University of Dundee

The misattribution of emotions and the error-related negativity

Elkins-Brown, Nathaniel; Saunders, Blair; Inzlicht, Michael

Published in:
Cortex

DOI:
[10.1016/j.cortex.2018.08.017](https://doi.org/10.1016/j.cortex.2018.08.017)

Publication date:
2018

Licence:
CC BY-NC-ND

Document Version
Peer reviewed version

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

Elkins-Brown, N., Saunders, B., & Inzlicht, M. (2018). The misattribution of emotions and the error-related negativity: a registered report. *Cortex*, 109, 124-140. <https://doi.org/10.1016/j.cortex.2018.08.017>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

**The misattribution of emotions and the error-related negativity: a
Registered Report**

Nathaniel Elkins-Brown¹, Blair Saunders², & Michael Inzlicht^{1,3}

¹Department of Psychology, University of Toronto, Canada

²Psychology, School of Social Sciences, University of Dundee, Scotland

³Rotman School of Management, University of Toronto, Canada

Please send correspondence to:
Nathaniel Elkins-Brown
University of Toronto, Department of Psychology
1265 Military Trail
Toronto, Ontario M1C 1A4, Canada
Email: nat.brown@mail.utoronto.ca
Telephone: 647-710-4950

Abstract

A growing body of work in social and affective neuroscience suggests that emotion plays an instrumental role in error monitoring processes, rather than only a moderating one. High-powered replications of studies that support this idea, however, are lacking. Here, we attempted a preregistered replication of our own study that had provided evidence for the functional role of emotions in error monitoring: that a neural signal of error monitoring—the error-related negativity—is reduced when participants undergo a misattribution of arousal procedure (Inzlicht & Al-Khindi, 2012). Like a previous replication attempt (Rodilla, Beauducel, & Leue, 2016), our misattribution procedure failed to reduce the amplitude of the ERN. However, it also failed its manipulation check to reduce state anxiety, limiting the conclusions we can draw. Nonetheless, these findings are consistent with the view that our original study may have been a false positive. We discuss these findings in the context of the replication crisis in psychology and of work on the emotional properties of the ERN.

Keywords: ERN; emotion; misattribution; anxiety; replication

Introduction

Monitoring performance for errors is important for day-to-day functioning. Errors can impede or prevent progress toward desired goal states, and failures to detect or adjust for errors in performance can have consequences of physical, social, and economic import. Errors while driving vehicles or performing surgery, for example, can have fatal consequences. Although research over the past two decades has done much to advance our understanding of the cognitive and neural systems that support error monitoring, there remains a dearth of literature on the role of emotions and motivation in error detection and compensation. Here, we seek to replicate Inzlicht and Al-Khindi (2012), our own previous study that provided experimental evidence that emotions contribute to error monitoring: the finding that a neural signal called the error-related negativity (ERN) is reduced when participants misattribute their emotions to a placebo.

The ERN

In electroencephalography, the ERN is a negative deflection in the event-related potential (ERP) that occurs within 100 ms of error responses over frontocentral electrode sites, and is thought to be generated in the dorsal regions of the medial frontal cortex (Dehaene, Posner, & Tucker, 1994; Pourtois et al., 2010; Van Veen & Carter, 2002). While the functional significance of the ERN remains a topic of debate, a number of computational theories have proposed that the component represents conflict monitoring (Yeung, Botvinick, & Cohen, 2004), reinforcement learning (Holroyd & Coles, 2002), or error likelihood (Brown & Braver, 2005). Recent work has also described ERN amplitudes as

representing to the affective, motivational, or evaluative aspects of errors (Aarts, De Houwier, & Pourtois, 2013; Proudfit, Inzlicht, & Menin, 2013; Weinberg, Riesel, & Hajcak, 2012). Consistent with this perspective, ERN amplitude become larger when task performance is incentivized or punished (Legault & Inzlicht, 2013; Pailing & Segalowitz, 2004; Riesel, Weinberg, Endrass, Kathmann, & Hajcak, 2012), and when participants focus on affect during performance (Saunders, Rodrigo & Inzlicht, 2016). Conversely, ERN amplitude is reduced when negative emotions are attenuated through cognitive reappraisal (Hobson, Saunders, Al-Khindi, & Inzlicht, 2014) or through the consumption of alcohol (Ridderinkhof et al., 2002; Bartholow, Henry, Lust, Sauls, & Wood, 2012). ERN amplitude also increases as a function of trait anxiety (Hajcak, McDonald, & Simons, 2003; Meyer, Weinberg, Klein, & Hajcak, 2012; Pourtois et al., 2010) and trait negative affect (e.g., Luu, Collins, & Tucker, 2000; Luu, Flaisch, & Tucker, 2000), and either increases or decreases as a function of clinical status (e.g., Gehring, Himle, & Nisenson, 2000; Weinberg, Olvet, & Hajcak, 2010; Xiao et al., 2010). These findings have led researchers to propose that the ERN reflects—at least in part—the degree to which errors are motivationally significant or endogenously threatening (Aarts et al., 2013; Proudfit et al., 2013; Weinberg et al., 2012; Weinberg et al., 2016).

Misattribution and the ERN

In our original study (Inzlicht and Al-Khindi, 2012), we had investigated how ERN amplitude and task performance would be influenced by a misattribution of arousal paradigm. Since the 1960s, researchers have shown

1 that the source of emotions can be misattributed, and that misattribution can alter
2 the quality, magnitude, and labeling of emotional experiences (Reisenzein, 1983;
3 Schacter & Singer, 1962). Indeed, classic studies from social psychology have
4 shown that the intensity of emotional experiences is reduced when individuals
5 misattribute their emotions to a benign, external source (Ross, Rodin, &
6 Zimbardo, 1969; Zanna & Cooper, 1974). When patients with insomnia were
7 instructed to consume a sugar pill, for example, those who were told the pill
8 would make them feel anxious fell asleep faster than those who were told the pill
9 would make them feel relaxed (Storms & Nisbett, 1970). Ostensibly, patients in
10 the former group experienced less anxiety because they could attribute at least
11 some of their restlessness to the pill, while patients in the latter group could only
12 attribute their restlessness to their own thoughts and feelings.

13 Accordingly, the original study investigated whether ERN amplitude would
14 be reduced under the effects of misattribution. We reasoned that if the ERN
15 reflects an affective response to mistakes, its amplitude would be reduced when
16 participants had the opportunity to misattribute their emotions to something other
17 than their own task performance (i.e., the placebo). In an initial pilot study, we
18 found that state anxiety was increased after performing a task, and that this
19 increase was smaller for the misattribution group ($M = 1.97$, $SD = 0.58$)
20 compared to the control group ($M = 2.29$, $SD = 0.39$), $t(46) = 2.16$, $p = 0.036$, $d =$
21 0.64 . In the main study, we repeated the misattribution paradigm while
22 measuring the ERN. We found that ERN amplitudes were smaller in the
23 misattribution group ($M = -3.92 \mu V$, $SD = 2.18$) compared to the control group (M

1 = -6.52 μ V, $SD = 3.96$), $F(1, 34) = 6.12$, $p = 0.019$, $\eta^2_p = .15$. We also found that
2 error rates were negatively correlated with the magnitude of the Δ ERN (the ERN
3 minus the correct trial equivalent; the correct-related negativity, CRN) in the
4 control group ($r = .54$, $p = 0.02$) and not the misattribution group ($r = .19$, $p = .44$),
5 although this difference was not significant between groups. In both the pilot and
6 main studies, task performance and performance expectations were not
7 significantly different between groups.

8 With less ambiguity than previous studies, these findings provided
9 evidence of the emotional characteristics of the ERN. Without altering task
10 performance or performance expectations, the study's misattribution paradigm
11 was successful in reducing both state anxiety and ERN amplitudes. Further, a
12 significant relationship between ERN amplitudes and error rates was present in
13 the control group, but not in the misattribution group. Although the relationship
14 between the ERN amplitudes and behavioral measures is unclear (see Weinberg
15 et al., 2012), these findings implied that the ERN's association with task
16 performance may be dependent on whatever processes are altered by
17 misattribution; at least some of these processes are probably emotional in
18 nature. In short, we had found compelling evidence that the ERN has emotional
19 properties.

20 **The Present Study**

21 The results from Inzlicht and Al-Khindi (2012) should be replicated for
22 several conceptual, statistical, and practical reasons. Firstly, despite the original

1 paper having more than 80 citations to date¹, only one recent study constitutes a
2 replication attempt of the original study (Rodilla, Beauducel, & Leue, 2016).
3 Given the recent attention to replication attempts of many seminal findings in
4 psychology (e.g., Open Science Collaboration, 2015), it would be prudent to try
5 further replicating one of the first studies to clearly demonstrate the emotional
6 properties of the ERN. In the single replication attempt, Rodilla et al. (2016) failed
7 to find any differences in ERN amplitudes across misattribution and control
8 groups while using a between-subjects design with high power. They also did not
9 observe any differences across groups for self-reported anxiety; both
10 misattribution and control groups experienced an equal increase in anxiety from
11 before the task to after the task. This failure to replicate may have occurred
12 because the authors used only the information available in the original
13 publication, and did not contact the original authors for complete experimental
14 protocols and procedures. If there is a true difference in ERN amplitudes
15 between misattribution and control conditions—and the replication attempt did
16 not constitute a Type II error (i.e., a false negative)—it may be the case that the
17 divergent findings between these studies are the result of differences in their
18 respective methodologies. If this is the case, then a replication attempt with full
19 information should be made to elucidate what factors might be influencing this
20 divergence.

21 Our original experiment should also be replicated because of its low
22 statistical power. We made use of a between-subjects design with less than 20

¹ As assessed on Google Scholar, June 2018.

1 individuals in each group. This low sample size limited the study's ability to detect
2 any true small- or medium-size effect, while simultaneously being more
3 susceptible to experimenter degrees of freedom that could inflate the probability
4 of false positive results.

5 Lastly, new perspectives on the function of the ERN may make an even
6 stronger argument for the results reported in the original study than the reasoning
7 we originally provided for it. If the ERN at least partially signifies the motivational
8 significance of errors (Aarts, De Houwer, Pourtois, 2013; Proudfit, Inzlicht, &
9 Mennin, 2013) or trait defensive reactivity to errors (Weinberg et al., 2012), then
10 the misattribution paradigm may constitute an exceptionally pertinent
11 manipulation of how endogenously threatening participants find errors to be
12 during task performance. Under the conditions of misattribution, emotional
13 experiences that would typically be attributed to the self—like failures and
14 successes in task performance—are instead attributed to an exogenous, benign
15 placebo. Consequently, errors under misattribution should be perceived as less
16 self-caused and less threatening than typical errors. Accordingly, the efficacy of
17 the misattribution paradigm may depend on individual differences in body
18 consciousness (Brockner & Swap, 1983; Heatherton, Polivy, & Herman, 1989)
19 and self-perception (Duncan & Laird, 1980), where high self-awareness is related
20 to a typical effect of a misattribution, while low self-awareness is related to
21 nonsignificant or reversed effects of misattribution (i.e., a typical placebo effect).
22 Thus, to the extent that ERN amplitudes represents an affective or defensive
23 response to errors, it should be reduced when a misattribution paradigm

more negative) when participants perform under the perceived effects of the placebo, compared to when they perform while not under the perceived effects of the placebo.

2. State anxiety will rise after participants perform the task under the effects of the placebo, but this increase will be reduced for participants in the Misattribution condition compared to the Control condition.

In the original pilot study, participants in both the Misattribution and Control groups experienced greater anxiety after performing the task. However, this rise was smaller in the Misattribution group compared to the Control group. Accordingly, we expect that participants will have greater state anxiety after performing the task compared to before it, and this rise will not be as large for participants in the Misattribution group compared to the Control group. Further, participants will perform the task twice in a counterbalanced fashion, so the second time they perform the task they will be less likely to experience a change in anxiety from before the task to after it. Thus, we will also ask participants about their state emotions related *specifically* to task performance, with the same expectations as those in the original study: Participants in both Control and Misattribution groups will experience a rise in state anxiety *about task performance* after the task compared to before it, and this increase will be reduced for participants in the Misattribution group compared to the Control group.

3. There will be no direct effect of group or condition on behavioural performance (i.e., no-go error rates) or performance expectations.

1 **However, there may be an *indirect* effect of group on error rates through**
2 **the ERN/ Δ ERN, such that reduced ERN/ Δ ERN amplitudes in the**
3 **Misattribution group will predict higher error rates. In general, we anticipate**
4 **that ERN/ Δ ERN amplitudes will be negatively correlated with error rates,**
5 **but significantly less so for participants under the effects of the placebo in**
6 **the Misattribution group compared to performance under the effects of the**
7 **placebo in the Control group.**

8 In the original experiment, participants in both groups had no significant
9 differences in no-go error rates and performance expectations, despite significant
10 differences in state anxiety and ERN amplitudes. Δ ERN amplitudes were also
11 negatively correlated with error rates in the Control group and not the
12 Misattribution group, although difference between the groups for this relationship
13 was not significant. Accordingly, we hypothesize that no-go error rates and
14 performance expectations will not be significantly different across groups and
15 conditions, and that ERN/ Δ ERN amplitudes will be correlated with error rates in
16 the Control group and either not correlated or less correlated in the Misattribution
17 group. However, we also extend this further and hypothesize that there will be an
18 indirect effect of group on error rates through the ERN/ Δ ERN, such that reduced
19 ERN/ Δ ERN amplitudes in the Misattribution group will be related to higher error
20 rates.

21 **4. The relationships between variables in the present replication—namely,**
22 **between (1) group and the ERN/ Δ ERN, between (2) group and state anxiety,**

1 **and between (3) group and error rates via the ERN/ Δ ERN—may be**
2 **moderated by trait anxiety.**

3 Unpublished work with a misattribution paradigm has suggested that the
4 amplifying effect of trait anxiety on ERN amplitudes (e.g., Moser, Moran,
5 Schroder, Donnellan, & Yeung, 2013) is dependent on the correct attribution of
6 state anxiety, such that misattribution eliminates the relationship between trait
7 anxiety and Δ ERN amplitudes (Rodilla, Beauducel, & Leue, 2015). These
8 findings highlight the importance of both state and trait anxiety in varying ERN
9 and Δ ERN amplitudes. Because both ERN/ Δ ERN amplitudes and state anxiety
10 may be reduced in individuals who are low in trait anxiety (and vice-versa),
11 relationships between variables in the present experiment may be weaker or fail
12 to reach statistical significance when not accounting for trait anxiety. Thus, we
13 hypothesize trait anxiety may enhance the relationships between group and the
14 ERN/ Δ ERN, between group and state anxiety, and between group and error
15 rates via the ERN/ Δ ERN.

Methods

Power and Participants

Prior to collecting data, we conducted an *a priori* power analyses in G*Power (Erdfelder, Faul, & Buchner, 1996) to determine the sample size for our study. Although the effect size of the original study was medium-to-large ($\eta^2_p=.17$), the original study did not have the power to accurately detect a small or medium-sized effect, and so the original effect size may be inflated (Button et al., 2013). Thus, we chose a small-to-medium effect size estimate (Cohen's $f = 0.15$) to accurately detect a smaller but still meaningful difference across conditions. This analysis determined that 74 participants would be needed to detect an effect of this size with 90% power using the aforementioned repeated-measures ANOVA, probing for between-within interactions (see { HYPERLINK "https://osf.io/su58j/files" } for details of the power analysis). Because of our subsequent use of primarily multilevel models without random slopes in the present study, this estimated sample size provided either accurate or conservative power for all analyses (Quené & van der Bergh, 2004).

Thus, prior to collecting data, we aimed to recruit 84 participants in total, expecting that some participants (10-15%) would be excluded entirely because of computer malfunction, too few usable error trials, suspicion about the effects of the placebo, or task disengagement. If after exclusions there were fewer than 19 participants in any of the four counterbalanced groups (see below), recruitment would continue until there were at least 19 participants in each group.

1 In the beginning of recruitment, the first three participants were “pilots”
2 used only to ensure our data collection procedures and preprocessing pipeline
3 were free from errors. The first two of these pilot participants had computer errors
4 that prevented complete data collection, which were fixed for the third participant.
5 The third participant had acceptable data, so all participants *after* the third
6 participant were considered valid to be included in the study. We did not
7 preregister the use of pilot participants to ensure error-free data; however, data
8 from these pilot participants were not analyzed, so they did not influence our
9 hypotheses, procedures, or analyses.

10 At the end of data collection, we recruited 82 participants in total following
11 the third pilot participant. Based on preregistered exclusion criteria (see Data
12 Removal Summary below), we completely excluded the data of three participants
13 and partially excluding the data of two participants. This produced a final sample
14 size of 79 participants for multilevel models and 77 participants for all univariate
15 models.

16 All participants were recruited at the University of Toronto Scarborough
17 and received course credit or \$10 payment as compensation for their
18 involvement in the study. Advertisement for recruitment specified that participants
19 needed to have normal or corrected-to-normal vision and must have been
20 speaking English fluently for at least 10 years. All participants provided informed
21 consent before participating.

22 **Procedure**

Participants were told that the purpose of the experiment was to investigate the effects of an herbal supplement (called *panax senticosus*) on brain responses during a self-control task. However, they were only ever administered a placebo pill that contained sucrose. A specific protocol was used to increase the plausibility of this herbal supplement cover story to participants, the details of which are outlined below and described in full detail on OSF (see Cover Story Procedures at { [HYPERLINK "https://osf.io/su58j/files"](https://osf.io/su58j/files) }). Participants were randomly assigned to one of two supplement groups (Misattribution or Control), and one of two counterbalanced ingestion groups (Supplement First or Supplement Second). The supplement group was used as a between-subjects factor in later statistical analyses, while participants in the two ingestion groups were combined in all preregistered analyses. All participants contributed two separate sets of data (the within-subjects “placebo” condition) over the course of the experiment: one set in which they believed they were under the effects of the supplement (either a “Misattribution” or “Control” supplement, depending upon their randomly assigned group), and one set in which they believed they were not. In using this mixed design, we sought to increase statistical power from the original study by having participants act as their own “control” group, i.e., performing the task both with and without the placebo pill.

After giving informed consent, participants were provided information about the *panax senticosus* supplement and its ostensible role in the experiment on the computer screen. This information is provided verbatim below, and was

only minimally altered to address the differences in protocol between the original study and the present one. These differences are denoted in square brackets:

As mentioned earlier, the purpose of this study is to investigate the effects of an herbal supplement, Panax senticosus, on cognitive performance. Panax senticosus grows along the west coast of North America (from British Columbia to California). Panax senticosus contains compounds that strengthen bones (Rhodes and Schwartz, 2007), and is increasingly being used in the elderly population (Black et al., 2009). It is commonly available at health food stores and is completely safe.

Although Panax senticosus has been marketed widely over the past 5 years, few studies have investigated its effects on cognitive performance. The present study was designed to address this question. In a few moments, the experimenter will give you [two pills] containing Panax senticosus [and 150 ml of filtered water in a cup]. We ask that you please [take the pill and] consume the entire [cup of water].

Misattribution Group. Following these two paragraphs, participants in the Misattribution group—but not the Control group—were then told that the herbal supplement could have minor side effects:

Please note that there are certain side effects associated with Panax senticosus. Studies have documented that Panax senticosus can activate the body's sympathetic nervous system, which governs arousal (Selariu et al., 2006). Past participants have reported feeling a number of side effects, including tenseness, anxiety, increased heart rate and a racing mind. From our experience, the effects of Panax senticosus begin to occur approximately [20] minutes after ingestion and last approximately [35-40] minutes

1 *[thereafter]. Please do not be alarmed if you feel [any] side effects [—] they are*
2 *temporary and will disappear [before] the conclusion of today's study session.*

3
4 **Control Group.** Conversely, participants in the Control group—and not in
5 the Misattribution group—were told that the herbal supplement is not known to
6 have any side effects:

7
8 *Please note that there are no physical side effects associated with Panax*
9 *senticosus. [From our experience, the effects of Panax senticosus on the brain*
10 *begin to occur approximately 20 minutes after ingestion and last approximately*
11 *35-40 minutes thereafter.]*

12
13 **Supplement First Group.** After being informed about the supplements on
14 the computer, participants in the Supplement First group were verbally reminded
15 of whether the supplement would have side effects, and then asked to consume
16 a pill containing the supplement. After consuming the pill, participants were again
17 told that they had to wait 20 minutes for it to take effect, during which time they
18 were prepared for the EEG recording. Two minutes before the waiting period
19 ended, participants were asked to provide a small saliva sample to verify that the
20 supplement had “begun working”. This saliva sample was mixed into a vial that
21 appeared to participants to contain water and starch; however, it actually
22 contained water and iodine, and a sleight-of-hand procedure was used to switch
23 the saliva sample with one dipped in starch. Participants saw this mixture turned
24 blue, and were then told that this indicated that the supplement had begun
25 working.

Participants were then asked to complete 3 sets of questions: a 6-item version of the State Anxiety Inventory (Marteau & Becker, 1992; Spielberger, 1983), 2 questions that specifically assessed state anxiety about task performance, and a modified version of the Intrinsic Motivation Inventory (McAuley, Duncan, & Tammen, 1989). The 4-point State Anxiety Inventory is a reliable measure of state anxiety ($\alpha = .82$) that includes items about feeling nervous or relaxed (e.g. “I am worried”; “I feel tense”). These items were presented with two additional items that assessed state emotions about expected task performance (i.e., “I feel good about the upcoming task”; “I am anxious about how I will perform on the task”). The modified version of the Intrinsic Motivation Inventory (IMI; McAuley, Duncan, & Tammen, 1989) contained only a set of items from the Perceived Competence ($\alpha = .80$) and Effort-Importance subscales. This set of questions included four adjusted items from the former subscale that assessed expectations about the upcoming task (e.g., “I think I will be pretty good at this task”; “I think I will be satisfied with my performance at the upcoming task”), and which was used to assess performance expectations. The remaining items from the Effort-Importance subscale were present only to reduce demand characteristics. Following the completion of these questions and at the 20-minute mark, participants were instructed to begin the Go/No-Go task.

The protocols of this task did not differ from that of the original study. In this task, participants responded using the M key on a DirectIN PCB keyboard (Empirisoft, New York, NY) in response to two stimuli on screen, the letters *M* and *W*. The presentation probability for each stimulus presented randomly was

1 asymmetric, giving a correspondingly asymmetric response ratio of 85:15 for M
2 and W, respectively. Participants were required to press the M key when they
3 saw the “go” stimulus (i.e., the letter M) and to refrain from pressing the M key
4 when they saw the “no-go” stimulus (i.e., the letter W). Each trial began with a
5 fixation cross (“+”) presented randomly for 300-700 ms, which was followed by
6 either a go or no-go stimulus for 100 ms. The maximum time allowed for a
7 response was 500 ms, while the minimum time allowed for a response was 100
8 ms. Participants first completed a practice block of 20 trials. These trials included
9 response feedback for errors of omission (i.e., “respond faster!”) and errors of
10 commission (i.e., “wrong response!”) to ensure participants understood the task.
11 Practice trials were followed by five experimental blocks without feedback, each
12 consisting of 85 go trials and 15 no-go trials. With this task, we measured the
13 average reaction time for correct (i.e., key presses on “go” trials) and incorrect
14 trials (i.e., key presses on “no-go” trials), the number of errors of commission
15 (pressing M during a no-go trial), and the number of errors of omission (not
16 pressing M during a go trial).

17 After completing the first Go/No-Go task, participants then immediately
18 completed another State Anxiety Inventory with the two questions about
19 emotions and task performance, with the latter modified to assess current
20 emotions about previous task performance (i.e., “I feel good about how I
21 performed”; “I am anxious about how I performed on the task”). Participants then
22 then waited 15 minutes for the supplement to ostensibly “wear off”. During this
23 waiting period, participants read some provided magazines. Two minutes before

the end of this second waiting period, participants again provided a saliva sample to verify that the supplement had worn off. This sample was mixed into a vial that appeared to participants to contain water and starch. Unlike before, this mixture did not turn blue, and participants were told that this indicated that the supplement had worn off. Participants were then asked to complete a third State Anxiety Inventory with the two additional questions about emotions and task performance, and a second modified IMI Scale. Following these scales, participants were then instructed to begin the second Go/No-Go task, which contained the same number of blocks and trials as the first Go/No-Go task.

After the completion of the second Go/No-Go task, participants were immediately asked to complete a fourth State Anxiety Inventory with two additional questions about emotions and task performance. This scale was followed by a BIS/BAS scale (Carver & White, 1994)—our measure of trait anxiety—and a funneled debriefing procedure (Bargh & Chartrand, 2000) to ensure that participants believed the cover story. In this procedure, participants were asked to freely respond to a series of questions that gauge their beliefs and knowledge about the experiment, its procedures, and their understanding of its intentions. The questions began by asking about the experiment very broadly, and then became increasingly specific.

(1) *What do you believe the experiment was about?*

(2) *Was there anything you believed the experimenters did not tell you about the experiment?*

(3) *What effects did you think the supplement had on you?*

1 (4) *At any point in this experiment did you think the experimenters may*
2 *have deceived you, and that the supplement may have been fake? If so, at what*
3 *point during the experiment did you think this?*

4
5 Participants were then asked about their demographic characteristics, and
6 to verify that they had been speaking English fluently for 10 years. Finally,
7 participants were fully debriefed, compensated, and thanked for their
8 participation in the study.

9 **Supplement Second Group.** Participants randomized into the
10 Supplement Second group engaged in identical procedures to those in the
11 Supplement First group, but ingested the supplement after the first Go/No-Go
12 task instead of before it. Participants began with preparations for EEG recording,
13 and completed a State Anxiety Inventory and modified IMI before the first Go/No-
14 Go task. After the first Go/No-Go task, they immediately completed a second
15 State Anxiety Inventory and IMI. They were then informed about the supplement
16 and its side effects, ingested the supplement, and waited 20 minutes for it to take
17 effect. They then completed a third State Anxiety Inventory and IMI before the
18 second Go/No-Go task, and a fourth State Anxiety Inventory and IMI after it.
19 Procedures thereafter were the same as the Supplement First group.

20 **Blinding Procedures and Behavioral Exclusion Criteria.** For research
21 assistants to verbally remind participants of whether or not the placebo had side
22 effects, they were forced to be aware of the group (Misattribution group or
23 Control group) to which a participant had been assigned. However, they were not
24 made aware of any of the dependent variables or hypotheses of the experiment

1 until its completion. They were asked not to read the debriefing form over the
2 course of data collection. In general, because most instructions and participant
3 task performance are computer-administered, experimenter effects on participant
4 behavior were small.

5 Only participants who explicitly reported having believed the placebo pill
6 was fake had their “supplement” data excluded, and only the portion of their data
7 during the time in which they believed the placebo was fake. Participants who
8 expressed suspicion about certain aspects of the procedures—but not
9 specifically suspicion about the placebo—did not have their data excluded.
10 Participants who only expressed some level of suspicion about the placebo did
11 not have their data excluded. Participants in the Misattribution group who
12 reported believing that the placebo had no effect on them also did not have their
13 data excluded; unregistered, exploratory analyses excluding these participants
14 did not change any of our results.

15 Participants had their data excluded if they had an overall commission or
16 omission error rate greater than 45% for all trials, collapsed over placebo
17 condition. Further, even if a participant verbally reported understanding the
18 directions and had been speaking English for more than 10 years, they would
19 have had their data excluded if there was substantial evidence that they were no
20 longer following directions or had disengaged from the task. Evidence that
21 participants had disengaged from the task included pressing the response key
22 without looking at the screen, which experimenters checked for by observing
23 participants surreptitiously on each block of trials.

1 **Neurophysiological Recordings and Artifact Removal.** Continuous
2 EEG activity was measured over the cortical midline (Fz, FCz, Cz, CPz, Pz, Oz)
3 using 6 Ag/AgCL electrodes embedded in a stretch Lycra cap (Electro-Cap
4 International, Eaton, OH). Vertical electrooculography (VEOG) was recorded
5 using a supra-to-suborbital bipolar montage placed around the right eye. EEG
6 activity was amplified using an ANT Refa8 TMSi (Advanced Neuro Technology,
7 Enschede, The Netherlands) device. The continuous EEG signal was grounded
8 to the forehead and referenced online to the average of all electrodes. Offline,
9 the EEG signal was re-referenced to the average of two bilaterally placed ear
10 lobe electrodes. Impedances were kept below 5 k Ω for all recordings. Recordings
11 were digitized at 512 Hz using Advanced Source Analysis 4.9.2 software.

12 All filtering, artifact corrections, and segmentation of EEG data were then
13 implemented in BrainVision Analyzer 2.0 (Brain Products GmbH, Gilching,
14 Germany). EEG data were digitally filtered between 0.1 and 20 Hz (24 dB/oct,
15 zero phase-shift Butterworth filter), and corrected for blink artifacts using an
16 independent components analysis procedure (Makeig, Bell, Jung, & Sejnowski,
17 1996). Automatic procedures were then used to reject EEG artifacts from
18 individual channels according to the following criteria: voltage steps of more than
19 25 μ V between sample points, a voltage difference of 150 μ V within 150-ms
20 intervals, voltages above 100 μ V and below -100 μ V, and a maximum voltage
21 difference of less than 0.50 μ V within 100-m intervals. Participants with more
22 than 35% of their EEG epochs rejected this way within placebo condition or trial
23 type—regardless of whether they have enough artifact free trials remaining to

1 reach a dependability threshold (described below)—had their data excluded from
2 analyses involving ERP data.

3 Participants had averaged cells of their data excluded if they did not have
4 at least 6 artifact-free trials within them for an internally reliable signal. Previous
5 research suggests that ERN amplitudes becomes moderately reliable ($\alpha > .50$) in
6 6 trials in undergraduates (Meyer, Riesel, & Proudfit, 2013; Olvet & Hajcak,
7 2009).

8 As reliability is dependent on each specific sample and condition of any
9 experiment, dependability estimates (a generalizability theory [G-theory]
10 analogue of reliability) were also calculated for each supplement group, placebo
11 condition, and trial type using the ERP Reliability Analysis Toolbox (Clayson &
12 Miller, 2017a; 2017b). This toolbox calculates ERP reliability based on algorithms
13 from generalizability theory (see Baldwin, Larson, & Clayson, 2015 for review)
14 and used CmdStan v2.17.1 (Stan Development Team, 2016) to implement the
15 analyses in Stan (Carpenter et al., in press). Overall dependability estimates and
16 their 95% credible intervals (CIs) for cell averages used the mean number of
17 trials retained for each supplement group, placebo condition, and trial type.

18 Participants whose data would have been excluded only because of too
19 few trials remaining after artifact-removal had their missing data estimated using
20 multilevel models (see Analysis Procedures and below). Conversely, participants
21 who did not have enough artifact-free trials because of naturally low error rates
22 would not have had their data estimated this way, and would have been removed

1 from analyses entirely. However, no participants were excluded either for having
2 too few trials after artifact-removal or for having naturally low error rates.

3 Epochs were defined as between 200 ms before and 800 ms after the
4 response, and baseline corrected using a -150 to -50 ms pre-response window.
5 For statistical analyses, the Δ ERN was operationalized using a collapsed
6 localizer method (cf., Luck & Gaspelin, 2017). In this method, we collapsed data
7 across all groups and conditions to determine the time window and scalp
8 distribution that best characterized the Δ ERN. The mean amplitude measure
9 taken from the chosen electrode was then used to analyze the non-collapsed
10 data for the all hypothesis tests. The same time window and electrode site
11 determined for the Δ ERN was also used to operationalize the ERN and its
12 correct-trial counterpart (the correct-response negativity; CRN). Using the
13 collapsed localizer method, we operationalized our ERPs with a time-window of
14 15 ms pre-response to 115 ms post-response over the FCz electrode.

15 **Data Removal Summary.** Participant exclusions and recruitment, artifact
16 removal, and data organization for statistical tests occurred in the following
17 ordered steps:

- 18 1. Participants or portions of participant data were excluded based on
19 belief that the placebo was fake, too high error rates, and evidence for
20 task disengagement. If this procedure had reduced the number of
21 participants in a counterbalanced group below 19, recruitment and
22 exclusion would have continued until there were at least 19
23 participants in each group. Using this procedure, one participant was

1 removed for belief that the placebo was fake, and two participants
2 were removed for too high omission error rates.

3 2. Individual trial epochs for remaining participants were removed based
4 on artifact rejection criteria.

5 3. Portions of participant ERP data were excluded for having more than
6 35% of all EEG epochs removed from that placebo condition and trial
7 type. Using this procedure, two participants had half of their ERP data
8 removed.

9 4. Portions of participant ERP data were excluded for having less than 6
10 artifact-free trials within a cell average. No data were removed using
11 this procedure.

12 5. For multilevel models, participants with missing data for a given cell
13 were excluded entirely from analyses that require full data, unless data
14 was missing exclusively because of artifact rejection or computer
15 errors (Step 3). No participants were excluded from analyses this way.
16 Participants with missing cell average trial numbers that would have
17 been 6 or greater (Step 4) had they not been rejected due to artifacts
18 (Step 3) were included in all multilevel models.

19 **Missing Data.** Because of a computer error, all state self-report data for
20 the first 32 participants had to be recovered manually from recorded key presses
21 in our physiological data. Because we did not always begin recording
22 physiological data immediately when participants started responding to self-
23 reports, 16 participants had missing data for at least one of four measurement

times of three possible variables: state anxiety, task-specific anxiety, or performance expectations. Of all measurement points for all three of these variables, only 5.2% of data in total was missing. Because no participants had entirely missing data for all measurement points of any variable, all missing data were capable of estimation using multilevel models. Because our missing data qualified as Missing Completely At Random (Little & Rubin, 2002), these models provided unbiased estimates of missing data.

Analysis Procedures

To account for heteroskedastic, non-spherical, and excluded data (Quené & van der Bergh, 2004), multilevel models were calculated in SPSS (v23) for almost all analyses. Because tests of the relationship between ERP amplitudes and error rates did not contain any within-subjects factors, we decided before analyzing data that they would be calculated using univariate analyses of variance (ANOVAs). All analyses were conducted in SPSS using the MIXED or UNIANOVA functions, and all main and interaction effects were evaluated using a Type III sums of squares approach.

For the multilevel models, we used a restriction maximum likelihood method for fitting, and an unstructured covariance matrix and Satterthwaite method to estimate random intercepts for each participant for all fixed effects. For ANOVAs, we decided before analyzing data that if residuals were found to be non-Normal based on significant results from Shapiro-Wilk or Kolmogorov-Smirnov tests, we would log-transform or natural-log transform dependent variables (i.e., commission error rates) until these tests were nonsignificant. The

1 residuals for all tested models were found to be sufficiently Normal (all $ps > .20$),
2 so error rates were not subsequently transformed.

3 For statistical calculations, effect sizes are denoted using either
4 semipartial R^2 (R^2_{β} ; Edwards, Muller, Wolfinger, Qaqish, & Schabenberger, 2008)
5 or partial η^2 (η_p^2 ; Cohen, 1973). To reduce the incidence of Type I error from
6 multiple comparisons, tests for different dependent variables were corrected
7 using a false discovery rate procedure (Benjamini & Hochberg, 1995; Benjamini
8 & Yekutieli, 2001).

9 **Statistical Tests & Results**

10 Statistical tests divided by category are described below, followed by their
11 results. Interaction terms are denoted using an asterisk (*) between factors. Our
12 models assessed the effect of group (misattribution vs. control), placebo
13 condition (on vs. off the placebo), trial type (error vs. correct trials), and time
14 (before vs. after the task) on our behavioral, physiological, and self-report
15 measures. In our multilevel models, these categorical variables were effect
16 coded such that data from the misattribution group were = 1, while data from the
17 control group were = -1; data under the effects of a placebo were = 1, while data
18 not under the effects of a placebo were = -1; data from error trials were = 1, while
19 data from correct trials were = -1, and data from before the task were = -1, while
20 data from after the task were = 1.

21 For all tests, only interactions with p -values below 0.05 were followed up
22 by simple effect or simple slope analyses, and only significant p -values of interest
23 were corrected using a false discovery rate procedure for all tests within the

1 same dependent variable (so if a main effect is significant but is not of direct
2 interest to our hypotheses, its p -value would not be included in a correction
3 procedure). Please note that all models conducted with trait anxiety and all post-
4 hoc analyses were exploratory, and may be underpowered.

5 **The CRN, ERN, and Δ ERN**

6 One multilevel model was calculated to assess CRN and ERN amplitude
7 across groups, placebo condition, and trial type, and another calculated to
8 assess Δ ERN amplitude across groups and condition. For the first model, the
9 between-subjects factor of group (Misattribution and Control), within-subjects
10 factor of placebo condition (Placebo and Non-Placebo), within-subjects factor of
11 trial type (Correct and Error), and three-way group*condition*trial type interaction
12 term were used to predict ERN amplitudes. These predictors were also
13 accompanied by all other possible two-way interactions between factors that
14 were not of central interest to our hypotheses. For the second model, the same
15 predictors without a trial type factor—and with a group*condition interaction term
16 instead of the three-way interaction term—were used to predict Δ ERN amplitude.

17 For the ERN, we hypothesized that there would be a significant three-way
18 interaction between group, placebo condition, and trial type. A simple effects
19 analysis of this interaction would reveal that ERP amplitudes would be smaller
20 (i.e., less negative) in the Misattribution group than in the Control group,
21 specifically for Placebo data and for error trials. For the Δ ERN, we hypothesized
22 that there would be a significant two-way interaction between group and
23 condition, and that a simple effects analysis of this interaction would reveal that

- 1 ERP amplitudes would be smaller in the Misattribution group than in the Control
- 2 group, specifically for Placebo data.

Figure 1. Error- and correct-related ERP amplitudes over the frontocentral electrode (FCz) across placebo condition and supplement group. The ERN, CRN, and Δ ERN were operationalized in all statistical tests as the mean activity between 15 and 115 ms.

Results. Overall dependability estimates and trial characteristics of the

CRN and ERN are depicted in Tables 1 and 2.

Indicating the presence of a robust ERN, error trials had more negative amplitudes ($M = -4.69$, $SE = 0.49$) compared to correct trials ($M = 5.53$, $SE = 0.49$), $b = 11.81$, $SE = 0.96$, $F(1,227.808) = 454.053$, $p < .001$, $R^2_{\beta} = 0.67$ (see Figure 1). For both the ERN and Δ ERN, there were no significant main effects of group, $F_s < 1.986$, $p_s > .163$, nor of condition, $F_s < 1.480$, $p_s > .228$. Contrary to our predictions, we did not observe a significant three-way interaction between trial type, group, and condition for the ERN, $F = 0.341$, $p = .560$, nor a significant

two-way interaction between group and condition for the Δ ERN, $F = 1.230$, $p = .271$. These findings suggest that ERP amplitudes were not influenced by the type of placebo participants consumed, nor by whether participants were on or off the placebo.

Exploratory Analyses. The null findings for ERP amplitude may have been confounded by order, such that participants' emotions could have been altered when they performed the Go/No-Go task for the second time. To account for these effects, we conducted the same multilevel analysis on ERN amplitudes as specified above, but only included Placebo data for participants who consumed the placebo the first time they performed the task. For the Δ ERN, we similarly conducted a one-way ANOVA for the effect of group on ERP amplitudes. These two analyses can be considered conceptually equivalent to the between-subjects tests conducted by Inzlicht & Al-Khindi (2012), with $n = 19$ for both Misattribution and Control groups. It should be noted that these analyses are underpowered, and so results should be interpreted with caution.

Like our standard analyses, we found a significant effect of trial type on ERP amplitudes, $F(1,36) = 116.198$, $p < .001$, $R^2_\beta = 0.76$, but no significant main effects of group or significant interactions between trial type and group for the Δ ERN or ERN, respectively, $F_s < 0.734$, $p_s > .397$. These findings suggest that even after eliminating potential order effects, ERP amplitudes were not influenced by the type of placebo participants consumed.

Table 1

Estimates of dependability and 95% credible intervals of error- and correct-related ERP amplitudes across groups and placebo condition

	Control	Misattribution
--	---------	----------------

MISATTRIBUTION REPLICATION

	Off Placebo			On Placebo			Off Placebo			On Placebo		
	Est	LB	UB	Est	LB	UB	Est	LB	UB	Est	LB	UB
<i>ERN</i>												
OD	0.85	0.77	0.91	0.90	0.85	0.94	0.88	0.82	0.93	0.92	0.88	0.95
BSSD	4.95	3.69	6.54	6.35	4.86	8.16	5.52	4.31	7.12	7.06	5.41	9.22
WSSD	11.82	11.35	12.32	11.93	11.46	12.40	11.37	10.93	11.82	11.89	11.45	12.36
ICC	0.15	0.09	0.24	0.22	0.14	0.32	0.19	0.12	0.28	0.26	0.17	0.38
<i>CRN</i>												
OD	0.98	0.98	0.99	0.98	0.97	0.99	0.99	0.98	0.99	0.99	0.99	1.00
BSSD	4.28	3.39	5.48	3.79	3.05	4.78	4.96	3.97	6.23	6.56	5.23	8.30
WSSD	10.27	10.16	10.39	10.50	10.38	10.61	10.32	10.20	10.43	10.77	10.65	10.89
ICC	0.15	0.10	0.22	0.12	0.08	0.17	0.19	0.13	0.27	0.27	0.19	0.37

Note. ERN = error-related negativity; CRN = correct-related negativity; Est = point estimate; LB = lower-bound; UB = upper-bound; OD = overall dependability; BSSD = between-subjects standard deviation; WSSD = within-subjects standard deviation; ICC = Intra-class correlation.

Table 2

Trial characteristics of ERP amplitude data submitted to dependability analyses across groups and placebo condition.

	Control				Misattribution			
	Off Placebo		On Placebo		Off Placebo		On Placebo	
	ERN	CRN	ERN	CRN	ERN	CRN	ERN	CRN
Mean # Trials	34.47	388.00	33.33	386.74	32.00	387.28	34.10	385.98
Min # Trials	19	316	9	272	11	129	11	37
Max # Trials	58	423	60	425	65	424	63	422
Trial Cutoff	14	14	9	13	11	11	7	9
N Included	38	39	39	39	39	40	39	40
N Excluded	1	0	0	0	1	0	0	0

Note. ERN = error-related negativity; CRN = correct-related negativity. Trial Cutoff refers to the minimum number of trials necessary to reach dependability of 0.70; results of present study remained unchanged when participants with data beneath the minimum cutoff were excluded.

State and Task-Related Anxiety

Two multilevel models were calculated to assess state and task-related anxiety across groups and time, looking only at data within the Placebo condition. For both models, state and task-related anxiety were predicted by group

(Misattribution and Control), time (Pre-task and Post-task), and a group*time interaction term.

We hypothesized that there would be significant interactions between group and time for both measures of anxiety. Simple effects analyses of these interactions would reveal that state and task-related anxiety would rise from pre-task to post-task more for the Control group than for the Misattribution group.

Figure 2a & 2b. State and task-related anxiety across time and group while participants were on the placebo. Error bars denote 95% between-subjects confidence intervals. Scale responses ranged from 1 to 4.

Results. When participants were on the placebo, state anxiety increased significantly from before the task ($M = 1.80$, $SE = 0.07$) to after the task ($M = 2.07$, $SE = 0.06$), $b = -0.27$, $SE = 0.09$, $F(1,66.736) = 17.947$, $p < .001$, $R^2_\beta = .21$, indicating that our task induced a mild increase in state anxiety (see Figure 2a). This increase was also true of task-related anxiety from before the task ($M = 1.99$, $SE = 0.08$) to after the task ($M = 2.51$, $SE = 0.07$), $F(1,71.543) = 32.355$, $p < .001$, $R^2_\beta = .31$ (see Figure 2b). These findings replicate those of our original study. There were no significant effects of group on state or task-related anxiety, $F_s < 0.781$, $ps > .380$.

Contrary to our main predictions, there were no significant two-way interactions between group and time, $F_s < .510$, $p_s > .478$. These findings indicate a crucial manipulation failure: the misattribution placebo did not influence the increase in state anxiety from pre- to post-task compared to the control placebo. Because our manipulation did not influence affect, our data cannot test the theory that the ERN has emotional properties.

Exploratory Analyses. Because state anxiety was low in general for the present study, misattribution may have only occurred in participants who experienced a larger rise in state or task-related anxiety when not on the placebo. We performed two additional multilevel models for the Δ ERN to see whether the interaction between group and condition was moderated by the difference in state and task-related anxiety from before to after the task. In these models, Δ ERN amplitudes were predicted by group, placebo condition, the difference in anxiety from before the task to after the task while *off* the placebo, and all two- and three-way interactions for these predictors. In both models, no predictors were significantly related to the Δ ERN, all $F_s < 2.553$, all $p_s > .117$. These findings show that changes in anxiety when off the placebo were altogether unrelated to the effects of group and condition on ERP amplitudes.

Behavior and Performance Expectations

Five multilevel models were calculated to assess the similarity in behavioral performance and performance expectations across groups and placebo condition. For all five models, each of the following dependent variables were predicted by group and placebo condition: commission error rates, omission

1 error rates, correct-trial reaction times, error-trial reaction times, and self-reported
2 performance expectations. We hypothesized that there will be no significant
3 interaction between group and placebo condition for all five tests, so that we
4 would fail to reject the null hypothesis that there was a significant difference in
5 behavioral performance and performance expectations across conditions and
6 within groups.

7 **Results.** We found no significant effects of group, condition, or their
8 interaction for any of our five measures, all $F_s < 3.908$, all $p_s > .052$ (see Table
9 2). Because our first test on performance expectations combined data from
10 before and after the task, we conducted a second test of performance
11 expectations to distinguish these sets of data. In this model, performance
12 expectations were predicted by time, group, condition, their three-way interaction,
13 and all two-way interactions not of interest to our hypotheses. We found a
14 significant effect of time, $b = 0.46$, $SE = 0.10$, $F(1, 76) = 119.564$, $p < .001$, $R^2_\beta =$
15 $.61$, such that participants rated their performance expectations before the task
16 more positively ($M = 2.67$, $SE = 0.06$) than after the task ($M = 2.12$, $SE = 0.06$).
17 There were no significant effects of group, condition, or the three-way interaction
18 between group, condition, and time on performance expectations, all $F_s < 1.335$,
19 all $p_s > 0.249$. These findings show that regardless of group or placebo condition,
20 participants' performance expectations dropped over time, reflecting a tendency
21 among participants to overestimate their competence before performing the task.

22 In line with our predictions, these findings show that our misattribution
23 paradigm did not significantly influence behavior or performance expectations

from before to after the task, either as a function of the type of placebo consumed or as a function of being on or off the placebo.

Table 3

Means and 95% between-subjects confidence intervals of behavior, performance expectations, and ERP amplitudes across groups and condition.

	Control (n = 39)						Misattribution (n = 40)					
	Off Placebo			On Placebo			Off Placebo			On Placebo		
	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB
Om ER (%)	5.48	3.60	7.35	5.20	3.46	6.94	5.05	2.94	7.16	4.19	2.79	5.60
Com ER (%)	46.67	42.15	51.28	45.73	40.72	50.74	42.68	37.05	48.32	46.80	41.07	52.53
RT Correct	292	285	299	293	283	302	294	282	306	289	276	302
RT Error	250	244	256	246	239	254	247	240	255	246	238	354
Pre-task PE	2.71	2.51	2.91	2.73	2.55	2.90	2.60	2.38	2.81	2.64	2.39	2.89
Post-task PE	2.05	1.83	2.28	2.13	1.96	2.29	2.08	1.85	2.30	2.18	1.97	2.38
ERN	-4.24	-5.92	-2.57	-3.94	-6.07	-1.80	-4.78	-6.58	-2.98	-5.68	-7.88	-3.48
Δ ERN	-9.62	-11.50	-7.74	-9.67	-11.89	-7.44	-11.09	-13.06	-9.13	-12.21	-14.43	-10.00

Note. LB = lower-bound; UB = upper-bound; Om ER = omission error rates; Com ER = commission error rates; PE = performance expectations. RTs are denoted in milliseconds and ERP amplitudes are denoted in microvolts.

The Brain and Behavior Relationship

Two ANOVAs were used to assess the relationship between ERP amplitudes and commission error rates as a function of group: one for the ERN, and another for the Δ ERN. These tests were conducted only on Placebo data. For both models, error rates were predicted by ERP amplitude, group, and an amplitude*group interaction term. For both the ERN and Δ ERN, we hypothesized that there would be significant interactions between ERP amplitude and group. Simple slopes analyses of these interactions would reveal that ERP amplitudes on error trials would be negatively correlated with error rates, but only for the Control group and not the Misattribution group (or at least more for the Control group than for the Misattribution group).

We also hypothesized that ERP amplitudes would mediate the relationship between group and commission error rates in Placebo data, such that reduced ERN/ Δ ERN amplitudes in the Misattribution group would be related to increased error rates. In conducting these mediation models, we assumed there would be a significant effect of group on ERP amplitudes; however, because group was not significantly related to ERP amplitudes in earlier tests, we did not conduct these models.

Figure 3a & 3b. Correlations between ERP amplitudes and commission error rates as a function of group.

Results. When participants were on the placebo, we found significant main effects of ERN amplitudes on commission error rates, $b = 0.01$, $SE < 0.01$, $F(1, 39) = 18.807$, $p < .001$, $\eta_p^2 = .20$, such that ERN amplitudes became larger (i.e., more negative) as error rates decreased (see Figure 3a). This same main effect was also found for Δ ERN amplitudes in the same direction, $b = 0.01$, $SE < 0.01$, $F(1, 39) = 7.986$, $p = 0.006$, $\eta_p^2 = .10$ (see Figure 3b). However, there were no significant main effects of group nor significant two-way interactions between ERP amplitudes and group on error rates, all $F_s < 1$, $p_s > .381$. Contrary to our predictions, we did not observe a reduction in the correlation between ERP

1 amplitudes and error rates for participants under the effects of the misattribution
2 placebo compared to those under the effects of the control placebo. These
3 findings show that the misattribution paradigm failed to disentangle the
4 relationship between brain and behavior, unlike our original study.

5 **Trait Anxiety**

6 Each preregistered analysis above was accompanied by a model that
7 included an additional main effect of trait anxiety, and all possible two-, three-, or
8 four-way interactions between trait anxiety and other predictors. These models—
9 preregistered but exploratory—assessed whether any principal interactions of
10 interest were moderated by trait anxiety, such that trait anxiety would confound
11 significant interactions or suppress non-significant ones.

12 **Results.** Trait anxiety did not moderate any of our interactions of interest,
13 all $F_s < 1.402$, all $p_s > .206$. It also did not have a significant main effect on ERP
14 amplitudes or state anxiety, $F_s < 1.609$, $p_s > .209$, failing to replicate positive
15 correlations found in previous studies (see Moser et al., 2013; Weinberg et al.,
16 2012). Trait anxiety did not suppress the effect of misattribution on ERP
17 amplitudes, state anxiety, or the relationship between brain and behavior, nor did
18 it confound significant increases in state anxiety from before to after the task.

19 However, trait anxiety was positively correlated with task-related anxiety, b
20 $= 0.07$, $SE = 0.02$, $F(1, 69.171) = 10.120$, $p = .002$, $R^2_\beta = .13$, showing that while
21 anxious participants did not report greater anxiety in general, they did report
22 greater anxiety about their task performance.

1 In the present study, we modified the design and analyses of the original
2 study to increase statistical power, but subsequently failed to replicate its three
3 principal findings. Although our paradigm did not alter behavior or performance
4 expectations as expected, it also did not reduce state anxiety, the amplitude of
5 the ERN/ Δ ERN, or the negative correlation between ERP amplitudes and error
6 rates. Further, trait anxiety did not significantly moderate any of the interactions
7 between these measured variables, and our results were unchanged when order
8 effects and changes in state anxiety were controlled for in exploratory analyses.

9 It is unclear why our paradigm failed to reduce state anxiety or ERN
10 amplitudes. Although participants believed in the effects of the placebo, we did
11 not ask them about the degree to which they may have misattributed their
12 emotions to it. Participants who did not report the placebo having a strong effect
13 may have felt that it was too mild to impact their anxiety. Conversely, participants
14 who believed the placebo had a strong effect on them may have viewed it as a
15 threat to their performance, and thus enhanced their own levels of performance-
16 monitoring to compensate for errors “caused by” the placebo. Given the range of
17 participants’ possible responses to our paradigm, the present study may have
18 benefited from collecting and coding participants’ construals of the placebo in
19 greater detail.

20 Because the present study failed to manipulate state anxiety through
21 misattribution, our data cannot address the question of the ERN’s emotional
22 properties. Our measure of state anxiety served as a manipulation check for
23 misattribution, where we expected the rise in anxiety from pre-to-post task to be

1 blunted for participants taking the misattribution placebo compared to those
2 taking the control placebo. Our data failed this manipulation check, with changes
3 in anxiety being the same across groups. Because of this manipulation failure,
4 our study cannot provide evidence for or against the hypothesis that the ERN has
5 emotional properties.

6 Despite these findings, the present study still contributes important
7 information toward research on emotions, misattribution, and the ERN. In the
8 following sections, we will interpret the present findings in the context of the two
9 previous related studies and discuss future research in this area.

10 **The Present Study in Context**

11 The failure of our paradigm to manipulate emotions aligns with the findings
12 of Rodilla et al. (2016), the only other published replication attempt of Inzlicht and
13 Al-Khindi (2012). Using a high-powered version of the original study, the authors'
14 misattribution paradigm failed to reduce state anxiety and ERN amplitudes, and
15 they did not find evidence for or against replication of the correlation between
16 Δ ERN amplitudes and error rates. This failure to replicate may have occurred
17 because of unreported differences in methodology between the two studies,
18 which could have influenced the paradigm or participants' behavior. For example,
19 for reasons that are unclear, reaction times in the replication attempt were faster
20 than those in the original study. However, equality-of-effect Bayes factors
21 (Bayarri & Mayoral, 2002) indicated that the mean differences in behavioral
22 variables across groups within each study were comparable, such that each
23 study's misattribution paradigm exerted equivalent nonsignificant effects on error

rates and reaction times. This showed that despite methodological differences between the studies, their respective misattribution paradigms were at least equivalent in their influence on behavior.

When considered together, the present study and replication attempt cast doubt on the replicability of the original study. This is for three major reasons. The first reason is the fundamental similarities between the present study and original study. Except for alterations made to increase statistical power, the materials and sampling populations of the two studies were nearly identical (e.g., same written directions, same laboratory environment, same recruiting procedures, etc.). 95% between-subject confidence intervals across the two studies overlapped for all measures except the ERN and Δ ERN (see Table 4). Results from the present study were also unchanged when split to eliminate order effects and replicate the precise tests of the original study. Thus, it is unlikely that the different findings across the two studies can be attributed to differences in methodology.

Table 4

Means and 95% between-subject confidence intervals of behavior, performance expectations, and ERP amplitudes across studies for On Placebo data.

	Present Study						Inzlicht & Al-Khindi (2012)					
	Control			Misattribution			Control			Misattribution		
	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB
Om ER (%)	5.20	3.46	6.94	4.19	2.79	5.60	3.83	2.28	5.38	4.89	2.69	7.09
Com ER (%)	45.73	40.72	50.74	46.80	41.07	52.53	47.01	39.84	54.17	47.79	41.06	54.52
RT Correct	293	283	302	289	276	302	288	278	297	283	266	258
RT Error	246	239	254	246	238	354	245	238	253	245	232	258
Pre-task PE	2.73	2.55	2.90	2.64	2.39	2.89	2.67 _a	1.84 _a	3.50 _a	3.12 _a	2.11 _a	4.13 _a
Post-task PE	2.13	1.96	2.29	2.18	1.97	2.38	n/a	n/a	n/a	n/a	n/a	n/a
ERN	-3.94	-6.07	-1.80	-5.68	-7.88	-3.48	-6.52	-8.40	-4.64	-3.92	-4.90	-2.94
Δ ERN	-9.67	-11.89	-7.44	-12.21	-14.43	-10.00	-4.25	-6.86	-1.64	-1.87	-3.36	-0.38

Note. LB = lower-bound; UB = upper-bound; Om ER = omission error rates; Com ER = commission error rates; PE = performance expectations. Subscripts denote data from pilot study. This table only depicts data while participants were under the effects of the placebo.

The second reason is that both the present study and the replication attempt had high power, while the tests of the original study had medium or low power. If we assume that we accurately determined the true effect size of misattribution on ERN amplitude ($\eta_p^2 = 0.17$) in the original study, we had only a 75% chance of successfully capturing that effect (i.e., 75% power). Conversely, both the replication attempt and present study had *a priori* power of 98% or greater to detect an effect of this same size. Further, if the small-to-medium effect size calculated for the present study had been the true effect size (about $\eta_p^2 = 0.02$), our original study would have had only a 14% chance of capturing it (i.e., 14% power). This was also true for the power of the original study's manipulation check; with a sample size of 47 in the pilot study, the test for the difference in state anxiety across groups had only a 61% chance of capturing its acquired medium-to-large effect ($d = .66$) and would have had only a 17% chance of capturing that effect had its true size been small-to-medium ($d = .35$).

The third reason is that the present study was fully preregistered, while our original study was not. Because all primary analyses and exclusion criteria were determined before data collection took place, the present study had few experimenter degrees of freedom that could have inflated the rate of false positives. In contrast, our original study had undisclosed flexibility in statistical analyses, exclusion criteria, and the content of hypotheses, increasing the chance that at least some of those positive findings were false. Indeed, several

1 procedures in the original study lacked details or explicit justification, giving us
2 the opportunity to inadvertently adjust data in support of our hypotheses. For
3 example, we did not justify log-transforming commission error rates before
4 correlating them with Δ ERN amplitudes, and did not perform this same
5 transformation for error rates when they were compared across groups.
6 Inconsistent transformation procedures without explicit justification are not
7 evidence of *p*-hacking, but do reveal the analytic flexibility that was available to
8 us in our statistical procedures at the time.

9 Two high-powered replication attempts with failed manipulation checks are
10 reasonable grounds to question the validity of the paradigm used in the original
11 study. However, these findings should not be misinterpreted as providing
12 evidence that misattribution cannot reduce the amplitude of the ERN, nor that a
13 successful manipulation of the ERN requires a substantially different approach to
14 misattribution than the approach employed in the original study. Until a
15 misattribution paradigm is developed that passes manipulation checks with
16 greater reliability, researchers should treat Inzlicht & Al-Khindi (2012) with more
17 caution as experimental evidence for the emotional properties of the ERN.

18 **The Future of Misattribution and the ERN**

19 Our findings should also not be interpreted as evidence against the idea
20 that the ERN has emotional properties, nor as evidence against the efficacy of
21 misattribution of arousal paradigms in general. There are many independent lines
22 of experimental and non-experimental evidence expounding on the role of affect
23 in error monitoring (e.g., Koban & Pourtois, 2014; Saunders, Lin, Milyavskaya, &

Inzlicht, 2017; Weinberg et al., 2012), and misattribution paradigms have shown promise in other areas of psychology (e.g., Heerdink, van Kleef, Homan, & Fischer, 2015; Huang & Gong, 2018; Yeung, Sharpe, Glozier, Hackett, & Colagiuri, 2017).

Instead, our findings should highlight the critical role of preregistered replications in creating a comprehensive and reproducible literature on a topic. Many studies conducted prior to the replication crisis (see Pashler & Wagnemakers, 2012) are underpowered, only a small proportion of studies in psychology have received published direct replications, and no published study in experimental psychology was preregistered before 2013. Indeed, these facts all pertain to the five most-cited publications that use a misattribution of arousal procedure (Cantor, Zillman & Bryant; 1975; Payne, Cheng, Govorun, & Stewart, 2005; Schwarz & Clore, 1983; Storms & Nisbett, 1970; Zanna & Cooper, 1974), which across eleven studies have a mean of only 18.58 ($SD = 4.69$) participants per cell². The present study adds nuance to this literature, remedying potential publication bias and drawing attention to findings that may require greater scrutiny.

The present study should also serve as a model for studies investigating the emotional properties of error monitoring signals that have yet to receive published replications. This is particularly the case for experimental studies that control for the cognitive and behavioral aspects of error monitoring. For example, Spunt and colleagues (2012) found that error-related activity in the dorsal medial

² As assessed on Web of Science and Google Scholar, June 2018.

1 frontal cortex tracked subjective reports of frustration; these findings held after
2 controlling for cognitive engagement and task performance. Pfabigan et al.
3 (2013) found that experimentally induced helplessness enhanced Δ ERN
4 amplitudes without influencing task-related behavior. And Ganuschak and
5 Schiller (2008) found that monetary incentives increased ERN amplitudes and
6 latencies in a verbal picture-naming task without altering reaction times or error
7 rates. For researchers concerned with the emotional properties of the ERN,
8 these studies represent opportune cases for direct replication: they can be well-
9 powered at relatively low sample sizes, they are methodologically simple to
10 implement, and they can be easily adapted to meet evolving standards in
11 statistical analyses.

12 Influential studies that are underpowered and analytically flexible will
13 continue to be published and cited until power analyses and preregistration
14 receive wider adoption among peer-reviewed journals. To reduce opportunity
15 costs from attempting to extend such studies, researchers should examine the
16 potential replicability of both their own paradigms and those of others when
17 preparing projects. This can be accomplished in several ways, depending upon
18 the heterogeneity of the relevant literature. When coupled with bias-correction
19 tools (Inzlicht, Gervais, & Berkman, 2015), meta-analytic procedures such as the
20 p-curve (Simonsohn, Nelson, & Simmons, 2014) or mini meta-analysis (Goh, Hall
21 & Rosenthal, 2016) can be used to determine the evidential value of a study or
22 line of research. If the supporting evidence is weak, such analyses can reveal
23 whether certain studies are worth trying to replicate at all.

Far from being uninteresting or redundant, replication attempts like the present study can greatly enrich our understanding and application of findings in psychology. Successive and independent replications elaborate on the generalizability, conceptual boundaries, and underlying ontology of phenomena like the ERN and paradigms like the misattribution of arousal. They help confirm the most plausible components of our theories and paradigms, while directing attention toward components that may be superfluous, incoherent, or in need of change. They will also be invaluable in shielding us from unreplicable programs of research that waste resources, effort, and time. Lastly, as the ERN receives increasing attention among clinical researchers as a possible biomarker of pathological anxiety (Kessel et al., 2016; Meyer, 2017), high-powered replications will be critical for developing an understanding of the signal that can be translated into clinical practice. By deliberately reserving resources to conduct replications, researchers will do much to speed up—rather than slow down—the development of new and effective ways to diagnose and treat mental illness.

Conclusion.

We conducted a methodologically precise and statistically powerful replication attempt of Inzlicht & Al-Khindi (2012), using nearly the exact same materials as the original study; however, our paradigm subsequently failed to manipulate emotions, and thus the present study cannot address the question of the ERN's emotional properties. Further studies with successful manipulations must be conducted to evaluate the degree to which variance in ERN amplitudes can be attributed to variance in state and trait emotions.

1 Nevertheless, the present study does provide some evidence against
2 replicability of Inzlicht & Al-Khindi's (2012) misattribution of arousal paradigm in
3 concurrently reducing state anxiety and the amplitude of the ERN. Researchers
4 should be more cautious in considering Inzlicht & Al-Khindi (2012) as evidence
5 that the ERN has fundamentally emotional properties, and may benefit from
6 seeking out alternative approaches to misattribution that manipulate emotions
7 with greater reliability.

8 In general, researchers have much to gain in pursuing and replicating
9 paradigms that manipulate emotions while controlling for the cognitive and
10 behavioral aspects of error monitoring. Establishing what paradigms are
11 reproducible in this area of the literature—and in psychology in general—will
12 greatly speed our acquisition of knowledge, especially for those aiming to
13 translate psychophysiological findings into the clinical domain.

Acknowledgements

We would like to thank all the members of the Toronto Laboratory for Social Neuroscience for their valuable input throughout the development of this project. This research will be made possible by grants from Canada's Natural Sciences and Engineering Research and Social Sciences and Humanities Research Councils to MI.

References

- Aarts, K., De Houwer, J., & Pourtois, G. (2013). Erroneous and correct actions have a different affective valence: Evidence from ERPs. *Emotion*, 13(5), 960.
- Aarts, K., De Houwer, J., & Pourtois, G. (2013). Erroneous and correct actions have a different affective valence: Evidence from ERPs. *Emotion*, 13(5), 960. doi:10.1037/a0032808
- Baldwin, S.A., Larson, M.J., Clayson, P.E., 2015. The dependability of electrophysiological measurements of performance monitoring in a clinical sample: a generalizability and decision analysis of the ERN and Pe. *Psychophysiology*, 52:790–800. doi:10.1111/psyp.12401
- Bartholow, B. D., Henry, E. A., Lust, S. A., Sauls, J. S., & Wood, P. K. (2012). Alcohol effects on performance monitoring and adjustment: affect modulation and impairment of evaluative cognitive control. *Journal of Abnormal Psychology*, 121(1), 173. doi:10.1037/a0023664
- Bayarri, M. J., & Mayoral, A. M. (2002). Bayesian analysis and design for comparison of effect-sizes. *Journal of Statistical Planning and Inference*, 103(1-2), 225–243. doi:10.1016/s0378-3758(01)00223-3
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*, 289-300.
- Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 1165-1188.

- 1 Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001).
2 Conflict monitoring and cognitive control. *Psychological Review*, 108(3),
3 624. doi:10.1037/0033-295X.108.3.624
- 4 Brockner, J., & Swap, W. C. (1983). Resolving the relationships between
5 placebos, misattribution, and insomnia: An individual-differences
6 perspective. *Journal of Personality and Social Psychology*, 45(1), 32–42.
7 doi:10.1037/0022-3514.45.1.32
- 8 Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the
9 anterior cingulate cortex. *Science*, 307(5712), 1118-1121.
10 doi:10.1126/science.1105783
- 11 Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E.
12 S., & Munafò, M. R. (2013). Power failure: why small sample size
13 undermines the reliability of neuroscience. *Nature Reviews Neuroscience*,
14 14(5), 365-376. doi:10.1038/nrn3475
- 15 Cantor, J. R., Zillmann, D., & Bryant, J. (1975). Enhancement of experienced
16 sexual arousal in response to erotic stimuli through misattribution of
17 unrelated residual excitation. *Journal of Personality and Social*
18 *Psychology*, 32(1), 69. doi:10.1037/h0076784
- 19 Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M.,
20 ..., Riddell, A. (n.d.). Stan: a probabilistic programming language. *Journal*
21 *of Statical Software* (in press).
- 22 Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation,
23 and affective responses to impending reward and punishment: The

- 1 BIS/BAS Scales. *Journal of Personality and Social Psychology*, 67(2),
2 319–333. doi:10.1037/0022-3514.67.2.319
- 3 Chartrand, T. L., & Bargh, J. A. (n.d.). Nonconscious motivations: Their
4 activation, operation, and consequences. *Self and Motivation: Emerging*
5 *Psychological Perspectives*, 13–41. doi:10.1037/10448-001
- 6 Clayson, P. E., & Miller, G. A. (2017a). ERP Reliability Analysis (ERA) Toolbox:
7 An open-source toolbox for analyzing the reliability of event-related brain
8 potentials. *International Journal of Psychophysiology*. doi:
9 10.1016/j.ijpsycho.2016.10.012
- 10 Clayson, P. E., & Miller, G. A. (2017b). Psychometric considerations in the
11 measurement of event-related brain potentials: Guidelines for
12 measurement and reporting. *International Journal of Psychophysiology*.
13 doi:10.1016/j.ijpsycho.2016.09.005
- 14 Cohen, J. (1973). Eta-Squared and Partial Eta-Squared in Fixed Factor ANOVA
15 Designs. *Educational and Psychological Measurement*, 33(1), 107-112.
16 doi:10.1177/001316447303300111
- 17 Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural
18 system for error detection and compensation. *Psychological Science*, 303-
19 305. doi:10.1016/j.tics.2006.03.007
- 20 Duncan, J. W., & Laird, J. D. (1980). Positive and reverse placebo effects as a
21 function of differences in cues used in self-perception. *Journal of*
22 *Personality and Social Psychology*, 39(6), 1024–1036.
23 doi:10.1037/h0077721

- 1 Edwards, L. J., Muller, K. E., Wolfinger, R. D., Qaqish, B. F., & Schabenberger,
2 O. (2008). An R2 statistic for fixed effects in the linear mixed model.
3 *Statistics In Medicine*, 27, 6137-6157. doi:10.1002/sim.3429
- 4 Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power
5 analysis program. *Behavior Research Methods, Instruments, &*
6 *Computers*, 28(1), 1–11. doi:10.3758/bf03203630
- 7 Ganushchak, L. Y., & Schiller, N. O. (2008). Motivation and semantic context
8 affect brain error-monitoring activity: an event-related brain potentials
9 study. *Neuroimage*, 39(1), 395-405.
10 doi:10.1016/j.neuroimage.2007.09.001
- 11 Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A
12 neural system for error detection and compensation. *Psychological*
13 *Science*, 4(6), 385-390. doi:10.1111/j.1467-9280.1993.tb00586.x
- 14 Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own
15 studies: Some arguments on why and a primer on how. *Social and*
16 *Personality Psychology Compass*, 10(10), 535-549.
17 doi:10.1111/spc3.12267
- 18 Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line
19 removal of ocular artifact. *Electroencephalography and clinical*
20 *neurophysiology*, 55(4), 468-484. doi:10.1016/0013-4694(83)90135-9
- 21 Hajcak, G., McDonald, N., & Simons, R. F. (2003). To err is autonomic: Error-
22 related brain potentials, ANS activity, and post- error compensatory

- 1 behavior. *Psychophysiology*, 40(6), 895-903. doi:10.1111/1469-
2 8986.00107
- 3 Heatherton, T. F., Polivy, J., & Herman, C. P. (1989). Restraint and internal
4 responsiveness: Effects of placebo manipulations of hunger state on
5 eating. *Journal of Abnormal Psychology*, 98(1), 89–92. doi:10.1037/0021-
6 843x.98.1.8
- 7 Heerdink, M. W., van Kleef, G. A., Homan, A. C., & Fischer, A. H. (2015).
8 Emotional expressions as social signals of rejection and acceptance:
9 evidence from the affect misattribution paradigm. *Journal of Experimental*
10 *Social Psychology*, 56, 60-68. Doi:10.1016/j.jesp.2014.09.004
- 11 Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error
12 processing: reinforcement learning, dopamine, and the error-related
13 negativity. *Psychological Review*, 109(4), 679. doi:10.1037/0033-
14 295X.109.4.679
- 15 Huang, Y., & Gong, H. (2018). The Minimal Deviation Effect: Numbers Just
16 Above a Categorical Boundary Enhance Consumer Desire. *Journal of*
17 *Consumer Research*. doi:10.1093/jcr/ucy048
- 18 Inzlicht, M., & Al-Khindi, T. (2012). ERN and the placebo: A misattribution
19 approach to studying the arousal properties of the error-related negativity.
20 *Journal of Experimental Psychology: General*, 141(4), 799.
21 doi:10.1037/a0027586
- 22 Inzlicht, M., Gervais, W., & Berkman, E. (2015). Bias-Correction Techniques
23 Alone Cannot Determine Whether Ego Depletion is Different from Zero:

- 1 Commentary on Carter, Kofler, Forster, & McCullough, 2015. Available at
2 SSRN: <http://ssrn.com/abstract=2659409>, doi:0.2139/ssrn.2659409
- 3 Kessel, E. M., Meyer, A., Hajcak, G., Dougherty, L. R., Torpey-Newman, D. C.,
4 Carlson, G. A., & Klein, D. N. (2016). Transdiagnostic factors and
5 pathways to multifinality: The error-related negativity predicts whether
6 preschool irritability is associated with internalizing versus externalizing
7 symptoms at age 9. *Development and Psychopathology*, 28(4pt1), 913–
8 926. doi:10.1017/s0954579416000626
- 9 Koban, L., & Pourtois, G. (2014). Brain systems underlying the affective and
10 social monitoring of actions: an integrative review. *Neuroscience &*
11 *Biobehavioral Reviews*, 46, 71-84. doi: 10.1016/j.neubiorev.2014.02.014
- 12 Legault, L., & Inzlicht, M. (2013). Self-determination, self-regulation, and the
13 brain: Autonomy improves performance by enhancing neuroaffective
14 responsiveness to self-regulation failure. *Journal of Personality and Social*
15 *Psychology*, 105(1), 123. doi:10.1037/a0030426
- 16 Little, R.J.A. and Rubin, D.B. (2002). Statistical Analysis with Missing Data, 2nd
17 edition, New York: John Wiley. doi:10.1002/9781119013563
- 18 Luck, S.J., & Gaspelin, N. (2017). How to get statistically significant effects in any
19 ERP experiment (and why you shouldn't). *Psychophysiology*, 54(1), 146-
20 157. doi:10.1111/psyp.12639
- 21 Luu, P., Collins, P., & Tucker, D. M. (2000). Mood, personality, and self-
22 monitoring: Negative affect and emotionality in relation to frontal lobe
23 mechanisms of error monitoring. *Journal of Experimental Psychology:*

- 1 *General*, 129, 43–60. doi:10.1037/0096-3445.129.1.43
- 2 Luu, P., Tucker, D. M., Derryberry, D., Reed, M., & Poulsen, C. (2003).
- 3 Electrophysiologic responses to errors and feedback in the process of
- 4 action regulation. *Psychological Science*, 14, 47–53. doi:10.1111/1467-
- 5 9280.01417
- 6 Makeig, S., Bell, A. J., Jung, T. P., & Sejnowski, T. J. (1996). Independent
- 7 component analysis of electroencephalographic data. *Advances in Neural*
- 8 *Information Processing Systems*, 145-151.
- 9 Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of
- 10 the state scale of the Spielberger State—Trait Anxiety Inventory (STAI).
- 11 *British Journal of Clinical Psychology*, 31(3), 301-306. doi:10.1111/j.2044-
- 12 8260.1992.tb00997.x
- 13 McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric Properties of
- 14 the Intrinsic Motivation Inventory in a Competitive Sport Setting: A
- 15 Confirmatory Factor Analysis. *Research Quarterly for Exercise and Sport*,
- 16 60(1), 48–58. doi:10.1080/02701367.1989.10607413
- 17 Meyer, A. (2017). A biomarker of anxiety in children and adolescents: A review
- 18 focusing on the error-related negativity (ERN) and anxiety across
- 19 development. *Developmental Cognitive Neuroscience*, 27, 58–68.
- 20 doi:10.1016/j.dcn.2017.08.001
- 21 Meyer, A., Weinberg, A., Klein, D. N., & Hajcak, G. (2012). The development of
- 22 the error-related negativity (ERN) and its relationship with anxiety:

- 1 Evidence from 8 to 13 year-olds. *Developmental Cognitive Neuroscience*,
2 2(1), 152–161. doi:10.1016/j.dcn.2011.09.005
- 3 Moser, J. S., Moran, T. P., Schroder, H. S., Donnellan, M. B., & Yeung, N.
4 (2013). On the relationship between anxiety and error monitoring: a meta-
5 analysis and conceptual framework. *Frontiers in Human Neuroscience*, 7.
6 doi:10.3389/fnhum.2013.00466
- 7 Olvet, D. M., & Hajcak, G. (2009). The stability of error- related brain activity with
8 increasing trials. *Psychophysiology*, 46(5), 957-961. doi:10.1111/j.1469-
9 8986.2009.00848.x
- 10 Open Science Collaboration. (2015). Estimating the reproducibility of
11 psychological science. *Science*, 349(6251). doi:10.1126/science.aac4716
12 aac4716
- 13 Pailing, P. E., & Segalowitz, S. J. (2004). The effects of uncertainty in error
14 monitoring on associated ERPs. *Brain and Cognition*, 56(2), 215-233.
15 doi:10.1016/j.bandc.2004.06.005
- 16 Pashler, H., & Wagenmakers, E.J. (2012). Editors' introduction to the special
17 section on replicability in psychological science: A crisis of confidence?
18 *Perspectives on Psychological Science*, 7, 528–
19 530.doi:10.1177/1745691612465253
- 20 Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for
21 attitudes: Affect misattribution as implicit measurement. *Journal of*
22 *Personality and Social Psychology*, 89(3), 277–293. doi:10.1037/0022-
23 3514.89.3.277

- 1 Pfabigan, D. M., Pintzinger, N. M., Siedek, D. R., Lamm, C., Derntl, B., & Sailer,
2 U. (2013). Feelings of helplessness increase ERN amplitudes in healthy
3 individuals. *Neuropsychologia*, 51(4), 613–621.
4 doi:10.1016/j.neuropsychologia.2012.12.008
- 5 Pourtois, G., Vocat, R., N'diaye, K., Spinelli, L., Seeck, M., & Vuilleumier, P.
6 (2010). Errors recruit both cognitive and emotional monitoring systems:
7 simultaneous intracranial recordings in the dorsal anterior cingulate gyrus
8 and amygdala combined with fMRI. *Neuropsychologia*, 48(4), 1144-1159.
9 doi:10.1016/j.neuropsychologia
- 10 Proudfit, G. H., Inzlicht, M., & Mennin, D. S. (2013). Anxiety and error monitoring:
11 the importance of motivation and emotion. *Frontiers in Human*
12 *Neuroscience*, 7, 636. doi:10.3389/fnhum.2013.00636
- 13 Quené, H., & van der Bergh, H. (2004). On multi-level modeling of data from
14 repeated measures designs: a tutorial. *Speech Communication*, 43(1-2),
15 103-121. doi:10.1016/j.specom.2004.02.004
- 16 Reisenzein, R. (1983). The Schachter theory of emotion: Two decades later.
17 *Psychological Bulletin*, 94(2), 239–264. doi:10.1037/0033-2909.94.2.239
- 18 Ridderinkhof, K. R., de Vlugt, Y., Bramlage, A., Spaan, M., Elton, M., Snel, J., &
19 Band, G. P. (2002). Alcohol consumption impairs detection of performance
20 errors in mediofrontal cortex. *Science*, 298(5601), 2209-2211. doi:
21 10.1126/science.1076929

- 1 Riesel, A., Weinberg, A., Endrass, T., Kathmann, N., & Hajcak, G. (2012).
2 Punishment has a lasting impact on error- related brain activity.
3 *Psychophysiology*, 49(2), 239-247. doi:10.1111/j.1469-8986.2011.01298.x
- 4 Rodilla, C.C., Beauducel, A., & Leue, A. (2015). *Misattribution of Anxious Arousal*
5 *Reduces the Effect of Trait-Anxiety on the ERN*. Poster session presented
6 at the 55th Annual Meeting of the Society for Psychophysiological
7 Research, Seattle, OR.
- 8 Rodilla, C.C., Beauducel, A., & Leue, A. (2016). Error-Related Negativity and the
9 Misattribution of State-Anxiety Following Errors: On the Reproducibility of
10 Inzlicht and Al-Khindi (2012). *Frontiers in Human Neuroscience*, 10, 475.
11 doi: 10.3389/fnhum.2016.00475
- 12 Ross, L., Rodin, J., & Zimbardo, P. G. (1969). Toward an attribution therapy: The
13 reduction of fear through induced cognitive–emotional misattribution.
14 *Journal of Personality and Social Psychology*, 12, 279–288.
15 doi:10.1037/h0027800
- 16 Saunders, B., Lin, H., Milyavskaya, M., & Inzlicht, M. (2017). The emotive nature
17 of conflict monitoring in the medial prefrontal cortex. *International Journal*
18 *of Psychophysiology*, 119, 31-40. doi:10.1016/j.ijpsycho.2017.01.004
- 19 Saunders, B., Rodrigo, A. H., & Inzlicht, M. (2016). Mindful Awareness of
20 Feelings Increases Neural Performance Monitoring. *Cognitive, Affective, &*
21 *Behavioral Neuroscience*, 16, 93-105. doi:10.3758/s13415-015-0375-2

- Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 69(5), 379–399. doi:10.1037/h0046234
- Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, 45(3), 513–523. doi:10.1037/0022-3514.45.3.513
- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science*, 9(6), 666–681. doi:10.1177/1745691614553988
- Spielberger, C. D. (2010). *State-Trait Anxiety Inventory*. John Wiley & Sons, Inc.
- Spunt, R. P., Lieberman, M. D., Cohen, J. R., & Eisenberger, N. I. (2012). The Phenomenology of Error Processing: The Dorsal ACC Response to Stop-signal Errors Tracks Reports of Negative Affect. *Journal of Cognitive Neuroscience*, 24(8), 1753–1765. doi:10.1162/jocn_a_00242
- Storms, M. D., & Nisbett, R. E. (1970). Insomnia and the attribution process. *Journal of Personality and Social Psychology*, 16, 319–328. doi:10.1037/h0029835
- Team, S.D. (2016). CmdStan: the command-line interface to Stan, version 2.10.0. Retrieved from <http://mc-stan.org/>.
- Van Veen, V., & Carter, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiology & Behavior*, 77(4), 477–482. doi:10.1016/S0031-9384(02)00930-7

- 1 Yeung, V., Sharpe, L., Glozier, N., Hackett, M. L., & Colagiuri, B. (2017). A
2 systematic review and meta-analysis of placebo versus no treatment for
3 insomnia symptoms. *Sleep Medicine Reviews*, 38, 17-27.
4 doi:10.1016/j.smr.2017.03.006
- 5 Weinberg, A., Riesel, A., & Hajcak, G. (2012). Integrating multiple perspectives
6 on error-related brain activity: the ERN as a neural indicator of trait
7 defensive reactivity. *Motivation and Emotion*, 36(1), 84-100.
8 doi:10.1007/s11031-011-9269-y
- 9 Weinberg, A., Meyer, A., Hale-Rude, E., Perlman, G., Kotov, R., Klein, D. N., &
10 Hajcak, G. (2016). Error-related negativity (ERN) and sustained threat:
11 Conceptual framework and empirical evaluation in an adolescent sample.
12 *Psychophysiology*, 53(3), 372–385. doi:10.1111/psyp.12538
- 13 Weinberg, A., Olvet, D. M., & Hajcak, G. (2010). Increased error-related brain
14 activity in generalized anxiety disorder. *Biological Psychology*, 85(3), 472–
15 480. doi:10.1016/j.biopsycho.2010.09.011
- 16 Xiao, Z., Wang, J., Zhang, M., Li, H., Tang, Y., Wang, Y., ... Fromson, J. A.
17 (2011). Error-related negativity abnormalities in generalized anxiety
18 disorder and obsessive–compulsive disorder. *Progress in Neuro-*
19 *Psychopharmacology and Biological Psychiatry*, 35(1), 265–272.
20 doi:10.1016/j.pnpbp.2010.11.022
- 21 Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error
22 detection: conflict monitoring and the error-related negativity.
23 *Psychological Review*, 111(4), 931.

- 1 Zanna, M. P., & Cooper, J. (1974). Dissonance and the pill: An attribution
- 2 approach to studying the arousal properties of dissonance. *Journal of*
- 3 *Personality and Social Psychology*, 29, 703–709. doi:10.1037/h0036651