Erschienen in: Cerebral Cortex. Oxford University Press (OUP). 2023, 33(6), S. 2919-2930
https://doi.org/10.1093/cercor/bhac250

Case-by-case: neural markers of emotion and task stimulus significance

Harald T. Schupp[1,2,*], Karl-Philipp Flösch[1,2], Ursula Kirmse[1]

[1]Department of Psychology, University of Konstanz, 78457 Konstanz, Germany,
[2]Centre for the Advanced Study of Collective Behaviour, University of Konstanz, 78457 Konstanz, Germany
*Corresponding author: Department of Psychology, University of Konstanz, PO Box 36 78457 Konstanz (Germany), Phone: +49 7531-88 2504;
Email: harald.schupp@uni-konstanz.de

The present study assessed the hypothesis that electrophysiological markers of emotional and task stimulus significance can be demonstrated in concert at the level of the individual case. Participants ($n = 18$, 9 females) viewed low and high-arousing pictures selected from behavior systems of sexual reproduction, disease avoidance, and predator fear. Furthermore, to concurrently manipulate task relevance, participants performed an explicit emotion categorization task with either low or high-arousing pictures alternating as target stimuli in separate experimental blocks. Pooled across behavior systems, event-related components sensitive to emotional significance reached statistical significance in 100% of the tests for the early posterior negativity and in 96% of the tests for the late positive potential. Regarding explicit task relevance, the target P3 effect was significant in 96% of the tests. These findings demonstrate that neural markers of stimulus significance driven by emotional picture content and explicit task demands can be assessed at the individual level. Replicating an effect case-after-case provides strong support for an effect common-to-all and may support individual inferences. Contributions of the case-by-case approach to reveal reproducible effects and implications for the development of neural biomarkers for specific affective and cognitive component processes are discussed.

Key words: affect; attention; ERP; EPN; LPP.

Introduction

The series of replication studies in psychological science has cast doubt on the reproducibility of study findings (Open Science Collaboration 2015), and a similar state of affairs has been indicated in the field of neuroscience (Ioannidis 2005; Button et al. 2013; Open Science Collaboration 2015; Poldrack et al. 2017). Many suggestions for improvements have been made including, for instance, larger samples, replication studies, and the pre-registration of experimental protocols (Lindsay 2015; Munafò et al. 2017; Shrout and Rodgers 2018). However, with the group study as the default mode, an empirical regularity is determined with respect to a “hypothetical person,” i.e. the group mean average, thus requiring a leap of faith when making inferences for the individual (Barlow and Nock 2009; Buckholtz and Faigman 2014). An alternative approach, often chosen by the pioneers in psychology, is the case-by-case approach analyzing empirical effects at the level of the individual case (Danziger 1990; Lamiell 1998). Replicating an effect case-after-case demonstrates the generality of the effect (Sidman 1960). Hence, using the individual case as unit of analysis provides a complementary approach to reveal robust and reproducible findings.

Recent studies explored the potential of the case-by-case approach in the domain of affective stimulus evaluation. Every stimulus encountered in the environment is evaluated according to its affective significance and the fast and reliable detection of positive and negative stimuli facilitates adaptive behavior by guiding selective attention processes (Cacioppo et al. 1993; Lang et al. 1997; Öhman et al. 2000). In a first assessment (Schupp and Kirmse 2021), participants passively viewed pictures drawn from behavior systems of predator fear (Study 1), disease avoidance (Study 2), and sexual reproduction (Study 3) while event-related potentials (ERP) were measured. Pooled across stimulus categories, the early posterior negativity (EPN), an ERP component occurring between 150 and 350 ms, was larger for high- than low-arousing stimuli in the majority of participants (92%). Furthermore, the late positive potential (LPP), occurring between 350 and 700 ms over centro-parietal scalp areas, was larger for high- than low-arousing stimuli in almost all participants (98%). Intersubject replication of the effects within (direct replication) and across (systematic replication) these 3 studies provided evidence for generality of these ERP effects of affective stimulus evaluation (cf. Sidman 1960).

While intersubject replication is a powerful tool to demonstrate an empirical regularity, intraindividual replication allows a more thorough assessment of affective stimulus evaluation effects within the individual case. Accordingly, a follow-up study used a within-subject paradigm to assess the consistency of the effect for behavior systems of sexual reproduction, disease avoidance, and predator fear within each individual case (Schupp and Kirmse 2022). A total of 13 out of 16 participants showed significant EPN and LPP tests across all 3 categories of emotion experience. Furthermore, nonsignificant tests in the remaining 3 cases were limited to specific behavior systems. Thus, interpretation of individual cases was facilitated by intraindividual replication allowing to infer a content-specific rather than general lack of effect within these cases.

Building upon these findings, the present study aimed to expand the scope of the case-by-case approach by assessing multiple facets of stimulus significance within the individual case. Specifically, in addition to emotional stimulus significance, humans are tuned to pick out stimuli that relate to current goals and task demands. Numerous group studies have shown that stimuli designated as targets according to selected stimulus features, such as color, form, or semantic category, elicit a larger P3 wave between 300 and 700 ms compared with nontarget stimuli (Johnson 1988; Codispoti et al. 2006). Accordingly, rather than relying on passive viewing, presenting the emotional stimuli within an explicit categorization task may allow to probe effects of implicit emotional and explicit task significance in concert.

The present study assessed the hypothesis that electrophysiological markers of emotional and task-driven stimulus significance can be demonstrated simultaneously at the level of the individual case, i.e. within 1 experimental design. To assess emotional stimulus significance, participants viewed low- and high-arousing pictures selected from behavior systems of sexual reproduction, disease avoidance, and predator fear in separate blocks. To probe effects of explicit stimulus significance, participants performed a categorization task by pressing a button whenever an exemplar from the target category (high or low arousal) was seen. Regarding inter- and intra-subject replication, we predicted significant single-case results for the 3 behavior systems for (i) the modulation of the EPN and LPP by emotional significance and (ii) the modulation of the P3 by task relevance. Further analyses explored the consistency of the effects by analyzing emotional EPN and LPP effects separately for targets and nontargets and the P3 target effect separately for high- and low-arousing stimuli. Finally, the influence of trial number on emotion and task effects was explored by progressively using fewer trials in the single-case analysis.

Materials and methods

Participants

Eighteen healthy volunteers (9 ♂/9 ♀) with a mean age of 20.6 years (SD = 1.9) were recruited on the campus of the University of Konstanz. All participants had normal or corrected-to-normal vision. All participants were healthy at the time of testing and reported no history of neurological or psychiatric disorders. Participants received monetary compensation or course credit for participation. The ethical committee of the University of Konstanz approved the experimental procedure in accordance with the regulations of the Declaration of Helsinki, and all methods were carried out in full compliance with the approved guidelines. All participants provided informed consent and were debriefed after the experiment.

Stimuli

Stimulus materials were selected from behavior systems of sexual reproduction, disease avoidance, and predator fear (Schupp and Kirmse 2021). For each behavior system, stimuli were selected to comprise 10 images either high or low in emotional arousal. Regarding sexual reproduction (Domjan 1994), the high-arousing emotion category comprised pictures showing couples in an explicit erotic posture. The low-arousing control category contained pictures showing couples in a romantic pose, i.e. hugging or kissing. To engage disease avoidance (Marks 1988; Neuberg et al. 2011), high-arousing stimuli were selected to show bleeding or injured and/or deformed human bodies (mutilation and injury). The low-arousing control category comprised pictures showing uninjured humans in neutral poses. Regarding predator fear (Öhman et al. 2000), the high-arousing stimulus category comprised pictures of wild animals in dangerous, threatening poses (e.g. tiger, shark, alligator),whereas the low-arousing control category included pictures of harmless animals in nonthreatening poses (e.g. cat, sheep, lizard). The International Affective Picture System (IAPS, Lang et al. 2008) and public domain sources were used to select pictures that were similar in overall picture composition for high- and low-arousing stimuli. IAPS pictures 4610 and 4658, 3051 and 2200, and 1301 and 1500 show representative exemplars for the 3 behavior systems. To minimize physical stimulus differences between categories, images were standardized with respect to brightness and contrast in the red, green, and blue channel (see Supplementary Material for details). Furthermore, pictures were presented in the original as well as horizontally flipped direction to control for horizontal eye movements, resulting in 40 pictures for each behavior system (20 high- and 20 low-arousing pictures).

Stimuli from the 3 behavior systems were presented in 3 separate experimental blocks. After each experimental block, participants evaluated the stimulus materials according to emotional dimensions of valence and arousal ratings using the Self-Assessment Manikin (Bradley and Lang 1994). Regarding arousal, erotica (M = 5.52, SD = 1.39) compared with romantic pictures (M = 3.18, SD = 1.06), mutilations (M = 6.17, SD = 1.61) compared with neutral people (M = 2.61, SD = 1.00), and threatening (M = 6.01, SD = 1.74) compared with safe animals (M = 2.80, SD = 1.24) were evaluated as more arousing, $t_s(17) \geq 7.8$, $P < 0.001$, Cohen's $d_s \geq 1.8$. Regarding valence, there was no significant difference between erotic (M = 5.82, SD = 1.48) and romantic pictures ($M = 6.14$, $SD = 0.82$), $t(17) = -0.8$, ns. Mutilations ($M = 2.46$, $SD = 0.89$) compared with neutral people ($M = 5.75$, $SD = 0.70$), and threatening ($M = 3.62$, $SD = 0.79$) compared with safe animals ($M = 6.59$, $SD = 0.78$) were rated as more unpleasant, $t_s \leq -10.3$, $P < 0.001$, Cohen's $d \geq 2.4$.

Experimental design

In separate experimental blocks per behavior system, each picture from the picture set consisting of low- and high-arousing stimuli was presented 30 times in a pseudo-randomized order for each participant. This resulted in a total of 1200 trials and 600 presentations for the high- versus low-arousing category per behavior system. No more than 3 consecutive presentations of the high-/low-arousing pictures were allowed, and transition frequencies between picture categories were controlled. Breaks between and within blocks, i.e. after 600 trials, allowed for participant's posture adjustments and rest.

Participants were asked to press a button as quickly and accurately as possible whenever a picture from the low or high arousal target category was seen. The target category was indicated to participants via on-screen instructions and alternated every 200 trials, resulting in 300 target and nontarget trials for high- and low-arousal categories. Before each experimental block, participants were familiarized with the pictures of the 2 stimulus categories.

Pictures were displayed for 117 ms, preceded by a fixation cross shown for 117 ms. The intertrial interval (ITI) varied pseudo-randomly between 733 and 1083 ms ($M = 907$ ms). Participants could respond until the end of the ITI. When making an error, a red X was shown for 150 ms on a black background followed by a black screen for 217 ms before the start of the next trial. The order of experimental blocks was permuted across participants and gender. Electroencephalogram (EEG) data recording lasted ~120 minutes including breaks. After the picture rating following each behavior system block, breaks (~5–8 minutes) allowed for participants' rest, and impedances were checked.

Task performance was high with on average <2% errors. Error rates were not different for erotic ($M = 1.4\%$, $SD = 4.2\%$) compared with romantic pictures ($M = 2.3\%$, $SD = 3.3\%$), and threatening ($M = 1.4\%$, $SD = 2.5\%$) compared with safe animals ($M = 1.8\%$, $SD = 2.4\%$), $t_s(17) = -2.0$ and -1.0, $P > 0.06$. However, less errors occurred for mutilations ($M = 1.0\%$, $SD = 1.3\%$) compared with neutral people ($M = 1.8\%$, $SD = 2.0\%$), $t(17) = -2.5$, $P = 0.022$, Cohen's $d = 0.596$. Furthermore, reaction times were significantly faster for erotica ($M = 393.05$ ms, $SD = 54.12$ ms) compared with romantic pictures ($M = 428.34$ ms, $SD = 56.30$ ms), mutilations ($M = 393.99$ ms, $SD = 43.54$ ms) compared with neutral people ($M = 417.98$ ms, $SD = 44.84$ ms), and threatening ($M = 416.60$ ms, $SD = 49.69$ ms) compared with safe animals ($M = 456.67$ ms, $SD = 57.64$ ms), $t_s(17) < -7.5$, $P < 0.001$, Cohen's $d < -1.7$.

EEG data acquisition and preprocessing

Brain and ocular scalp potentials were measured with a 256-lead hydrocell geodesic sensor net, recorded DC, filtered on-line below 100 Hz, and sampled at 250 Hz using Netstation acquisition software and EGI amplifiers (Electrical Geodesics Inc., Eugene, OR). Electrode impedance was kept below 40 kΩ, as recommended for this type of EEG amplifier by EGI guidelines. Data were recorded continuously using the vertex sensor as a reference electrode and offline filtered using a digital low-pass filter with a half-power cut-off at 40 Hz (Butterworth IIR filter, order 19, stopband: −45 dB at 50 Hz) and a digital high-pass filter with a half-power cut-off at 0.06 Hz (Butterworth IIR filter, order 4, stopband: −18 dB at 0.05 Hz). The data were then corrected for ocular artifacts based on a multiple regression method, converted to an average reference, and baseline-adjusted (100 ms pre-stimulus) using EMEGS software (Peyk et al. 2011). Artifact rejection was performed based on an elaborate method for statistical control of artifacts, specifically tailored for the analyses of dense sensor EEG recordings (Junghöfer et al. 2000). To ensure equal representations of high- and low-arousing trials in the bootstrap analyses, trial numbers after artifact rejection were equated between conditions by randomly dropping trials from the condition with more trials. On average, bootstrap calculations were run with $M = 472$ ($SD = 58$) trials per condition for the analysis of the EPN and LPP emotion effects and with $M = 467$ ($SD = 59$) for the analysis of the P3 task effect.

Finally, EEG signal quality was determined visually by comparing the measured ERP waveform for each condition and sensor cluster of interest (see below) with respect to the "(±) reference" ERP. This procedure removes the event-related signal from the waveform by alternating the polarity of every second trial before averaging, thereby showing the magnitude of noise that is superimposed on the ERP calculated the standard way (Schimmel 1967). In addition, signal quality was quantitatively assessed by calculating signal-to-noise confidence intervals (see Parks et al. 2016, for details). Specifically, a bootstrap procedure resampled each participant's signal-to-noise ratio, determined as the ratio of root mean square post- (1 to 750 ms) versus pre-stimulus (−200 to 0 ms) activity expressed logarithmically in decibels, using the number of trials used in the main analyses for this subject. A total of 90% lower and higher confidence intervals of the signal-to-noise ratios were calculated and a minimum threshold of 3 dB was defined for acceptable signal quality (cf. Parks et al. 2016). The lower bound of the signal-to-noise confidence interval of the individual cases in the 3 behavioral system conditions ranged from 9.6 to 25.3 ($M = 17.6$, $SD = 3.4$) for the EPN cluster, from 9.2 to 23.7 ($M = 15.1$, $SD = 3.1$) for the LPP cluster, and from 7.9 to 23.8 ($M = 15.9$, $SD = 3.6$) for the P3 cluster. Thus, for each behavioral system and the 3 ERP clusters, signal-to-noise ratio exceeded the 3 dB minimum threshold of signal quality for each individual case (Parks et al. 2016).

ERP scoring

Sensor clusters

EPN and LPP clusters were a priori defined based on previous research (Schupp and Kirmse 2021, 2022). Accordingly, the EPN was scored as mean activity in an occipitoparietal sensor cluster comprising the following sensors (Supplementary Fig. S1): 106, 107, 108, 113, 114, 115, 116, 117, 121, 122, 123, 124, 125, 126, 133, 134, 135, 136, 137, 138, 139, 145, 146, 147, 148, 149, 150, 151, 156, 157, 158, 159, 160, 165, 166, 167, 168, 169, 174, 175, and 176. The LPP was scored in in a central cluster including the following sensors: 6, 7, 8, 9, 15, 16, 17, 23, 24, 30, 42, 43, 44, 45, 51, 52, 53, 59, 60, 79, 80, 81, 89, 130, 131, 132, 143, 144, 155, 183, 184, 185, 186, 196, 197, 198, 206, 207, 215, and 257. For the P3 component, findings from group analysis (see Supplementary Material) were used to determine the sensor cluster of interest. The P3 was scored as mean activity in a parietal sensor cluster comprising the following sensors: 78, 79, 80, 81, 87, 88, 89, 90, 98, 99, 100, 101, 109, 110, 118, 119, 127, 128, 129, 130, 131, 140, 141, 142, 143, 152, 153, 154.

Latency of the components

For single case analysis, the time window for each of the ERP components was allowed to vary between participants and emotion categories to acknowledge interindividual variability in functional brain organization. Time restrictions were a priori defined based on previous research for the emotion effects (EPN and LPP) and based on group analysis of the present data for the P3 (see Supplementary Material). Specifically, it was defined that an effect for the EPN should appear between 150 and 350 ms and for the LPP and P3 components between 350 and 750 ms. Within these temporal restrictions, a custom software determined for each individual case the time window showing the maximal average difference (EPN: negative, spanning 60 ms; LPP/P3: positive, spanning 100 ms) for high- minus low-arousing pictures (EPN/LPP) or targets minus nontargets (P3).

Single subject bootstrap analysis

EPN, LPP, and P3 data from each individual case were submitted to a bootstrap analysis (Efron and Tibshirani 1993; Di Nocera and Ferlazzo 2000; Rosenfeld 2020). Specifically, with 50.000 bootstrap repetitions, each case's EPN, LPP, and P3 mean data were resampled by randomly (re-)assigning a case's trials to the categories of interest (i.e. high- vs. low-arousing for emotion effects and target vs. nontarget for the task effect) with replacement. The mean difference between the resulting ERPs was calculated for each bootstrap run. Significance ($P < 0.05$, 1-sided) on the individual case level was determined as the proportion of results in the empirical probability distribution that were equal or more extreme than the de facto measured EPN/LPP difference. P-values of $\leq$0.05, $\leq$0.01, $\leq$0.001, and <0.00002, respectively, indicate that less than or equal to 2.500, 500, 50, and 0 out of 50.000 randomized calculations yielded an equal or more extreme result.

Results

Emotional stimulus significance: the EPN effect

The emotional modulation of the EPN is shown for each case in Fig. 1A, separately for each behavior system, as a difference scalp map (high–low arousal). For comparability, a ± 3 μV scale has been chosen for display, resulting in the truncation of effects for individuals with large amplitude differences. For all participants a prototypical pattern of emotional modulation of the EPN was indicated. However, the display shows considerable variation between individuals.

Moreover, the amplitude difference of the EPN effect was larger in many cases for sexual reproduction compared with disease avoidance and predator fear.

Findings from single-subject bootstrap analysis are displayed in Fig. 1B. The figure shows a representation of each case's bootstrap distribution (whiskers indicate the 5th and 95th percentiles, thus representing $P \leq 0.05$, 1-sided), the actual observed amplitude difference and information about the level of significance. For all 3 behavior systems, all cases (100%) showed significantly larger EPNs to high- than low-arousing stimuli. Findings at the group level are provided in the Supplementary Material.

Emotional stimulus significance: the LPP effect

Similar to the EPN, consistent effects were observed for the LPP at the level of the individual case. As shown in Fig. 2A, most participants showed a prototypical pattern of emotional modulation of the LPP, with considerable variation between individuals. Furthermore, in most participants, smaller LPP amplitude differences were observed in the disease avoidance and predator fear condition than in the sexual reproduction condition.

Results of the single-subject bootstrap analysis for the LPP are shown in Fig. 2B. For the behavior domains of sexual reproduction and disease avoidance, all cases (100%) showed significantly larger LPP amplitudes to high- than low-arousing stimuli. For predator fear, 16 out of 18 cases (89%) showed a significant LPP effect for threatening versus harmless animals.

Task stimulus significance: the P3 target effect

As shown in Fig. 3A, the P3 target effect was observed at the individual level in each of the 3 emotional behavior systems. Statistical analysis corroborated this impression (Fig. 3B). In 17 out of 18 participants (94%), a significantly larger P3 to target than nontarget trials was observed in each of the 3 behavior systems. The remaining participant (case #2) showed the P3 target effect only for predator fear, but not for sexual reproduction and disease avoidance.

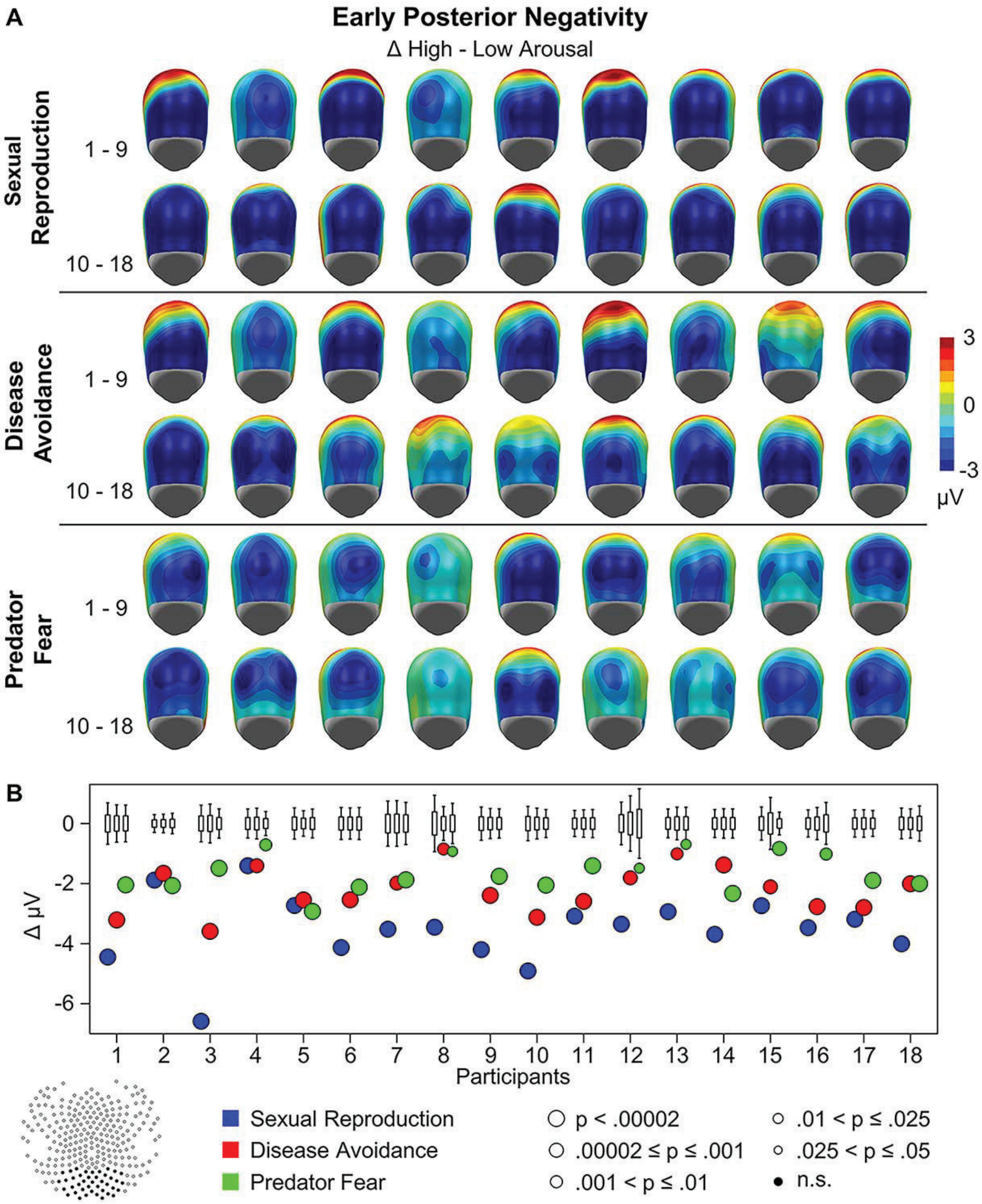

Fig. 1. A) Emotional modulation of the EPN for each individual case and behavior system. Please note the common scale (+/−3 μV), which led to truncation of effects in individuals showing large effects. B) Outcome of the case-by-case statistics for the emotional modulation of the EPN. For each case, a boxplot displays the bootstrap distribution with the bottom and top edges of the box indicating the 25th and 75th percentiles and the whiskers indicating the 5th and 95th percentiles. The dots indicate the de facto measured amplitude difference (high−low arousal). Dots outside the range indicated by the whiskers represent significant effects ($P < 0.05$) with size and color indicating different P-levels. The sensor cluster used to score the EPN is illustrated on the bottom left.

Consistency analysis

Consistency of the emotion and task effects was further assessed by the separate analysis of targets and nontarget stimuli for the EPN and LPP effect (cf. Supplementary Fig. S3) and of low- and high-arousing stimuli for the P3 target effect (cf. Supplementary Fig. S4).

When analyzing only target stimuli, significant EPN results were observed for all 18 cases for sexual reproduction (100%), for 17 cases for disease avoidance (94%), and for 16 cases for predator fear (89%). Analyzing only nontargets, significant EPN results were observed for all 18 cases for sexual reproduction (100%), for 16 cases for

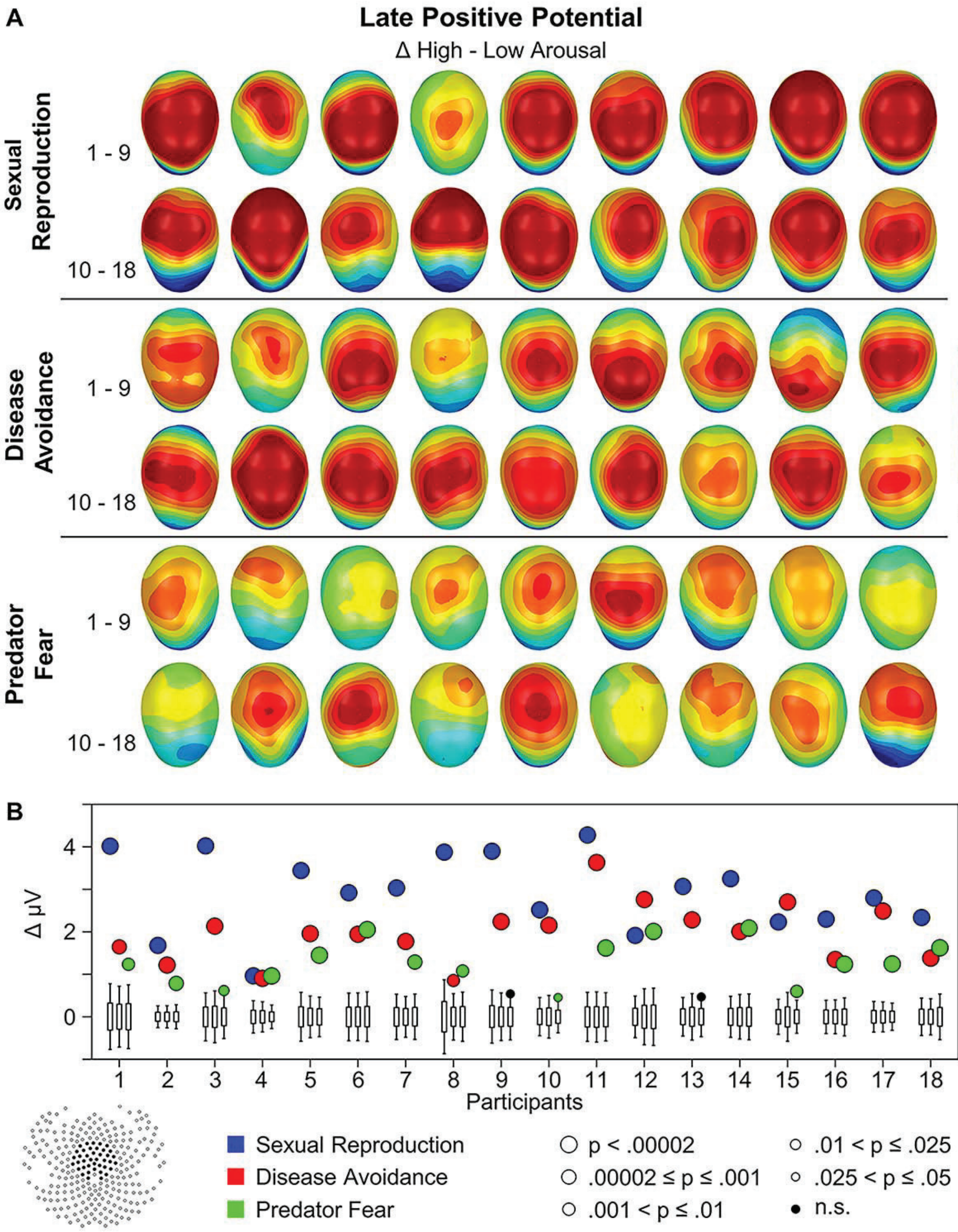

Fig. 2. Emotional modulation of the LPP component for each individual case and behavior system with respect to scalp difference maps (A) and outcome of the single subject bootstrap analysis (B). The dots indicate the de facto measured amplitude difference (high—low arousal).

disease avoidance (89%), and for 15 cases for predator fear (83%).

For the LPP, significant effects for target stimuli were observed in all 18 cases for sexual reproduction (100%) as well as disease avoidance (100%), and in 15 out of 18 cases for predator fear (83%). For nontarget stimuli, the LPP effect was significant in 17 cases for sexual reproduction (94%), in all 18 cases for disease avoidance (100%), and in 14 cases for predator fear (77%). The somewhat lower proportion of significant effects for the predator fear condition presumably reflects the smaller amplitude of the EPN and LPP effects for this behavior system.

Significant bootstrap results for the target P3 were observed for 16 (89%), 17 (94%), and 16 (89%) cases for high-arousing stimuli and for 16 (89%), 18 (100%), and 18 (100%) cases for low-arousing stimuli in the behavior domains of sexual reproduction, disease avoidance, and predator fear, respectively. Case #2 showed no significant

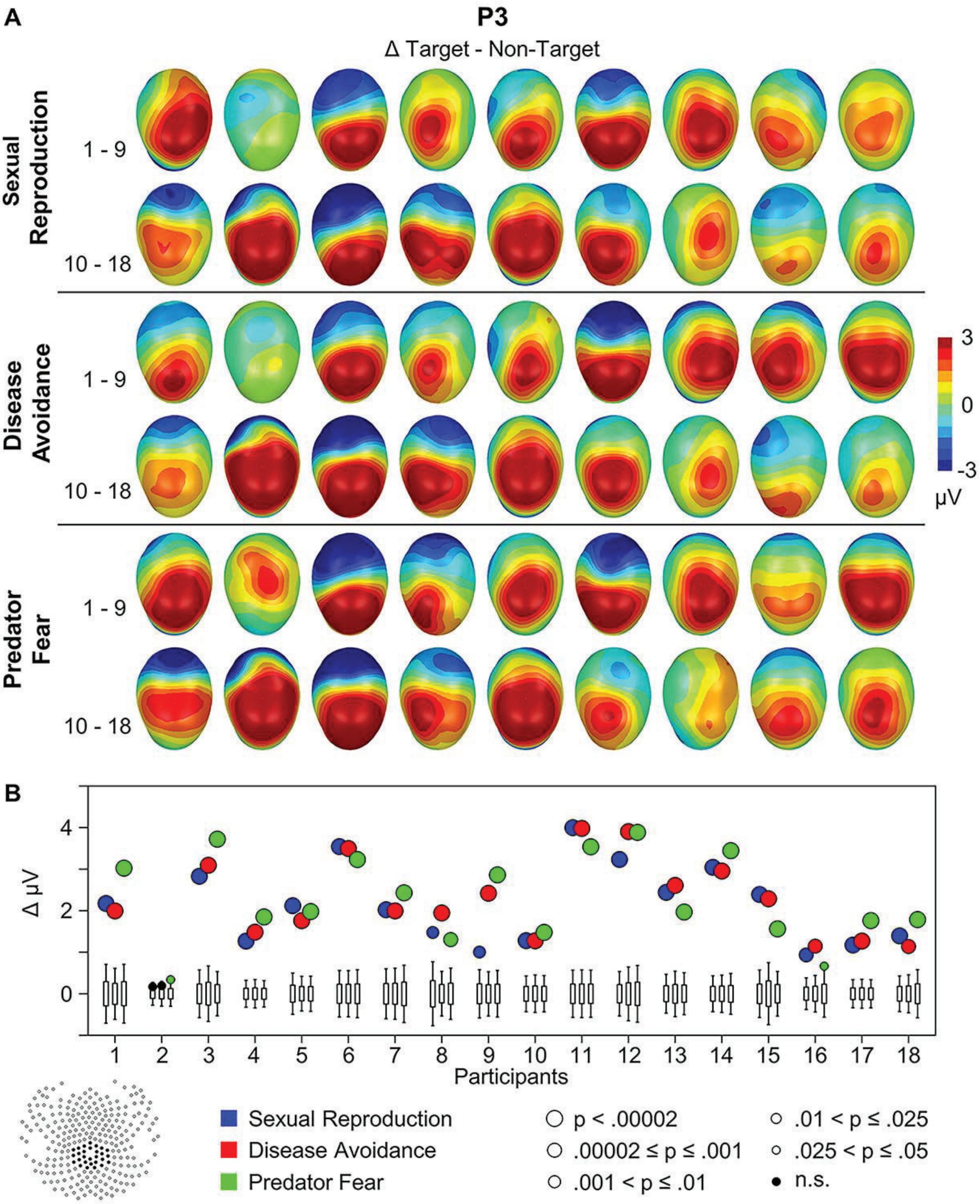

Fig. 3. The target P3 effect for each individual case and behavior system with respect to scalp difference maps (A) and outcome of the single subject bootstrap analysis (B). The dots indicate the de facto measured amplitude difference (target−nontarget).

target P3 effect for high-arousing stimuli in any of the 3 behavior domains, whereas the effect reached significance for low-arousing stimuli for disease avoidance and predator fear.

Emotion and task effects as a function of trial number

How the proportion of significant tests varies as a function of trial number was assessed in analyses progressively downsampling the number of trials used to calculate the bootstrap statistic, i.e. using the first 300, 200, 100, 50 trials per condition (total $n = 600$, 400, 200, and 100, respectively). Findings are illustrated in Fig. 4 separately for EPN and LPP emotion and P3 target effects and summarized in Tables S1–S3 in the Supplementary Material. Subsampling trial numbers systematically affected the proportion of significant EPN and LPP effects as a function of the robustness of effects and the

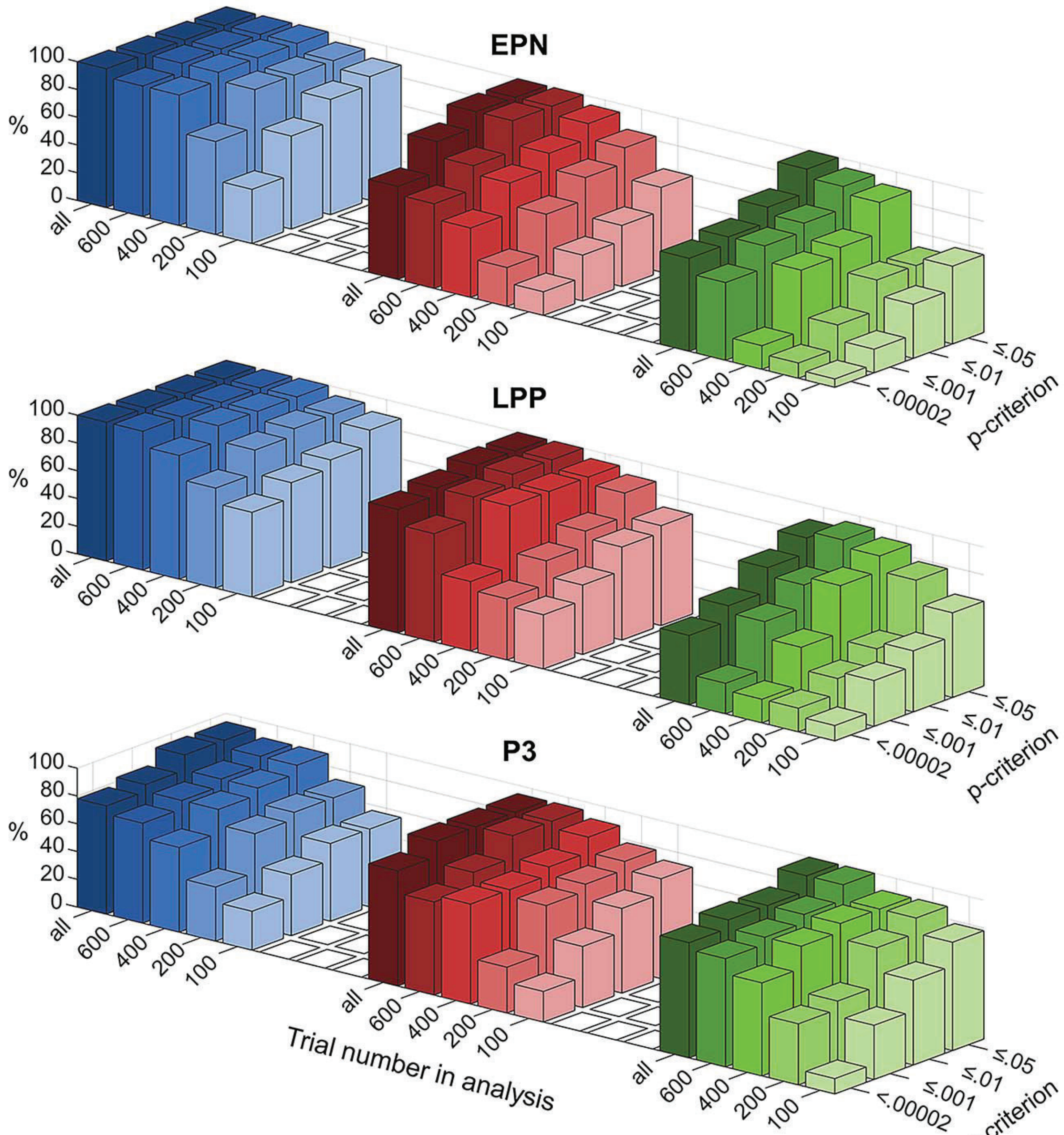

Fig. 4. Proportion (%) of significant cases as a function of *P*-criterion and number of trials used for the calculation of the bootstrap sampling distribution. Blue, red, and green colors indicate behavior domains of sexual reproduction, disease avoidance, and predator fear, respectively.

chosen *P*-criterion. For instance, regarding a significance criterion of $P \leq 0.05$, 89% of the tests conducted for sexual reproduction were still significant for the EPN and LPP when 50 trials per condition ($n = 100$) were used for bootstrap analysis. However, the decline of significant cases was substantially larger for behavior domains of disease avoidance and predator fear in the $n = 100$ trials analysis. Specifically, the proportion of significant cases dropped from 100% (EPN and LPP) to 61% (EPN) and 72% (LPP) for disease avoidance and from 100% (EPN) and 89% (LPP) to 56% (EPN) and 61% (LPP) for predator fear. Apparently, these 2 stimulus categories required a total of 400 trials to retain a higher proportion of significant cases, i.e. 94% and 89% for the EPN as well as LPP, respectively. Furthermore, the proportion of significant tests progressively declined with trial reduction for stricter *P*-criteria. For instance, in the $n = 100$ analysis, 67%, 33%, and 17% of cases for sexual reproduction, diseases avoidance, and predator fear were significant at the $P \leq 0.001$ criterion for the EPN and 72%, 50%, and 33% for the LPP. For the target P3 effect, the pattern of findings was similar across the 3 behavior systems, with a systematic decline in the proportion of significant cases as a function of both, downsampling trials, and reliance on stricter *P*-criteria.

Discussion

The main aim of the present research was to explore whether electrophysiological markers of implicit emotional and explicit task-driven stimulus significance can be assessed in concert at the level of the individual case. Toward this end, in separate blocks, participants were asked to press a button to either high- or low-arousal target stimuli drawn from emotional domains of sexual reproduction, disease avoidance, and predator fear. The findings demonstrate inter- and intrasubject replication of the modulation of the EPN and LPP by emotional significance as well as the modulation of the P3 by task relevance.

The present findings lend further support to the notion that electrophysiological markers of emotional stimulus significance can be assessed at the level of the individual case. Similar to previous findings relying on passive picture viewing (Schupp and Kirmse 2021, 2022), 16 out of 18 cases showed a significant emotional modulation of the EPN and LPP in each of the 3 emotion domains. Based on a total of 6 tests, the presence of affective stimulus evaluation revealed by the EPN and LPP may be inferred in each of these cases. The remaining 2 cases showed strong EPN and LPP effects for categories of sexual reproduction and disease avoidance and a specific lack of effect for the LPP in the predator fear condition. This pattern of findings suggests that the nonsignificant effects are domain-specific rather than indicating a general lack of affective stimulus evaluation effects. Furthermore, as participants performed an explicit categorization task, the possibility arises that emotional modulation effects differ for target and nontarget stimuli. While the outcome of the bootstrap statistic in the consistency analysis is expectedly less robust due to the lower trial numbers, emotional EPN and LPP effects appeared similar for target and nontarget stimuli. Overall, emotional modulation of the EPN and LPP within the individual case can be assessed within the context of an active categorization task.

Beyond emotional stimulus significance, the present data also explored whether the target P3 effect, a robust and reliable measure of task relevance in group research (Johnson 1988; Codispoti et al. 2006), can be assessed at the level of the individual case. Noteworthy, 17 out of 18 participants (94%) showed a significant target P3 effect across all 3 emotion domains. For the remaining case, 2 nonsignificant tests for sexual reproduction and disease avoidance were accompanied by a significant test for predator fear ($P = 0.0325$) just meeting the conventional *P*-criterion. These results are not due to low EEG signal quality, low trial numbers, or poor behavioral performance, as these measures were squarely in the range of the research sample. Thus, although multiple tests are available, interpretation of this case remains ambiguous and would need further testing using a broader range of tasks and stimulus materials. Overall, the consistent demonstration of the P3 target effect in the majority of cases supports the case-wise assessment of effects associated with task-related stimulus significance.

Can 2 different forms of stimulus significance be probed in the individual case within 1 experimental paradigm? Demonstrating intrasubject replication, 15 out of 18 cases showed significant EPN, LPP, and P3 effects across the 3 emotion domains (i.e. 9 tests altogether). From a 2-stage processing perspective, stimulus processing includes a first large capacity perceptual scanning stage providing a more or less complete analysis of sensory information. Upon detection of significant stimuli, the perceptual system may emit a call for processing resources in a capacity-limited second stage of stimulus processing acting as gateway to focused attention and conscious recognition (Öhman 1979, 1986). Relatedly, Potter (Potter 2012; see also Chun and Potter 1995) suggested that stimulus recognition occurs rapidly within the first few hundred milliseconds, which is presumably captured by the EPN. However, for conscious stimulus recognition, this fleeting stage of processing needs to be followed by a second stage of consolidation in short-term memory. Larger LPP and P3 components relate to second-stage processing, presumably reflecting heightened attention to significant stimuli increasing the quality of and confidence in stimulus processing (Hillyard et al. 1971; Vogel et al. 1998; Sergent et al. 2005; Del Cul et al. 2007; Sergent et al. 2021). According to this perspective, the electrophysiological markers assessed in the present study relate to perceptual processing that bring the organisms into contact with the environment, setting the stage for meaningful interaction and preparation of behaviors.

Conceptually, the 2 forms of stimulus significance examined in the present study differ in the mechanisms of how perceptual stimulus processing effects are brought out. Regarding emotion, the effect arises due to activation of stored long-term memory representations, which in the case of emotional stimuli contain more and/or stronger interconnected elements (Lang 1993). Accordingly, emotional significance operates spontaneously and in the absence of explicit processing goals, hence, refers to an implicit process (Öhman et al. 2000). In contrast, goal or task relevance is presumed to reflect the expectancy of certain objects, i.e. the temporary activation of long-term memories (Öhman 1979). The question arises whether the experimental paradigm captures both forms of stimulus significance in comparable ways or whether there is an asymmetry favoring one or the other form of significance. Relying on the $P \leq 0.05$ significance criteria, the proportion of significant cases for emotion and task significance appears highly comparable for the EPN (100%), LPP (96%), and P3 (96%) components. Building upon these findings, a broader assessment of the attention capture of significant stimuli appears possible and new directions of future research may assess emotion and task stimulus significance as a function of spatial location in multi-stimulus displays.

However, in group research, reliance on a $P < 0.05$ criterion has been questioned to determine replicable study findings (Kapur et al. 2012). Being proportional to the confidence interval of the estimated mean, the chosen *P*-criterion serves in group research as a graded measure of confidence in rejecting the null hypothesis (cf. Halsey et al. 2015). Confidence in rejecting the null hypothesis is also relevant in the context of the case-by-case approach and the interpretation of individual test results. The proportion of significant findings was assessed in the present study as a function of both, *P*-criterion and number of trials from which the sampling distribution is determined (see Fig. 4 and Tables S1–S3 in the Supplementary Material). Considering the analysis based on all available trials after artifact rejection, 100% of the cases showed significant EPN and LPP effects for sexual reproduction at the strongest *P*-criterion ($P < 0.00002$) and 89% (EPN) and 94% (LPP) of the tests reached significance for disease avoidance with a $P \leq 0.001$ criterion, indicating highly robust emotional modulation effects observed for these 2 behavior systems. In contrast, for predator fear, the number of significant cases showed a substantially larger decline when using $P \leq 0.001$ (EPN: 72%, LPP: 61%) or stronger *P*-criteria, suggesting to improve stimulus materials for this category. For task effects, significant P3 target modulations were observed for the $P \leq 0.05$ criterion (94%, 94%, and 100% for sexual reproduction, behavior disease, and predator fear, respectively), dropping at the stricter *P*-criterion of $P \leq 0.001$ to 83% (sexual reproduction), 94% (disease avoidance), and 89% (predator fear) of significant effects. Furthermore, analyses based on trial reduction revealed a tradeoff between trial number and robustness of findings. Specifically, the decline in the proportion of significant EPN and LPP effects for progressively smaller trial numbers varied as a function of the robustness of effects, i.e. showing the weakest decline for sexual reproduction and the strongest decline for predator fear, as well as *P*-criterion, i.e. stricter *P*-criteria being associated with a lower number of significant tests (Fig. 4). These findings demonstrate the need to optimize experimental designs to increase confidence in the outcome of case-by-case analyses.

While the present study provides a proof-of-principle that neural markers of emotion and task stimulus significance can be assessed simultaneously at the level of the individual case, some limitations need to be acknowledged. The sample consisted of young, educated college students and it is of particular relevance is to study the neural signature of emotion and task stimulus significance across the full age range. There is evidence that the target P3 effect can be reliably observed in participants at higher age (Friedman 2003). Furthermore, regarding emotional stimulus processing, recent studies observed differential ERP responses to emotional and neutral pictures in a sample of older adults, i.e. ~72 years (Renfroe et al. 2016), and in middle aged women, i.e. 41–62 years (Wirkner et al. 2021). However, these studies also indicated effects of age on emotion processing by observing somewhat attenuated LPP responses to high-arousing pleasant and unpleasant stimulus materials as compared with young participants (Renfroe et al. 2016) and reduced LPP amplitudes to unpleasant contents in middle-aged women (Wirkner et al. 2021). Overall, a case-by-case approach may be useful to probe effects of age on the capture of attention by implicit and explicit stimulus significance.

The present findings may also have implications for biomarker development. There has been agreement that the development of neural biomarkers related to health has been difficult (Kapur et al. 2012; Woo et al. 2017). Woo and colleagues (Woo et al. 2017) attributed this limited progress to problems in experimental methodology such as "significance chasing with underpowered studies" and "approximate replications", and therefore suggested a shift in research practices in order to improve the identification of reliable biomarkers. The core of this approach is constituted by large-scale consortia studies, in the order of hundreds or thousands of participants, allowing predictive modeling and generalization across samples, contexts, and populations. While this approach seems promising to identify biomarkers related to diagnostic categories and pathological symptoms, there is also a call for neural biomarkers measuring general affective and cognitive component processes (Marteau et al. 2012; Insel and Cuthbert 2015; Woo et al. 2017). For instance, ERP indices have been proposed as neural indices for top–down attentional control (Luck et al. 2012), the mismatch negativity as a biomarker for the processing of unattended changes in background stimulation (Light and Näätänen 2013), and a brain-activation map as a neural signature of physical pain (Wager et al. 2013). Here, we aim to demonstrate the feasibility of developing experimental protocols using distinct electrophysiological measures as neural markers for stimulus significance and attentive processing at the level of the individual case. There are several reasons to assume that neural process biomarkers related to stimulus significance are a candidate for translation to the clinical domain. First, there is evidence for selective attention deficits to significant stimuli in mental disorders, involving for instance depression, anxiety spectrum disorders, eating disorders, and drug addiction (Mogg and Bradley 1998; De Houwer et al. 2004; Shafran et al. 2007). Second, changes in the emotional modulation of the EPN and LPP were observed for various mental disorders, e.g. psychopathy, depression, generalized anxiety disorder, and trauma-related psychopathology (MacNamara et al. 2016; Sill et al. 2020; Vallet et al. 2020; Klawohn et al. 2021). Furthermore, the Research Domain of Criteria (RDoC) project by the National Institute of Mental Health aims to identify new ways of classifying mental disorders based on the development of experimental protocols and measures assessing specific emotional brain functions and processes (Cuthbert 2014; Insel and Cuthbert 2015). Overall, the case-by-case approach may contribute to biomarker development by focusing on the individual person.

Funding

This research was supported in part by the Centre for the Advanced Study of Collective Behaviour, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC2117-422037984 granted to HS).

Conflict of interest statement: None declared.

References

Barlow DH, Nock MK. Why can't we be more idiographic in our research? *Perspect Psychol Sci*. 2009:4(1):19–21.

Bradley MM, Lang PJ. Measuring emotion: the self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry*. 1994:25(1):49–59.

Buckholtz JW, Faigman DL. Promises, promises for neuroscience and law. *Curr Biol*. 2014:24(18):R861–R867.

Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, Munafò MR. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*. 2013:14(5): 365–376.

Cacioppo JT, Crites SL Jr, Berntson GG, G. H. Coles M. If attitudes affect how stimuli are processed should they not affect the event-related brain potential. *Psychol Sci*. 1993:4(2):108–112.

Chun MM, Potter MC. A two-stage model for multiple target detection in rapid serial visual presentation. *J Exp Psychol Hum Percept Perform*. 1995:21(1):109–127.

Codispoti M, Ferrari V, Junghöfer M, Schupp HT. The categorization of natural scenes: brain attention networks revealed by dense sensor ERPs. *NeuroImage*. 2006:32(2):583–591.

Cuthbert BN. The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry*. 2014:13(1):28–35.

Danziger K. *Constructing the subject: historical origins of psychological research*. Cambridge (UK): Cambridge University Press; 1990

De Houwer J, Crombez G, Koster EH, De Beul N. Implicit alcohol-related cognitions in a clinical sample of heavy drinkers. *J Behav Ther Exp Psychiatry*. 2004:35(4):275–286.

Del Cul A, Baillet S, Dehaene S. Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biol*. 2007:5(10):e260.

Di Nocera F, Ferlazzo F. Resampling approach to statistical inference: bootstrapping from event-related potentials data. *Behav Res Methods Instrum Comput*. 2000:32(1):111–119.

Domjan M. Formulation of a behavior system for sexual conditioning. *Psychon Bull Rev*. 1994:1(4):421–428.

Efron B, Tibshirani R. *An introduction to the bootstrap*. New York (NY): Chapman & Hall; 1993

Friedman D. Cognition and aging: a highly selective overview of event-related potential (ERP) data. *J Clin Exp Neuropsychol*. 2003:25(5):702–720.

Halsey LG, Curran-Everett D, Vowler SL, Drummond GB. The fickle P value generates irreproducible results. *Nat Methods*. 2015:12(3): 179–185.

Hillyard SA, Squires KC, Bauer JW, Lindsay PH. Evoked potential correlates of auditory signal detection. *Science*. 1971:172(3990): 1357–1360.

Insel TR, Cuthbert BN. Medicine. Brain disorders? Precisely. *Science*. 2015:348(6234):499–500.

Ioannidis JPA. Why most published research findings are false. *PLoS Med*. 2005:2(8):e124.

Johnson R Jr. The amplitude of the P300 component of the event-related potential: review and synthesis. In: Ackles PK, Jennings JR, Coles MGH, editors. *Advances in psychophysiology: a research annual*. Greenwich (CT): JAI Press; 1988. pp. 69–137

Junghöfer M, Elbert T, Tucker DM, Rockstroh B. Statistical control of artifacts in dense array EEG/MEG studies. *Psychophysiology*. 2000:37(4):523–532.

Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry*. 2012:17(12):1174–1179.

Klawohn J, Burani K, Bruchnak A, Santopetro N, Hajcak G. Reduced neural response to reward and pleasant pictures independently relate to depression. *Psychol Med*. 2021:51(5): 741–749.

Lamiell JT. 'Nomothetic' and 'idiographic': contrasting Windelband's understanding with contemporary usage. *Theory Psychol*. 1998:8(1):23–38.

Lang PJ. The network model of emotion: motivational connections. In: Wyer RS, Srull TK, editors. *Perspectives on anger and emotion*. Hillsdale (NJ): Erlbaum; 1993. pp. 109–133

Lang PJ, Bradley MM, Cuthbert BN. Motivated attention: affect, activation, and action. In: Lang PJ, Simons RF, Balaban M, editors. *Attention and orienting: sensory and motivational processes*. Mahwah (NJ): Erlbaum; 1997. pp. 97–135

Lang PJ, Bradley MM, Cuthbert BN. International affective picture system (IAPS): affective ratings of pictures and instruction manual. In: *Technical report A-8*. Gainesville, FL: University of Florida; 2008

Light GA, Näätänen R. Mismatch negativity is a breakthrough biomarker for understanding and treating psychotic disorders. *PNAS*. 2013:110(38):15175–15176.

Lindsay DS. Replication in psychological science. *Psychol Sci*. 2015:26(12):1827–1832.

Luck SJ, Ford J, Sarter M, Lustig C. CNTRICS final biomarker selection: control of attention. *Schizophr Bull*. 2012:38(1):53–61.

MacNamara A, Kotov R, Hajcak G. Diagnostic and symptom-based predictors of emotional processing in generalized anxiety disorder and major depressive disorder: an event-related potential study. *Cognit Ther Res*. 2016:40(3):275–289.

Marks IM. Blood-injury phobia: a review. *Am J Psychiatry*. 1988:145(10):1207–1213.

Marteau TM, Hollands GJ, Fletcher PC. Changing human behavior to prevent disease: the importance of targeting automatic processes. *Science*. 2012:337(6101):1492–1495.

Mogg K, Bradley BP. A cognitive-motivational analysis of anxiety. *Behav Res Ther*. 1998:36(9):809–848.

Munafò MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, Percie du Sert N, Simonsohn U, Wagenmakers EJ, Ware JJ, Ioannidis JPA. A manifesto for reproducible science. *Nat Hum Behav*. 2017:1(1):0021.

Neuberg SL, Kenrick DT, Schaller M. Human threat management systems: self-protection and disease avoidance. *Neurosci Biobehav Rev*. 2011:35(4):1042–1051.

Öhman A. The orienting response, attention, and learning: an information-processing perspective. In: Kimmel HD, van, Olst EH, Orlebek JF, editors. *The orienting reflex in humans*. Hillsday (NJ): Erlbaum; 1979. pp. 443–471

Öhman A. Face the beast and fear the face: animal and social fears as prototypes for evolutionary analyses of emotion. *Psychophysiology*. 1986:23(2):123–145.

Öhman A, Flykt A, Lundqvist D. Unconscious emotion: evolutionary perspectives, psychophysiological data and neuropsychological mechanisms. In: Lane RD, Nadel L, editors. *Cognitive neuroscience of emotion*. New York (NY): Oxford University Press; 2000. pp. 296–327

Open Science Collaboration. Estimating the reproducibility of psychological science. *Science*. 2015:349(6251):aac4716.

Parks NA, Gannon MA, Long SM, Young ME. Bootstrap signal-to-noise confidence intervals: an objective method for subject exclusion and quality control in ERP studies. *Front Hum Neurosci*. 2016: 10:50.

Peyk P, De Cesarei A, Junghöfer M. ElectroMagnetoEncephalography software: overview and integration with other EEG/MEG toolboxes. *Comput Intell Neurosci*. 2011:2011:861705–861710.

Poldrack RA, Baker CI, Durnez J, Gorgolewski KJ, Matthews PM, Munafò MR, Nichols TE, Poline J-B, Vul E, Yarkoni T. Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nat Rev Neurosci*. 2017:18(2):115–126.

Potter MC. Conceptual short term memory in perception and thought. *Front Psychol*. 2012:3:113.

Renfroe JB, Bradley MM, Sege CT, Bowers D. Emotional modulation of the late positive potential during picture free viewing in older and young adults. *PLoS One*. 2016:11(9):e0162323.

Rosenfeld JP. P300 in detecting concealed information and deception: a review. *Psychophysiology*. 2020:57(7):e13362.

Schimmel H. The (+−) reference: accuracy of estimated mean components in average response studies. *Science*. 1967:157(3784): 92–94.

Schupp HT, Kirmse UM. Case-by-case: emotional stimulus significance and the modulation of the EPN and LPP. *Psychophysiology*. 2021:58(4):e13766.

Schupp HT, Kirmse U. Neural correlates of affective stimulus evaluation: a case-by-case analysis. *Soc Cogn Affect Neurosci*. 2022:17(3): 300–310.

Sergent C, Baillet S, Dehaene S. Timing of the brain events underlying access to consciousness during the attentional blink. *Nat Neurosci*. 2005:8(10):1391–1400.

Sergent C, Corazzol M, Labouret G, Stockart F, Wexler M, King JR, Meyniel F, Pressnitzer D. Bifurcation in brain dynamics reveals a signature of conscious processing independent of report. *Nat Commun*. 2021:12(1):1146.

Shafran R, Lee M, Cooper Z, Palmer RL, Fairburn CG. Attentional bias in eating disorders. *Int J Eat Disord*. 2007:40(4):369–380.

Shrout PE, Rodgers JL. Psychology, science, and knowledge construction: broadening perspectives from the replication crisis. *Annu Rev Psychol*. 2018:69(1):487–510.

Sidman M. *Tactics of scientific research: evaluating experimental data in psychology*. New York (NY): Basic Books; 1960

Sill J, Popov T, Schauer M, Elbert T. Rapid brain responses to affective pictures indicate dimensions of trauma-related psychopathology in adolescents. *Psychophysiology*. 2020:57(1):e13353.

Vallet W, Hone-Blanchet A, Brunelin J. Abnormalities of the late positive potential during emotional processing in individuals with psychopathic traits: a meta-analysis. *Psychol Med*. 2020:50(12): 2085–2095.

Vogel EK, Luck SJ, Shapiro KL. Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *J Exp Psychol Hum Percept Perform*. 1998:24(6):1656–1674.

Wager TD, Atlas LY, Lindquist MA, Roy M, Woo CW, Kross E. An fMRI-based neurologic signature of physical pain. *New Eng J Med*. 2013:368(15):1388–1397.

Wirkner J, Ventura-Bort C, Hamm AO, Weymar M. Glancing at the "sunny side of life"? Emotion processing and memory in middle-aged women. *Psychol Neurosci*. 2021. https://doi.org/10.1037/pne0000277.

Woo C-W, Chang LJ, Lindquist MA, Wager TD. Building better biomarkers: brain models in translational neuroimaging. *Nat Neurosci*. 2017:20(3):365–377.