# A Predictive Processing-Based Explanatory Model of Role Exchange

**Authors:** Ye Fang, Qiu Huilin, Jiang Ke, Li Changjin, Jiang Ke, Li Changjin

**Date:** 2023-01-14T00:00:00+00:00

## Abstract

Role reversal is one of the fundamental techniques in psychodrama therapy. Based on Moreno' s classical theory, a role constitutes the self-expression of the client. While contemporary psychodrama therapy practice has accumulated rich clinical experience building upon Moreno' s theory, it lacks an explanatory framework that can integratively account for the working mechanisms of psychodrama therapy. Predictive processing theory, through analyzing the role reversal experience, elucidates the interactive and constructive relationship between the 'protagonist' in the individual' s own position and the 'auxiliary' in the reversed position during the role reversal process. According to predictive processing theory, role reversal is a self-construction process. At different levels ranging from a single role reversal therapeutic session to the entire therapeutic process, this self-construction process is manifested respectively as: (1) role reversal produces changes in signal types; (2) role reversal brings about model updating; (3) new interpersonal models are generated through interactive construction. In summary, employing predictive processing theory to explain the structured characteristics of role reversal will facilitate this technique in achieving better therapeutic effects in clinical practice.

## Full Text

## Preamble

### Interpretation Model of Role Reversal Based on Predictive Processing Theory

Ye Fang[1,2,3], Qiu Huilin[4], Jiang Ke[1,2,3], Li Changjin[1,2,3]
[1]School of Mental Health, Wenzhou Medical University, Wenzhou 325035, China
[2]Key Laboratory of Alzheimer' s Disease of Zhejiang Province, Institute of Aging, Wenzhou Medical University, Wenzhou 325035, China

[3]Corning Hospital Affiliated to Wenzhou Medical University, Wenzhou 325000, China
[4]College of Foreign Languages and Cultures, Sichuan University, Chengdu 610207, China

## Abstract

Role reversal is one of the fundamental techniques in psychodrama therapy. According to Jacob Levy Moreno' s classic theory, the self emerges through the roles we enact. Contemporary psychodrama practice has accumulated rich clinical experience based on Moreno' s theory, yet lacks an explanatory framework to comprehensively elucidate its therapeutic mechanisms. Predictive processing theory, starting from an analysis of the role reversal experience, explains the interactive and constructive relationship between the "protagonist"in the individual' s own position and the "antagonist" in the exchanged position. According to predictive processing theory, role reversal constitutes a self-construction process. Across different levels—from a single role reversal session to the entire therapeutic process—this self-construction manifests as: (1) changes in signal types, (2) model updating, and (3) generation of new interpersonal models through interactive construction. In summary, applying predictive processing theory to explain the structural features of role reversal can enhance the effectiveness of this technique in clinical practice.

**Keywords:** psychodrama, role reversal, role construction, self, predictive processing

Role reversal is a core technique employed in psychodrama counseling, primarily applied to interpersonal relationship difficulties. In interpersonal contexts, the client is termed the "protagonist," while the person with whom they experience conflict is called the "antagonist." In clinical practice, the client first enacts their own role in a specific scenario, then shifts positions to portray the antagonist' s role. While role reversal belongs to the broader category of role-playing techniques, it adds a crucial transformational step that requires the actor to adopt the antagonist' s perspective and mode of thinking and acting.

Jacob Levy Moreno (1946/1961), the founder of psychodrama, conceptualized "role" as a form of interpersonal experience through which individuals understand relationships and ultimately achieve greater self-understanding. Moreno' s role theory distinguishes three types: psychosomatic roles (the earliest roles developed after birth, encompassing basic physiological self-experiences and somatic responses triggered by psychological experiences), social roles (emerging from interactions with others such as mothers or friends), and psychodramatic roles (where individual experiences become objectified as externalized roles that can dialogue with the self, such as personifying one' s "pain" to converse with it). Moreno argued that roles contain different action components—collective action components (socially and culturally prescribed patterns) and personal action components (creative patterns formed from individual experience). He

maintained that the more roles an individual possesses, the better their quality of life (Moreno, 1946/1969).

Psychodrama emphasizes three factors: space, time, and action (Blatner, 2000). First, psychodrama requires a safe space where individuals can enact life roles without fear of punishment (Moreno, 1946; Moreno, 1965). Second, temporally, role reversal can bring past scenes, present states, or future possibilities onto the psychodrama stage as needed. Regardless of when the role enactment occurs, individuals may exhibit the same behavioral patterns in the here-and-now—a phenomenon Moreno called "psychological time travel."Third, Moreno valued the power of action, famously telling Freud: "You analyze people' s dreams; I teach them the courage to dream again" (Moreno, 1969). Through psychodramatic action coordination, imagined roles can be realized in lived reality, creating what Moreno termed "surplus reality" (Moreno, 1969).

Numerous studies confirm psychodrama's close relationship with self-experience, demonstrating its effectiveness in enhancing self-awareness (Boroomandian et al., 2020; Dogan, 2018), self-acceptance (Kaya et al., 2021), feelings of inferiority (Tümlü & Şimşek, 2021), and self-compassion (Abeditehrani et al., 2020). Research also links psychodrama to emotional capacity, showing it can reduce perceived stress (Boroomandian et al., 2020), enhance empathy (Kipper & Ritchie, 2003; Soysal, 2021), and improve emotional intelligence (Beauvais et al., 2019), particularly in alleviating anxiety (Dorothea, 2016; Rudokaite & Indriuniene, 2019). Studies indicate psychodrama' s efficacy in reducing symptoms across various conditions, including social anxiety disorder (Abeditehrani et al., 2020), schizophrenia (Miller et al., 2021), ADHD (Mojahed et al., 2021), phobias (Tarashoeva et al., 2017), substance dependence (Giacomucci & Marquit, 2020), and epilepsy (Chu et al., 2009; Han & Chu, 2007; Pan, 2015). As an expressive arts therapy, psychodrama has increasingly been combined with other interventions such as cognitive-behavioral therapy (CBT) and family therapy (Abeditehrani et al., 2020; Blatner, 1997; Farmer et al., 2018). Beyond clinical applications, Moreno' s role theory has made significant contributions to sociological theory and practice (Žurić-Jakovina & Jakovina, 2017; Moreno, 1961).

Current research predominantly treats psychodrama as an intervention method, focusing on pre-post differences through self-report measures to demonstrate efficacy (Chen et al., 2019). However, because psychodramatists typically combine and repeatedly use multiple techniques in a single session, few studies have investigated the mechanisms of single techniques.

As one of psychodrama' s most important techniques—called "the engine of psychodrama" —role reversal involves multiple relational experiences including empathy, compassion, and self-reflection (Karp et al., 1998). Clinical case analyses and research demonstrate its success as an intervention (Boies, 1972), and it can be applied in individual counseling without auxiliary egos (Blatner, 2007). While various operational models exist, research faces three challenges: (1) difficulty controlling for individual differences in personality and experience (Boies,

1972), (2) lack of clear mechanisms explaining how role reversal modifies behavioral patterns, making outcome assessment problematic, and (3) absence of a unified theory to integrate empirical findings and provide scientific explanation (Orkibi & Feniger-Schaal, 2019).

Role reversal' s effectiveness stems from enhancing the structural organization of individuals' psychological experiences of self and others during role-playing (Carlson-Sabelli, 1989). This paper introduces predictive processing theory to explain role reversal's mechanisms, clarifying the relationship between predictive mind and self-knowledge, and using self-construction as the main thread to explain how role reversal facilitates psychological structuring through predictive modeling.

## 2.1 Self-Knowledge Through the Predictive Mind

Extensive research demonstrates psychodrama' s close association with self-experience, showing its effectiveness in promoting structured development of self-concept (Butler & Fuhriman, 1980; Rohrbaugh & Bartels, 1975; Yalom, 1970). To explore the internal processes through which role reversal affects self-experience, we must first clarify the concept of "self-knowledge."

Self-knowledge has long been a "hard problem" in philosophy of mind and psychology (Fei, 2018), with diverse interpretations of self and self-experience. Analytic philosophy argued that individuals know themselves through interactional objects rather than consciousness or introspection (Ryle, 1992). These objects include not only others but also the knowing subject themselves. When individuals interact with others or themselves, they come to know others by observing their performances and experiences, and thereby know themselves. Certain actions—such as observing, opposing, imitating, or mocking—involve or are superimposed upon other actions, constituting "higher-order actions." By developing the capacity to take higher-order actions toward interactional objects and applying this capacity to self-interaction, individuals achieve self-knowledge through self-reflexivity (Mead, 1934), a distinguishing feature of humans (Baumeister, 1998).

Shavelson et al.' s (1976) multi-faceted hierarchical model of self remains widely accepted, defining self-concept as an organized, hierarchical, multidimensional, and distinguishable set of self-perceptions, and describing its generation from interpersonal interactions and perceptual experiences of self-attributes and social environments. However, how interpersonal interactions generate self-knowledge requires further theoretical explanation.

Clark's (2013) predictive processing theory has generated strong responses across philosophy of mind and cognitive science. The theory posits that in an uncertain world, individuals must construct models composed of probability density distributions of worldly regularities. Using these models, they generate top-down predictive signals that continuously match actual input signals. The discrepancy between them is called prediction error (PE), which feeds back to the

brain through bottom-up pathways, driving action to minimize PE. Through continuous model revision to coordinate with the real world, individuals reduce uncertainty (Clark, 2013/2017). As a new paradigm for understanding cognitive processes, predictive processing integrates neuroscience, philosophy, biology, and information theory, viewing the mind as an active "Bayesian inference engine" and offering the best current explanation of cognition (Wang, 2021).

The predictive processing approach (hereafter "predictive mind" ) provides a framework for explaining self-knowledge (Wang, 2021), but we must address how the predictive mind functions in self-knowledge acquisition.

Predictive processing theory maintains that modeling one' s own body during environmental interaction is a natural consequence of prediction error minimization (Hohwy & Michael, 2017). For example, when walking on a street, individuals must not only predict others' direction and speed but also have accurate bodily self-knowledge to maintain appropriate boundaries and avoid collisions.

According to predictive processing theory, the self is a hierarchical causal model hidden within the body, with lower-level selves (sensory experiences, bodily knowledge) regulated by higher-level selves (personality traits, beliefs). Prediction error feeds back from lower to higher levels, constructing stable self-features. An individual' s predictive signals constitute concrete self-representations, with higher-level signals depending on stable patterns in lower-level signals (Kwisthout et al., 2017). Based on the free energy principle, individuals achieve self-consistency and continuity through prediction error minimization (Hohwy & Michael, 2017). In short, the predictive mind knows itself through: (1) self-representation via predictive signal generation, and (2) new self-representational modeling needs arising from matching between predictive and input signals. These processes cycle repeatedly until prediction error is minimized.

Self-knowledge is crucial for mental and physical health. Erikson' s identity theory posits that individuals must continuously integrate experiences of self and object, id and superego, to develop a sense of internal consistency and continuity, and adaptive responses to social environments (Erikson, 1959). Incomplete self-knowledge development, characterized by unclear or discontinuous self-concept, can lead to interpersonal conflicts, negative emotions like anxiety and disappointment, and even psychological disorders (Han, 2021; Kellermann, 1994). Negative mood states may trigger cognitive immunity—denying or weakening prediction errors from unreasonable expectations and refusing to modify entrenched generative models. For example, in major depressive disorder, patients with initial negative expectations may maintain them by cognitively immunizing against incongruent positive signals, while those with initial positive expectations may amplify prediction errors from negative signals, shifting expectations negatively. This attentional bias reduces accurate self-knowledge, creating a negative cycle (Kube et al., 2020).

Psychodrama's role reversal technique helps individuals accept objectively exist-

ing prediction errors through dialogue between protagonist and antagonist. For patients' stereotypical negative expectations in negative mood states, role reversal breaks the negative cycle of cognitive immunity, facilitating timely updating of negative expectations. For pre-existing positive expectations, role reversal guides individuals to spontaneously generate diverse evaluations of negative events while coordinating input signals with external reality, reducing negative experiences from discrepancies between prior positive expectations and posterior negative events, ultimately integrating diverse experiences into a unified self and enhancing mental health.

## 2.2 Role Reversal as Predictive Model Construction

The theoretical foundation of role reversal practice rests on psychodrama' s role theory, whose basic views on self-concept align closely with predictive processing theory (Li, 2021; Hohwy, 2020), enabling theoretical explanation of psychodrama' s mechanisms.

First, individuals' internal generative models are organized by roles. Moreno defined role as a symbolic representation of individual functional forms—the tangible aspects of self that actually and concretely exist within the self (Moreno, 1946/1961). When introducing themselves or learning about others, people tend to report roles such as "student" or "mother." Roles provide the brain with possibilities for model construction, including prior knowledge about the role, generation of prior hypotheses or predictions about others, and establishment of adaptive interpersonal relationships through prediction error minimization (Lee et al., 2021). Role stereotypes exemplify the predictive mind in action. Inadequate role knowledge or limited role repertoires can cause mental health problems. For instance, Levin' s (2021) case analysis shows that mothers experiencing rapid role and life changes during postpartum care may develop postpartum depression if they fail to manage inter-role relationships and transitions. In workplace settings, accurate self-knowledge and role positioning are essential for professional women' s mental health (Chen, 2012; Ma & Zhang, 1999). Role reversal enables individuals to try different roles, promoting self-knowledge and facilitating decision-making when role conflicts arise.

Second, the relationship between generative models and the world is fundamentally constructive. Self-constructiveness was nascent in early psychoanalytic theory—Freud' s classic theory viewed the ego as formed through id-environment interaction, a compromise between desire and parental constraints. Humanistic self-psychology similarly derives psychological and social selves from individuals projecting actions onto others or environments and receiving feedback. Despite varied descriptions across theoretical schools, none reject the constructivist premise that the self emerges from individual action-environment interaction. Moreno explicitly stated that the self develops from socially ascribed roles, not vice versa (Moreno, 1961). Roles are interdependent, with each role having a counterpart (e.g., husband-wife, boss-employee). Self-knowledge of a particular role depends on interaction with its counterpart. For example, a child constantly

receiving negative maternal evaluations ( “You’ re stupid, you can’ t do anything right” ) likely internalizes this self-concept. Similarly, father-daughter relationships construct internal predictive models that profoundly influence adult intimate relationships, as daughters may seek partners resembling or opposite to their fathers based on these models.

In summary, individuals connect their inner and outer worlds through the predictive mind, and role reversal constitutes a predictive model construction process. Role reversal encourages each role to understand the counterpart’ s perspective and find peaceful coexistence, promoting mental health development. Moreno (1972) considered role reversal ability crucial for children’ s socialization, helping them move beyond egocentrism. Mann’ s (1959) research showed that role reversal groups demonstrated significant improvements over control groups in “attractiveness as a friend,” “helping group achieve goals,” and “cooperativeness.” Role reversal facilitates interpersonal regulation—Abeditehrani et al. (2020) applied it to social anxiety disorder, targeting negative cognitions about others, and found significant reductions in “confidence in negative cognitions,” “probability of negative evaluation,” and “cost of negative evaluation” compared to role-playing without reversal, supporting its effectiveness.

Psychodrama’ s successful application of role reversal across populations demonstrates its therapeutic value (Boies, 1972), yet lacks adequate theoretical explanation. Orkibi and Feniger-Schaal’ s (2019) systematic review of psychodrama research from 2007-2017 critically evaluated methodologies and emphasized the need for theoretical explanations of psychodrama’ s specific mechanisms. This paper aims to introduce predictive processing theory to explain role reversal’ s unique therapeutic effects and provide a hypothetical framework for future empirical research.

## 3 Predictive Processing Theory’ s Explanation of Role Reversal

Predictive processing theory posits that the brain does not passively receive external stimuli but actively generates expectations about upcoming input signals (Clark, 2018; Kwisthout et al., 2017). The comparison between predictive and input signals creates prediction error. Following the free energy principle (Friston, 2010; Ransom et al., 2020; Wiese & Metzinger, 2017), individuals minimize prediction error to enhance environmental adaptation. When facing prediction error, individuals either act to satisfy their needs or adjust expectations to match the environment, depending on precision weighting of predictive and input signals. Higher precision means less uncertainty, influencing how individuals balance top-down and bottom-up processes (Clark, 2013/2017).

Role reversal occurs in reciprocal relationships between protagonist and antagonist, each possessing models for understanding reality that are often inconsistent or contradictory. Individuals generate predictive signals from their own role models, receive input signals from the other, and experience cognitive disso-

nance when top-down predictions and bottom-up inputs mismatch, making self-integration difficult. Role reversal aims to help individuals structure prediction errors through the "protagonist-antagonist-protagonist" transformation process, creating opportunities for model reconstruction, promoting self-awareness and growth, and ultimately eliminating prediction error.

This paper proposes a three-level explanatory model of role reversal from signal, model, and interaction perspectives, addressing both single-session effects on the predictive mind and cumulative effects of multiple sessions on model updating. The model explains how models change within a single psychodrama session and across the entire therapeutic process, exploring the mechanisms underlying role reversal' s effectiveness.

### 3.1 Role Reversal Induces Signal Changes

Daily life predictions occur in two forms (Ryle, 1992). First, when individuals consider explicit propositions, action outcomes closely match expectations, reducing surprise. Second, predictions can manifest as a mood where individuals remain aware of their actions without explicit propositions, yet conclusions don' t surprise them. In this second form, higher-order actions aren' t always implied, so prediction errors don' t reduce surprise, leaving individuals unable to adopt effective strategies for minimizing error when facing familiar interpersonal situations, ultimately causing emotional problems.

Due to the elusive nature of the "I" system, I can chase my past self but only grasp the fluttering hem of my coat (Sartre, 1957). The distinction between "I" and "you" is that I can never grasp my current self but can grasp yours (the antagonist). Role reversal enacts specific life situations where the protagonist' s signals are predictive and received antagonist signals are input. When shifting to the antagonist position, past input signals become predictive signals, and past predictive signals become input signals. The object of examination shifts from self to other, enabling the individual to know their own existence and behavior as an object, thereby elevating self-knowledge. This process is illustrated in Figure 1 [Figure 1: see original paper].

Consider a common parent-child relationship example: The protagonist is a daughter whose antagonist is her mother. The presenting problem is that after leaving home for university, the mother calls multiple times daily to check on her, which the daughter experiences as intrusive, prompting arguments. The daughter holds a model of parent-child relationships including appropriate boundaries and communication frequency. When this model functions normally, information feels familiar and non-problematic. However, when she leaves for university, discrepancies between her model and her mother' s model disrupt her equilibrium, creating prediction error. Her input signal is "Mother is disturbing my life, she doesn' t care about me or my feelings, only her own needs," while her predictive signal is "Mother should care about me, consider my feelings, and not disturb my life with constant calls" (see Figure 1a). Due to low spontaneity

and limited role repertoire, her received input signals are often biased.

When she shifts to the antagonist (mother) position, input becomes predictive signal. Now "My daughter doesn't understand me; my care is rejected" becomes input signal, while "My daughter has always been physically weak, I must remind her to take care of herself, calling more is necessary, she should understand my concern" becomes predictive signal (see Figure 1b). At session conclusion, when she returns to protagonist, the predictive signal becomes input signal, transformed from "Mother constantly calls, disturbing my life, she doesn't care" to "Mother calls repeatedly because she worries, she doesn't know this disturbs me, she does it out of love" (see Figure 1c). This modified input signal aligns with the predictive signal "Mother should care about me, consider my feelings, and not disturb my life with multiple daily calls," reducing prediction error and fostering adaptive relationships that decrease interpersonal distress. Role reversal helps clarify how differences between self and maternal roles implicitly influence interaction patterns, understand conflict origins, and identify appropriate problem-solving strategies.

To effectively achieve signal transformation, counselors must facilitate adequate warm-up before role reversal. Following Kellermann's (1994) three-stage model, signal change occurs in three steps: First, the client physically mimics the antagonist's movements and language to inhabit the unfamiliar role—empathic role taking, the first step in signal change. Second, the counselor collects case-relevant information through relational questioning, presenting background information about the antagonist's predictive signals. The final stage is role feedback, where input signals from the antagonist position include "how I (as antagonist) perceive you (protagonist)," while predictive signals include "how I (as antagonist) perceive you (protagonist) perceiving me (antagonist)." These three steps interactively reveal implicit conflict patterns through signal transformation.

**Figure 1. Signal Type Changes in Role Reversal**
*Note: (a) shows the client in protagonist position, (b) after shifting to antagonist, (c) after returning to protagonist. Signal types differ across role reversal stages. Generative Model' refers to the antagonist' s best guess about the world, distinct from the protagonist' s generative model.*

### 3.2 Role Reversal Leads to Model Updating

Role reversal doesn't require completely abandoning one's model but, following creative principles, integrates adaptive aspects of the antagonist's model into one's own (Kellermann, 1994). Clients must simultaneously activate representations of self and other. Shifting to the antagonist position requires temporarily suppressing self-feelings to present experiences and actions different from one's authentic feelings—a creative process demanding high-level volitional control, unlike simple role-playing (Yaniv, 2012). Role reversal provides a real-world link between protagonist and antagonist models, raising the question: How is

prediction error minimized during this process?

Rothenberg (1971) proposed the concept of homospatial process to clarify the distinction between creativity and psychotic symptoms. Homospatial process involves two concrete entities simultaneously occupying one space to produce a new entity. For example, German organic chemist Kekulé discovered benzene' s structure by superimposing an image of a snake biting its tail onto atomic representation. Dani Yaniv (2012) suggests homospatial process may underlie role reversal, with protagonist and antagonist models coexisting in the predictive mind. Moreno believed individuals could create adaptive behavioral patterns at high spontaneity levels, ultimately enriching their cultural conserve. Individuals generate predictive signals from their models, compare them with environmental input signals, and either consolidate models when they match or update models or change the environment when they mismatch. The protagonist brings one model to the antagonist position, initiating homospatial process. The predictive mechanism in the antagonist position mirrors that in the protagonist position. Since prediction errors in both positions arise from the same interactional context, controlling either party' s prediction error changes the other' s error and model (see Figure 2 [Figure 2: see original paper]). In the parent-child example, matching predictive and input signals in the mother role reduces prediction error. The homospatial process of both models creates momentum for error reduction, so when the client returns to the daughter role, the generative model transforms, generating new cognition: "Mother calls repeatedly because she worries, she loves me." Ultimately, prediction error in the protagonist position is also controlled. Role reversal helps clients move beyond past experiences, forming new experiences from the confusion of coexisting models and escaping stereotyped patterns.

When do we update our models? Precision weighting determines reliance on environmental evidence versus prior models (Yon & Frith, 2021). With ambiguous environmental information, we depend more on past experiences and established models. With high-precision input, we decrease prior probability weighting and trust posterior evidence more. Interpersonal precision weighting works similarly: individuals assign different weights to different people' s statements. Role reversal can be viewed as a precision weighting process. When antagonist and protagonist models coexist, clients' responses to relational questions supplement knowledge and details about both models, regulating their precision. Higher detail levels yield higher model precision (Kwisthout et al., 2017). Counselors then facilitate direct dialogue between roles, promoting model weighing and comparison. Specifically, they invite clients to shift to an empty chair to enact their mother, pose relational questions ( "What worries you when your daughter is away at school?" "What do you think when she doesn' t answer your calls?" ), and guide them to imagine the daughter present and express maternal perceptions. When assessing relational uncertainty, clients typically report integrated perceptions (e.g., "Mother is always strict,""Mother is nagging" ). Through role reversal, higher-order actions (e.g., metacognition) actively participate in prediction error regulation, making subconscious processes conscious.

Clients collect more information about protagonist-antagonist and other relevant roles, enabling more precise representations (e.g., "What are mother' s main concerns?" "Which specific statements trigger my unhappiness?" ). Only after understanding representation precision can clients make choices based on self-experience.

Predictive processing theory emphasizes cognition-action unity (Friston et al., 2010; Hohwy, 2020). Individuals can reduce prediction error through two means: updating generative models or acting to change the environment. Environmental stimuli generating input signals constitute the self-dependent world, endowed with meaning corresponding to the self, with self constructed through environmental interaction. The antagonist constitutes part of environmental stimuli.

If clients choose to control prediction error by changing the environment, this implies two levels: (1) in real life, subject-self and object-world interact through behavioral interfaces to change the object-world and satisfy expectations; (2) in role reversal, environment includes specific auxiliary egos, directors, and therapeutic settings. Clients have some opportunity to modify the environment—for example, counselors may invite them to adjust chair distance and orientation based on their relationship model to reduce prediction error.

However, in single-session role reversal, effectiveness more likely manifests in helping clients update generative models rather than changing the environment. We must therefore address how the mind makes predictions and the mechanism of predictive mind throughout the psychodrama process, dynamically explaining how role reversal promotes self-healing and growth to facilitate smoother individual-environment interaction.

**Figure 2. Prediction Error and Model Updating Process**
*Note: When predictive and input signals match in the protagonist position, the model consolidates; mismatch stimulates model updating or environmental change. The antagonist position mechanism works similarly. Protagonist predictive/input signals become input/predictive signals in the antagonist position. Position shifting from protagonist to antagonist promotes prediction error reduction through model adjustment and better matching between models. "Changing environment" in single-session role reversal has multiple meanings and is omitted from the figure.*

### 3.3 Models as Interactive Constructions

In predictive processing theory, action processes create the need for self-construction. Environmental uncertainty necessitates world modeling for timely prediction and appropriate response to meet survival needs (Kelly et al., 2019).

The mind' s predictive pathway approximates Bayesian inference (Clark, 2013/2017). The objective world contains numerous discrete random variables whose stochastic causal relationships form causal Bayesian networks (Pearl,

2000). Both world knowledge and self-knowledge are hierarchical causal models, with each level describable by a causal Bayesian network containing three variable types: hypothesis variables (generating predictions), prediction variables (generated from hypotheses), and intermediate variables (affecting hypothesis-prediction relationships) (Kwisthout et al., 2017). Logically, hypothesis variables precede prediction variables at the same level, with upper-level prediction variables equaling lower-level hypothesis variables. The idealized parent-child example illustrates this hierarchical structure. At level n-1 in the protagonist model, the prediction variable is "Mother won' t keep calling me" and the hypothesis variable is "Mother won' t overly interfere in my life." At level n, prediction is "Mother should trust I can care for myself" and hypothesis is "Mother should trust my abilities." At level n+1, prediction is "Mother should care about me and consider my feelings" and hypothesis is "Mother loves me." Through role reversal, the antagonist' s hierarchical model is constructed interactively: at level n-1, prediction is "Calling more is necessary" and hypothesis is "My daughter has always been physically weak, I must remind her to care for herself" ; at level n, prediction is "My daughter needs my reminders" and hypothesis is "I worry about her, fearing she can' t care for herself" ; at level n+1, prediction is "I care for my daughter" and hypothesis is "I love my daughter deeply."

The relationship between prediction and hypothesis variables is complex. Within psychodrama, social systems, cultural contexts, and different norms constitute intermediate variables. Each individual exists within culture, which subtly influences internal conflict models and their external expressions. Psychodrama emphasizes social and cultural sources underlying roles, with collective action components resulting from generations of human interaction, significantly impacting case conceptualization (Moreno, 1961/1969). Both protagonist and antagonist possess multi-layered Bayesian models that interact to form a larger causal Bayesian network (see Figure 3 [Figure 3: see original paper]b). Thus, protagonist-antagonist interaction represents the interactive construction of their predictive processing models within this larger network.

**Figure 3. Interactive Network Model of Role Reversal**
*Note: (a) Protagonist and antagonist models have hierarchical structures. Dashed lines indicate that multiple role reversals lead to mutual construction of corresponding levels. Throughout the process, models follow a clear conceptual hierarchy from n-1 to n+1. Individual roles develop from role taking to role creating. (b) Causal Bayesian network showing same-level model interaction. H1, H2 are protagonist hypothesis variables; H3, H4 are antagonist hypothesis variables. P1, P2 are protagonist prediction variables; P3, P4 are antagonist prediction variables. i1-i10 are intermediate variables.*

Role reversal therapy typically involves multiple exchanges within a session according to client needs. Through repeated role reversals, individuals experience a developmental process from role learning to spontaneity: role taking, role playing, and role creating, expanding their role repertoire and improving life

quality (Moreno, 1961). When the antagonist is the client' s mother, shifting to her position first yields general maternal traits (care, devotion)—role taking. Next, by imitating mother' s speech patterns, tone, and expressions, the client engages in role playing—higher-level model processing. Finally, in role creating, the client brings new insights into the antagonist model, performing unspoken words or undone actions at high spontaneity levels. This developmental process follows hierarchical model construction rules from lower to higher levels. Through role shifting, clients construct the antagonist model while clarifying their own predictive model, updating self-knowledge from lower to higher levels. Multiple role reversals align corresponding levels of protagonist and antagonist models, mutually constructing them through interaction.

Causal Bayesian networks depend not only on static causal relationships but also dynamic temporal interactions. Current predictive signals depend on both current hypotheses and previous time-step predictions (Kwisthout et al., 2017). Over longer timescales, models update not only across multiple exchanges within a session but also across multiple sessions (see Figure 4 [Figure 4: see original paper]). Goldman and Momson's (1984) psychodrama spiral describes the entire therapeutic process of predictive processing and model transformation. Starting from current problems, psychodrama enacts specific past scenes, connects to other relevant scenes, and links to earlier childhood experiences, aiming for integrative catharsis and role training for better life adaptation. The psychodrama spiral works from surface prediction errors to deeper predictive processing models, ultimately returning to current conflict situations. Through continuous upward spiraling from past to future, individuals gain more structured knowledge about self and others, enabling co-creation of effective strategies for handling prediction errors. This makes psychodrama potentially more practical than psychoanalysis (Kadyrov, 1997). In actual practice, the psychodrama spiral transforms dynamically, with role interactions determined by clients' present states, reflecting psychodrama' s vitality and integration of spontaneity and creative growth.

**Figure 4. Psychodrama Spiral Based on Predictive Processing Theory**
*Note: Adapted from Goldman & Momson (1984) and Kwisthout et al. (2017).*

## 4 Summary and Outlook

Role reversal presents potential conflicts in concrete form, promoting change in non-adaptive behavioral patterns. From the predictive processing perspective, role reversal is a self-construction process through which individuals acquire structured knowledge about self and others, reduce prediction error, and guide positive life-world interaction. The three explanatory models proposed here elaborate role reversal' s mechanisms at different levels, with a progressive internal logic: the transformational step first induces signal change, then internal model updating and hierarchical transformation, and finally dynamic model updating through the psychodrama spiral across multiple role reversals.

A clarification is needed: predictive processing theory, as a constructivist framework, treats the relationships between self-model-generated predictions, prediction-input matching, and model construction/updating as logical rather than temporal sequences. That is, top-down predictive processing based on representation and bottom-up predictive processing based on experience, along with model generation and construction, occur simultaneously. Protagonist and antagonist models construct each other interactively without temporal sequence, as the antagonist constitutes the world in which the protagonist' s self exists. The free energy principle emphasizes adaptivity as life' s fundamental principle, with continuous error trade-offs between top-down and bottom-up predictions (Friston, 2010). In practice, we clearly observe sequential "protagonist-antagonist-protagonist" position changes, but internally, role creation—acquiring the antagonist model (psychosomatic, social, or psychodramatic)—results from prediction error trade-offs. Early psychodrama work typically progresses through role taking and playing before reaching role creating. As therapy advances and spontaneity increases, later stages may skip phases and move directly to creation. Thus, role development doesn' t map one-to-one onto model levels, but overall follows the "taking-playing-creating" trajectory.

Another question arises: Does the antagonist role and model enacted by the client match the actual antagonist' s model? Since "I" cannot output what isn' t in my model, the enacted antagonist model remains part of "me." This understanding is partially valid. Internally, the antagonist model acquired through role reversal relates to subjective reality—interpersonal conflicts reflect internal model conflicts. However, we must not overlook psychodrama' s interpersonal perspective and the connection between intra-personal and inter-personal views. Models are constructed from the environment; both self-evaluations and others' evaluations are learned through relationships and culture. Others do evaluate us in certain ways, and our internal models respond by identifying with or rejecting these evaluations. With "I" necessarily comes "non-I," and "I" gains construction of "I" within "non-I."

Spontaneity and creativity are core concepts in Moreno' s theory. During self-construction through role reversal, multiple role models may conflict. How does role creation emerge in predictive processing? Does the degree of correspondence between input signal recognition and objective reality affect therapeutic outcomes? Counselors can help clients'input signals approximate reality through guided questioning, relational queries, and facilitating direct role dialogue. Additionally, we must return to Moreno' s view of human nature—he believed human creativity reflects God' s infinite creativity, and everyone can create and modify their environment for continuous adaptation. Rigid interaction patterns change through action' s power, enabling clients to develop new responses to old situations or adaptive behaviors for new situations.

Applying predictive processing theory to role reversal provides a framework for future intervention practice. Counselors can concretize prediction errors aris-

ing from client-antagonist interactions and collaboratively explore strategies for minimizing error in specific issues. If prediction error stems from the world' s inherent, irreducible properties, clients must modify hypothesis variables. If error arises from insufficient antagonist model knowledge, clients must modify random causal relationships or introduce new variables. Other psychodrama techniques (e.g., mirroring) can resample input signals to minimize prediction error. Psychodrama's theoretical foundation also resonates with traditional Chinese culture—concepts like introspective awareness and sincere intention from "benevolence, righteousness, propriety, wisdom, and harmony" culture show internal connections and potential for comparative work with predictive processing theory. Future directions should emphasize psychodrama' s core features of spontaneity and creativity while integrating Chinese cultural elements to develop localized, life-oriented, and effective psychodrama techniques adapted to Chinese characteristics, thereby enhancing psychological healing and national health promotion.

## References

Abeditehrani, H., Dijk, C., Toghchi, M. S., & Arntz, A. (2020). Integrating cognitive behavioral group therapy and psychodrama for social anxiety disorder: An intervention description and an uncontrolled pilot trial. *Clinical Psychology in Europe, 2*(1), 1–21.

Baumeister, R. F. (1998). The self. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 680–740). Cambridge: Oxford University Press.

Beauvais, A. M., Özbaş, A. A., & Wheeler, K. (2019). End-of-life psychodrama: Influencing nursing students' communication skills, attitudes, emotional intelligence and self-reflection. *Journal of Psychiatric Nursing, 10*(3), 103–110.

Blatner, A. (1997). Psychodrama: The state of the art. *The Arts in Psychotherapy, 24*(1), 23–30.

Blatner, A. (2000). *Foundations of psychodrama: History, theory, and practice* (4th ed.). New York, NY: Springer.

Blatner, A. (2007). Morenean approaches: Recognizing psychodrama' s many facets. *Journal of Group Psychotherapy, Psychodrama and Sociometry, 59*(4), 159–170.

Boies, K. G. (1972). Role playing as a behavior change technique: Review of the empirical literature. *Psychotherapy: Theory, Research and Practice, 9*(2), 185–192.

Boroomandian, N., Mohammadi, N., & Rahimi Taghanaki, C. (2020). Comparing the effectiveness of group therapy by "interaction and psychodrama" on emotional awareness of adolescent girls with psychological distress. *Journal of Psychological Achievements, 27*(1), 1–22.

Butler, T., & Fuhriman, A. (1980). Patient perspective on the curative process: A comparison of day treatment and outpatient psychotherapy groups. *Small Group Behavior, 11*(4), 371–388.

Carlson–Sabelli, L. (1989). Role reversal: A concept analysis and reinterpretation of the research literature. *Journal of Group Psychotherapy, Psychodrama and Sociometry, 42*, 139–152.

Chen, J., Qu, Y., & Hu, H. (2019). Research progress on the mechanism and effect of psychodrama techniques in depression intervention. *China Modern Medicine, 26*(08), 25–28.

Chen, S. (2012). Research on problems and countermeasures for professional women. *Enterprise Technology Development, 31*(34), 121–122.

Chu, P., Yan, J., & Li, D. (2009). Effect of psychodrama therapy on negative psychology and coping styles in adolescent epilepsy patients. *International Journal of Nursing, 28*(3), 322–324.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral & Brain Sciences, 36*, 181–204.

Clark, A. (2017). Predictions, precision, and agentive attention. *Consciousness and Cognition, 56*, 115–119.

Clark, A. (2018). A nice surprise? Predictive processing and the active pursuit of novelty. *Phenomenology and the Cognitive Sciences, 17*(3), 521–534.

Dogan, T. (2018). The effects of the psychodrama in instilling empathy and self-awareness: A pilot study. *PsyCh Journal, 7*(4), 227–238.

Dorothea, G. (2016). Effectiveness of psychodrama group therapy on pupils with mathematics anxiety. In C. Stadler, M. Wieser, & K. Kirk (Eds.), *Psychodrama Empirical Research and Science 2* (pp. 197–215). Wiesbaden: Springer.

Erikson, E. H. (1959). *Identity and life cycle*. New York: Norton.

Farmer, C., Moreno, Z. T., Campbell, D., & Draper, R. (2018). *Psychodrama and systemic therapy*. New York: Routledge.

Fei, D. (2018). *Mind-body relationship research*. Beijing: The Commercial Press.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11*(2), 127–138.

Friston, K. J., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010). Action and behavior: A free-energy formulation. *Biological Cybernetics, 102*(3), 227–260.

Giacomucci, S., & Marquit, J. (2020). The effectiveness of trauma-focused psychodrama in the treatment of PTSD in inpatient substance abuse treatment. *Frontiers in Psychology, 11*, 896.

Goldman, E., & Momson, D. (1984). *Psychodrama: Experience and process*. Dubuque: Kendall Hunt.

Han, J. (2021). A study on the relationship among social support, self-concept and mental health of college students (Master's thesis). Heilongjiang University, Harbin.

Han, K., & Chu, P. (2010). Effect of psychodrama therapy on shame and quality of life in children with epilepsy. *Chinese Journal of Practical Nursing: Mid-volume, (8)*, 30–32.

Hohwy, J. (2020). New directions in predictive processing. *Mind & Language, 35*(2), 209–223.

Hohwy, J., & Michael, J. (2017). Why should any body have a self? In F. de Vignemont & A. Alsmith (Eds.), *The Body and the Self, Revisited* (pp. 363–391). Cambridge, MA: The MIT Press.

Kadyrov, I. M. (1997). Psychodrama and psychoanalysis: Two theaters for mental drama. *Journal of Russian & East European Psychology, 35*(6), 55–72.

Karp, M., Holmes, P., & Tauvon, K. B. (Eds.). (1998). *The handbook of psychodrama*. New York: Psychology Press.

Kaya, Y., Güç, E., Kırlangıç-Şimşek, B., & Cesur-Atintaş, M. (2021). The effect of psychodrama group practices nursing students' self-liking, self-competence, self-acceptance: A quasi-experimental study. *Perspectives in Psychiatric Care, 59*, 1–9.

Kellermann, P. F. (1994). Role reversal in psychodrama. In P. Holmes, M. Karp, & M. Watson (Eds.), *Psychodrama since Moreno: Innovations in theory and practice* (pp. 263–279). London: Routledge.

Kelly, M. P., Kriznik, N. M., Kinmonth, A. L., & Fletcher, P. C. (2019). The brain, self and society: A social-neuroscience model of predictive processing. *Social Neuroscience, 14*(3), 266–276.

Kipper, D. A., & Ritchie, T. D. (2003). The effectiveness of psychodramatic techniques: A meta-analysis. *Group Dynamics: Theory, Research, and Practice, 7*(1), 13.

Kube, T., Schwarting, R., Rozenkrantz, L., Glombiewski, J. A., & Rief, W. (2020). Distorted cognitive processes in major depression: A predictive processing perspective. *Biological Psychiatry, 87*(5), 388–398.

Kwisthout, J., Bekkering, H., & Van Rooij, I. (2017). To be precise, the details don' t matter: On predictive processing, precision, and level of detail of predictions. *Brain and Cognition, 112*, 84–91.

Lee, K. M., Ferreira-Santos, F., & Satpute, A. B. (2021). Predictive processing models and affective neuroscience. *Neuroscience & Biobehavioral Reviews, 131*, 211–228.

Levin, K. R. (2021). *Mommy or me: Perceptions of mothers with postpartum depression on occupational disruption and role conflict*. New York, NY: Ithaca

College.

Li, H. (2021). The life view of mind: From pragmatism to predictive mind. *Social Science Front, (09)*, 17-25.

Liu, X. (2021). The effect of mood on predictive sentence processing by older adults. *SAGE Open, 11*(2), 1-10.

Ma, J., & Zhang, T. (1999). Analysis of the relationship between role cognition, sense of control and work stress. *Ergonomics, (04)*, 15-19.

Mann, J. H., & Mann, C. H. (1959). Role playing experience and interpersonal adjustment. *Journal of Counseling Psychology, 6*(2), 148.

Mead, G. H. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.

Miller, R., Mason, S., & Skolnik, S. (2021). Psychodrama for Schizophrenia: Adaptive techniques in mutual aid groups. *Social Work with Groups, 44*(3), 273-287.

Mojahed, A., Zaheri, Y., & Moqaddam, M. F. (2021). Effectiveness of group psychodrama on aggression and social anxiety of children with attention-deficit/hyperactivity disorder: A randomized clinical trial. *The Arts in Psychotherapy, 73*, 101756.

Moreno, J. L. (1946). *Psychodrama* (Vol. 1). New York: Beacon House.

Moreno, J. L. (1961). The role concept: A bridge between psychiatry and sociology. In J. Fox (Ed.), *The essential Moreno* (pp. 60-66). New York, NY: Springer.

Moreno, J. L. (1969). *Psychodrama* (Vol. 3). New York: Beacon House.

Moreno, J. L. (1972). *Psychodrama*. New York: Beacon House.

Moreno, Z. T. (1965). Psychodramatic rules, techniques, and adjunctive methods. *Group Psychotherapy, 18*, 73-86.

Orkibi, H., & Feniger-Schaal, R. (2019). Integrative systematic review of psychodrama psychotherapy research: Trends and methodological implications. *PLOS ONE, 14*(2), 0212575.

Pan, H. (2015). Effect of psychodrama therapy on psychology and hope in school-age children with epilepsy. *International Journal of Nursing, (10)*, 1389-1391.

Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge, MA: The MIT Press.

Ransom, M., Fazelpour, S., Markovic, J., Kryklywy, J., Thompson, E. T., & Todd, R. M. (2020). Affect-biased attention and predictive processing. *Cognition, 203*, 104370.

Rohrbaugh, M., & Bartels, B. D. (1975). Participants' perceptions of "curative factors" in therapy and growth groups. *Small Group Behavior, 6*, 430–456.

Rothenberg, A. (1971). The process of Janusian thinking in creativity. *Archives of General Psychiatry, 24*, 195–205.

Rothenberg, A. (1995). Creative cognitive processes in Kekule' s discovery of the structure of the benzene molecule. *The American Journal of Psychology, 109*, 419–438.

Rudokaite, D., & Indriuniene, V. (2019). Effectiveness of psychodrama for mitigating school fears among senior secondary school students. *Zeitschrift für Psychodrama und Soziometrie, 18*(2), 369–385.

Ryle, G. (1992). *The concept of mind* (Xu Dajian, Trans.). Beijing: The Commercial Press. (Original work published 1949)

Sartre, J. P. (1957). *The transcendence of the ego: An existentialist theory of consciousness*. London: Macmillan.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*(3), 407–441.

Su, J., & Ye, H. (2021). "Enactive cognition,""interaction theory,"and "predictive processing model"—New thinking on the "problem of other minds." *Psychological Science, 44*(6), 1519–1527.

Soysal, F. S. Ö. (2021). The effects of psychodrama on emotion regulation skills in emerging adults. *Current Psychology, 34*, 1–14.

Tarashoeva, G., Marinova-Djambazova, P., & Kojuharov, H. (2017). Effectiveness of psychodrama therapy in patients with panic disorders: Final results. *International Journal of Psychotherapy, 21*(2), 55–66.

Tümlü, G. Ü., & Şimşek, B. K. (2021). The effects of psychodrama groups on feelings of inferiority, flourishing, and self-compassion in research assistants. *The Arts in Psychotherapy, 73*, 101763.

Wang, Q. (2021). The "prediction" concept in predictive mind. *Fujian Forum (Humanities and Social Sciences), (09)*, 15–24.

Wiese, W., & Metzinger, T. (2017). Vanilla PP for philosophers: A primer on predictive processing. In T. Metzinger & W. Wiese (Eds.), *Philosophy and predictive processing* (pp. 1–18). Frankfurt am Main: MIND Group.

Yalom, I. (1970). *The theory and practice of group psychotherapy*. New York: Basic Books.

Yaniv, D. (2012). Dynamics of creativity and empathy in role reversal: Contributions from neuroscience. *Review of General Psychology, 16*(1), 70–77.

Yon, D., & Frith, C. D. (2021). Precision and the Bayesian brain. *Current Biology, 31*(17), 1026–1032.

Žurić Jakovina, I., & Jakovina, T. (2017). Role theory and role analysis in psychodrama: A contribution to sociology. *Socijalna ekologija: časopis za ekološku misao i sociologijska istraživanja okoline, 26*(3), 247–267.

*Note: Figure translations are in progress. See original paper for figures.*