

Reframe Your Life Story: Interactive Narrative Therapist and Innovative Moment Assessment with Large Language Models

Yi Feng, Jiaqi Wang, Wenzuan Zhang, Zhuang Chen, Shen Yutong, Xiyao Xiao, Minlie Huang*, Liping Jing*, Jian Yu

Beijing Jiaotong University, Singapore University of Techonology and Design, Tsinghua University, Lingxin AI

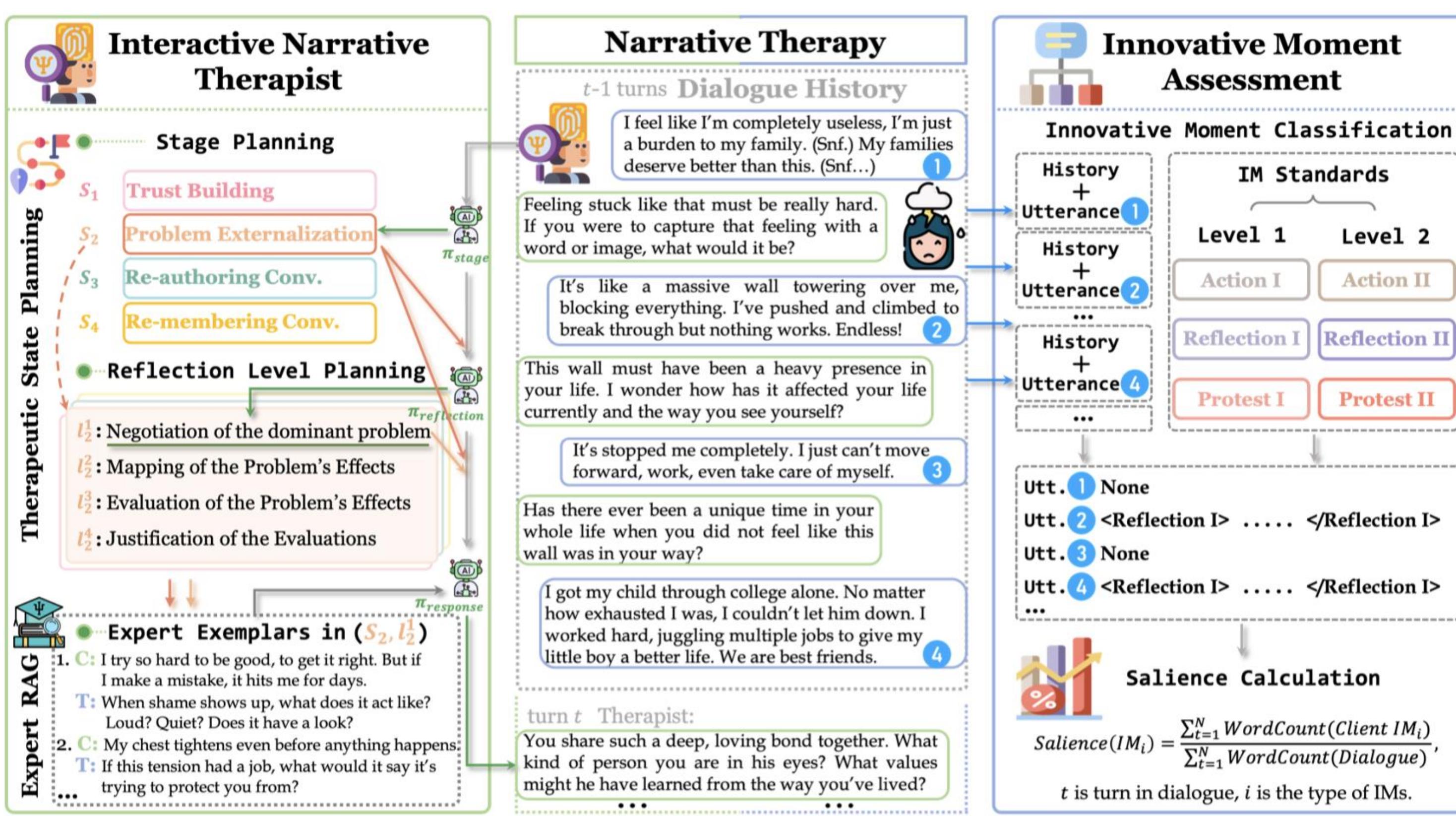
Introduction

Recent progress in large language models (LLMs) has opened new possibilities for mental health support, yet current approaches lack realism in simulating specialized psychotherapy and fail to capture therapeutic progression over time. Narrative therapy, which helps individuals transform problematic life stories into empowering alternatives, remains underutilized due to limited access and social stigma. We address these limitations through a comprehensive framework with two core components.

First, **INT** (Interactive Narrative Therapist) simulates expert narrative therapists by planning therapeutic stages, guiding reflection levels, and generating contextually appropriate expertlike responses.

Second, **IMA** (Innovative Moment Assessment) provides a therapy-centric evaluation method that quantifies effectiveness by tracking “Innovative Moments” (IMs), critical narrative shifts in client speech signaling therapy progress. Experimental results on 260 simulated clients and 230 human participants reveal that INT consistently outperforms standard LLMs in therapeutic quality and depth. We further demonstrate the effectiveness of INT in synthesizing high-quality support conversations to facilitate social applications.

Framework Overview



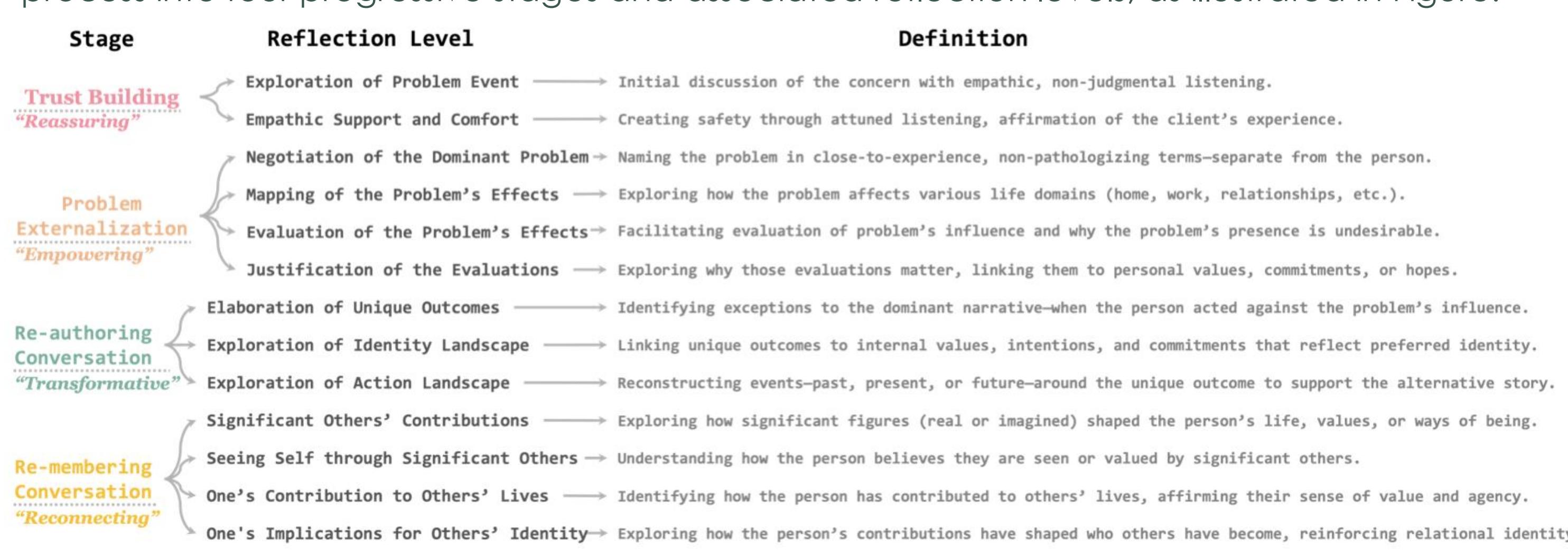
The Overview of our framework, comprising **INT** (left, green box) for narrative therapist simulation and **IMA** (right, blue box) for evaluating therapeutic progression.

We define C_i and T_i as the client utterance and therapist's response at the turn i of the therapeutic conversation $D = (C_i, T_i)_{i=1}^N$.

- INT** plans the therapeutic state (stage, reflection level) and generates responses aligned with expert style using retrieved exemplars. The INT responds with T_t given the history $H_t = (C_i, T_i)_{i=1}^{t-1}$ and the current client speech C_t .
- IMA** classifies each client utterance into six IM types (e.g., Action I) across two levels and quantifies therapeutic progression via salience calculation. The IMA classifies the Innovative Moments of client utterances $C_{i=1}^N$ given the entire session D .

Interactive Narrative Therapist

Maintaining an appropriate therapeutic pace is critical: premature progression may ruin client trust, whereas stagnation impedes therapeutic progress. To navigate this balance between client readiness and therapeutic progression, we translate narrative therapy principles into a computational framework INT by systematically formalizing the therapeutic process into four progressive stages and associated reflection levels, as illustrated in Figure.



Innovative Moment Assessment

Innovative Moments (IMs) are episodes where clients express thoughts, feelings, or behaviors that contradict their problem-saturated narratives (White, 2007). The table shows the Classification of six Innovative Moments (IMs) across Level 1 and Level 2 in narrative therapy.

IM Type	Definition & Contents			
	Level 1: Creating Distance from the Problem			
Action I	Contents: New behavioral strategies to overcome the problem, active exploration of solutions and information about the problem. Problem stories: Client got nervous and refused to go to public places after experiencing domestic violence from partner. Client utterance: <Action I> Yesterday I went out to the cinema for the first time this month to watch a movie. </Action I>			
Reflection I	Contents: New understandings of the problem, intention to fight (CONTEST) the problem's demands, and references to self-worth. Problem stories: Client let the depression take over his/her life for a long time. Client utterance: <Reflection I> It wants to control my entire life, eventually taking it all away. </Reflection I>			
Protest I	Contents: Rejecting or objecting to the problem, critique of those who support it, and critique of problematic facets of the self. Problem stories: Client must live according to his/her parents' expectations. Client utterance: <Protest I> Parents should love their children, not constantly judge them. I've really had enough. </Protest I>			
Level 2: Centered on the Change				
Action II	Contents: Generalization into the future and other life dimensions of good outcomes (performed or projected actions). Problem stories: Client was afraid to say no even he/she was uncomfortable. Client utterance: <Action II> I'll also bring this boundary awareness to work, like no longer working overtime silently. </Action II>			
Reflection II	Contents: Contrasting self (what changed?) or self-transformation (how/why change occurred?). Problem stories: Client exhibited excessive anxiety when coping with daily pressure. Client utterance: <Reflection II> Before, when I encountered any problem, I would spend the whole day anxious, self-critical, even wanting to escape. ...I suddenly realized I'm not so easily defeated anymore. </Reflection II>			
Protest II	Contents: Centering on the self, affirming personal rights, needs, and values. Problem stories: Client exhibited the pattern of deriving self-worth from prioritizing others' needs. Client utterance: <Protest II> I think my feelings are important too. I have the right to say 'no', the right to rest when tired, rather than constantly pleasing others. I want to start living for myself, not according to others' expectations. </Protest II>			

Experimental Results

We conduct comprehensive evaluations via both automated assessment (with 260 simulated clients) and extensive human interactive evaluation (involving 230 participants in the main (200) and follow-up (30) studies, plus expert annotators). Results consistently demonstrate that our INT outperforms all standard LLMs based on direct role-playing, with significant gains across core therapeutic dimensions and nearly doubling the elicitation of advanced narrative transformation markers according to IMA.

Therapeutic Dimension Assessment: We evaluate systems across 4 therapeutic dimensions using a 5-point Likert scale: Reassuring (creating safety for client disclosure, Reas.), Empowering (facilitating problem externalization, Emp.), Transformative (uncovering alternative narratives, Trans.), Reconnecting (strengthening significant relationships, Recon.). The Average score is (Avg.).

Innovative Moment Assessment: Following the established protocols from Gonçalves et al. (2011), each client utterance is annotated with a subset of six IM categories (Table 1), or labeled as “None” if no IM is present. For co-occurrence cases, we follow explicit coding rules in psychotherapy: when “Action” and “Reflection” markers co-occur, both are coded; when either co-occurs with “Protest”, the utterance is coded as “Protest”.

Model-only Assessment

Model	Therapeutic Dimensions				Innovative Moment Assessment(Salience)							
	Reas.	Emp.	Trans.	Recon.	Avg.	Action I	Reflection I	Protest I	Action II	Reflection II	Protest II	SUM
Claude-3.7-sonnet	3.13	3.29	3.12	2.96	3.13	2.459%	6.796%	0.036%	4.762%	8.971%	0.100%	23.124%
Gemini-2.5-pro	2.18	2.47	2.84	2.63	2.53	3.982%	7.656%	0.027%	8.782%	15.738%	0.117%	36.302%
Qwen-2.5	3.51	3.35	3.08	3.10	3.26	3.740%	7.460%	0.011%	7.328%	12.819%	0.051%	31.409%
GLM-4-plus	2.93	3.58	3.23	3.17	3.23	4.602%	8.933%	0.062%	8.169%	15.504%	0.148%	37.418%
Deepseek-V3	3.31	3.80	3.71	3.45	3.57	3.824%	9.388%	0.092%	8.099%	14.760%	0.067%	36.234%
Douba-1.5-pro	2.80	3.23	3.00	2.95	3.00	4.866%	8.489%	0.082%	10.606%	17.988%	0.079%	42.110%
GPT-4o	3.34	3.52	3.19	3.19	3.31	3.115%	7.480%	0.037%	6.819%	11.770%	0.127%	29.348%
INT	3.60	3.87	3.84	3.51	3.71	1.594%	3.092%	0.096%	8.730%	9.680%	0.998%	29.698%

Human interactive evaluation

Model	Therapeutic Dimensions				Innovative Moment Assessment(Salience)							
	Reas.	Emp.	Trans.	Recon.	Avg.	Action I	Reflection I	Protest I	Action II	Reflection II	Protest II	SUM
Role-playing												
Claude-3.7-sonnet	3.08	2.77	2.56	2.40	2.70	3.539%	6.895%	0.629%	4.059%	5.919%	0.794%	21.835%
Gemini-2.5-pro	3.01	2.14	2.01	1.94	2.28	3.458%	6.934%	0.486%	3.281%	4.647%	0.613%	19.419%
Qwen-2.5	2.76	2.37	2.15	2.10	2.35	3.171%	6.355%	0.538%	3.551%	4.971%	0.679%	19.265%
GLM-4-plus	2.83	2.70	2.27	2.27	2.52	3.251%	6.606%	0.613%	3.839%	5.248%	0.774%	20.331%
Deepseek-V3	2.73	2.54	2.46	2.61	2.59	3.136%	6.722%	0.577%	3.737%	5.225%	0.728%	20.125%
Douba-1.5-pro	2.66	2.45	2.10	2.10	2.33	3.056%	6.528%	0.556%	3.551%	4.855%	0.702%	19.248%
GPT-4o	3.11	2.75	2.52	2.49	2.72	3.513%	6.895%	0.624%	4.211%	5.827%	0.788%	21.858%
INT	3.09	3.11	3.42	3.37	3.25	2.794%	6.834%	0.662%	8.730%	9.680%	0.998%	29.698%
w/o RAG	3.13	2.92	2.74	2.69	2.87	3.573%	8.333%	0.610%	4.235%	9.438%	0.803%	26.992%
w/o RAGRL	3.16	2.83	2.65									