How do humans process contradictory cues - an EEG-ERP study testing hypotheses derived from classical and non-classical logics

Abstract

Research on the cognitive processing of contradictions has predominantly focused on rhetorical or context-dependent inconsistencies rather than strict logical contradictions of the form $(p \wedge \neg p)$. Consequently, empirical evidence regarding how the human cognitive system handles genuinely inconsistent information remains limited. This study addresses this gap by investigating the neural processing of true logical contradictions within a predictive coding framework, where logical conclusions are operationalized as sensory predictions. We derived competing hypotheses from classical and non-classical logic: (H1) *the principle of explosion* from classical logic, which posits that from a contradiction anything follows, implying the brain generates multiple, mutually exclusive predictions when facing contradictory premises; and (H2) *the principle of implosion* from the non-Fregean logic of content, which suggests that contradictions lead to a suspension of inference and an inability to form any clear predictions.

We conducted an EEG experiment with N = 30 participants using a visual oddball task. Participants learned associations between letter cues and target shapes (e.g., symbol "A" predicted a smooth shape, while "B" predicted a spiky shape). The experiment included consistent cues ("AA", "BB"), contradictory cues ("AB", "BA"), and novel/meaningless cues ("C"). Brain activity was recorded via 32-channel EEG, focusing on event-related potentials (ERPs) indicative of prediction errors and cognitive load: the visual mismatch negativity (vMMN), P3b, and the Late Positive Component (LPC).

The results showed that targets following consistent cues but violating learned associations elicited a significant vMMN at right-frontal electrodes (F8, FT10), reflecting a standard prediction error. Crucially, targets following contradictory cues did not elicit a vMMN, supporting the principle of implosion (H2) where the predictive system fails to generate specific expectations. Interestingly, novel-/meaningless cues did elicit a significant vMMN, suggesting they were treated as cues for potential future events rather than total inferential dead-ends. Furthermore, both contradictory and meaningless cues elicited significantly enhanced P3b and LPC amplitudes compared to consistent trials, indicating that while contradictions may not generate predictions, they prompt the allocation of additional conscious resources for evaluation and working memory processing.

Exploratory analysis of individual differences revealed that a subset of participants (n = 13) employed explicit strategies to sidestep contradictions, such as focusing on only one part of the "AB" cue. For these individuals, neural activity

revealed a significant vMMN when the subsequent target violated their self-reported strategy. This indicates that these participants effectively "removed" the contradiction to restore consistency and form single-cue predictions. However, even for strategy-congruent targets, a weakened prediction error was observed, suggesting that the effort to ignore contradictory information was only partially successful.
These findings suggest that for intramodal visual information, the human cognitive system tends toward implosion rather than explosion. While cross-modal studies suggest that conflicting predictions can coexist across different sensory streams, our results indicate that contradictions within a single modality disrupt the formation of predictive models. This supports the utility of non-classical frameworks, such as the logic of content, for modeling human information processing under inconsistency.
Beyond the specific empirical findings regarding contradictions, this study serves as a proof-of-concept for the broader feasibility of "psychologistic logic" - a framework that utilizes formal logical systems as descriptive tools to model the architecture of human information processing. By adopting a broad definition of reasoning that encompasses both deduction and unconscious sensory processing, the research demonstrates how logic can bridge the historical rift between formal philosophy and cognitive science. Central to this approach is the methodological use of paradoxes and contradictions as "edge cases". Much like extreme phenomena in natural sciences, logical anomalies in psychology can allow researchers to identify which specific logical properties - such as paraconsistency or the principle of explosion - accurately reflect the constraints of human cognitive architecture.

Keywords: contradiction, EEG, principle of explosion, principle of implosion, predictive coding, logic, psychologism