Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

How Negative Evidence Influences Inductive Reasoning: An ERP Study

Jiantao Zhang
Hubei University of Medicine
Juan Wang
wang.juan@hbmu.edu.cn
Hubei University of Medicine
Xi Wang
Taihe Hospital
Peipei Li
Hubei University of Medicine

Research Article

Keywords: general arguments, negative arguments, P300–N400, P2, N2, SN

Posted Date: August 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10492358/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Negative evidence could increase the strength of specific arguments in semantic category-based inductive reasoning, but whether the effect still exists for general arguments, more importantly, the mechanism of the cognitive processing of general arguments, remains unclear. In the current study, event-related potentials (ERPs) were used to investigate this effect by controlling the category type (related vs. unrelated) and argument type (positive vs. negative) in general arguments. The behavioral results demonstrated that under the unrelated-category conditions, negative arguments had a greater argument strength and were associated with a shorter reaction time than positive arguments. The ERP results indicated that in the premise 2 presentation phase (a) unrelated categories elicited larger P300 and N400 amplitude than did related categories, reflecting the categorization process, and (b) positive properties elicited larger P600 amplitudes than did negative properties, demonstrating the process of property judgment. In the conclusion phase (c) negative arguments elicited larger P2 and N2 amplitudes than did positive arguments, which embodied cognitive control and conflict detection, and (d) negative arguments also elicited a SN than did positive arguments under the related-category condition, which illustrated the property inference. This study provides robust evidence that negative evidence can increase the argument strength in general arguments, and that the temporal course of negative evidence influences the inductive strength.

1. Introduction

Category-based inductive reasoning is a process in which people infer, through known knowledge that a category has certain properties, that one or a set of categories also has those properties (Feeney, 2007; Osherson et al., 1990; Sloman, 1993). For example, if it is known that both crows and sparrows have sesamoid bones, it can be inferred that birds have sesamoid bones because both crows and sparrows are birds (Osherson et al., 1990).

Previous studies on category-based inductive reasoning have generally involved positive evidence, for example, “the premise: sparrows have sesamoid bones and crows also have sesamoid bones; the conclusion: swallows have sesamoid bones,” which means that a premise referring to a certain little acknowledged blank property could be extrapolated to the conclusion (Osherson et al., 1990). In real life, however, people usually encounter not only plenty of positive evidence, but also abundant negative evidence that may be contrary to previously established views. For example, some animals that look gentle are actually very dangerous; thus, how can negative and positive evidence be integrated to implement new inductive reasoning? Some studies have shown that negative evidence does not decrease, but actually increases argument strength. These studies have found that when the categories of negative evidence originate from various categories of positive premises and conclusions, the argument strength will be enhanced (Heussen, et al., 2011; Heussen, Voorspoels, et al., 2011; Kalish & Lawson, 2007; Lee et al., 2019; Voorspoels et al., 2015). Given that the findings noted above were commonly based on specific arguments (both the premise and conclusion refer to subordinate categories), however, whether the same conclusion can be drawn for general arguments unknown. In

addition, to explore how negative evidence affects the strength of the argument in inductive reasoning, previous researchers compared mixed arguments (positive-negative arguments) with a single-premise positive argument. Although this method can reveal the change of the inductive strength, it is affected by the number of premises, which is defined as the premise monotonicity effect (Cui et al., 2018; Osherson et al., 1990). Thus, whether the change of the argument strength is caused by increased premises or negative evidence remains uncertain. Most importantly, in general arguments, the temporal course of negative evidence affecting the argument strength remains unclear. The main purpose of this study was, therefore, to explore how negative evidence influences the argument strength of general arguments as well as the related cognitive mechanism during semantic category-based induction.

In previous studies neuroimaging tools, such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), were used to examine inductive and deductive reasoning to determine the brain mechanism of categorical inductive reasoning and discover whether various types of reasoning activate specific brain regions. For example, Goel et al. (1997) applied PET and found that inductive reasoning activates the left medial prefrontal region (BA 8/9); similarly, Goel and Dolan (2004) used fMRI and found the activation of the dorsal prefrontal lobe. These results provided evidence not only for the mechanisms of various types of reasoning but also for the debate between single-process reasoning (Rips, 2001) and dual-process reasoning (Evans, 2003; Sloman, 1993). Partly because of the technique involved, however, the various stages of processing during categorical inductive reasoning could not be distinguished. Fortunately, event-related potentials (ERPs) can overcome these shortcomings and eventually individuate the sequences of a cognitive process.

Some researchers have used ERPs to reveal the neural basis of categorical inductive reasoning (Lei et al., 2019; Li et al., 2016; Liang et al., 2010) and some effects during category-based induction (Cui et al., 2018; Liang et al., 2016; Liang, Xiao, Lei, et al., 2020; Long et al., 2015; Wang et al., 2016). The results indicated that N2, P300, and N400 as well as the late positive component (LPC) or late negative component (LNC), are related to the reasoning process when performing conditional reasoning (Bonnefond & Henst, 2009, 2013; Bonnefond, Kaliuzhna, et al., 2014), transitive reasoning (Bonnefond, Castelain, et al., 2014), and category-based induction (Lei et al., 2017; Li et al., 2016; Liang, Xiao, Zhu, et al., 2020). For example, N2 reflects the detection of conflict in the early stage of categorical inductive reasoning (Liang et al., 2010), and inconsistent properties elicit a larger N2 amplitude than consistent properties in semantic category-based induction (Li et al., 2016). In addition, P300 reflects the reasoning expectation in categorical inductive reasoning, and the more consistent the reasoning result with the reasoning expectation, the greater the elicited P300 amplitude (Lei et al., 2019). Moreover, N400 is closely associated with semantic processing and very sensitive to category membership, so a greater N400 amplitude is usually induced by stimuli with semantic mismatch (Folstein & Van Petten, 2007; Li et al., 2016; Núñez-Peña & Honrubia-Serrano, 2005). When the categories of the conclusion are related to those of the premises, inconsistent properties violate the reasoning logic and rules, thus eliciting greater positive deflection than consistent properties. This process is related to LPC or P600, which are considered to be related to rule violation (Chen et al., 2013; Núñez-Peña & Honrubia-Serrano, 2004). These studies were not, however, aimed at the mechanism of how negative evidence affects categorical

inductive reasoning; they all focused solely on positive evidence. Furthermore, these researchers focused solely on the processing of the conclusion but did not consider the processing of the premises. Goel (2007) believed that argument premises and conclusions constitute two parts of the reasoning process, so analysis should occur in both of them during reasoning. Based on this perspective, the main purpose of the current study was to investigate the neural processes from the presentation of the premises to the evaluation of the conclusion, thus allowing the entire course of the influence of negative evidence on semantic inductive reasoning to be divided into more detailed steps.

In this study, the number of premises was controlled by manipulating the category type (related vs. unrelated) and argument type (positive vs. negative), after which the reaction time (RT) and argument strength during the conclusion presentation were measured. According to previous studies, categories with a far semantic distance require a longer reaction time than those with a close semantic distance (Collins & Loftus, 1975); furthermore, negative processing requires more cognitive resources than positive processing. We, therefore, hypothesized that the RT of negative arguments would be longer than that of positive arguments and also that the RT of related-category arguments would be longer than that of unrelated-category arguments. Regarding argument strength, consistent with previous studies (Heussen, Voorspoels, et al., 2011; Kalish & Lawson, 2007; Voorspoels et al., 2015), we hypothesized that the strength of positive arguments would be stronger than that of negative arguments under the related-category condition, and that the strength of negative arguments would be stronger than that of positive arguments under the unrelated-category condition.

Regarding the ERPs, in previous studies, participants formed a hypothesis when the first premise was presented. If the second premise and conclusion were inconsistent with their hypothesis, the subjects would change their hypothesis and form a new one (Feeney et al., 2010; Medin et al., 2003). On the contrary, when unrelated-category negative evidence and related-category positive evidence were received, the subjects did not change their reasoning hypothesis or expectation. In accordance with the aforementioned findings, the unrelated-category negative arguments and related-category positive arguments were assumed to elicit greater N2 and P300 amplitudes than unrelated-category positive arguments and related-category negative arguments, both in the premise and conclusion presentation stages. Similarly, N400 reflects the categorization process and the associative strength of the category in the reasoning process (Núñez-Peña & Honrubia-Serrano, 2005). In the current study the related and unrelated categories had different associative strengths; thus, we assumed that compared with related-category arguments, unrelated-category arguments would induce a greater N400 amplitude in the premise and conclusion presentation phases. In addition, because the LPC (or P600) reflects the violation of reasoning rules, we assumed that in the premise stage negative properties would induce a greater LPC amplitude than would positive properties and that in the conclusion presentation stage negative arguments would induce a greater LPC amplitude than would positive arguments.

2. Methods

2.1. Participants

G-Power software was used, in which the forecast effect size was 0.25, the significance level was 0.05, and the statistical power was 0.8. A total of 24 subjects were required for this experiment, and 31 undergraduate students were recruited. Of these recruits, one completed fewer than the required number of trials, the experimental data of one was not completely recorded by the EEG recording system, and two interrupted the experiment. Thus, ultimately, the data of only 27 participants (19 female participants; Mean age = 23.56, SD = 2.52, Max = 29.00, Min = 19.00) were included in the statistical analysis. All participants were self-reported as being right-handed and native speakers of Mandarin Chinese. They had normal or corrected-to-normal vision, no hair dye, and no history of mental illness, and they had not used psychotropic drugs in the past three months. All experimental protocols were approved by the Academic Ethics Committee of Capital Normal University, and signed written informed consent was obtained. After the experiments, the subjects were paid 90¥ for their participation.

2.2 Design

In the current study, a 2 (category type: related vs. unrelated) ×2 (argument type: positive vs. negative) repeated measures design was adopted. A related category was defined as the situation in which the categories of the first and second premises belonged to the conclusion category (e.g., premise 1 category: apple; premise 2 category: pear; conclusion: fruits); similarly, unrelated categories meant that the categories of the first and second premises did not belong to the conclusion category (e.g., premise 1 category: apple; premise 2 category: coriander; conclusion: fruits). In addition, the types of arguments included positive and negative arguments. Similarly, negative evidence was defined as the second premise category not having the to-be-projected property, and positive evidence was defined as the second premise category having some of to-be-projected property (Heussen, Heussen, et al., 2011; e.g., premise 1: apples have a property X1; premise 2: pears have/do not have a property X1; conclusion: fruits have a property X1). The main experiment, therefore, included four types of arguments: (a) related-category positive arguments (R-P); (b) unrelated-category positive arguments (UR-P); (c) related-category negative arguments (R-N); (d) unrelated-category negative arguments (UR-N).

2.3 Materials

To create arguments, six natural categories (birds, flowers, vegetables, insects, mammals and fruits) were selected as stimuli, each of which consisted of 30 members and every member was named by 2 or 3 Chinese characters (Cui et al., 2018; Lei et al., 2019). Because of the typical effect in categorical inductive reasoning (Feeney & Heit, 2011; Rein et al., 2010; Zhong et al., 2014), 20 subjects (10 female subjects; Mean age = 24.95, SD = 1.64) evaluated the familiarity and typicality of all the materials before the experiment on a 7-point scale, ranging from 1 (*familiar/typical*) to 7 (*familiar/typical*); categories with an average score of less than 4 for typicality and familiarity were excluded. The mean degrees of familiarity and typicality were respectively 5.80 ± 0.75(M ± SD) and 5.45 ± 0.79. To avoid the influence of the participants' background knowledge of categorical inductive reasoning (Heit, 2000; Osherson et al., 1990), the property was represented by mixing a capital letter ranging from A-Z with an Arabic number ranging from 0 to 9, which served as the blank property. For example, "Tigers have A3" meant that tigers have a property A3 (Cui et al., 2018; Liang et al., 2016). The category and property were spaced by one

character. At the end of the conclusion, a question mark (?) was presented with the intention of requiring the participants to judge the plausibility of the conclusion according to the premise. In this study, the categories of premise 1 and premise 2 were subordinate categories (e.g., premise: apple, pear) and the category of the conclusion was a basic-level category (e.g., conclusion: fruits). The basic-level category was, therefore, presented 24 times under each condition. The order of the trials was pseudo-randomly presented, and trials under the same conditions could not occur three times in succession.

2.4 Procedure

The materials were presented on a computer screen using Presentation software (Neurobehavioral Systems Inc., Berkeley, CA, USA). The formal experiment occurred as shown in Fig. 1.

First, a white fixation cross (+) appeared in the center of the screen for 500 ms at the beginning of each trial, after which premise 1 was presented for 1400 ms, a blank screen was presented 500—800 ms premise 2 was presented for 1,400 ms, and a blank screen was presented for 500—800 ms. Then, the conclusion appeared, and the participants were required to respond as rapidly and accurately as possible to judge the strength of the conclusion by pressing the “F” key (“strong”) or “J” key (“weak”) with their left or right index finger, respectively, which was counterbalanced across participants. The responses of the subjects were followed by the disappearance of the conclusion and then the presentation of a blank screen for 1000 ms. Finally, a 7-point scale, ranging from 1 (*weak*) to 7 (*strong*) appeared on the screen. The participants were required to score the argument strength from 1 to 7. After completing a block, they rested for 120 s before starting the next block. The Chinese characters contained in the premises and conclusions were all presented in Song typeface 40. The color of the background was black and that of the characters was white. The stimuli subtended a visual angle of 1.91° vertically and a mean of 8.58° horizontally.

A practice phase was completed before the formal test experiment, during which six trials were arranged for practice under each condition, with a total of 24 practice trials not included in the test experiment. The test phase consisted of 240 trials in the main experiment (60 trials for each condition).

2.5 EEG data acquisition and data analysis

Brain electrical activity was recorded with a 64-tin Ag/Ag CI cap based on the extended International 10–20 System (Compumedics Neuroscan Inc., El Paso, TX, USA). A ground electrode was placed at the midpoint of FPz and Fz and was re-referenced offline on the left and right mastoids. The EEG signal was amplified to 0.05-100Hz with a sampling rate of 500Hz. The vertical electroencephalogram was recorded at 1cm above and below the left eye, and the horizontal electroencephalogram was recorded at 1cm outside the corner of both eyes. All impedances were kept lower than 5 KΩ.

The off-line data analysis was performed by using EEGLAB toolbox in MATLAB software (Delorme & Makeig, 2004). The signal was passed through a 0.05 Hz high-pass filter to a 30 Hz low-pass filter for offline analysis. Independent component analysis (ICA) was subsequently performed to correct for the component associated with eye movements and blinking. Trials exceeding ± 75μ0V were excluded from

the ERP average, which included amplifier blocking and muscle potentials. The average trials under each condition were as follows: in the premise 2 premise presentation phase, R-P: 54.88 ± 6.09; UR-P: 55.19 ± 5.47; R-N: 55.65 ± 5.85; UR-N: 55.69 ± 5.83; in the conclusion presentation phase, R-P: 52.96 ± 8.49; UR-P: 54.33 ± 6.42; R-N: 53.44 ± 7.62; UR-N: 54.22 ± 7.00.

Several studies showed EEG signals at the time of conclusion presentation (Cui et al., 2018; X. Liang, Xiao, Lei, et al., 2020; Long et al., 2017) as well as during the presentation of premise 2. The average epoch is 1200 ms, including a 200 ms pre-stimulus for baseline correction. Via the observation of the waveforms and topographic maps obtained under various conditions, in the premise 2 phase, the mean amplitude was analyzed at 330–390 ms, 400–460 ms and 550–750 ms epochs. In the conclusion phase the mean amplitude was analyzed at 150–190 ms, 330–370 ms, and 600–1000 ms epochs. A total of nine electrodes of the 10–20 System were selected to define different scalp regions (frontal: F3, Fz, F4; central: C3, Cz, C4; parietal: P3, Pz, P4). Four-factor repeated measures analyses of variance (ANOVAs) for the category type (related, unrelated) × argument type (positive, negative) × brain region (anterior: F3, Fz, F4; central: C3, Cz, C4; parietal: P3, Pz, P4) × lateralization (left, middle, right) were executed to compare the EEG results. For all analyses, the Bonferroni criterion and Greenhouse-Geisser correction were used to correct the violation of the sphericity assumption and each comparison.

3. Results

3.1 Behavioral results

The RT for each condition was entered into a repeated measures ANOVA. The results shown in Fig. 2 indicated that the interaction of the category type and argument type was significant, $F_{(1, 26)} = 9.80$, $p < 0.01$, η2 p = 0.27. When the factor of the category type was entered into the follow-up analysis, the results showed that the RT of negative arguments was longer than that of positive arguments under the related-category condition, $F_{(1, 26)} = 8.89$, $p = 0.004$, η2 p = 0.26. When the factor of the argument type was entered into the follow-up analysis, the results suggested that the RT of unrelated category arguments was longer than that of related-category arguments under positive conditions, $F_{(1, 26)} = 30.42$, $p < 0.001$, η2 p = 0.54 and that the RT of related-category arguments was longer than that of unrelated-category arguments under negative conditions, $F_{(1, 26)} = 8.56$, $p = 0.007$, η2 p = 0.25.

Regarding the argument strength, the interaction of the category type and argument type was significant, $F_{(1, 26)} = 135.92$, $p < 0.001$, η2 p = 0.84. Simple effect analyses showed that the subjects' responses to positive arguments were stronger than those to negative arguments under the related-category condition, $F_{(1, 26)} = 164.26$, $p < 0.001$, η2 p = 0.86, and that their responses to positive arguments were weaker than those to negative arguments under the unrelated-category condition, $F_{(1, 26)} = 20.44$, $p < 0.001$, η2 p = 0.44. The results also suggested that related-category arguments were evaluated as stronger than unrelated-category arguments under positive conditions, $F_{(1, 26)} = 115.53$, $p < 0.001$, η2 p =

0.82, and unrelated-category arguments were evaluated as stronger than related-category arguments under negative conditions, $F_{(1, 26)} = 73.59$, $p < 0.001$, η2 p = 0.74.

3.2 ERP results

3.2.1 ERP results of the premise 2 presentation phase

P300 (330–390 ms)

The interaction of lateralization and the category type reached statistical significance during the 330–390 ms time window, $F_{(1, 26)} = 3.70$, $p = 0.039$, η2 p = 0.12. Simple effect analyses showed that related-category premises elicited a larger P300 amplitude than did unrelated-category premises in the left hemisphere, $F_{(1, 26)} = 4.43$, $p = 0.045$, η2 p = 0.15; in the middle, $F_{(1, 26)} = 5.58$, $p = 0.026$, η2 p = 0.18; and in the right hemisphere, $F_{(1, 26)} = 7.48$, $p = 0.011$, η2 p = 0.22. The main effect of the category type was significant $F_{(1, 26)} = 6.08$, $p = 0.021$, η2 p = 0.19, Post-hoc tests showed that related-category premises elicited a larger P300 amplitude than did unrelated-category.

N400 (400–460 ms)

As shown in Fig. 3, the main effect of the category type was significant, $F_{(1, 26)} = 16.97$, $p < 0.001$, η2 p = 0.40. Post-hoc tests showed that the N400 amplitude elicited by related category-premises was larger than that elicited by unrelated-category premises. The main effect of lateralization was significant, $F_{(1, 26)} = 9.60$, $p < 0.001$, η2 p = 0.27. Post-hoc tests showed that the activation of N400 in the right hemisphere was larger than in the middle and in the left hemisphere.

P600 (550–750ms)

As shown in Fig. 3, during the 550–750 ms time window, the main effect of argument types was significant, $F_{(1, 26)} = 8.36$, $p = 0.008$, η2 p = 0.24. Post-hoc tests suggested that negative premises elicited a larger P600 amplitude than positive premises. The interaction of lateralization and the brain region was significant, $F_{(1, 26)} = 3.75$, $p = 0.035$, η2 p = 0.13. Simple effect analyses showed that the activation of the right hemisphere was larger than that of the middle and the left hemisphere in the frontal region, $F_{(1, 26)} = 17.35$, $p < 0.001$, η2 p = 0.58, and central region, $F_{(1, 26)} = 11.30$, $p < 0.001$, η2 p = 0.48.

3.2.2 ERP results of the conclusion presentation phase

P2 (150–210 ms)

As shown in Fig. 4, during the 150–210 ms time window, the main effect of the argument type was significant, $F_{(1, 26)} = 6.20$, $p = 0.02$, η2 p = 0.19. Post-hoc tests suggested that positive arguments elicited a larger P2 amplitude than did negative arguments. The result also showed that the interaction of lateralization and the category type was significant, $F_{(1, 26)} = 3.43$, $p = 0.048$, η2 p = 0.22. Simple effect analyses showed that positive arguments elicited a larger P2 amplitude than did negative arguments in

the right hemisphere, $F_{(1, 26)}$ = 4.75, p = 0.039, η2 p = 0.15; in the middle, $F_{(1, 26)}$ = 7.58, p = 0.011, η2 p = 0.23, and in the left hemisphere, $F_{(1, 26)}$ = 4.97,p = 0.035, η2 p = 0.16.

N2 (250–310 ms)

As shown in Fig. 4, during the 250–310 ms time window, the interaction of lateralization, the category type, and the argument type was significant, $F_{(1, 26)}$ = 7.09, p = 0.002, η2 p = 0.21. When the factor of the category type was entered into the follow-up analysis, the results showed that the N2 amplitude induced by positive arguments was larger than that induced by negative arguments in both the right hemisphere $F_{(1, 26)}$ = 5.71, p = 0.024, η2 p = 0.18 and in the middle, $F_{(1, 26)}$ = 4.93, p = 0.035, η2 p = 0.16.

SN (400–600 ms)

Figure 4 also shows that during the 400–600 ms time window, the interaction of the brain region, category type, and argument type was significant, $F_{(1, 26)}$ = 5.19, p

= 0.022, η2 p = 0.17. Simple effect analyses showed that under the related-category condition, the SN amplitude elicited by negative arguments was larger than that elicited by positive arguments in the frontal region, $F_{(1, 26)}$ = 4.85, p = 0.037, η2 p = 0.16. The interaction of the brain region and category type was also significant, $F_{(1, 26)}$ = 5.82, p = 0.014, η2 p = 0.18. Simple effect analyses suggested that negative arguments elicited a larger SN amplitude than did positive arguments in the frontal region, $F_{(1, 26)}$ = 4.39, p = 0.046, η2 p = 0.14.

As shown in Fig. 4, during the 600–800 ms time window, the interaction of the brain region and argument type was significant, $F_{(1, 26)}$ = 8.17, p = 0.005, η2 p = 0.24. Simple effect analyses suggested that the SN amplitude elicited by negative arguments was larger than that elicited by positive arguments in the frontal region, $F_{(1, 26)}$ = 6.95, p = 0.014, η2 p = 0.21.

As shown in Fig. 4, during the 800–1000 ms time window, the interaction of the brain region, lateralization, the category type, and the argument type was significant, $F_{(1, 26)}$ = 2.88, p = 0.046, η2 p = 0.10. When the frontal region was fixed in the analysis, the results showed that the interaction of the lateralization, the category type, and the argument type was significant, $F_{(1, 26)}$ = 5.37, p = 0.004, η2 p = 0.19. Simple effect analyses suggested that negative arguments elicited a larger SN amplitude than did positive arguments under the related-category condition in the left hemisphere, $F_{(1, 26)}$ = 15.41, p = 0.001, η2 p = 0.37, and in the middle, $F_{(1, 26)}$ = 7.21, p = 0.021, η2 p = 0.22. When the central region was fixed in the analysis, the results suggested that the main effect of the category type was significant, $F_{(1, 26)}$ = 4.76, p = 0.038, η2 p = 0.16. Post-hoc tests showed that negative arguments elicited a larger SN amplitude than did positive arguments.

4. Discussion

4.1 The analysis of the behavioral results

The influence of negative evidence on semantic category-based induction in general arguments was explored based on previous studies. We found that under the related-category condition, the RT of negative arguments was significantly longer than that of positive arguments; on the contrary, under the unrelated-category condition, no remarkable difference emerged in RT between negative and positive arguments. In addition, under positive conditions, the RT of unrelated-category arguments was longer than that of related-category arguments; contrarily, under negative conditions, the RT of related-category arguments was longer than that of unrelated-category arguments. The results were consistent with those of previous studies, which indicated that categories with a far semantic distance require a longer RT than those with a close semantic distance (Janiszewski & Wyer, 2014; Lei et al., 2019; Liang, Xiao, Lei, et al., 2020). The spreading activation model in semantic networks explains this phenomenon as follows. Related categories are stored in close semantic nodes, but unrelated categories are stored in distant nodes; therefore, compared with unrelated categories, related categories have a stronger semantic connection and are more easily extracted from memory, which means that related categories require less cognitive resources and a shorter RT (Collins & Loftus, 1975).

In addition, previous studies have shown that people need more cognitive resources for affirmation than negation in category verifications. The results of this experiment also prove that in categorical inductive reasoning, the process of negative arguments had a longer RT than that of positive arguments, which may be so because the processing of negative arguments consumes more cognitive resources than that of positive arguments.

Regarding argument strength, the number of premises and the category relation were controlled, and we found that in general arguments, negative evidence could increase the argument strength under unrelated-category condition. This is consistent with the findings of previous research (Osherson et al., 1990). To some extent, the findings support relevance theory, which proposes that negative evidence could enhance argument strength because of negative and positive premises as well as conclusions that share correlation dimensions. This reduces the cognitive effort during cognitive input and enhances the cognitive effect, thus enhancing the induction of arguments (Heussen, Voorspoels, et al., 2011; Medin et al., 2003).

4.2 The analysis of the premise 2 presentation phase

Previous studies showed that P300 is related to semantic distance of the categories; compared with categories with far semantic distances, categories with close semantic distances induce a greater P300 amplitude (Lei et al., 2017, 2019). In this experiment, the category of premise 2 was related to the category of premise 1 in two ways, namely the related or unrelated. The P300 amplitude induced by an unrelated category was greater than that induced by a related category, thus reflecting the effect of the semantic distance between categories. Long et al. (2015) also found that N2-P3b-N400 reflects the process of categorization in semantic category-based inductive reasoning. Some studies have shown, however, that P300 is the result of the N2—N400 continuous negative wave (Li et al., 2016); this reflects

the process of subjects comparing categories and properties simultaneously, because it would induce a continuous N2–N400 negative wave as compared to more complex dimensions (Bennett et al., 2014). In the present study, similar effects were also found in the premise 2 presentation phase. Thus, we posited that the findings not only reflect that P300 was the result of the N2–N400 continuous negative wave but also that the semantic distance between the categories had an effect.

N400 is generally considered to be related to the process of semantic categorization in inductive reasoning, and greater a N400 amplitude is considered to be elicited by unrelated categories as compared to related categories (Liang, Xiao, Lei, et al., 2020; Long et al., 2015). The results of this experiment showed that unrelated categories triggered a larger N400 amplitude than did related categories, which reflects the process of semantic categorization in the premise 2 presentation phase. Moreover, N400 is related to the associative strength of the category (Núñez-Peña & Honrubia-Serrano, 2005), which indicates that the associative strength of related categories is stronger than that of unrelated categories.

During the presentation of premise 2, negative arguments were found to trigger a greater P600 amplitude than did positive arguments; that is, the main effect of the arguments was significant. Previous studies showed that the LPC in the central and parietal regions reflects an illogical conclusion that induces a greater rear positive deflection than the logical conclusion in the task of transitive reasoning (Bonnefond, Castelain, et al., 2014; Bonnefond & Henst, 2013). This parietal positive deflection is similar to P600 and reflects the process of rule violation. Some studies have also indicated that P600 is related to rule violation (Núñez-Peña & Honrubia-Serrano, 2004). Niedeggen et al. (1999) found that the violation of mathematical rules induces an LPC more similar to P600 than does the compliance with mathematical rules. In addition, P600 has been shown to be related to the syntactic process or the semantic violation between verbs and opinions in a specific context, including when the subject relationship is involved (Hoeks et al., 2004; Kim & Osterhout, 2005). The results of the current research suggest that negative premises elicited a larger P600 amplitude than did positive premises, which indicates that the negative premises violated the inductive rule derived from premise 1. Moreover, the typical position of P600 generally appears in the central parietal region, and the maximum activation of P600 found in this experiment also appeared in the parietal region. The results also show that negative premises may change the semantic background induced by premise 1, especially the theme or background relationship between premise 1 and premise 2.

4.3 The analysis of the conclusion presentation phase

In the conclusion presentation phase, the difference between positive and negative arguments was observed in the P2 time window, and the maximum activation of P2 occurred in the frontal and parietal regions. According to previous studies P2 in the frontal region is more sensitive to differences in the comparison of features and relationships, and P2 in the frontal region reflects the activation of the cognitive matching system for the comparison of the current input and the information stored in memory (Wang et al., 2016). In the current experiment, when the conclusion was presented, a cognitive

mismatch appeared with premise 2 in the negative argument, which contributed to negative arguments eliciting a smaller P2 amplitude than did positive ones.

In addition, some studies have revealed that P2 is correlated with cognitive control; for example, Liang et al. (2016) found that the P2 amplitude induced by a typical category is larger than that induced by an atypical category. They posited that in categorical inductive reasoning, compared with typical category members, atypical category members require more cognitive control. In the current study, negative arguments elicited a smaller P2 amplitude than did positive arguments, which may be related to the need for more cognitive control in the processing of negative arguments. Negative arguments have more organizational semantic requirements than positive arguments, so they require more cognitive resources.

According to previous studies N2 may be associated with perceptual or rule mismatch in the reasoning process in inductive reasoning tasks. Perceptual or rule mismatch would produce a greater N2 amplitude than perceptual or rule matching (Prado et al., 2011), which reflects the early perceptual stimulation in the process of inductive reasoning (Liang et al., 2010). Long et al. (2015) found that in categorical inductive reasoning tasks, regardless of whether the conclusion category is related or unrelated to the premise category, inconsistent properties induce the classical frontal N2 as compared to consistent properties, reflecting the perceptual mismatch of inconsistent properties. The current study also shows that the amplitude of N2 induced by negative arguments was greater than that induced by positive arguments; that is, negative arguments needed more cognitive resources and attention consumption than positive arguments.

The ERP results suggest that at the conclusion presentation stage, the SN amplitude elicited by negative arguments was larger than that elicited by positive arguments during the 400–600 ms and 800–1000 ms time windows. We also found that the negative deflections elicited by negative arguments were greater than those elicited by positive arguments under the related-category condition. According to previous studies, SN is related to the more complex information integration and interpretation process driven by reasoning. For example, Pijnacker et al. (2011) found that in semantic reasoning, conclusions inconsistent with background information induce greater SN deflections than conclusions consistent with background information. They also believed that this component reflects the reasoning-driven information integration and interpretation process. In addition, the LNC is related to the formed hypothesis (Liang et al., 2010), the test hypothesis and the property conflict reasoning process (Long et al., 2015). Similar to the LNC, SN reflects that the subjects need to integrate categories and properties and finally judge the plausibility of arguments.

In summary, the ERP characteristics exhibited in the premise 2 presentation phase, the P300–N400 effect reflected in the categorization process, and P600 were found to account for the perceptual process of the property. In the conclusion phase, P2 and N2 reflected the early perceptual process of positive and negative properties while SN reflected the process of attribute reasoning. Moreover, P2, N2, and P300 are generally produced during the early stage of information input and processing; thus, the

three components could very possibly be caused by the contradiction between the previously formed anticipation and newly input knowledge. According to the EEG results, the subjects' inductive reasoning process was neither passive nor a gradual course of information integration following new input, but an active interaction between the existing expectation and new evidence. In other words, when reasoning, people not only expect the new evidence but also constantly try to form a new cognitive anticipation that is especially true at a given stage of reasoning. In the meantime, the formed cognitive anticipation may be further used as a mindset to filter and preprocess the new information. In the current study, this interactive model of "anticipation-input" was found to be strengthened by manipulating negative evidence, thus increasing the prominence of P2, N2, and P300 reflecting cognitive conflict and control in early information processing.

This study has some limitations. For instance, the category and property were presented simultaneously, so the categorization process could not be distinguished from the property inference process of reasoning. In future research exploring the ERP characteristics of the categorization and property inference processes would be worthwhile when the categories and attributes of reasoning are presented separately.

5. Conclusion

In this experiment the influence of negative evidence on categorical inductive reasoning was investigated by manipulating category and argument types. The results demonstrated that in general argument conditions, the RT of negative arguments was longer than that of positive arguments, the RT of different types of arguments was longer than that of the same type of argument; moreover, regarding argument strength we show that negative evidence could increase the argument strength under the unrelated-category condition, which is the same finding as that under the specific argument condition. Regarding the ERP characteristics, we show that an unrelated-category premise 2 elicited larger P300–N400 amplitudes, which may be correlated with the categorization process. In addition, our study shows that a negative premise 2 elicited a larger P600 amplitude than did positive premises, which may reflect the perceptual process of the property. Interestingly, in the conclusion presentation phase, compared with positive arguments, negative arguments elicited not only a larger P2 amplitude, which means that the premises and conclusions were matched, but also a larger N2 amplitude, which reflects the participants' early perceptual process of properties while SN reflected the process of property inferences.

Declarations

Declaration of competing interest

The authors declared that they had no conflict of interest concerning their authorship or the publication of this article.

Ethical approval and consent to participate

The studies involving human participants were reviewed and approved by the Science and Technology Ethics Committee of Hubei University of Medicine (No. : 2025-RE-033) following the Declaration of Helsinki; all participants signed informed consent forms in this study.

Consent for publication

Not Applicable.

Clinical trial number

Not applicable.

Funding

The present study is Supported by Key Research Center for Humanities and Social Sciences in Hubei Province (Hubei University of Medicine) (2024ZD003, 2025ZD001), China, and Faculty Development Grants from Hubei University of Medicine (2023QDJRW01), China.

Author Contribution

J.Z. conceived and designed the study, collected the data, performed the formal analysis, and drafted the manuscript. J.W. supervised the study, contributed to the study design, interpreted the data, and critically revised the manuscript. X.W. assisted with data collection and experimental implementation. P.L. contributed to data analysis and manuscript revision. All authors read and approved the final manuscript.

Acknowledgement

We thank all individuals who contributed to this research but are not listed as authors.

Data Availability

The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

References

1. Bennett MA, Duke PA, Fuggetta G. Event-related potential N270 delayed and enhanced by the conjunction of relevant and irrelevant perceptual mismatch: ERPs of perceptual mismatch. Psychophysiology. 2014;51(5):456–63. https://doi.org/10.1111/psyp.12192.
2. Bonnefond M, Castelain T, Cheylus A, Van der Henst J-B. Reasoning from transitive premises: An EEG study. Brain Cogn. 2014;90:100–8. https://doi.org/10.1016/j.bandc.2014.06.010.
3. Bonnefond M, Henst J-BVder. Deduction electrified: ERPs elicited by the processing of words in conditional arguments. Brain Lang. 2013;124(3):244–56. https://doi.org/10.1016/j.bandl.2012.12.011.
4. Bonnefond M, Kaliuzhna M, Van der Henst J-B, De Neys W. Disabling conditional inferences: An EEG study. Neuropsychologia. 2014;56:255–62. https://doi.org/10.1016/j.neuropsychologia.2014.01.022.
5. Bonnefond M, Van der Henst J-B. What's behind an inference? An EEG study with conditional arguments. Neuropsychologia. 2009;47(14):3125–33. https://doi.org/10.1016/j.neuropsychologia.2009.07.014.
6. Chen Q, Li P, Xi L, Li F, Lei Y, Li H. How do taxonomic versus thematic relations impact similarity and difference judgments? An ERP study. Int J Psychophysiol. 2013;90(2):135–42. https://doi.org/10.1016/j.ijpsycho.2013.06.015.
7. Collins AM, Loftus EF. A spreading-activation theory of semantic processing. Psychol Rev. 1975;82(6):407–28. https://doi.org/10.1037/0033-295X.82.6.407.
8. Cui R, Liu Y, Long C. FN400 and sustained negativity reveal a premise monotonicity effect during semantic category-based induction. Int J Psychophysiol. 2018;134:108–19. https://doi.org/10.1016/j.ijpsycho.2018.10.011.
9. Delorme A, Makeig S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods. 2004;134(1):9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.
10. Evans J, St. BT. In two minds: Dual-process accounts of reasoning. Trends Cogn Sci. 2003;7(10):454–9. https://doi.org/10.1016/j.tics.2003.08.012.
11. Feeney A. How many processes underlie category-based induction? Effects of conclusion specificity and cognitive ability. Mem Cognit. 2007;35(7):1830–9. https://doi.org/10.3758/BF03193513.
12. Feeney A, Coley JD, Crisp A. The relevance framework for category-based induction: Evidence from garden-path arguments. J Exp Psychol Learn Mem Cogn. 2010;36(4):906–19. https://doi.org/10.1037/a0019762.
13. Feeney A, Heit E. Properties of the diversity effect in category-based inductive reasoning. Think Reason. 2011;17(2):156–81. https://doi.org/10.1080/13546783.2011.566703.
14. Folstein JR, Van Petten C. Influence of cognitive control and mismatch on the N2 component of the ERP: A review. Psychophysiology. 2007;070915195953001. https://doi.org/10.1111/j.1469-8986.2007.00602.x.

15. Goel V. Differential involvement of left prefrontal cortex in inductive and deductive reasoning. Cognition. 2004;93(3):B109–21. https://doi.org/10.1016/j.cognition.2004.03.001.
16. Goel V. Anatomy of deductive reasoning. Trends Cogn Sci. 2007;11(10):435–41. https://doi.org/10.1016/j.tics.2007.09.003.
17. Goel V, Gold B, Kapur S, Houle S. The seats of reason? An imaging study of deductive and inductive reasoning. NeuroReport. 1997;8(5):1305–10. https://doi.org/10.1097/00001756-199703240-00049.
18. Heit E. Properties of inductive reasoning. Psychon b Rev. 2000;7(4):569–92. https://doi.org/10.3758/BF03212996.
19. Heussen D, Heussen D, Voorspoels W, Voorspoels W. 2011. Can similarity-based models induction handle Negat evidence? 6.
20. Heussen D, Voorspoels W, Verheyen S, Storms G, Hampton JA. Raising argument strength using negative evidence: A constraint on models of induction. Mem Cogn. 2011;39(8):1496–507. https://doi.org/10.3758/s13421-011-0111-2.
21. Hoeks JCJ, Stowe LA, Doedens G. Seeing words in context: The interaction of lexical and sentence level information during reading. Int J Psychophysiol. 2004;19(1):59–73. https://doi.org/10.1016/j.cogbrainres.2003.10.022.
22. Isabel Núñez-Peña M, Luisa Honrubia-Serrano M. N400 and category exemplar associative strength. Int j Psychophysiol. 2005;56(1):45–54. https://doi.org/10.1016/j.ijpsycho.2004.09.006.
23. Janiszewski C, Wyer RS. Content and process priming: A review. J Consum Psychol. 2014;24(1):96–118. https://doi.org/10.1016/j.jcps.2013.05.006.
24. Kalish CW, Lawson CA. Negative evidence and inductive generalisation. Think Reason. 2007;13(4):394–425. https://doi.org/10.1080/13546780701273402.
25. Kim A, Osterhout L. The independence of combinatory semantic processing: Evidence from event-related potentials. J Mem Lang. 2005;52(2):205–25. https://doi.org/10.1016/j.jml.2004.10.002.
26. Lee JC, Lovibond PF, Hayes BK, Navarro DJ. Negative evidence and inductive reasoning in generalization of associative learning. J Exp Psychol Gen. 2019;148(2):289–303. https://doi.org/10.1037/xge0000496.
27. Lei Y, Liang X, Lin C. How do the hierarchical levels of premises affect category-based induction: Diverging effects from the P300 and N400. Sci Rep. 2017;7(1):11758. https://doi.org/10.1038/s41598-017-11560-y.
28. Lei Y, Wang J, Zhu Y, Chen Q, Li H. P300 and positive slow waves reveal the plausibility in inductive reasoning. Psychophysiology. 2019;56(6):e13337. https://doi.org/10.1111/psyp.13337.
29. Li J, Chen A, Chen J, Long C. The neural signs of categorization and property inferences during verbal category-based properties induction. Acta Physiol Sinica. 2016;48(11):1410. https://doi.org/10.3724/SP.J.1041.2016.01410.
30. Liang P, Zhong N, Lu S, Liu J. ERP characteristics of sentential inductive reasoning in time and frequency domains. Cogn Syst Res. 2010;11(1):67–73.

https://doi.org/10.1016/j.cogsys.2008.10.001.

31. Liang X, Chen Q, Lei Y, Li H. How types of premises modulate the typicality effect in category-based induction: Diverging evidence from the P2, P3, and LPC effects. Sci Rep. 2016;6(1):37890. https://doi.org/10.1038/srep37890.
32. Liang X, Xiao F, Lei Y, Li H, Chen Q. N400/frontal negativity reveals the controlled processes of taxonomic and thematic relationships in semantic priming for artifacts. Psychophysiology. 2020;57(2). https://doi.org/10.1111/psyp.13486.
33. Liang X, Xiao F, Zhu Y, Lei Y, Chen Q. How types of prior knowledge and task properties impact the category-based induction: Diverging evidence from the P2, N400, and LPC effects. Biol Psychol. 2020;156:107951. https://doi.org/10.1016/j.biopsycho.2020.107951.
34. Long C, Lei X, Chen J, Chang Y, Chen A, Li H. Event-related potential parameters of category and property violations during semantic category-based induction. Int j Psychophysiol. 2015;96(3):141–8. https://doi.org/10.1016/j.ijpsycho.2015.04.005.
35. Long C, Li X, Lei Y. Event-related potential parameters related to typicality effects during category-based induction. Sci Sin -Vitae. 2017;47(3):325–40. https://doi.org/10.1360/N052016-00297.
36. Medin DL, Coley JD, Storms G, Hayes BL. A relevance theory of induction. Psychon b Rev. 2003;10(3):517–32. https://doi.org/10.3758/BF03196515.
37. Niedeggen M, Rösler F, Jost K. Processing of incongruous mental calculation problems: Evidence for an arithmetic N400 effect. Psychophysiology. 1999;36(3):307–24. https://doi.org/10.1017/S0048577299980149.
38. Núñez-Peña MI, Honrubia-Serrano ML. P600 related to rule violation in an arithmetic task. Cogn Brain Res. 2004;18(2):130–41. https://doi.org/10.1016/j.cogbrainres.2003.09.010.
39. Osherson DN, Smith EE, Wilkie O, López A, Shafir E. Category-based induction. Psychol Rev. 1990;97(2):185–200. https://doi.org/10.1037/0033-295X.97.2.185.
40. Pijnacker J, Geurts B, van Lambalgen M, Buitelaar J, Hagoort P. Reasoning with Exceptions: An Event-related Brain Potentials Study. J Cogn Neurosci. 2011;23(2):471–80. https://doi.org/10.1162/jocn.2009.21360.
41. Prado J, Chadha A, Booth JR. The Brain Network for Deductive Reasoning: A Quantitative Meta-analysis of 28 Neuroimaging Studies. J Cogn Neurosci. 2011;23(11):3483–97. https://doi.org/10.1162/jocn_a_00063.
42. Rein JR, Goldwater MB, Markman AB. What is typical about the typicality effect in category-based induction? Mem Cognition. 2010;38(3):377–88. https://doi.org/10.3758/MC.38.3.377.
43. Rips LJ. Two kinds of reasoning. Psychol Sci. 2001;12(2):129–34. https://doi.org/10.1111/1467-9280.00322.
44. Sloman SA. Feature-Based Induction. Cogn Psychol. 1993;25(2):231–80. https://doi.org/10.1006/cogp.1993.1006.

45. Voorspoels W, Navarro DJ, Perfors A, Ransom K, Storms G. How do people learn from negative evidence? Non-monotonic generalizations and sampling assumptions in inductive reasoning. Cogn Psychol. 2015;81:1–25. https://doi.org/10.1016/j.cogpsych.2015.07.001.
46. Wang X, Tao Y, Tempel T, Xu Y, Li S, Tian Y, Li H. 2016. Categorization Method Affects the Typicality Effect: ERP Evidence from a Category-Inference Task. *Front Psycho*, *7*. https://doi.org/10.3389/fpsyg.2016.00184
47. Zhong L, Lee MS, Huang Y, Mo L. Diversity Effect in Category-Based Inductive Reasoning of Young Children: Evidence from Two Methods. Psychol Rep. 2014;114(1):198–215. https://doi.org/10.2466/10.11.PR0.114k17w3.

Figures

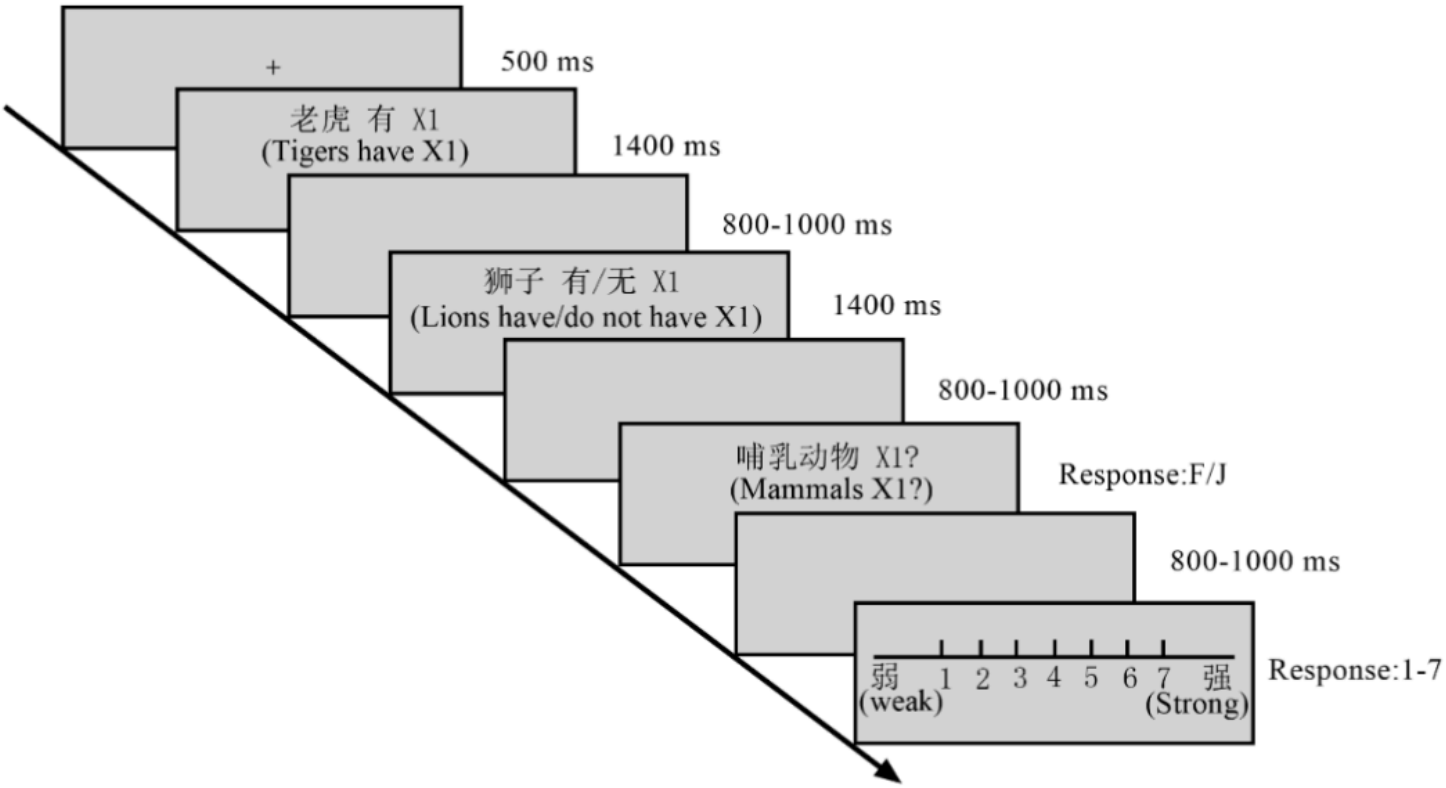


Figure 1

Experimental procedures for the category-based induction

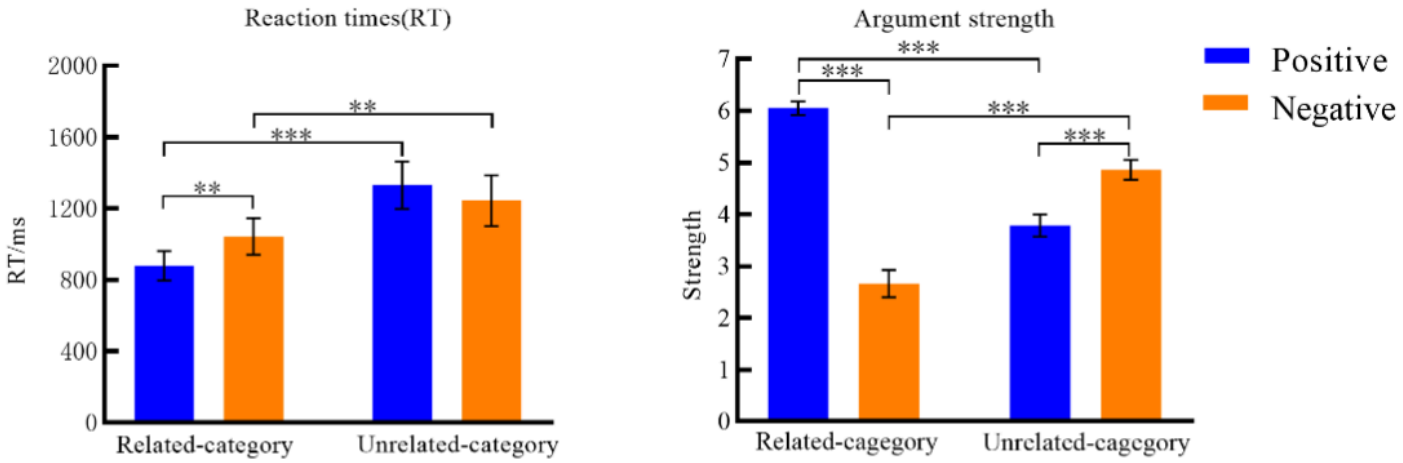


Figure 2

Mean RTs and argument strength in different conditions.

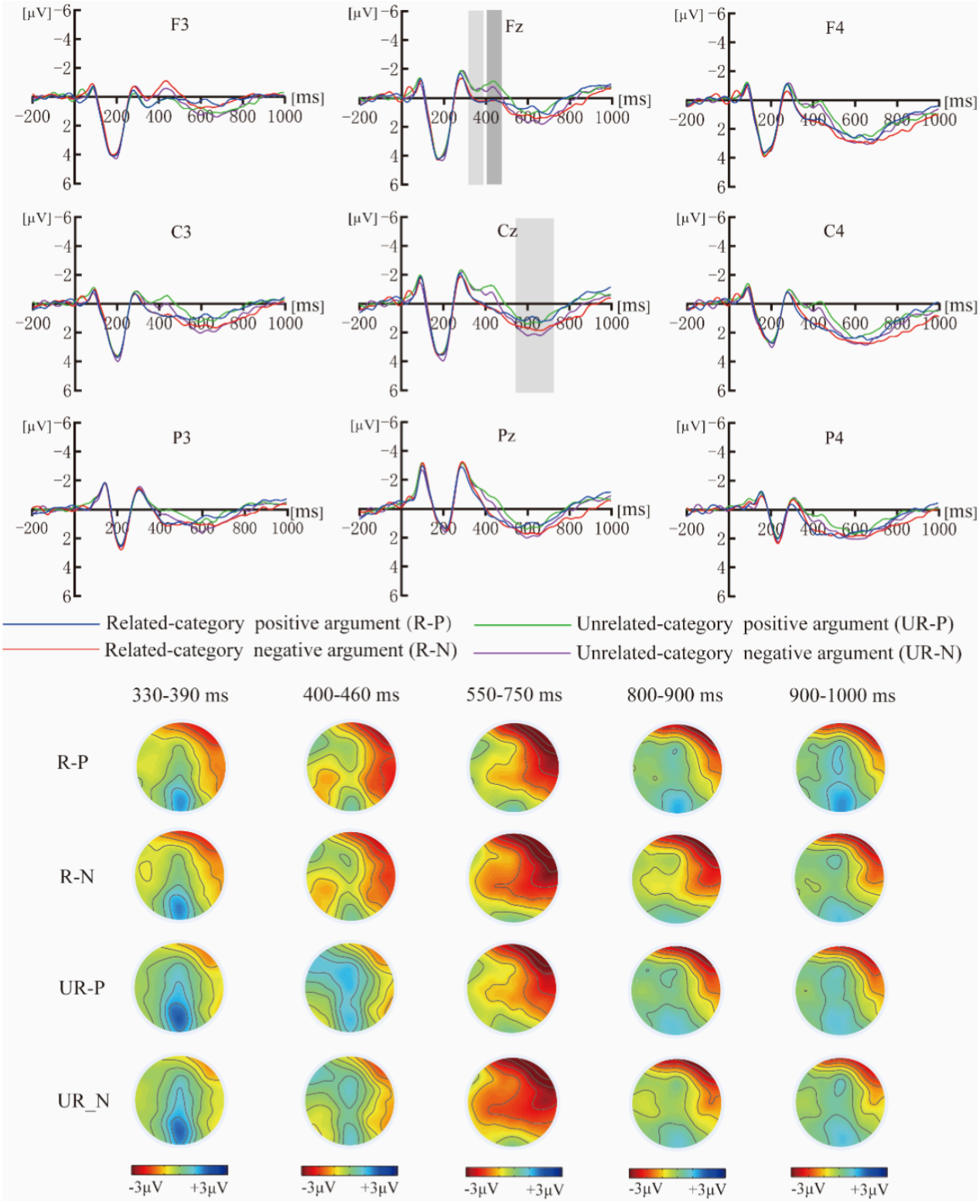


Figure 3

Waveforms and topographies elicited by the different conditions of the premise 2 phase. The original potentials for different conditions at F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, and topographies of P3, N400 and P600 for the R-P (related-category positive argument), R-N (related-category negative argument), UR-P (unrelated-category positive argument), UR-P (unrelated-category positive argument) conditions.

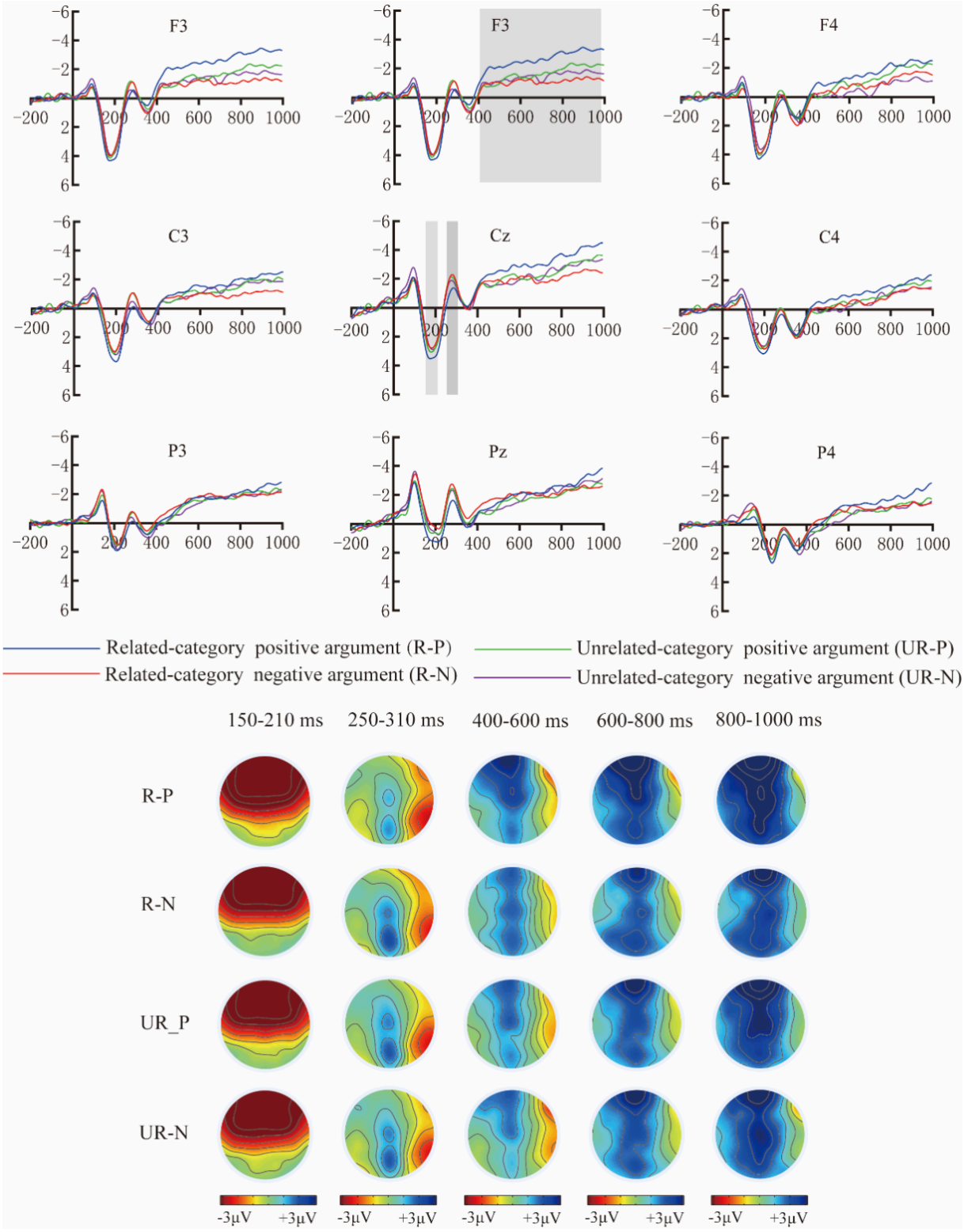


Figure 4

Waveforms and topographies elicited by the different conditions of the conclusion phase. The original potentials for different conditions at F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, and topographies of, N2 and SN for the R-P (related-category positive argument), R-N (related-category negative argument), UR-P (unrelated-category positive argument), UR-P (unrelated-category positive argument) conditions.