Suspending Causality:
Predictive Processing and Moral Valence in Narrative Comprehension

Erzsébet Szabó
Universtity of Szeged, Hungary

Recent cognitive reading experiments suggest that narrative works offer models of the social world via abstraction. (cf. Mar & Oathley 2008, Zunshine 2006, Willems 2015) By presenting multiple interacting characters, narratives function as social simulations; because these characters make emotionally, morally, and ideologically charged decisions, narratives simultaneously model choices, moral reasoning, and ideological processes - in short, the workings of emotional, value, and ideological systems.

Most theorists agree that narrative texts are tailored to the reader's mind, that is, narrative information is arranged to guide readers along a specific cognitive and emotional trajectory toward meaning (cf. Gerrig 1993). Yet considerable debate remains about which reader- or text driven mechanisms are decisive in this process. One of the most influential recent approaches is Karin Kukkonen's predictive processing model (2020), which integrates probabilistic causal inference with affective and value-based prediction within a unified framework.

Drawing on neuropsychology (Frith 2008; Friston 2008, 2012), cognitive and developmental psychology (Seth 2013; Garfinkel et al. 2015; Weisberg & Gopnik 2013; Tenenbaum et al. 2011; Dehaene 2013), and philosophy (Clark 2013, 2016; Hohwy 2013), Kukkonen argues that reading involves continuously generated predictions weighted by expected precision across multiple textual levels: words, sentences, rhythm, style, episodes, and plot. When incoming information either deviates from the predicted content or diverges from the expected reliability (precision) of that prediction, a prediction error occurs, triggering revisions. Literary texts, through their deliberately crafted probability design, make strategic use of both content- and precision-based mismatches to shape understanding and interpretation.

This model has several explanatory strengths. Chief among them is its integration of the emotional dimensions of comprehension. Kukkonen emphasizes that the predictive design of the text offers not only probabilistic guidance but also desire-based predictions. Readers do not merely attempt to predict what will happen; they also hope for or long for particular outcomes (e.g., the triumph of justice, the punishment of the guilty, or the survival of the innocent). These desire-based expectations can effectively override probabilistic assessments: the further readers come from the desired outcome, the more willing they become to accept unlikely, causally loose, or even impossible events as part of the narrative. Literary texts deliberately exploit this suspension of causal expectations; this is not a flaw, but a feature of how predictive and value-driven processing combine. Consequently, the reader's "explanatory system" functions as a complex cognitive mechanism that integrates cognitive, emotional, and moral components, accepting explanations even when the causal network is disrupted if those explanations satisfy moral or affective needs.

This paper examines the dynamics of predictive processing and causal suspension in Edgar Allan Poe's *The Pit and the Pendulum*, showing how the story manipulates readers 'moral desires and causal expectations to sustain tension and emotional engagement.