AUTHOR QUERY FORM

Journal Title: Psychological Science

Article Number: 654911

Dear Author/Editor,

Greetings, and thank you for publishing with SAGE Publications. Your article has been copyedited, and we have a few queries for you. Please respond to these queries when you submit your changes to the Production Editor.

Thank you for your time and effort.

Please assist us by clarifying the following queries:

Sl. No.	Query
1	Please check that all authors are listed in the proper order; clarify which part of each author's name is his or her surname; and verify that all author names are correctly spelled/punctuated and are presented in a manner consistent with any prior publications.
2	Please review the entire document for typographical errors, mathematical errors, and any other necessary corrections; check headings, tables, and figures.
3	Please confirm that you have sufficiently reviewed your proof and queries, and that you understand this is your FINAL opportunity to review your article before publication.

Research Article

Psychological Science
1–8
© The Author(s) 2016
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797616654911
pss.sagepub.com

The Bitter Truth About Sugar and Willpower: The Limited Evidential Value of the Glucose Model of Ego Depletion

Miguel A. Vadillo[1], Natalie Gold[2], and Magda Osman[3][AQ: 1][AQ: 2][AQ: 3]
[1]Department of Primary Care and Public Health Sciences, King's College London; [2]Department of Philosophy, King's College London; and [3]Department of Biological and Experimental Psychology, Queen Mary University of London

Abstract

Ego depletion is the hypothesis that willpower draws on a limited mental resource, so that engaging in an act of self-control impairs self-control in subsequent tasks. To present ego depletion as more than a convenient metaphor, some researchers have proposed that glucose is the limited resource that becomes depleted with self-control. However, there have been theoretical challenges to the proposed glucose mechanism, and the experiments that have tested it have found mixed results. We used a new meta-analytic tool, *p*-curve analysis, to examine the reliability of the evidence from these experiments. We found that the effect sizes reported in this literature are possibly influenced by publication or reporting bias and that, even within studies yielding significant results, the evidential value of this research is weak. In light of these results, and pending further evidence, researchers and policymakers should refrain from drawing any conclusions about the role of glucose in self-control.

Keywords

ego depletion, glucose, meta-analysis, *p*-curve, self-control, open data

Received 11/20/15; Revision accepted 5/25/16

According to the law of conservation of energy, the total amount of energy of an isolated system can never increase. In the domain of psychology, the idea that energy is a limited resource originated with Freud (1923/1961). Energy models have been little used in psychology since Freud, though, with the rare exception of the *ego-depletion* model developed by Baumeister, Bratslavsky, Muraven, and Tice (1998). According to this research, "the self's acts of volition draw on some limited resource, akin to strength or energy. . ., therefore, one act of volition will have a detrimental impact on subsequent volition" (Baumeister et al., 1998, p. 1252). Research on ego depletion has substantial implications. It has been claimed that reliably exerting self-control, either actively doing something "good" or avoiding the temptation to act on "bad" impulses, can greatly reduce many of the major ills that affect society and people's personal lives, such as "crime, violence, unwanted pregnancy, drug addiction, venereal diseases, bankruptcy, and premature deaths" (Baumeister, Muraven, & Tice, 2000, pp. 130). It is not surprising that the work of Baumeister et al. (1998) has affected a number of disciplines, including advertising, behavioral economics, business, consumerism, law, management, marketing, and medicine. In fact, it is fair to say that this seminal article by Baumeister et al. has become a classic: At the time of this writing, it has more than 1,250 citations in the *Web of Science*.

When ego depletion was first proposed, the idea of a limited resource was a convenient metaphor. Given how fundamental exerting self-control is thought to be, it is important to establish the energy source that is depleted and to provide a mechanism by which ego depletion occurs. The most popular explanation found in the literature involves glucose. Gailliot et al. (2007) presented nine studies supporting three main findings: (a) Blood glucose levels are reduced after performing a self-control

Corresponding Author:
Magda Osman, School of Biological and Chemical Sciences, Queen Mary University of London, Fogg Building, Mile End Rd., London E14NS, United Kingdom
E-mail: m.osman@qmul.ac.uk

task but not after performing a comparable cognitive task that does not require self-control; (b) low levels of blood glucose after a first self-control task predict behavioral deficits on a second self-control task; and (c) participants whose glucose levels are restored by ingesting a glucose drink after a self-control task perform better on a subsequent task than do participants who are given a diet drink between tasks. They concluded that self-control depletes blood glucose, which leads to decreased self-control on subsequent tasks, and restoring glucose levels replenishes the ability to exert self-control.

Although the conclusions drawn by Gailliot et al. (2007) have been extraordinarily influential, their glucose hypothesis remains controversial. The mechanism they propose has been challenged, and the reliability of their results has been disputed. Kurzban (2010) argued that the glucose mechanism, as presented by Gailliot et al., is biologically implausible. The mechanism invokes the idea that self-control tasks deplete glucose because of energy consumption by the brain, but the supporting evidence shows changes only in blood glucose levels. Kurzban cited evidence that the sort of self-control tasks used in the literature have little effect on brain metabolism and that changes in blood glucose are unlikely to reflect blood glucose uptake by the brain.

There are also concerns about the empirical evidence for the glucose mechanism. For instance, Schimmack (2012) showed that the number of significant results reported by Gailliot et al. (2007) is too large, given their average power. In other words, these results are likely to be influenced by publication bias or *p*-hacking (see Francis, 2012; Simmons, Nelson, & Simonsohn, 2011). Kurzban's (2010) concerns were supported by a reanalysis of the data from Gailliot et al., in which he found that self-control does not decrease blood glucose levels, and by recent failures to replicate the effect of glucose on self-control (Job, Walton, Bernecker, & Dweck, 2013; Kelly, Sünram-Lea, & Crawford, 2015; Lange & Eggert, 2014; Lange, Seer, Rapior, Rose, & Eggert, 2014).

The effect of glucose on self-control has been examined in three broad categories of studies:

- *Correlational studies* measure effects on blood glucose levels before and after self-control is exerted (Dvorak & Simons, 2009; Gaillot, 2012; Gailliot et al., 2007; Molden et al., 2012).
- *Glucose-ingestion studies* manipulate blood glucose levels by leaving enough time between ingestion and a control task for glucose to be absorbed into the bloodstream (Birnie, Smallwood, Reay, & Riby, 2015; Denson, von Hippel, Kemp, & Teo, 2010; DeWall, Baumeister, Gailliot, & Maner, 2008; Dickinson, McElroy, & Stroh, 2014; Gailliot et al., 2007; Gailliot, Peruche, Plant, & Baumeister, 2009; Howard & Marczinski, 2010; Job et al., 2013; Kelly et al., 2015; Lange & Eggert, 2014; Lange et al., 2014; Masicampo & Baumeister, 2008; Wang & Dvorak, 2010). Although the findings from these studies provide mixed support for the glucose hypothesis, the methods used imply a mechanism that is consistent with the proposals of Gailliot et al. (2007).
- *Glucose-rinsing studies* examine the impact of simply rinsing one's mouth with a glucose solution before exerting self-control (Hagger and Chatzisarantis, 2013; Lange & Eggert, 2014; Molden et al., 2012; Sanders, Shirk, Burgin, & Martin, 2012). Results from these studies suggest that the signal of glucose from the mouth to the brain is sufficient to neutralize the ego-depletion effect. This mechanism is consistent with the results of the ingestion manipulations but suggests that the effect does not depend on a metabolic explanation.

To help settle the growing concerns in the academic community regarding the reliability of the glucose mechanism, which in turn has implications for the ego-depletion hypothesis that it underpins, we sought to use a new meta-analytic tool, *p*-curve analysis, to investigate the presence of publication and reporting biases in this literature.

Method

Literature-search strategy

We looked for studies supporting the idea that sugar consumption is related to ego depletion and self-regulation. Specifically, we considered any study exploring the hypothesis that glucose ingestion or rinsing improves performance (e.g., overcoming an impulse, inhibiting an aggressive reaction, or controlling a cognitive process) or ameliorates the effect of an ego-depleting experience in these laboratory self-regulation tasks. We also included studies testing whether performance in laboratory tasks (again, specifically those that explicitly require self-regulation) is correlated with pre- or posttesting sugar levels. Studies in which participants were not asked to drink a sugary beverage but simply to rinse their mouths with it were also included in the present analyses; this literature also supports the idea that sugar consumption improves self-regulation (even if it challenges the specific hypothesis that such improvement is achieved through metabolic processes).

Given these criteria, we excluded experiments showing a relation between sugar consumption and cognitive processes (e.g., short-term memory or general cognition function) that prima facie do not seem to pose demands on self-regulation (e.g., Carter & McCullough, 2013;

Owen, Scholey, Finnegan, Hu, & Sünram-Lea, 2012). In addition, we also excluded studies that investigated the correlation between general glucose levels (or regular glucose ingestion) and self-regulated behavior in naturalistic settings over many days. This included, for instance, studies on the relationships between glucose ingestion and smoking cessation and studies on the relationship between diabetes and various psychological processes (see Gailliot & Baumeister, 2007). These studies rely on measures that differ substantially from the dependent variables gathered in laboratory-based ego-depletion tasks, and the lack of experimental control makes the results amenable to alternative explanations that bear little or no relation to ego depletion and self-regulation.

We began our search by inspecting a small set of studies that had included an exhaustive literature review. These included a meta-analysis by Hagger, Wood, Stiff, and Chatzisarantis (2010) on the general ego-depletion literature, a study by Job et al. (2013) exploring individual differences in the impact of glucose on self-control, and Lange and Eggert's (2014) recent attempt to replicate the effect of sugar consumption or rinsing on ego depletion. Then, to make sure that we included all relevant studies, we conducted a systematic search in *Web of Science* and *Google Scholar* for the term "glucose" along with "ego depletion," "self-control," or "self-regulation." This strategy allowed us to identify 18 articles with one or more eligible studies. All these studies are listed in Table 1 and are also marked with asterisks in the reference list. Furthermore, we found out that one of our selected studies (Masicampo & Baumeister, 2008) had been included in the famous project on the reproducibility of psychological science (Open Science Collaboration, 2015). This replication was also included in our analyses, which resulted in a total of 19 articles.

P-curve analysis

To assess the reliability of this set of studies, we used *p*-curve analysis, a recently designed meta-analytic tool that allows for the exploration of various biases solely by examining the distribution of significant *p* values (Simonsohn, Nelson, & Simmons, 2014). Imagine a set of studies exploring an effect that does not exist. Occasionally, these studies will yield a significant result (i.e., a *p* value lower than .05) just by chance. In this scenario, all *p* values will be equally likely: 5% of studies will have *p* values lower than .05, 4% of studies will have *p* values lower than .04, and so on. Consequently, the *p* values of a set of studies exploring a nonexistent effect should typically follow a flat distribution. Note that this is not the case if the studies are exploring a true effect: In that case, significant *p* values should follow a right-skewed distribution in which small *p* values (e.g., $p < .01$) are more prevalent than larger *p* values (e.g., *p* between .04 and .05). As explained by Simonsohn et al. (2014), this can be easily understood if one imagines an experimenter exploring a very large effect with a large sample of participants. Most likely, the experimenter will observe a very low *p* value. Experiments with smaller effect sizes and smaller samples are simply less extreme versions of this ideal scenario. Even for low-powered studies, the distribution of *p* values should be right skewed. This implies that, in principle, one can know whether a set of experiments is exploring true effects or null effects simply by checking whether their *p* values follow a right-skewed distribution or a rather flat distribution. An interesting feature of this approach is that it focuses exclusively on significant *p* values (i.e., studies for which $p < .05$); consequently, its results are unaffected by publication bias.

Simonsohn et al. (2014) designed an online application (available at http://www.p-curve.com) that allows researchers to test whether an observed distribution of *p* values is significantly right skewed or suspiciously flat, which could suggest that the significant results are false positives. A simple way to test whether the distribution of *p* values is significantly right skewed is to compare the number of significant *p* values lower than .025 with the number of *p* values between .025 and .05 by using a binomial test. A potential shortcoming of this approach is that this binomial test gives the same weight to exceptionally small *p* values (e.g., .00001) as to *p* values barely smaller than .025 (e.g., .024). To overcome this limitation, the latest versions of the *p*-curve application conduct not only a binomial test but also an alternative analysis, known as a continuous test, that is sensitive to the exact *p* values.

If the distribution of *p* values is not significantly right skewed, this might mean that the studies lack any evidential value or, in other words, that the significant results could be false positives. However, failure to find a significant right-skewed distribution might also be due to a lack of statistical power (e.g., if the analysis includes a very small number of studies). Simonsohn et al. (2014) suggested that in order to determine whether the distribution of *p* values is too flat, one should test whether the *p*-curve is flatter than the theoretical distribution that one would observe in a set of studies with 33% statistical power. If the *p*-curve is significantly flatter than this very flat standard, a common conclusion is that the set of studies might lack evidential value and that they might be the product of publication bias, selective reporting, or *p*-hacking.

Selection of statistical contrasts

We selected the key statistical contrasts of each study following the guidelines offered by Simonsohn et al. (2014). In the case of correlational studies or experiments with just two groups, we registered the target correlation

Table 1. Studies Included in the Analyses and Their Key Statistical Contrasts

Study	Sugar rinsing?	Key statistical contrast	*p*
Birnie, Smallwood, Reay, and Riby (2015)	No	$t(15) = 2.469$[a]	.02605
Denson, von Hippel, Kemp, and Teo (2010) Study 1	No	$t(67) = -2.19$	.03201
Denson et al. (2010) Study 2	No	$t(151) = 2.24$	.02655
DeWall, Baumeister, Gailliot, and Maner (2008) Study 2	No	$F(1, 55) = 6.64$	.01268
Dickinson, McElroy, and Stroh (2014)	No	$z = 1.88$	.06011
Dvorak and Simons (2009)	No	$F(1, 177) = 5.63$	.01873
Gailliot (2012)	No	$r(50) = -.30$	.03071
Gailliot et al. (2007) Study 1	No	$F(1, 100) = 6.08$	.01537
Gailliot et al. (2007) Study 2	No	$t(33) = 2.20$	.03492
Gailliot et al. (2007) Study 3	No	$r(14) = -.62$[a]	.01041
Gailliot et al. (2007) Study 4	No	$r(10) = .56$	.05828
Gailliot et al. (2007) Study 5	No	$r(21) = .45$	.03120
Gailliot et al. (2007) Study 6	No	$r(15) = .43$	.08493
Gailliot et al. (2007) Study 7	No	$F(1, 57) = 5.04$	.02866
Gailliot et al. (2007) Study 8	No	$F(1, 69) = 5.45$	.02249
Gailliot et al. (2007) Study 9	No	$t(16) = 3.13$	.00646
Gailliot, Peruche, Plant, and Baumeister (2009)	No	$t(47) = 2.21$[a]	.03201
Hagger and Chatzisarantis (2013) Study 1	Yes	$F(1, 24) = 8.42$	.00783
Hagger and Chatzisarantis (2013) Study 2	Yes	$F(1, 30) = 6.12$	.01925
Hagger and Chatzisarantis (2013) Study 3	Yes	$F(1, 32) = 4.06$	.05238
Hagger and Chatzisarantis (2013) Study 4	Yes	$F(1, 40) = 10.32$	.00260
Hagger and Chatzisarantis (2013) Study 5	Yes	$F(1, 36) = 7.28$	.01055
Howard and Marczinski (2010)	No	$F(4, 75) = 2.95$	.02544
Job, Walton, Bernecker, and Dweck (2013) Study 1	No	$t(78) = 2.10$	.03896
Job et al. (2013) Study 2	No	$F(1, 58) = 5.16$	.02684
Job et al. (2013) Study 3	No	$F(1, 139) = 5.28$	.02306
Kelly, Sünram-Lea, and Crawford (2015)	No	$F(1, 67) = 0.80$	.37430
Lange and Eggert (2014) Study 1	No	$F(1, 68) = 1.12$	.29366
Lange and Eggert (2014) Study 2	No	$F(1, 110) = 0.01$[a]	.92053
Lange, Seer, Rapior, Rose, and Eggert (2014)	No	$t(68) = 0.05$[a]	.96027
Masicampo and Baumeister (2008)	No	$F(1, 111) = 5.311$	.02305
Molden et al. (2012) Study 1	No	$F(1, 83) = 2.05$	.15596
Molden et al. (2012) Study 2	Yes	$F(1, 39) = 4.54$	.03947
Molden et al. (2012) Study 3	Yes	$F(1, 28) = 5.02$	.03317
Open Science Collaboration (2015); replication of Masicampo and Baumeister (2008)	No	$F(1, 158) = 0.379$	.53902
Sanders, Shirk, Burgin, and Martin (2012)	Yes	$t(49) = -2.129$	.03831
Wang and Dvorak (2010) Simple effect 1	No	$t(31) = 2.55$	.01593
Wang and Dvorak (2010) Simple effect 2	No	$t(32) = 3.12$	.00381

Note: The statistical contrasts were selected according to the guidelines of Simonsohn, Nelson, and Simmons (2014). A complete *p*-curve disclosure table justifying the selection of each statistical contrast is available at Open Science Framework (https://osf.io/yf8p3/).

[a]These statistical contrasts were replaced by alternative contrasts in the robustness test (for further details, see the *p*-curve disclosure table at the Open Science Framework).

coefficient, the statistic testing the regression slope, or the statistic testing the difference of means. In complex factorial designs, if researchers expected the ego-depletion effect to disappear in a specific condition, then we registered the statistic testing the interaction. In contrast, if they expected to find a complete cross-over interaction, we registered the statistics for the two simple effects. A total of 38 statistical contrasts were included in the analyses. In accordance with the recommendations of Simonsohn et al. (2014), when two statistics were equally valid, we used one of them in the main analysis and the other one in a second analysis that we refer to as a *robustness test*. In most cases (four out of five), we adopted the general rule of selecting the first one to

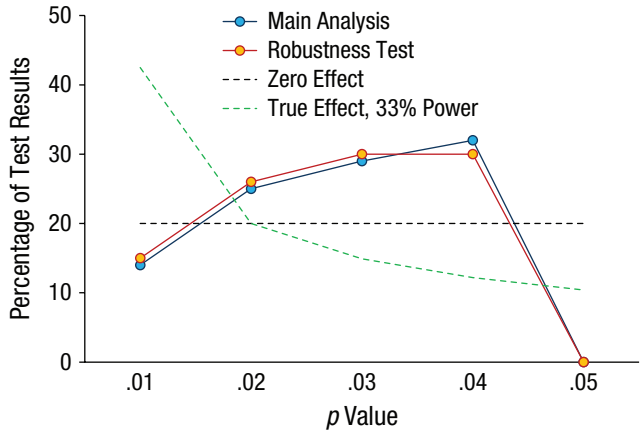

Fig. 1. Distribution of observed *p* values for both the main analysis and the robustness test, along with the expected distribution of *p* values if the null hypothesis is true (zero effect) and if the alternative hypothesis is true but the experiments lack sufficient power (true effect, 33% power).

appear in the text for the main analysis and the second one to appear for the robustness test. However, on one occasion (Birnie et al., 2015), we broke this rule because the conclusions of the authors relied more heavily on one of the statistics than on the other. In this particular case, we selected the more appropriate statistic for the main analysis and the other one for the robustness test. We found no studies in which three or more statistics were equally valid for *p*-curve analysis. A *p*-curve disclosure table with all the selected statistics and the justification for our choices are available at Open Science Framework (https://osf.io/yf8p3/).

Results

The key statistical contrasts of the studies included in our analysis are shown in Table 1. Figure 1 plots the frequency of each range of *p* values among these studies. As the figure shows, the main *p*-curve did not fit with the right-skewed distribution that one would expect if these studies were exploring a true effect. Although there were no significant results immediately below .05, *p* values in the interval between 0 and .04 show, if anything, a left-skewed distribution. Not surprisingly, the statistical contrast testing the right skewness of the *p*-curve was nonsignificant (binomial test comparing the proportions of contrasts with *p* values < .025 and those with *p* values between .025 and .05: $p = .575$; continuous test: $z = -0.83$, $p = .204$). In other words, the distribution of *p* values was not significantly different from what would be expected if the null hypothesis (in this case, that the average effect size is zero) were true. Furthermore, the observed distribution is flatter than would be expected if the studies were simply underpowered. Statistical analyses confirmed that the *p*-curve was significantly flatter than would be expected if there were an effect but the studies had only 33% power on average (binomial test: $p = .019$; continuous test: $z = -2.08$, $p = .019$). Thus, we can reject the hypothesis that although there was an effect, the studies had an average power of only 33% to detect it. The results were very similar for the robustness test, which also failed to find significant evidence of right skewness (binomial test: $p = .500$; continuous test: $z = -0.88$, p = .190) and, in fact, detected that the *p*-curve was significantly flatter than would be expected if there were an effect but the studies had only 33% power on average (binomial test: $p = .033$; continuous test: $z = -1.99$, $p = .023$).

Note that when we removed from the analyses all the studies that explored the effects of glucose rinsing (because these have a somewhat different theoretical background), *p*-curve results did not improve. After the glucose-rinsing experiments were removed, neither the continuous test ($z = -0.57$, $p = .285$) nor the binomial test ($p = .668$) suggested that the remaining set of studies had any evidential value. Furthermore, both tests (continuous: $z = -1.94$, $p = .026$; binomial: $p = .024$) suggested that the *p*-curve was significantly flatter than would be expected if there were an effect but the studies had only 33% power on average. In other words, the poor results of the previous tests cannot be attributed to the inclusion of glucose-rinsing studies in the analyses.

Discussion

The results of our analyses suggest that the relationship between glucose levels and self-control behaviors might be unreliable. Figure 1 shows that the key *p* values of the 19 studies included in the present analyses follow a surprisingly flat distribution. This is exactly the pattern of results that one would expect to find if those results were false positives. These results remain unchanged regardless of whether glucose-rinsing studies are included or excluded from the sample.

These results may not come as a complete surprise given the empirical challenges to the glucose hypothesis suggested by failed replications (Job et al., 2013; Kelly et al., 2015; Lange & Eggert, 2014; Lange et al., 2014) and theoretical critiques regarding its biological plausibility (Kurzban, 2010; Osman, 2014). Furthermore, a detailed analysis of the seminal article suggesting the glucose hypothesis showed that the number of significant findings reported in that article was too large, given the low power of each study (Schimmack, 2012). In other words, the results were too good to be true (Francis, 2012).

Nevertheless, the findings from the present study are a surprise in the context of the wide acceptance of the glucose hypothesis in general scientific research and its popularity, as evidenced by the number of citations of

Gailliot et al. (2007) in the literature and the continued influence of this hypothesis in recent reviews on ego depletion (e.g., Baumeister, 2014; Baumeister & Alghamdi, 2015). Moreover, the hypothesis has intuitive and seemingly practical appeal. If one accepts that a failure of self-control in regulating actions contributes to the many personal and societal problems that people face (Baumeister et al., 2000), then glucose supplements would provide a simple means to enhance willpower and ameliorate these problems (Baumeister & Tierny, 2011). In light of our results, it is doubtful that such a recommendation will work in the real world. This conclusion converges with recent evidence that glucose might have little or no impact on domain-general decision-making tasks (Orquin & Kurzban, 2016) and with an intriguing series of meta-analyses and preregistered replications suggesting that the ego-depletion effect itself might be less robust than previously thought (Carter, Kofler, Forster, & McCullough, 2015; Hagger et al., in press).

Previous criticisms of the glucose model of ego depletion have typically focused on individual articles (e.g., Kurzban, 2010; Schimmack, 2012). Article-level analyses such as those by Francis (2012) are ideal in some respects because they ensure that all the studies under scrutiny are grounded in the same theoretical view and rely on very similar research methods. Unfortunately, only a couple of the articles included in the present review contain a sufficiently large number of studies to allow this type of analysis (Gailliot et al., 2007; possibly Hagger and Chatzisarantis, 2013). An examination of the wider literature, such as the one offered in the present article, must necessarily collate studies with heterogeneous methods and theoretical backgrounds. In exchange, this approach allows researchers to check for publication and reporting biases in areas of research in which articles with a small number of studies are prevalent. In this sense, our study adds to the conclusions of article-level analyses by suggesting that the kinds of biases that have been detected in isolated studies might be representative of the wider area of research on the glucose model of ego depletion. In any case, the rest of the experiments included in the present analyses, with the possible exception of glucose-rinsing studies, share a common theoretical background.

It is worth noting that, as with any other statistical test, *p*-curve analysis is not a flawless indicator of bias (Bishop & Thompson, 2016; Bruns & Ioannidis, 2016; Lakens, 2015). Our results suggest that, on average, these studies have little or no evidential value, but they do not allow us to determine whether the significant results are due to publication bias, selective reporting of outcomes or analyses, *p*-hacking, or all of these. It is not impossible that some of these studies are exploring small but true effects and that their evidential value may be diluted by the biases that pervade the rest of the studies. Perhaps future research will show that glucose does play a role in ego-depletion effects, but our conclusions are based on the analysis of the extant literature in this area. Thus, our contribution must be seen as an additional piece of information in the wider context of attempts to verify the reliability of the glucose model of ego depletion. Note that the kind of biases explored in the present study are prevalent in other (but not all) areas of psychological research (e.g., Bakker, van Dijk, & Wicherts, 2012) and that low reproducibility is not exclusively a problem of psychological research (Camerer et al., 2016; Errington et al., 2014). In fact, it is fair to say that psychology is taking a leading role in the dissemination of open research practices (Open Science Collaboration, 2015). We hope that this new trend in psychological research will soon render meta-analytic studies unnecessary.

Action Editor

D. Stephen Lindsay served as action editor for this article.

Author Contributions

All the authors developed the study concept. The literature search was conducted by M. A. Vadillo and N. Gold. M. A. Vadillo performed the data analysis and interpretation. M. Osman drafted the manuscript, and M. A. Vadillo and N. Gold provided critical revisions. All the authors approved the final version of the manuscript for submission.

Acknowledgments

We are indebted to Greg Francis, David R. Shanks, Eric-Jan Wagenmakers, and two anonymous reviewers for their valuable comments on earlier versions of the manuscript.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This work was supported by European Research Council Grant 283849 (to N. Gold) under the European Union's Seventh Framework Programme (FP/2007–2013).

Open Practices

All data have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/4a6jk/. The complete Open Practices Disclosure for this article can be found at http://pss.sagepub.com/content/by/supplemental-data. This article has received the badge for Open Data. More information about the Open Practices badges can be found at https://osf.io/tvyxz/wiki/1.%20View%20the%20Badges/ and http://pss.sagepub.com/content/25/1/3.full.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science*, *7*, 543–554.

Baumeister, R. F. (2014). Self-regulation, ego depletion, and inhibition. *Neuropsychologia*, *65*, 313–319.

Baumeister, R. F., & Alghamdi, N. G. (2015). Role of self-control failure in immoral and unethical actions. *Current Opinion in Psychology*, *6*, 66–69.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*, 1252–1265.

Baumeister, R. F., Muraven, M., & Tice, D. M. (2000). Ego depletion: A resource model of volition, self-regulation, and controlled processing. *Social Cognition*, *18*, 130–150.

Baumeister, R. F., & Tierny, J. (2011). *Willpower: Rediscovering the greatest human strength*. New York, NY: Penguin Press.

*Birnie, L. H. W., Smallwood, J., Reay, J., & Riby, L. M. (2015). Glucose and the wandering mind: Not paying attention or simply out of fuel? *Psychopharmacology*, *232*, 2903–2910.

Bishop, D. V. M., & Thompson, P. A. (2016). Problems in using *p*-curve analysis and text-mining to detect rate of *p*-hacking and evidential value. *PeerJ*, *4*, Article e1715. doi:10.7717/peerj.1715

Bruns, S. B., & Ioannidis, J. P. A. (2016). *p*-Curve and *p*-hacking in observational research. *PLoS ONE*, *11*, Article e0149144. doi:10.1371/journal.pone.0149144

Camerer, C. F., Dreber, A., Forsell, E., Ho, T. H., Huber, J., Johannesson, M., . . . Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, *351*, 1433–1436.

Carter, E. C., Kofler, L. M., Forster, D. E., & McCullough, M. E. (2015). A series of meta-analytic tests of the depletion effect: Self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General*, *144*, 796–815.

Carter, E. C., & McCullough, M. E. (2013). After a pair of self-control-intensive tasks, sucrose swishing improves subsequent working memory performance. *BMC Psychology*, *1*, Article 22. doi:10.1186/2050-7283-1-22

*Denson, T. F., von Hippel, W., Kemp, R. I., & Teo, L. S. (2010). Glucose consumption decreases impulsive aggression in response to provocation in aggressive individuals. *Journal of Experimental Social Psychology*, *46*, 1023–1028.

*DeWall, C. N., Baumeister, R. F., Gailliot, M. T., & Maner, J. K. (2008). Depletion makes the heart grow less helpful: Helping as a function of self-regulatory energy and genetic relatedness. *Personality and Social Psychology Bulletin*, *34*, 1653–1662.

*Dickinson, D. L., McElroy, T., & Stroh, N. (2014). Impact of glucose on Bayesian versus heuristic-based decision making. *Journal of Neuroscience, Psychology, and Economics*, *7*, 237–247.

*Dvorak, R. D., & Simons, J. S. (2009). Moderation of resource depletion in the self-control strength model: Differing effects of two modes of self-control. *Personality and Social Psychology Bulletin*, *35*, 572–583.

Errington, T. M., Iorns, E., Gunn, W., Tan, F. E., Lomax, J., & Nosek, B. A. (2014). An open investigation of the reproducibility of cancer biology research. *eLife*, *3*, Article e04333. doi:10.7554/eLife.04333

Francis, G. (2012). Too good to be true: Publication bias in two prominent studies from experimental psychology. *Psychonomic Bulletin & Review*, *19*, 151–156.

Freud, S. (1961). The ego and the id. In J. Strachey (Ed. & Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 19, pp. 1–66). London, England: Hogarth Press. (Original work published 1923)

*Gailliot, M. T. (2012). Improved self-control associated with using relatively large amounts of glucose: Learning self-control is metabolically expensive. *Psychology*, *3*, 987–990.

Gailliot, M. T., & Baumeister, R. F. (2007). The physiology of willpower: Linking blood glucose to self-control. *Personality and Social Psychology Review*, *11*, 303–327.

*Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D., . . . Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source: Willpower is more than a metaphor. *Journal of Personality and Social Psychology*, *92*, 325–336.

*Gailliot, M. T., Peruche, B. M., Plant, E. A., & Baumeister, R. F. (2009). Stereotypes and prejudice in the blood: Sucrose drinks reduce prejudice and stereotyping. *Journal of Experimental Social Psychology*, *45*, 288–290.

Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A., . . . Zwienenberg, M. (in press). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science*.

*Hagger, M. S., & Chatzisarantis, N. L. D. (2013). The sweet taste of success: The presence of glucose in the oral cavity moderates the depletion of self-control resources. *Personality and Social Psychology Bulletin*, *39*, 28–42.

Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. (2010). Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin*, *136*, 495–525.

*Howard, M. A., & Marczinski, C. A. (2010). Acute effects of a glucose energy drink on behavioral control. *Experimental and Clinical Psychopharmacology*, *18*, 553–561.

*Job, V., Walton, G. M., Bernecker, K., & Dweck, C. S. (2013). Beliefs about willpower determine the impact of glucose on self-control. *Proceedings of the National Academy of Sciences, USA*, *110*, 14837–14842.

*Kelly, C. L., Sünram-Lea, S. I., & Crawford, T. J. (2015). The role of motivation, glucose and self-control in the antisaccade task. *PLoS ONE*, *10*, Article e0122218. doi:10.1371/journal.pone.0122218

Kurzban, R. (2010). Does the brain consume additional glucose during self-control tasks? *Evolutionary Psychology*, *8*, 244–259.

Lakens, D. (2015). What *p*-hacking really looks like: A comment on Masicampo and LaLande (2012). *Quarterly Journal of Experimental Psychology*, *68*, 829–832.

*Lange, F., & Eggert, F. (2014). Sweet delusion. Glucose drinks fail to counteract ego depletion. *Appetite*, *75*, 54–63.

*Lange, F., Seer, C., Rapior, M., Rose, J., & Eggert, F. (2014). Turn it all you want: Still no effect of sugar consumption on ego depletion. *Journal of European Psychology Students*, *5*(3), 1–8. doi:10.5334/jeps.cc

*Masicampo, E. J., & Baumeister, R. F. (2008). Toward a physiology of dual-process reasoning and judgment: Lemonade, willpower, and expensive rule-based analysis. *Psychological Science*, *19*, 255–260.

*Molden, D. C., Hui, C. M., Scholer, A. A., Meier, B. P., Noreen, E. E., D'Agostino, P. R., & Martin, V. (2012). Motivational versus metabolic effects of carbohydrates on self-control. *Psychological Science*, *23*, 1137–1144.

*Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*, 943. doi:10.1126/science.aac4716

Orquin, J. L., & Kurzban, R. (2016). A meta-analysis of blood glucose effects on human decision making. *Psychological Bulletin*, *142*, 546–567.

Osman, M. (2014). *Future-minded: The psychology of agency and control*. Basingstoke, England: Palgrave Macmillan.

Owen, L., Scholey, A. B., Finnegan, Y., Hu, H., & Sünram-Lea, S. I. (2012). The effect of glucose dose and fasting interval on cognitive function: A double-blind, placebo-controlled, six-way crossover study. *Psychopharmacology*, *220*, 577–589.

*Sanders, M. A., Shirk, S. D., Burgin, C. J., & Martin, L. L. (2012). The gargle effect: Rinsing the mouth with glucose enhances self-control. *Psychological Science*, *23*, 1470–1472.

Schimmack, U. (2012). The ironic effect of significant results on the credibility of multiple-study articles. *Psychological Methods*, *17*, 551–566.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *11*, 1359–1366.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). *P*-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, *143*, 534–547.

*Wang, X. T., & Dvorak, R. D. (2010). Sweet future fluctuating blood glucose levels affect future discounting. *Psychological Science*, *21*, 183–188.