WK, Higgins ST, Badger GJ. Caffeine self-administration and withdrawal: incidence, individual differences and interrelationships. *Drug Alcohol Depend* 1993;32:239–46.
- James JE. Caffeine and health. London: Academic Press; 1991. p. 430.
- James JE. Does caffeine enhance or merely restore degraded psychomotor performance? *Neuropsychobiology* 1994;30:124–5.
- James JE. Understanding caffeine: a biobehavioral analysis. Thousand Oaks, CA: Sage; 1997.
- James JE. Acute and chronic effects of caffeine on performance, mood, headache, and sleep. *Neuropsychobiology* 1998;38:32–41.
- James JE, Gregg ME. Effects of dietary caffeine on mood when rested and sleep restricted. *Hum Psychopharmacol Clin Exp* 2004a;19:333–41.
- James JE, Gregg ME. Hemodynamic effects of dietary caffeine, sleep restriction, and laboratory stress. *Psychophysiology* 2004b;41:914–23.
- James JE, Rogers PJ. Effects of caffeine on performance and mood: withdrawal reversal is the most plausible explanation. *Psychopharmacology (Berl)* 2005;182:1–8.
- James JE, Gregg ME, Kane M, Harte F. Dietary caffeine, performance and mood: enhancing and restorative effects after controlling for withdrawal relief. *Neuropsychobiology* 2005;52:1–10.
- James JE, Paull I, Cameron-Traub E, Mines JO, Lelo A, Birkett DJ. Biochemical validation of self-reported caffeine consumption during caffeine fading. *J Behav Med* 1988;11:15–30.
- Judelson DA, Armstrong LE, Sokmen B, Roti MW, Casa DJ, Kellogg MD. Effect of chronic caffeine intake on choice reaction time, mood, and visual vigilance. *Physiol Behav* 2005;85:629–34.
- Juliano LM, Griffiths RR. A critical review of caffeine withdrawal: empirical validation of symptoms and signs, incidence, severity, and associated features. *Psychopharmacology (Berl)* 2004;176:1–29.
- Lelo A, Miners JO, Robson RA, Birkett DJ. Assessment of caffeine exposure: caffeine content of beverages, caffeine intake, and plasma concentrations of methylxanthines. *Clin Pharmacol Ther* 1986a;39:54–9.
- Lelo A, Miners JO, Robson RA, Birkett DJ. Quantitative assessment of caffeine partial clearances in man. *Br J Clin Pharmacol* 1986b;22:183–6.
- Lieberman HR, Wurtman RJ, Emde GG, Roberts C, Coviella ILG. The effects of low doses of caffeine on human performance and mood. *Psychopharmacology (Berl)* 1987;92:308–12.
- Pfeifer RW, Notari RE. Predicting caffeine plasma concentrations resulting from consumption of food or beverages: a simple method and its origin. *Drug Intell Clin Pharmacol* 1988;22:953–9.
- Rodriguez del Rey Z, Granek EF, Sylvester S. Occurrence and concentration of caffeine in Oregon coastal waters. *Mar Pollut Bull* 2012;6:1417–24.
- Rogers PJ. Caffeine and alertness: the debate about withdrawal reversal. *J Caffeine Res* 2014;4:3–8.
- Rogers PJ, Richardson NJ, Elliman NA. Overnight caffeine abstinence and negative reinforcement of preference for caffeine-containing drinks. *Psychopharmacology (Berl)* 1995;120:457–62.
- Rogers PJ, Martin J, Smith C, Heatherley SV, Smit HJ. Absence of reinforcing, mood and psychomotor performance effects of caffeine in habitual non-consumers of caffeine. *Psychopharmacology (Berl)* 2003;167:54–62.
- Rogers PJ, Heatherley SV, Hayward RC, Seers HE, Hill J, Kane M. Effects of caffeine and caffeine withdrawal on mood and cognitive performance degraded by sleep restriction. *Psychopharmacology (Berl)* 2005;179:742–52.
- Rogers PJ, Heatherley SV, Mullings EL, Smith JE. Faster but not smarter: effects of caffeine and caffeine withdrawal on alertness and performance. *Psychopharmacology (Berl)* 2013. <http://dx.doi.org/10.1007/s00213-012-2889-4>.
- Smit HJ, Rogers PJ. Effects of caffeine on cognitive performance, mood and thirst in lower and higher caffeine consumers. *Psychopharmacology (Berl)* 2000;152:167–73.
- Smith AP, Christopher G, Sutherland D. Effects of caffeine in overnight-withdrawn consumers and non-consumers. *Nutr Neurosci* 2006;9:63–71.
- Smith AP, Christopher G, Sutherland D. Acute effects of caffeine on attention: a comparison of non-consumers and withdrawn consumers. *J Psychopharmacol* 2013;27:77–83.
- Van't Hoff W. Caffeine in pregnancy. *Lancet* 1982;2:1020.
- Walther H, Banditt P, Kohler E. Significance of caffeine values in serum, saliva and urine: determination of pharmacokinetic data by non-invasive methods in psychopharmacologic studies. *Pharmacopsychiatry* 1983;16:166–70.
- Yassa, M.A. (2014). Personal communication, 15 January.
- Yeomans MR, Ripley T, Davies LH, Rusted JM, Rogers PJ. Effects of caffeine on performance and mood depend on the level of caffeine abstinence. *Psychopharmacology (Berl)* 2002;164:241–9.