International Journal of Psychophysiology 111 (2017) 17–32

Contents lists available at ScienceDirect

International Journal of Psychophysiology

journal homepage: www.elsevier.com/locate/ijpsycho

Meta-analysis and psychophysiology: A tutorial using depression and action-monitoring event-related potentials

Tim P. Moran [a,*], Hans S. Schroder [b], Chelsea Kneip [b], Jason S. Moser [b]

[a] School of Psychology, Georgia Institute of Technology, United States
[b] Department of Psychology, Michigan State University, United States

ARTICLE INFO

Article history:
Received 15 February 2016
Received in revised form 9 June 2016
Accepted 1 July 2016
Available online 1 July 2016

Keywords:
Meta-analysis
Psychophysiology
Depression
Error-related negativity
Feedback negativity

ABSTRACT

Meta-analyses are regularly used to quantitatively integrate the findings of a field, assess the consistency of an effect and make decisions based on extant research. The current article presents an overview and step-by-step tutorial of meta-analysis aimed at psychophysiological researchers. We also describe best-practices and steps that researchers can take to facilitate future meta-analysis in their sub-discipline. Lastly, we illustrate each of the steps by presenting a novel meta-analysis on the relationship between depression and action-monitoring event-related potentials – the error-related negativity (ERN) and the feedback negativity (FN). This meta-analysis found that the literature on depression and the ERN is contaminated by publication bias. With respect to the FN, the meta-analysis found that depression does predict the magnitude of the FN; however, this effect was dependent on the type of task used by the study.

1. Introduction

Most scientific questions are addressed by multiple studies conducted by independent research teams using a diverse range of methods rather than by a single study. Researchers understand and accept that the results of these studies will often vary and, in some cases, may directly contradict each other. Yet researchers also want to be able to use these varied and conflicting findings to come to a consensus regarding a body of work – for example, it is often desirable to determine whether the predictions of a theory have been supported or whether a finding has practical applications. For the greater part of the previous century, researchers from a number of fields including physics, psychology, ecology, zoology, archaeology, astronomy and medicine (Birge, 1929, 1932; Haidich, 2010; Petticrew, 2001) have relied on meta-analysis to quantitatively summarize a body of work and draw conclusions. “Meta-analysis” refers to a set of procedures that statistically analyze the results of primary studies (i.e. the original research) in order to synthesize the findings (Glass, 1976).

The purpose of this article is to provide a broad overview of what meta-analysis is as well as a practical tutorial aimed at psychophysiologists. This article is organized around a series of steps that nearly all meta-analyses will follow (adapted from Cooper, 2010; Cumming, 2012): formulating the problem, conducting the literature search, coding studies and extracting data, synthesizing effect sizes and assessing for heterogeneity, and assessing for threats to validity. Each of these sections will present tips, strategies and best-practices for conducting a meta-analysis. Each of the five sections will end with an illustrative example from a novel meta-analysis we performed on the relationship between depression and action-monitoring event-related potentials (ERPs), namely the error-related negativity (ERN) and the feedback negativity (FN). We conclude by identifying challenges to conducting robust meta-analyses, and offer some possible solutions for psychophysiologists to take up in planning, executing, and reporting on future studies.

2. Step 1: formulate the problem

Conducting a meta-analysis can take a great deal of time and effort. For example, one of the authors recently completed a meta-analysis which required approximately 16 months (Moran, in press). Given the work involved, one could legitimately wonder if summarizing the existing literature with a meta-analysis is a better use of one's time than trying to address an existing question with a new primary study or by summarizing the literature with a narrative review. The type of study that one conducts should, of course, be dependent on one's goals. The following are goals that meta-analysis is particularly well-suited, or uniquely-suited, to addressing: 1) Determine if an effect is “real.” Psychologists of all stripes must deal with findings that

* Corresponding author at: Georgia Institute of Technology, 654 Cherry St., Atlanta, GA 30332, United States.
E-mail address: timothy.moran@psych.gatech.edu (T.P. Moran).

http://dx.doi.org/10.1016/j.ijpsycho.2016.07.001
0167-8760/

occasionally fail to replicate – whether due to false positives or simply sampling error and low power. A meta-analysis can help test an effect, often with far greater power than any single study (see Section 6.6 for an example). 2) Determine the consistency of an effect. Psychologists are accustomed to the fact that some effects may be dependent on a particular population/setting/design etc. – that is, an effect might be moderated by some other variable. Given that even two studies testing the same hypotheses can differ markedly in these variables, a highly inclusive meta-analysis is well-suited to testing the role of moderating variables – and with much greater power than any single study. Additionally, meta-analyses can test moderators that no single study is capable of addressing. For example, a researcher may wish to determine if the effects in a given field are shrinking over time – the so-called "decline effect" (Schooler, 2011). Obviously, no single study is capable of addressing this; but a meta-analysis involves summarizing the results of many studies published over several years. 3) Increase the precision of an estimate. In some cases, a psychologist may wish to go beyond statistical significance and compute a highly precise estimate of the magnitude of an effect. This is likely to be most important in applied settings where precise estimates of a measure's predictive validity are highly desired. 4) Assess the literature for publication bias. Publication bias, described below, occurs when the published literature systematically differs from the population of all studies conducted on a topic. Thus, publication bias affects an entire literature, not an individual study. A meta-analysis may be able to determine whether a literature is contaminated by publication bias whereas a primary study is not.

That having been said, meta-analyses are not appropriate for all situations. In particular, a meta-analysis cannot fix a "broken" literature. For example, if a researcher believes that a given literature is full of poorly conducted studies – e.g. invalid instruments, poor experimental designs, etc. – a meta-analysis will not be able to produce a meaningful summary of that literature. Meta-analyses obey the "law of conservation of garbage" – i.e. garbage in, garbage out. If a meta-analysis includes low-quality studies with questionable findings, the results of the meta-analysis will also be questionable. In these situations, a meta-analysis might be able to directly compare effects from low- and high-quality studies (although, the low/high distinction must be determined by the meta-analyst), but it will not be able to correct for poor design. When a literature is contaminated with poorly conducted studies, a new primary study which address the limitations of the literature may be a more prudent use of time.

Once it has been determined that a meta-analysis is the appropriate design for a given question, the researcher must carefully formulate the problem they wish to address. This may be as simple as wanting to know if one variable predicts another or as complicated as testing the predictions of a theoretical model. The meta-analyst must also carefully define the scope of their investigation. For example, a psychophysiologist who wants to study the N2 ERP component must decide whether to include the N2a, the N2b, the N2pc etc. This is not a strictly linear process. It is possible that the scope of a meta-analysis may need to be refined as research reports are located. Given that the process of formulating a research question is likely to be familiar to most researchers, we will not discuss these issues in depth.

In the following section, we formulate the research problem regarding the relationship between depression and action-monitoring ERPs – i.e., the ERN and the FN. We briefly review the relevant literature and describe why a meta-analysis is a useful way to proceed.

2.1. Example using depression and the ERN/FN

Depression is among the most common psychiatric conditions and is associated with a high rate of recurrence and significant personal and societal cost (Greden, 2001; Lai, 2011; Lopez et al., 2006; World Health Organization, Switzerland, 2011). For example, depression is associated with increased healthcare costs and service utilization, missed work, impaired academic and social functioning, recurrent depressive episodes and increased risk for suicide. Given these findings, the last few years have seen increased effort to identify "biomarkers" which indicate risk for the development of depression thereby facilitating early diagnosis and preventative care.

For depression, two event-related potentials have shown promise as candidate biomarkers: the error-related negativity (ERN) and the feedback negativity (FN). The ERN is a negative deflection in the human event-related potential (ERP) that occurs within 100 ms of the commission of an error during forced-choice reaction time tasks – e.g. the Flankers task (see Fig. 1). In the Flankers task, participants must identify a central stimulus that is surrounded by several flanking distracters (e.g. <<<<< or <<><<) as quickly as possible. In tasks such as this, participants often make quick incorrect responses due to lapses in attention or the incongruency between the central and flanking stimuli. The ERN is generated in the anterior cingulate cortex and surrounding motor areas and is often considered an error detection/correction (Carter and van Veen, 2007; Gehring et al., 2012; Holroyd and Coles, 2002; van Veen and Carter, 2002) or response-conflict (Yeung et al., 2004) signal. With respect to psychopathology, it has been hypothesized that the ERN may serve as a biomarker for all internalizing disorders including depression (Olvet and Hajcak, 2008). In support of this proposal, an enlarged ERN has been observed in individuals with major depression (e.g. Holmes and Pizzagalli, 2008, 2010) as well as undergraduates high in self-reported sadness (e.g. Dywan et al., 2008).

Unlike the ERN, which is elicited by an internal monitoring process, the FN is elicited by external feedback indicating an unfavorable response. The FN is most often elicited in a gambling/guessing task in which the participant must make a choice (e.g. determine which door has a prize behind it) and is then rewarded for a "correct" choice (e.g. money might be awarded for a correct guess and taken away for an incorrect guess). The FN is thought to originate in the ACC (Gehring et al., 2012; Holroyd and Coles, 2002) or the striatum (Foti et al., 2011); theorizing on the FN suggests that it signals that an event was worse than expected or that it signals a desired event (Holroyd and Coles, 2002; Foti et al., 2011). With respect to depression, Hajcak and colleagues have conducted a number of studies demonstrating that depression is related to an attenuated FN in undergraduates (Foti and Hajcak, 2009), children (Bress et al., 2012) and patients suffering from major depression (Foti et al., 2014).

The findings reviewed above suggest that depression is characterized by both an enhanced ERN and a blunted FN. However, this set of findings has proven somewhat difficult to replicate. For example, with respect to the ERN, a number of studies have found no difference between depressed individuals and controls in both adults (Olvet et al., 2010; Weinberg et al., 2012) and children (Bress et al., 2015) whereas others have found that depression is associated with a reduced ERN (Ladouceur et al., 2012; Ruchsow et al., 2004, 2006; Schrijvers et al., 2009). Similarly, although several studies have found evidence for a reduced FN in depression, other work has found evidence for an enlarged FN in depression (e.g. Mies et al., 2011; Mueller et al., 2015). Before the ERN/FN can be applied to clinical settings, it must be determined whether, and how strongly, they are associated with depression.

There are a number of possibilities that can potentially explain these disparate findings. For example, the differences in findings may be attributable to some untested moderator(s). Studies assessing the association between depression and the ERN/FN have employed a variety of different types of samples and tasks. Some have studied depressed undergraduates whereas others have studied patients suffering from major depressive disorder. Among the studies of major depressive disorder, some have focused on untreated patients only whereas others have included patients receiving medication. Additionally, a number of different tasks (e.g. Flanker, Stroop, gambling etc.) have been used to elicit the ERN/FN. Each of these factors may have a role in explaining these disparate findings. It is also possible that publication bias – the tendency for smaller studies to be published only if they produce positive findings and for larger studies to be published regardless of their