Reflection on “Trial and Error (-Related Negativity)”

JOTE
Journal of Trial and Error

Maarten Derksen¹

¹ Maarten Derksen, Theory & History of Psychology, University of Groningen

Correspondence: Maarten Derksen, Theory & History of Psychology, University of Groningen, Groningen, the Netherlands

Email: m.derksen@rug.nl

Companion Piece: This paper reflects on a previously published empirical article, [Traxler et al. \(2020\)](#). Read the abstract at the end of this article.

Abstract

‘Trial and Error (-related negativity)’ is a fascinating paper detailing the attempt to develop a new experimental paradigm to study the role of error-related negativity in the development of avoidance behavior. In my comments on this paper I will focus on the interaction between experimenters and participants as the former investigate various ways of designing the experiment, aiming to elicit the right kind of behavior from the participants. As in many psychological experiments, there is a fundamental tension here that experimenters must find a way to deal with: they must guide the subject to the proper performance, without the subject responding to the guidance as such. The performance must be natural, but within tight constraints. Recalcitrance or resistance of the subject must be prevented. Ultimately, the authors of ‘Trial and Error (-related negativity)’ failed in their attempt to do this. Their reflections on their failure are thorough and illuminating, but I will argue that they can be pushed slightly further.

KEYWORDS

tact, performance, recalcitrance, spontaneity, validity

Introduction: three kinds of errors

This paper is about error on three levels. First, it deals with research into the ERN, error-related negativity. The ERN is a negative deflection in the electroencephalography (EEG) signal, which tends to occur within 100 milliseconds of making an error. The authors hypothesize that physical pain can be considered as a bodily signal that a type of error has been committed: there is a “discrepancy between the actual and optimal/targeted state”, as the authors put it (Traxler et al., 2020, p.27). This raises the question whether the ERN is also associated with pain and the avoidance of pain, and if so how. More specifically the authors want to know whether people with an elevated ERN are more prone to avoidance behavior, which in turn can lead to chronic pain.

I am an historian of psychology with philosophical interests and have no expertise in clinical neuropsychology, so I will not comment on this hypothesis. But the paper also deals with error in two other ways, which I do feel able to reflect on. The authors describe their attempts to develop an experimental paradigm for the study of the role of the ERN in pain avoidance. In these attempts they make errors which they then try to correct in a further attempt, six task versions in total. This is the second way this paper deals with errors - those of the experimenters themselves. But there is a third level too: it is crucial for the experimental task to induce the participant to make the right number of errors - not too many, not too few. The second and third aspects are obviously related: the errors of the experimenters concern, among others, the number of errors the participants make.

The experiment as a psychological problem

The authors describe their challenge as an interdisciplinary one: they had to combine elements of neurophysiology (ERN) with clinical psychology (pain avoidance). Specifically: they had to somehow induce an ERN in the participants and elicit and measure some type of avoidance behavior at the same time. Moreover, to determine what each participant’s average ERN is, they needed at least six ERN measures per participant, and thus a minimum of six errors. The errors, finally, had to be “inhibition errors”, not errors due to lack of knowledge or skill. It was not clear to me why this was important, or what an inhibition error is in the first place, but this is no doubt due to my own lack of knowledge in this field.

All in all, the specifications of the task were narrow and demanding: not any type of error would do (only inhibition errors); the errors had to produce a proper ERN; a minimum of six was needed; the participants had to be aware of their error (otherwise they would not show avoidance behavior); and of course there had to be pain associated with the errors, but not so much pain that the ethics committee would reject the pilot study, or the subjects would refuse to participate. The researchers had to set up the experiment in such a way that this narrow target was met, and the participants were the most challenging factor: the experimenters had to elicit the right kind of performance from them, without creating suspicion or resistance. Human participants tend to form their own ideas about the experiment, they do not merely respond to the stimuli.

The first one to discuss this particularity of the psychological experiment

was the American psychologist Saul Rosenzweig. In an article titled ‘The experimental situation as a psychological problem’, Rosenzweig (1933) argued that an experiment, in any scientific discipline, normally combines three factors: experimental materials, one or more experimenters, and instruments. Ideally, Rosenzweig argued, these three factors remain separate: ‘(e)very factor stays within its own domain’ (Rosenzweig, 1933, p.338). In psychology, however, this is often not the case, and this creates problems. The experimenter may unwittingly act as a stimulus in the experiment (thus becoming part of the experimental materials), because his or her presence may have an influence on the behavior of the subject. Conversely, subjects may take on the role of experimenter when they become self-conscious of their role in the experiment and start to modify their responses accordingly. Managing these twin problems requires the experimenter to be tactful in how they guide the behavior of the participant.

In a recent article, Brenninkmeijer, Rietzschel, and I reported on a series of interviews we had conducted with researchers in psychology (Brenninkmeijer et al., 2019). We had asked them about their informal research practices, that is to say those practices that are not made explicit in the method section of a paper but are nevertheless considered important. We saw two themes in their answers. The first is a strong concern with professionalism, expressing itself among other things in an orderly lab and a smoothly running experiment, and respectful (but not amicable) conduct towards the participants. The second theme is a focus on producing good data by managing the performance of the participant. This second theme is relevant with regard to this paper on ‘Trial and Error’. Much of the work that is done in the lab, including the kind of informal, unwritten work that we explored in our interviews, is geared towards eliciting the right kind of behavior from the participant. What happens in the lab has a theatrical quality: a particular performance is expected from the participants, and they are guided to it by the staging and scripting (or choreography) of the experiment and the conduct of the experimenter. Our interviewees mentioned the importance of the instructions (clear, fool-proof, not too long), of keeping participants focused and motivated (particularly challenging in cognitive psychology, where tasks are often boring and repetitive), and of creating a certain psychological realism, so that the artificiality of the laboratory situation is less salient and the stimuli and tasks more life-like (this was especially a concern for social psychologists).

What makes this work so difficult is that despite the staging and the scripting, the participant’s behavior must be spontaneous and natural. Or rather: the spontaneity is elicited and facilitated by the staging, the instructions, the props, the stimuli, and the conduct of the experimenter. A psychological experiment aims to create natural behavior artificially, aims to produce spontaneity (see also Derksen (2001) about psychological tests). This paradoxical task demands that the experiment is crafted with subtlety and tact. The experimental situation and the experimenter must be forceful, yet unobtrusive. If the staging and the scripting become too prominent the participant risks become recalcitrant, and their behavior is no longer natural and spontaneous (Derksen, 2017; Lezaun, 2007). In designing an experiment, the management of the participants’ awareness is therefore often a major concern: they must be attentive to the stimulus, but not to the stimulus *as* stimulus, as part of an experiment in an artificial laboratory situation, aimed at probing their responses (Abma, 2020). They

must not become reflexive but remain natural and spontaneous.

There is thus a fundamental asymmetry in this endeavor: it is the experimenter who leads the participant according to a particular choreography and in an environment that the experimenter has created. The experiment is like a dance, with the participant following the lead of the experimenter. At the same time, it is essential that the experimental subjects participate in the dance freely and spontaneously, i.e. that they ‘remain themselves’. Everything in the experiment is aimed towards eliciting a response from the participant that is somehow ‘natural’ and ‘real’ rather than contrived and artificial - a response that represents real-life behavior, despite the fact that it occurred in a thoroughly artificial setting. This paradoxical challenge requires a great deal of ingenuity and tact from researchers.

Awareness and distraction

Interestingly, the authors of this paper designed their new experimental paradigm with the help of their twelve participants. Although they are described in the paper in the usual impersonal, technical terms – ‘twelve participants (8 females)’, ‘mean age of $M = 29.25$ ($SD = 10.64$)’ (Traxler et al., 2020, p.29) - the participants nonetheless had an active role in the pilot, almost as collaborators, specifically by supplying the researchers with their experience of the experiment. ‘(T)hey were asked whether they were able to perceive when they made errors, how difficult they found the task as well as initiating the avoidance response, and any additional comments they would like to provide’ (Traxler et al., 2020, p.29).

As noted in this quote, it was vital that the participants were aware they had made an error, because otherwise they would not “initiate an avoidance response”. The experimenters had to hit a very narrow target here. On the one hand, participants had to be induced to make a sufficient number of errors, so the task had to be difficult enough that participants did not always know the right response. On the other hand, once they had responded they had to be aware of having made an error, and thus *did* have to know the right response. This transition from not knowing to knowing had to occur within 1300 ms, that amount of time being the fixation period plus the time allowed for an avoidance response. This required precision engineering of the awareness of the participants.

It quickly became clear that the basic task - determining where on their lower back they felt a vibration from a “tactor” - did not induce enough errors. To make the task more difficult the researchers decided to distract the participants. The tool that they used to modulate the awareness and attention of the participants was the wonderfully named distractor tactor. The researchers tried various ways to distract the participants, aiming at just the right amount of distraction to induce errors, without distracting the participants so much that they no longer were aware of their errors. Their first attempt was creative, but unsuccessful. I suspect that working out which song is playing merely from the beat is difficult enough in normal circumstances, but if the beat is transmitted by vibrations on one’s lower back it becomes well-nigh impossible. I am not surprised that the one participant in this task version simply gave up trying and focused entirely on the main task of locating the main stimuli. Here we have an example of participant recalcitrance, in this case not in response to the mere fact of being manipulated, but in response to the difficulty of the distractor task.

The participant refused to be distracted. Moreover, this participant discovered they could hear the beat as well as feel it: the distractor tactor worked as a little speaker, and the ear plugs that the experimenters had given the participant did not stop the sound. In other words, the distractor task was at the same time too difficult and too easy: it was too difficult to perform in the intended way (by attending to the vibrations of the distractor tactor), and at the same time it was too easy to not be distracted by it, and too easy to cheat. The experimenters had failed to tactfully usher the participant into the right state of distraction, and instead met resistance. The participant did not perform as the experimenters had intended. Ultimately, the 100% coactivation task, which the experimenters had tried first, was the only one that met the requirements of a sufficient number of errors combined with awareness of having made an error after responding.

Stimulating avoidance

Regarding the other part of the experimental paradigm, the avoidance response, matters were even more complicated. To get a sufficient number of avoidance responses, the experimenters not only had to make sure that participants were aware of having made errors, but also that they were sufficiently motivated to avoid the punitive electrocutaneous stimulus (e-stim) that would follow an error. While in the error-induction part of the task the attention of the participants had to be modulated in such a way that they made errors but were also aware that they made errors, for the avoidance response a different balance had to be struck. The dynamic of control and resistance that is so typical of psychological experiments is quite prominent here.

First, the e-stim had to be calibrated for each individual participant, in such a way that it was 'painful and demanding some effort to tolerate' (Traxler et al., 2020, p.29). The researchers, in consultation with the participants, had to find the golden mean between not enough pain (and therefore no motivation to avoid) and too much pain (which would lead participants to reject the experiment altogether, if it would even pass the Ethics Committee). The participants were allowed to choose their preferred level of punishment, so to say, but the researchers were nonetheless afraid that participants would subvert the experiment by always performing the avoidance response (pressing the space bar on the keyboard) as a precaution, whether or not they thought they had made an error. Thus, there had to be a cost to avoidance. At the same time, non-avoidance should be costly, too, otherwise the researchers would not get the kind of responses they were interested in.

To direct their participants to this happy medium between too much and too little avoidance, the experimenters employed threats. Participants were told that they would be punished for avoiding punishment. Every time they pressed the space bar to prevent the e-stim, extra trials would be added to the experiment. This ploy failed: Again participants resisted, this time by simply foregoing avoidance and accepting punishment. The authors imply that participants had made a strategic calculation: If each avoidance response would lead to extra trials, the experiment would take longer to finish. Moreover, participants realized that extra trials would also mean more chances to make errors, which would either be punished with the e-stim, or, if avoided, would lengthen the experiment even more. In other words, the threats of the experimenters unwittingly focused the attention of the participants on the boundaries of the experiment, rather than on the task *within* the experiment. The experiment

became visible to the participants as a social situation that they were in, with a beginning and an end, and they were eager not to stay in it too long.

To 'balance the trade-off' (p.31) and make non-avoidance more costly the experimenters again issued a threat: occasionally a punishing stimulus with 'a slightly higher intensity' (Traxler et al., 2020, p.31) would be delivered after an error. This failed to strike sufficient fear into the participants and did not result in more avoidance responses. The authors are surprised and do not offer an explanation, but I suspect the reason may be similar to why the first threat failed. Participants perhaps reasoned that a 'slightly higher' intensity would still be bearable, bearing in mind that the Ethics Commission sets conservative limits on how much pain a participant may be made to bear. Participants did not attend to the task alone but were aware of the boundaries of the experiment (in this case the limits set by an Ethics Commission) and took these into consideration in the way they performed the task.

Above I wrote that, in psychology, the experimental situation and the experimenter must be forceful yet unobtrusive. It seems to me that, in as much as the pilot failed to elicit the right number of avoidance responses and at the same time increase their cost, it was due to the situation becoming too prominent relative to the task. The experimenters' threats made the participants attend to the experimental situation as such, as well as to the task within it, and their responses became colored by their thoughts about the boundaries of the situation they were in. It is not clear whether the researchers asked the participants about their experiences with the threats, as they said they did regarding other aspects of the task. It would have been interesting to learn the participants' reflections and see whether they confirm mine.

Ecological validity

Discussing the results of their pilot, the authors argue that the low-cost avoidance responses that they ended up with are nonetheless clinically relevant, since that kind of avoidance behavior, like always carrying pain medication, occurs in real life too. Given that they also emphasize the 'considerable costs' (Traxler et al., 2020, p.34) of avoidance behavior in real life this argument seems unconvincing. I am sure they would have preferred high-cost avoidance in their experiment. Perhaps this is an inherent limitation of studying this phenomenon experimentally. First, for ethical reasons it is difficult to make responses in an experiment really costly. Many if not most participants will realize this and understand that the cost they are threatened with is illusory, as indeed it was here. Secondly, if the cost of avoidance in real life concerns people's 'social life, physical functioning, and personal well-being', as the authors write (Traxler et al., 2020, p.34), an experiment may simply be too limited a situation; too small and brief to cause that kind of cost. The authors themselves note that 'the operationalization of avoidance behavior as a single button press may be considered simplistic.' (Traxler et al., 2020, p.34) But it is, I think, not just the simplicity of a button press compared to the complexity and variety of real-life avoidance that is the problem. It is also the brief duration and narrow spatial boundaries of an experiment. If I may be permitted one remark on the topic of study itself: It seems to me that the cost of avoidance behavior in real life builds up gradually over time, in a variety of situations. It has a history, one that is longer than the duration of an experiment and that plays out in more locations than a single laboratory. The failure of this

pilot study to elicit high-cost avoidance behavior may well be due to the fact that an experiment is simply too limited an event to successfully emulate the development of pathological avoidance behavior in real life. As became clear in this pilot study, participants know there is an end to the experiment and take it into account in their response strategy. The problem is thus not so much that this particular experiment lacked ecological validity because of the way it was designed, but that *any* experiment would fail to simulate the development of pathological avoidance behavior.

It might be wise to at least complement experimental studies with research that looks at avoidance behavior in a longitudinal perspective, as a phenomenon that develops over a longer period of time, in various situations. The experience sampling method might be useful here: participants regularly (several times a day, for example) report their current state of mind, behavior, or experiences, typically via an app on their mobile phone. Alternatively, qualitative methods such as interviews could allow a more in-depth exploration of pain and avoidance. Such approaches still require a ‘performance’ from the participant and tactful guidance from the researcher, but their relationship will be different from that in an experiment, if only because it lasts longer. More like a friendship, perhaps, than a single dance.

The authors’ tutorial proposes ‘five key elements to consider when developing a new experimental paradigm or merging existing ones’ (Traxler et al., 2020, p.35). Each of these is solid advice, in my opinion, but it might be worthwhile adding a sixth: an experimental paradigm may not be suitable for your topic of study, so be willing to consider other options.

Peer review

JOTE aims to make the peer review process accessible to its readers. Therefore, the initial submission with integrated peer review comments is available [here](#).

References

- Abma, R. (2020). Experiment and fail. *Journal of Trial and Error*, 1(1), 22–27. <https://doi.org/10.36850/r1>
- Brenninkmeijer, J., Derksen, M., & Rietzschel, E. (2019). Informal laboratory practices in psychology. *Collabra: Psychology*, 5(1), 45. <https://doi.org/10.1525/collabra.221>
- Derksen, M. (2001). Discipline, subjectivity and personality: An analysis of the manuals of four psychological tests. *History of the Human Sciences*, 14(1), 25–47. <https://doi.org/10.1177/095269510101400102>
- Derksen, M. (2017). Histories of human engineering: Tact and technology. <https://doi.org/10.1017/9781107414921>
- Lezaun, J. (2007). A market of opinions: The political epistemology of focus groups. *Sociological Review*, 55, 130–151. <https://doi.org/10.1111/j.1467-954X.2007.00733.x>
- Rosenzweig, S. (1933). The experimental situation as a psychological problem. *Psychological Review*, 40(4), 337–354. <https://doi.org/10.1037/h0074916>
- Traxler, J., Philips, R., von Leupoldt, A., & Vlaeyen, J. (2020). Trial and error (-Related negativity): An odyssey of integrating different experimental paradigms. *Journal of Trial and Error*, 1(1), 27–38. <https://doi.org/10.36850/e2>

Companion Piece

Abstract. Pain can be considered as a signal of “bodily error”: Errors – discrepancies between the actual and optimal/targeted state – can put organisms at danger and activate behavioral defensive systems. If the error relates to the body, pain is the warning signal that motivates protective action such as avoidance behavior to safeguard our body’s integrity. Hence, pain shares the functionality of errors. On the neural level, an important error processing component is the error-related negativity (ERN), a negative deflection in the electroencephalographic (EEG) signal generated primarily in the anterior cingulate cortex within 100 ms after error commission. Despite compelling evidence that the ERN plays an important role in the development of various psychopathologies and is implicated in learning and adjustment of behavior, its relation to pain-related avoidance has not yet been examined. Based on findings from anxiety research, it seems conceivable that individuals with elevated ERN amplitudes are more prone to engage in pain-related avoidance behavior, which may, under certain conditions, be a risk factor for developing chronic pain. Consequently, this new line of research promises to contribute to our understanding of human pain. As in most novel research areas, a first crucial step for integrating the scientific fields of ERN and pain is developing a paradigm suited to address the needs from both fields. The present manuscript presents the development and piloting of an experimental task measuring both ERN and avoidance behavior in response to painful mistakes, as well as the challenges encountered herein. A total of 12 participants underwent one of six different task versions. We describe in detail each of these versions, including their results, shortcomings, our solutions, and subsequent steps. Finally, we provide some advice for researchers aiming at developing novel paradigms. Traxler, J., Philips, R., von Leupoldt, A., & Vlaeyen, J. (2020). Trial and error (-Related negativity): An odyssey of integrating different experimental paradigms. *Journal of Trial and Error*, 1(1), 27–38. <https://doi.org/10.36850/e2>

License

Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020