bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Why does the error negativity peak higher when accuracy is emphasised over speed? The impact of noise and single-trial latency variability challenges traditional interpretations

André Mattes, Elisa Porth, Eva Niessen, Kilian Kummer, Markus Mück, Jutta Stahl

Department of Individual Differences and Psychological Assessment, University of Cologne

Author Note

This research was supported by the German Research Foundation (STA 1035/7-1).

Correspondence concerning this article should be addressed to André Mattes, Department of Individual Differences and Psychological Assessment, University of Cologne, Pohligstraße 1, 50969 Köln, Germany. E-Mail: andre.mattes@uni-koeln.de

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

Understanding human error processing is a highly relevant interdisciplinary goal. More than 30 years of research in this field have established the error negativity (Ne) as a fundamental electrophysiological marker of various types of erroneous decisions (e.g. perceptual, economic) and related clinically relevant variations. A common finding is that the Ne is more pronounced when participants are instructed to focus on response accuracy rather than response speed, an observation that has been interpreted as reflecting more thorough error processing. We challenge this wide-spread interpretation by demonstrating that when controlling for the level of non-event-related noise in the participant-average waveform and for single-trial peak latency variability, the significant speed-accuracy difference in the participant-average waveform vanishes. This suggests that the previously reported Ne differences may be mostly attributable to a more precise alignment of neuro-cognitive processes and not (only) to more intense error processing under accuracy instructions, opening up novel perspectives on previous findings.

Keywords: error negativity; event-related potential; single-trial analysis; accuracy instruction; speed instruction

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Why does the error negativity peak higher when accuracy is emphasised over speed? The impact of noise and single-trial latency variability challenges traditional interpretations

There are hardly any tasks that humans perform flawlessly, i.e. without making any errors. To decrease the number of undesired action outcomes, it is important for the cognitive control system to monitor responses and identify incorrect actions. One marker of response monitoring is the error negativity (Ne; Falkenstein et al., 1991, also error-related negativity, ERN; Gehring et al., 1990). The Ne is an event-related potential (ERP) component with a fronto-central distribution characterised by a negative peak up about 100 ms after an erroneous response. Although it is usually more pronounced after errors, it can also be observed after correct responses albeit with a smaller amplitude (Vidal et al., 2000). The Ne is thought to reflect early error processing (Falkenstein et al., 1991; Gehring et al., 1990) independent from actual error awareness (Wessel, 2012). One of the earliest findings related to the Ne is that its peak amplitude is larger (i.e., more negative) when participants are instructed to respond as accurately as possible compared to when the instruction emphasises speed over accuracy (Gehring et al., 1993). This finding has been replicated numerous times (e.g. Endrass et al., 2012; Riesel et al., 2019; Themanson et al., 2008; Themanson et al., 2011). Common explanations suggest that the error significance in the accuracy condition is increased compared to the speed condition, prompting participants to process errors more deeply when accuracy is emphasised (Gehring et al., 1993; Hajcak et al., 2005). A more thorough processing of errors may in turn enhance the participants chances of giving more accurate responses in the future (Wessel, 2018). Here, we present two additional explanations for Ne peak amplitude difference under accuracy and speed instructions that have found little consideration so far: (1) the level of noise and (2) single-trial latency variability.

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Non-event-related noise and averaging

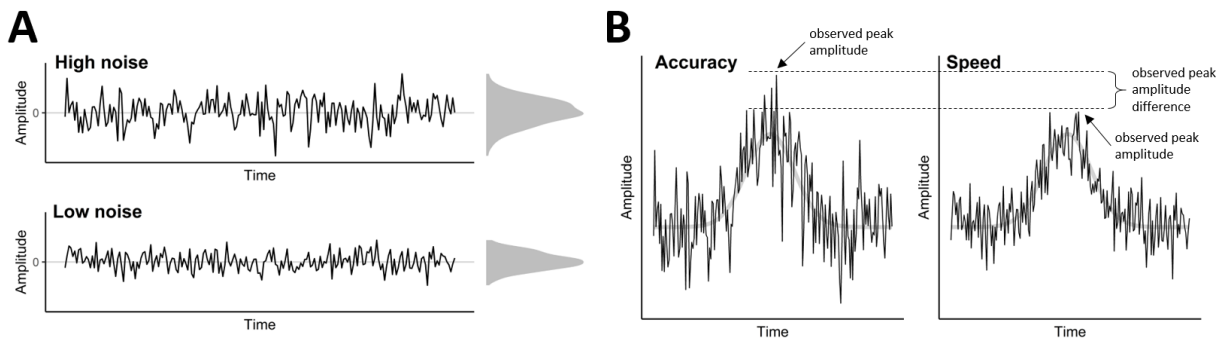
The difference in Ne peak amplitude may be partially attributable to differences in the level of noise included in the averaged waveform. In this context, by noise, we mean random, unsystematic and non-event-related noise which is normally distributed with a mean of zero and a standard deviation reflecting the magnitude of the noise (Figure 1A). This noise is superimposed on the true waveform (i.e. the systematic, event-related signal). The observed waveform thus reflects the sum of the signal and the noise. The principle aim of building a participant-average waveform is to reduce this noise (and thus to increase the signal-to-noise ratio): The more trials go into a participant-average waveform, the less noisy it is. When accuracy is emphasised over speed, participants usually make less errors (e.g. Gehring et al., 1993; Wickelgren, 1977). Hence, the participant-average waveform should be based on less trials than when speed is emphasised, and the remaining noise is larger (Figure 1B).

Often, the peak amplitude is defined as the most positive or most negative point in a certain time interval. This measure not only reflects differences in the true waveform between experimental conditions, but is also highly influenced by different levels of noise (e.g. Clayson et al., 2013; Luck, 2014). For example, in Figure 1B, we imposed different levels of random noise on the same true waveform. The peak amplitude in the left panel is much higher than the peak amplitude in the right panel. However, this difference merely reflects different numbers of trials and thus different levels of noise: A lower number of errors in the accuracy condition results in a noisier average waveform and a higher peak amplitude. A higher number of errors in the speed condition results in a less noisy average waveform and thus a lower peak amplitude because the random noise is averaged out.

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 1

The Impact of Different Noise Levels on the Ne Peak Amplitude



Note. A) Different levels of noise after averaging a different number of trials. Both levels have a mean of zero but the standard deviation of the distribution of high noise is larger than the standard deviation of the distribution of low noise. The noise distributions are shown to the right of the plots. B) When the noise is superimposed on the true waveform (light grey), the peak is higher when noise is high than when noise is low.

Effect of single-trial latency jitter on the peak amplitude

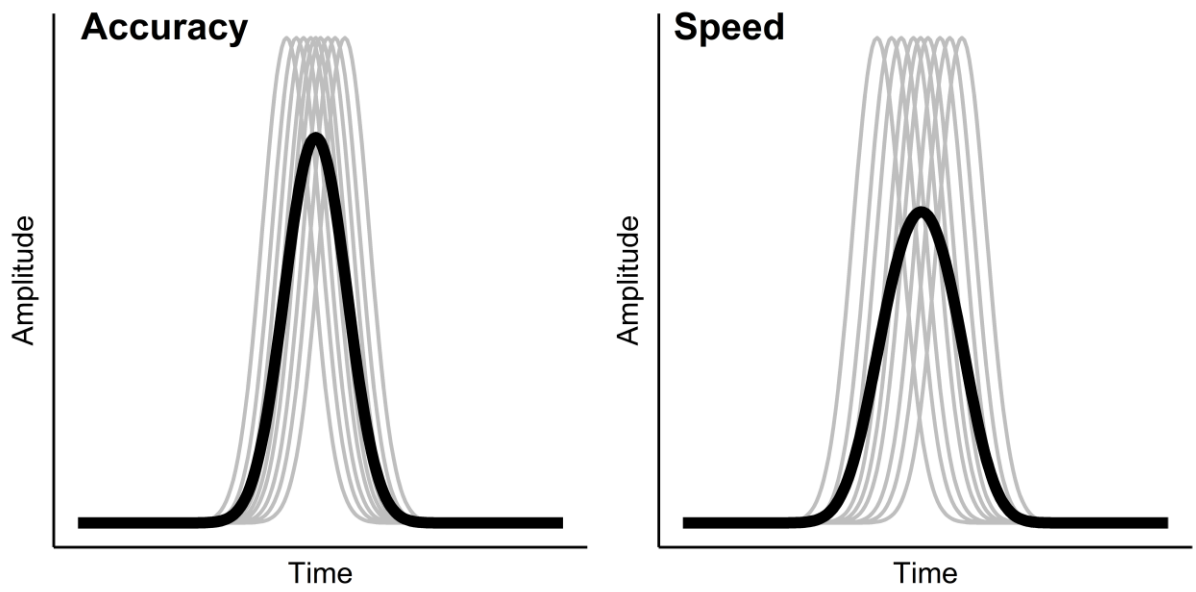
Differences in the peak amplitudes in the participant-average waveforms may also be explained by differences in single-trial latency variability. Peak amplitudes are usually quantified in the participant-average waveform in ERP research because single-trial waveforms are too noisy to reliably detect a peak (e.g. Luck, 2014). This approach, however, implicitly assumes that the single-trial waveforms all peak at the same point in time (Kappenman & Luck, 2012). In Figure 2, we illustrate what happens when this assumption is not met. If the single-trial peak latency variability is large, the average of the single-trial waveforms will have a smaller amplitude than when the latency variability is small. Hence, the observed Ne peak amplitude difference in the participant-average waveforms between the speed and the accuracy condition may also be a result of different levels of single-trial peak latency variability. Specifically, the single-trial waveforms should vary more in their peak latencies in the speed condition than in the

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

accuracy condition, resulting in a flatter participant-average waveform and thus producing a smaller peak amplitude in the average waveform for the speed condition (high latency variability) compared to the accuracy condition (low latency variability).

Figure 2

The Impact of Single-Trial Latency Variability on the Participant-Average Waveform



Note. The grey lines indicate single-trial waveforms, the black lines result from averaging the single-trial waveforms. In the left panel, the single-trial waveforms peak around the same point in time, hence the variability in single-trial peak latencies is small and the resulting average waveform has a comparatively high peak amplitude. In the right panel, the variability in single-trial latencies is much larger and as a consequence, the resulting average waveform has a lower peak amplitude.

Objectives

To sum up, we suspect that parts of the difference in the Ne peak amplitude between the accuracy and speed instruction assessed in the participant-average waveform can be explained by difference in the noise level and the single-trial variability in both conditions. Specifically, as we

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

outlined above, we expect that the level of noise is larger in the accuracy condition than in the speed condition, and that the extent of single-trial latency variability is larger in the speed condition than in the accuracy condition. As a consequence, we hypothesize that that the Ne amplitude is negatively related to the level of noise (the higher the noise, the more negative the Ne peak amplitude) and positively related with the extent of single-trial latency variability (the higher the latency variability, the less negative the Ne peak amplitude). (Note that the Ne has a negative peak such that a negative relationship between the Ne and another measure means that higher levels in this measure are accompanied by a more pronounced Ne.) We predict that when controlling for the level of noise and the extent of latency variability, the difference in the Ne peak amplitude between the two instructions should become smaller or even disappear altogether. To test our hypotheses, we analysed existing data of an experiment in which participants received a speed instruction in the first part of the experiment and an accuracy instruction in the second part of the experiment (Bode & Stahl, 2014; Kummer et al., 2020; Mattes et al., 2022) and resorted to a single-trial peak estimation technique to assess latency variability.

Results and Discussion

Testing Prerequisites

First, we ensured that the instruction had the expected effects in terms of response time and number of error trials. Responses were faster in the speed (M = 248 ms, SD = 101) than in the accuracy condition (M = 517 ms, SD = 116), $t(51) = 13.25$, $p < .001$, $d_z = 2.47$, and there was a smaller number of errors when accuracy (M = 23.71, SD = 18.35) was emphasised over speed (M = 67.40, SD = 28.03), $t(51) = 10.65$, $p < .001$, $d_z = 1.82$. Next, we examined our assumption that the averaged Ne peak amplitude was indeed more pronounced in the accuracy

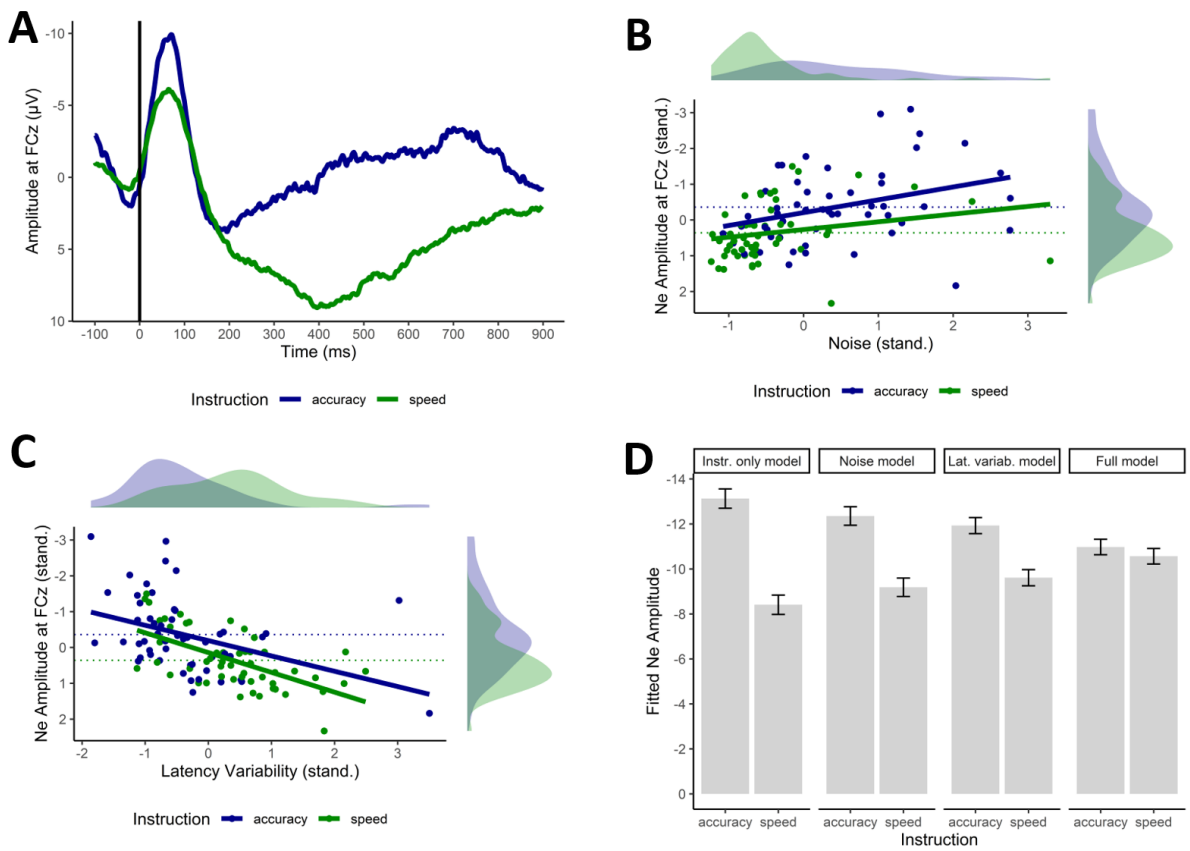
bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(M = -13.13 µV, SD = 6.91) than in the speed condition (M = -8.41 µV, SD = 5.28; see also Figure 3A), $t(51) = 5.16$, $p < .001$, $d_z = 0.76$. As expected, we also found that the level of noise was higher in the accuracy (M = 6.27, SD = 3.06) than in the speed condition (M = 3.56, SD = 2.60; see also the density plot at the top of Figure 3B), $t(51) = 5.02$, $p < .001$, $d_z = 0.95$. Further, the extend of single-trial latency variability was higher in the speed (M = 34.42, SD = 7.59) than in the accuracy condition (M = 27.65, SD = 8.33; see also the density plot at the top of Figure 3C), $t(51) = 5.40$, $p < .001$, $d_z = 0.85$. Hence, the instruction had the desired effect in terms of response time and number of error trials, and we replicated the commonly observed pattern of larger Ne peak amplitudes for accuracy vs. speed instructions. Importantly, we were able to demonstrate that the accuracy and speed instruction differed indeed in their level of noise and the extent of latency variability.

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 3

Illustration of the Results



Note. A) Grand-average waveform of the speed and accuracy condition at the electrode FCz. B) Relationship between the level of standardised noise (x-axis) and the standardised Ne amplitude (y-axis) for the accuracy and speed instruction. C) Relationship between the extent of single-trial latency variability (x-axis) and the Ne amplitude (y-axis) for the accuracy and speed instruction. D) The difference of the Ne amplitude between the accuracy and speed instruction when only considering the instruction ("Instruction only model"), when controlling for the level of noise ("Noise model"), when controlling for the extent of latency variability ("Lat. variab. model"), and when controlling for both noise and latency variability ("Full model").

Identifying the sources of speed-accuracy effects on Ne

Next, we tested whether these differences were able to explain the difference in peak amplitude quantified in the participant averages. To this end, we computed a series of mixed

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

models predicting the standardised Ne amplitude. The models are summarised in Table 1. In the first model, we predicted the Ne peak amplitude of the averaged waveform by the instruction only and found that the instruction significantly predicted the Ne peak amplitude, $b = -0.72$, $p < .001$. This analysis is similar to the *t*-test reported in the previous paragraph. Next, we added the level of noise to the mixed model. The level of noise significantly predicted the Ne peak amplitude, $b = -0.27$, $p = .002$. The higher the noise is, the higher (i.e., more negative) the Ne peak amplitude is (Figure 3B). The Ne peak amplitude still differed significantly between the speed and the accuracy instruction, although to a smaller extent, $b = -0.48$, $p = .002$. This finding lends initial support to our assumption that parts of the differences in the Ne amplitude between the instructions can be explained by the level of noise.

Table 1

Summary of the mixed models

Predictors	Model 1					Model 2					Model 3					Model 4				
	b	*SE*	*t*	*df*	*p*	*b*	*SE*	*t*	*df*	*p*	*b*	*SE*	*t*	*df*	*p*	*b*	*SE*	*t*	*df*	*p*
(Intercept)	0.00	0.11	0.00	50	>.999	0.00	0.11	0.00	49	>.999	0.00	0.10	0.00	49	>.999	0.00	0.09	0.00	48	>.999
Instruction	-0.72	0.14	-5.16	50	<.001	-0.48	0.16	-3.11	49	.002	-0.35	0.14	-2.46	49	.014	-0.06	0.15	-0.41	48	.680
Noise (z)						-0.27	0.09	-3.03	49	.002						-0.31	0.08	-4.00	48	<.001
Latency variability (z)											0.47	0.09	5.31	49	<.001	0.49	0.08	5.99	48	<.001

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

In the third model, we removed the level of noise as a predictor and entered the extent of single-trial latency variability. As expected, the latency variability positively predicted the Ne amplitude, $b = 0.47$, $p < .001$. The more the single-trial latencies varied, the smaller the Ne amplitude was (Figure 3C). As in the second model, the instruction was still a significant predictor of the Ne amplitude, $b = -0.35$, $p = .014$, although the difference between the instructions was smaller than in the first model, which did not include the latency variability as an additional predictor. This finding provides evidence for our assumption that parts of the differences in the Ne amplitude between the instructions can be explained by the extend of single-trial latency variability.

In the last model, we entered the instruction, the level of noise and the extent of latency variability as predictors. While the level of noise and the extent of latency variability still predicted the Ne amplitude significantly, $b = -0.31$, $p < .001$ and $b = 0.49$, $p < .001$, respectively, the instruction did not reach the level of significance anymore (Figure 3D), $b = -0.06$, $p = .680$. This suggests that the level of noise and the extent of single-trial latency variability were able to explain large parts of the variance in the Ne peak amplitude that is usually found to differ between speed and accuracy instructions. Note however, that the lack of statistical significance of the instruction in the last model does not mean that there is no effect of the instruction on the Ne after controlling for noise and latency variability. The insignificance may also reflect a lack of power to detect an instruction effect which is much smaller than when noise and latency variability are not controlled for ($d_z = 0.06$ vs. $d_z = 0.72$). Importantly, our results show that a higher Ne amplitude for accuracy instructions may not uniquely be interpreted as reflecting a deeper processing of errors due to an increased error significance (e.g. Gehring et al., 1993;

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Hajcak et al., 2005; Themanson et al., 2011). It may even be that error processing is just as deep for accuracy instructions as for speed instructions (Figure 3D).

Level of noise

The level of noise is most likely an artefact due to averaging different number of trials. To avoid artificial differences between the instructions that are uniquely driven by the level of noise, other peak measures which are not as vulnerable to unsystematic noise may be used. For example, when using the mean peak amplitude as the dependent variable in the first model, the effect size is descriptively smaller (peak amplitude: d_z = -0.72; mean peak amplitude (peak ± 20 data points): d_z = -0.51), indicating that the averaging across neighbouring data points reduced the noise in the peak measure. Clayson et al. (2013) have discussed the impact of noise on ERP measures extensively, so we will mostly focus on the effects of single-trial latency variability in the following. Unlike the level of noise – which we assume is unsystematic and exclusively due to the different numbers of trials – latency variability potentially provides insights into differences in psychological processes and a different interpretation of the observed speed-accuracy Ne effect.

Latency variability and process alignment

When accuracy is emphasised over speed, the cognitive processes involved in decision-making (perception, response selection, response execution, response monitoring, etc.) might be more “aligned”, perhaps because time restraints play a less important role. In this context, alignment means that *across trials*, the same processes occur at the same point in time. Hence at the time of response onset, processes related to perception and response selection are largely finished and response monitoring is executed. This alignment ultimately produces less peak latency variability at the single-trial level and a higher Ne amplitude in the accuracy condition

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

compared to the speed condition. When speed is emphasised, the processes are less aligned across trials, resulting more single-trial latency jittering and a smaller peak amplitude of the average waveform. Hence, different processes might be executed at response onset on each trial. As a consequence, the average waveform is flatter due to jittering in the single-trial peak latencies.

These analyses provide a different way of interpreting literature findings that build on the instruction effect. For example, Riesel et al. (2019) found that patients with obsessive compulsive disorder (OCD) showed a much smaller difference in the Ne amplitude between the speed and accuracy condition than healthy controls, suggesting that patients with OCD have difficulty adapting a response strategy that does not focus on accuracy. Our findings do not change the conclusions Riesel et al. draw from their data, but postulate different underlying mechanisms. Assuming that the same latency effect occurs in their data, we would postulate that while healthy controls seem to disengage from aligning cognitive processes in favour of faster responses when speed is stressed, patients with OCD may have difficulty in doing so. In another study, Endrass et al. (2012) found that the Ne amplitude was reduced in older adults compared to younger adults, especially under speed instructions. Hence, for older adults, the speed instruction might disrupt the alignment of cognitive processes even more than for younger adults. These examples demonstrate that speed and accuracy instructions are still useful to study error processing. However, Ne differences may not (only) reflect more thorough error processing or more error significance, but differences in the level of noise and single-trial latency variability. While the level of noise is probably a statistical artefact, the single-trial latency variability may be an alternative mechanism that is helpful in interpreting and understanding differences in the Ne amplitude between speed and accuracy instructions.

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Conclusion

Using single-trial ERP analyses, we could show that substantial parts of the well-known difference in Ne amplitude between speed and accuracy instructions can be explained by differences in single-trial latency variability. This finding shifts the interpretation of the Ne amplitude difference away from processing intensity and neural capacity, and towards temporal dynamics and the alignment of cognitive processes. Our method can easily be applied to research questions in other domains in which the origins of ERP amplitude differences are unknown and may pave the way to novel insights into cognitive mechanisms reflected by ERP amplitude differences.

Method

Participants

The study was approved by the ethics committee of the German Psychological Society (DGPs) and was conducted according to the Declaration of Helsinki. In the original study, the order of the speed and accuracy instruction was counterbalanced. As the order effect was not relevant for the current investigation, we only analysed the data of participants who completed the speed instruction first and the accuracy instruction second, which was the order with the higher number of participants. Of the 61 participants who completed the task in this order, 9 participants yielded less than six error trials in any of the speed or accuracy conditions (for the most part in the accuracy condition) and were excluded from the analyses (see Electrophysiological Data). The remaining 52 participants (28 male, 24 female) had a mean age of 25.40 years (*SD* = 5.54). All participants gave written informed consent prior to the study.

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Procedure and Experimental Task

Participants performed a modified version of the Flanker Task (Eriksen & Eriksen, 1974) in which participants had to determine whether a target digit (out of the numbers 1 to 8) was odd or even by pressing a key with one of the index fingers. The assignment of the odd/even responses to the left/right index fingers was counterbalanced across participants. At the left and right of the target digit, two identical distractor digits were presented. The stimulus comprising the three digits was presented for 67 ms and was followed by a response window of 1133 ms. Subsequently, a feedback was presented for 700 ms, followed by an inter-trial interval of 1500 ms. In the accuracy condition, the feedback informed participants whether the response was correct or erroneous. In the speed condition, the feedback additionally informed participants that the response was too slow when the response time limit was exceeded. The response time limits were set to 850 ms (accuracy condition) and 85 % of the average response time of the practice block (speed condition). Responses that were given after this response time limit but within the response window were recorded as “too slow”. Participants first completed 10 blocks of 40 trials each in which they were instructed to respond as fast as possible. Next, they completed another 10 blocks in which they were instructed to respond as accurately as possible. More information on the task can be found in Bode and Stahl (2014).

Electrophysiological Data

The EEG was recorded from 61 scalp electrode sites (Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, C3’, Cz, C4, C4’, T8, FPz, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, FCz, O1, Oz, O2, AF7, AF3, AF4, AF8, F5, F1, F2, F6, FT7, FC3, FC4, FT8, C5, IZ, C6, TP7, CP3,CPz, CP4, TP8, P5, P1, P2, P6, PO7, PO3, POz, PO4, PO8) with active Ag/AgCl electrodes (actiCAP, BrainProducts) which were referenced against the left mastoid. A DC converter

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(BrainAmp DC, BrainProducts) was used to digitise the continuous EEG signal recorded with a sampling rate of 500 Hz. An online band-pass filter (DC-70 Hz) was applied to the signal. The electrooculogram (EOG) was recorded from two electrodes placed at the outer sides of the eyes (horizontal EOG) and two electrodes placed above and below the left eye (vertical EOG).

The EEG data were preprocessed in the *BrainVision Analyzer* software (BrainProducts). First, a DC detrend was applied to the data. Next, the data were stimulus-locked (-100 ms to 2100 ms) and baseline-corrected. An ocular correction was performed (Gratton et al., 1983) and segments in which the signal exceeded ± 150 µV were removed. The stimulus-locked segments were used for the noise estimation (see next section). For the main analyses, the segments were response-locked (-100 ms to 900 ms) and baseline-corrected. After averaging the wavelines within participants, the Ne was defined as the most negative point at the FCz electrode site in the interval from response onset to 150 ms following the response. Only participants who had at least six trials in each of the conditions were considered to ensure a reliable quantification of the Ne (Olvet & Hajcak, 2009).

Noise Estimation

The remaining noise after averaging a given number of trials was estimated as the standard deviation of the participant-averaged EEG signal at the FCz electrode in the 100 ms pre-stimulus baseline. We chose the pre-stimulus baseline over the pre-response baseline because we cannot rule out that systematic processes such as response preparation or even response monitoring are already ongoing before a response is carried out (Bode & Stahl, 2014). Unlike the pre-response baseline, we assume that no systematic response-related processes are present in the pre-stimulus baseline and that variation in this interval reflects random noise. The stimulus-

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

locked epochs were baseline-corrected such that the ERP signal of the pre-stimulus interval in which the noise was determined had a mean of zero.

Latency Variability Estimation

We used a procedure introduced by Hu et al. (2010; 2011) to quantify peak latencies in single-trial waveforms and computed the standard deviation of single-trial latencies for each participant as an indicator of single-trial latency variability. The single-trial estimation was performed in MATLAB (MathWorks) using the *Letswave 6* toolbox (Moureaux et al., 2016).

In the following, we briefly describe the single-trial estimation procedure and refer to Hu et al. (2010; 2011) for a detailed description. The procedure consists of two major steps: (1) applying a wavelet filter to denoise the data and (2) modelling the filtered single-trial waveforms in a multiple linear regression approach. In the first step, a wavelet filter is computed by transforming each single-trial waveform from the time domain to the frequency domain and averaging all time-frequency transforms. The resulting average time-frequency transform indicates in which frequencies and at which time points there is stable activation across trials and serves as the basis for the wavelet filter. All single-trial signals are time-frequency transformed, the wavelet filter is applied to them, and they are transformed back to the time domain. The filtered signal is then modelled in a multiple linear regression approach by three regressors capturing the overall waveform, the latency, and the morphology (i.e., whether the waveform is wide or narrow). Finally, the peak is determined in the fitted single-trial waveform and the peak amplitude and latency can be quantified. In a simulation study, Hu et al. (2011) have demonstrated that the peak estimates derived from this procedure are accurate and unbiased. The technique has successfully been applied in a multitude of studies (e.g. Franz et al., 2015; Hu et

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

al., 2014; Mattes et al., 2022). The same trials were used for the participant-average waveform, the estimation of single-trial latency variability, and the estimation of noise.

Statistical Analyses

We used within-subject *t*-tests to investigate differences in response time, number of error trials, Ne peak amplitude, level of noise and latency variability between the speed and the accuracy condition. We then computed a series of mixed models predicting the Ne amplitude quantified in the participant-average waveform. The models differed in the predictors that they included. The instruction was entered as a contrast-coded dichotomous predictor (speed: -0.5; accuracy: 0.5). The level of noise and the extent of latency variability were entered as standardised continuous predictors. Furthermore, the Ne peak amplitude was also standardized. This allowed to interpret the regression coefficient of the instruction factor in terms of Cohen's *d* and the regression coefficients of the continuous predictors similar to partial correlations. Participants were included as random effects in the mixed models. We carried out all statistical analyses in R (R Core Team, 2017). We fitted the mixed models to the data using the *lme4* package (Bates et al., 2015) and tested the regression coefficients for significance using the *lmerTest* package (Kuznetsova et al., 2017). Data processing and plotting were done using functions provided by the *tidyverse* package (Wickham et al., 2019). The data and analysis script are available here: https://osf.io/c638e/ [anonymous link for review; will be made public upon acceptance of the manuscript]

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Declarations

Acknowledgements

We are grateful to Lina Beck for supporting us in the literature research.

Competing interests

The authors declare no competing interests.

Data availability

The datasets generated and/or analysed during the current study are available in the OSF repository, https://osf.io/c638e/ [anonymous link for review; will be made public upon acceptance of the manuscript]

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bode, S., & Stahl, J. (2014). Predicting errors from patterns of event-related potentials preceding an overt response. *Biological Psychology*, *103*, 357–369. https://doi.org/10.1016/j.biopsycho.2014.10.002

Clayson, P. E., Baldwin, S. A., & Larson, M. J. (2013). How does noise affect amplitude and latency measurement of event-related potentials (ERPs)? A methodological critique and simulation study. *Psychophysiology*, *50*(2), 174–186. https://doi.org/10.1111/psyp.12001

Endrass, T., Schreiber, M., & Kathmann, N. (2012). Speeding up older adults: Age-effects on error processing in speed and accuracy conditions. *Biological Psychology*, *89*(2), 426–432. https://doi.org/10.1016/j.biopsycho.2011.12.005

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, *78*(6), 447–455. https://doi.org/10.1016/0013-4694(91)90062-9

Franz, M., Nickel, M. M., Ritter, A., Miltner, W. H. R., & Weiss, T. (2015). Somatosensory spatial attention modulates amplitudes, latencies, and latency jitter of laser-evoked brain

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

potentials. *Journal of Neurophysiology*, *113*(7), 2760–2768. https://doi.org/10.1152/jn.00070.2015

Gehring, W. J., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1990). The error-related negativity: An event-related brain potentail accompanying errors. *Psychophysiology*, *27*, 34.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A Neural System for Error Detection and Compensation. *Psychological Science*, *4*(6), 385–390. https://doi.org/10.1111/j.1467-9280.1993.tb00586.x

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468–484. https://doi.org/10.1016/0013-4694(83)90135-9

Hajcak, G., Moser, J. S., Yeung, N., & Simons, R. F. (2005). On the ERN and the significance of errors. *Psychophysiology*, *42*(2), 151–160. https://doi.org/10.1111/j.1469-8986.2005.00270.x

Hu, L., Valentini, E., Zhang, Z. G., Liang, M., & Iannetti, G. D. (2014). The primary somatosensory cortex contributes to the latest part of the cortical response elicited by nociceptive somatosensory stimuli in humans. *NeuroImage*, *84*, 383–393. https://doi.org/10.1016/j.neuroimage.2013.08.057

Hu, L., Liang, M., Moureaux, A., Wise, R. G., Hu, Y., & Iannetti, G. (2011). Taking into account latency, amplitude, and morphology: Improved estimation of single-trial ERPs by wavelet filtering and multiple linear regression. *Journal of Neurophysiology*, *106*(6), 3216–3229. https://doi.org/10.1152/jn.00220.2011

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Hu, L., Moureaux, A., Hu, Y., & Iannetti, G. (2010). A novel approach for enhancing the signal-to-noise ratio and detecting automatically event-related potentials (ERPs) in single trials. *NeuroImage*, *50*(1), 99–111. https://doi.org/10.1016/j.neuroimage.2009.12.010

Kappenman, E. S., & Luck, S. J. (2012). ERP Components: The Ups and Downs of Brainwave Recordings. In E. S. Kappenman & S. J. Luck (Eds.), *The Oxford Handbook of Event-Related Potential Components.* Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0014

Kummer, K., Dummel, S., Bode, S., & Stahl, J. (2020). The gamma model analysis (GMA): Introducing a novel scoring method for the shape of components of the event-related potential. *Journal of Neuroscience Methods*, 108622. https://doi.org/10.1016/j.jneumeth.2020.108622

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). MIT Press.

Mattes, A., Porth, E., & Stahl, J. (2022). Linking Neurophysiological Processes of Action Monitoring to Post-Response Speed-Accuracy Adjustments in a Neuro-Cognitive Diffusion Model. *NeuroImage*, *247*, Article 118798. https://doi.org/10.1016/j.neuroimage.2021.118798

Moureaux, A., Rossin, B., Hu, L., & Iannetti, G. (2016). *Letswave 6 - MATLAB EEG Signal Processing Toolbox*. https://github.com/NOCIONS/letswave6

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Olvet, D. M., & Hajcak, G. (2009). The stability of error-related brain activity with increasing trials. *Psychophysiology*, *46*(5), 957–961. https://doi.org/10.1111/j.1469-8986.2009.00848.x

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria. R Foundation for Statistical Computing. https://www.R-project.org/

Riesel, A., Kathmann, N., & Klawohn, J. (2019). Flexibility of error-monitoring in obsessive-compulsive disorder under speed and accuracy instructions. *Journal of Abnormal Psychology*, *128*(7), 671–677. https://doi.org/10.1037/abn0000463

Themanson, J. R., Pontifex, M. B., & Hillman, C. H. (2008). Fitness and action monitoring: Evidence for improved cognitive flexibility in young adults. *Neuroscience*, *157*(2), 319–328. https://doi.org/10.1016/j.neuroscience.2008.09.014

Themanson, J. R., Pontifex, M. B., Hillman, C. H., & McAuley, E. (2011). The relation of self-efficacy and error-related self-regulation. *International Journal of Psychophysiology*, *80*(1), 1–10. https://doi.org/10.1016/j.ijpsycho.2011.01.005

Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the 'error negativity' specific to errors? *Biological Psychology*, *51*(2-3), 109–128. https://doi.org/10.1016/S0301-0511(99)00032-0

Wessel, J. R. (2012). Error awareness and the error-related negativity: Evaluating the first decade of evidence. *Frontiers in Human Neuroscience*, *6*, 88. https://doi.org/10.3389/fnhum.2012.00088

Wessel, J. R. (2018). An adaptive orienting theory of error processing. *Psychophysiology*, *55*(3). https://doi.org/10.1111/psyp.13041

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.19.504504; this version posted August 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, *41*(1), 67–85. https://doi.org/10.1016/0001-6918(77)90012-9

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., . . . Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686