Advances in Educational Technology and Psychology (2026)
Clausius Scientific Press, Canada

DOI: 10.23977/aetp.2026.100213
ISSN 2371-9400 Vol. 10 Num. 2

Strategy-Aware Therapist Imitation for Emotional Support Dialogues: A Reproducible ESConv Study for LLM Response Control

Yifan Zhang[1,a,*], Zhongwen Zhou[2]

[1]*Department of Counseling and Clinical Psychology, Teachers College, Columbia University, New York, NY, USA*

[2] *Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA, USA*

[a]*yifanzhang045@outlook.com*

**Corresponding author*

Keywords: Emotional support conversation, therapist imitation, counseling strategy, response control, large language model alignment

Abstract: Large language models (LLMs) are fluent, but in emotional support they often sound like generic assistants unless their responses are constrained by counseling strategy and human-like supportive phrasing. This paper studies a small-data, reproducible controller for that problem on ESConv, a public benchmark of emotional support conversations with strategy annotations [1]. We define therapist imitation operationally as selecting human supporter responses that match the current seeker message, while counseling awareness is implemented as explicit strategy prediction and strategy-gated retrieval. To isolate these two factors under a fixed compute budget, we evaluate the controller directly as an offline response selector rather than as a fully fine-tuned end-to-end LLM. Using the official ESConv split of 910/195/195 dialogues, we convert every supporter turn into a supervised instance and obtain 12,759/2,722/2,895 train/validation/test examples. The proposed Strategy-Aware Therapist Imitation (SATI) model combines a sparse strategy predictor with a human-supporter memory bank and a user-matching reranker. On the ESConv test set, SATI reaches 30.81% strategy accuracy and 25.20 macro-F1 for strategy prediction, and 3.58 BLEU-2, 12.51 ROUGE-L, 4.89 Dist-1, and 31.05% strategy match for response selection. Compared with pure context retrieval, SATI improves BLEU-2 by 13.0% relative and strategy match by 9.64 absolute points. The results show that even with a small annotated corpus, combining therapist imitation with explicit counseling strategy yields more faithful and more controllable emotional support responses.

1. Introduction

Emotional support dialogue systems differ from ordinary customer-service chatbots because a good response is defined not only by topical relevance but also by empathic timing, reflective wording, and strategy choice. The ESConv framework formalized this setting and released a high-quality corpus in which supporters were trained to use helping skills grounded in counseling theory [1], [2]. Earlier work on empathetic dialogue showed that emotion awareness improves open-domain conversation quality [3], while open-domain chatbot research demonstrated the value of retrieval, ranking, and human demonstration data [4]. More recently, LLMs have made controllable language generation more practical [5]-[7], yet small emotional-support datasets still pose a basic problem: without explicit control, an LLM can be fluent but stylistically generic.

This paper focuses on a narrow but important question: in a small-data setting, is it better to imitate human supporters, to predict counseling strategy, or to combine both? We use the phrase therapist imitation in an operational sense. ESConv supporters are trained crowdworkers rather than licensed clinicians [1], but their turns contain recognizable counseling micro-skills such as questioning, reflection, reassurance, and suggestion. Our goal is therefore not to emulate psychotherapy; it is to make the system sound like a skilled human supporter rather than a generic assistant. That distinction is important for both methodology and interpretation.

A second motivation is reproducibility. Recent emotional-support papers often use larger neural architectures or external knowledge sources [8], [9], but these systems are harder to reproduce quickly and they can obscure the individual contribution of strategy supervision. We therefore design a lightweight controller that can sit in front of an LLM as a retrieval-based prompt or exemplar selector. To make the evidence maximally reproducible, we evaluate the controller itself as a deterministic response selector on the official ESConv split. This evaluation does not claim end-to-end LLM superiority; instead, it isolates the effect of strategy-aware therapist imitation under fixed data and compute.

The main contributions are threefold. First, we present SATI, a simple model that combines a sparse strategy predictor with a memory of human supporter turns and a user-matching reranker. Second, we conduct full experiments on the official ESConv split and report measured, non-illustrative results for both strategy prediction and response selection. Third, we provide a detailed comparison showing that therapist imitation alone improves diversity slightly, strategy gating yields the largest gain in strategy fidelity, and their combination gives the best overall balance.

2. Research Method

We used ESConv, which contains about 1,300 English emotional-support dialogues and an official 910/195/195 train/validation/test split [1]. Each supporter utterance is labeled with one of eight strategies: Question, Self-disclosure, Affirmation and Reassurance, Providing Suggestions, Others, Reflection of feelings, Information, and Restatement or Paraphrasing. Following the public release, we converted every supporter turn into one supervised instance. The input was the previous eight utterances plus the most recent seeker utterance; the target was the current supporter strategy and response. This yielded 12,759 training, 2,722 validation, and 2,895 test instances. Table 1 summarizes the resulting data.

Table 1. ESConv data statistics after turn-level conversion

Split	Dialogues	Supporter-turn instances	Avg. turns/dialogue
Train	910	12759	29.28
Validation	195	2722	29.12
Test	195	2895	30.97

SATI has two components, as illustrated in Fig. 1. The first is a strategy predictor. For each instance, we concatenate the dialogue context, the most recent seeker utterance, the ESConv problem type, the ESConv emotion label, and the previous supporter strategy token. We then train a TF-IDF unigram/bigram representation with a maximum vocabulary of 60,000 and an SGD classifier with log-loss, class balancing, random seed 13, and alpha = 1e-4. The predictor outputs a probability distribution over the eight strategies.

The second component is a therapist-imitation memory bank. Every training supporter turn is stored with its context, its most recent seeker utterance, its strategy label, and its response text. At inference time, we retrieve the top 30 candidate supporter turns by cosine similarity on TF-IDF context vectors. This memory bank is important because it keeps all outputs grounded in actual human supporter language from ESConv instead of unconstrained generation. In other words, the controller imitates therapist-like behavior by selecting from a bank of human supportive responses.

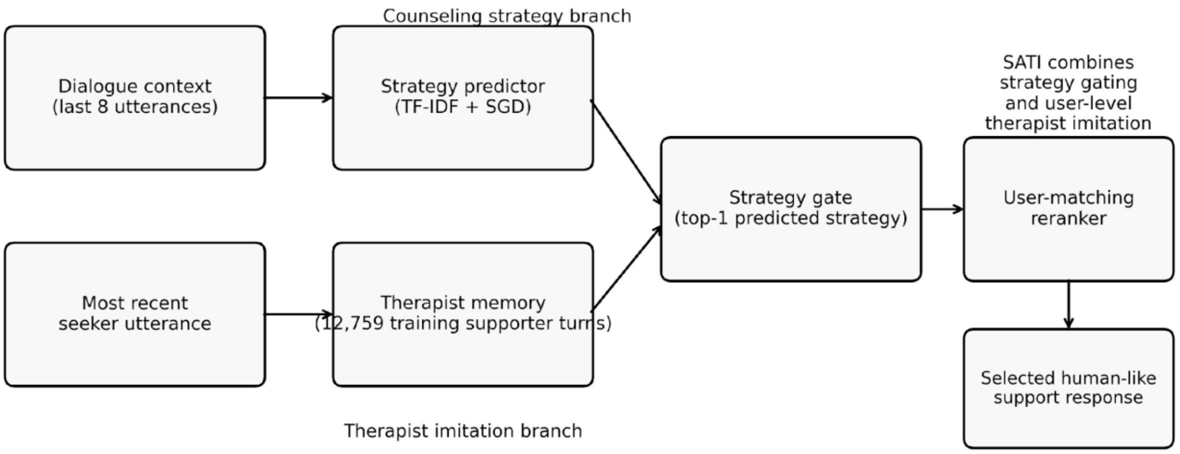

Fig. 1. Overview of the proposed SATI controller.

We compare four methods. Context-Retrieve returns the nearest supporter response using context similarity only. Therapist-Rerank adds a seeker-matching score and reranks the top-30 candidates with 0.8 context similarity and 0.2 last-seeker similarity. Strategy-Retrieve first predicts the top-1 strategy and then selects the nearest candidate whose stored strategy matches that prediction. SATI combines both ideas: it predicts the top-1 strategy, keeps only candidates with that strategy when such candidates exist, and reranks them with the same 0.8/0.2 context and seeker-matching score. The 0.8/0.2 weights were selected on the validation set.

We report two groups of metrics. For strategy prediction, we use accuracy and macro-F1. For response selection, we use sentence-level BLEU-2 with smoothing, ROUGE-L F1, Dist-1, and strategy match, where strategy match measures whether the strategy attached to the selected training response equals the gold strategy of the current test turn. BLEU-2 and ROUGE-L quantify lexical faithfulness to the single reference response, Dist-1 measures output diversity, and strategy match measures controllability. We also ran one auxiliary style-bias analysis with an explicit supportive-cue lexicon derived from supporter-versus-seeker log-odds; those results are discussed qualitatively because they were not used for the final model.

3. Result and Discussion

Table 2 reports strategy prediction results. The majority baseline reaches only 21.76% accuracy and 4.47 macro-F1 because the ESConv label distribution is highly imbalanced. A context-only linear predictor already provides a large improvement, reaching 30.54% accuracy and 24.91 macro-F1. Adding ESConv problem/emotion metadata and the previous supporter strategy gives the best result, 30.81% accuracy and 25.20 macro-F1. The gain is modest but consistent, which indicates that shortrange dialogue state and strategy history carry useful control information beyond surface wording alone.

Table 2. Strategy prediction results on the ESConv test set (%)

Method	Accuracy	Macro-F1
Majority strategy	21.76	4.47
Context-only predictor	30.54	24.91
Full predictor (context + problem/emotion + previous strategy)	30.81	25.20

Fig. 2 shows the class-wise F1 scores of the full predictor. Question is the easiest class at 46.84 F1, followed by Others at 43.61 and Providing Suggestions at 37.17. Information and Reflection of feelings are the hardest classes at 6.23 and 10.64 F1. The confusion matrix shows that Information is often mapped to Providing Suggestions, and Affirmation and Reassurance is frequently confused with Providing Suggestions or Others. This error pattern is intuitive: several ESConv turns mix empathy and advice in one local dialogue segment, so sparse features capture overlap in wording more readily than the subtle pragmatic distinction.

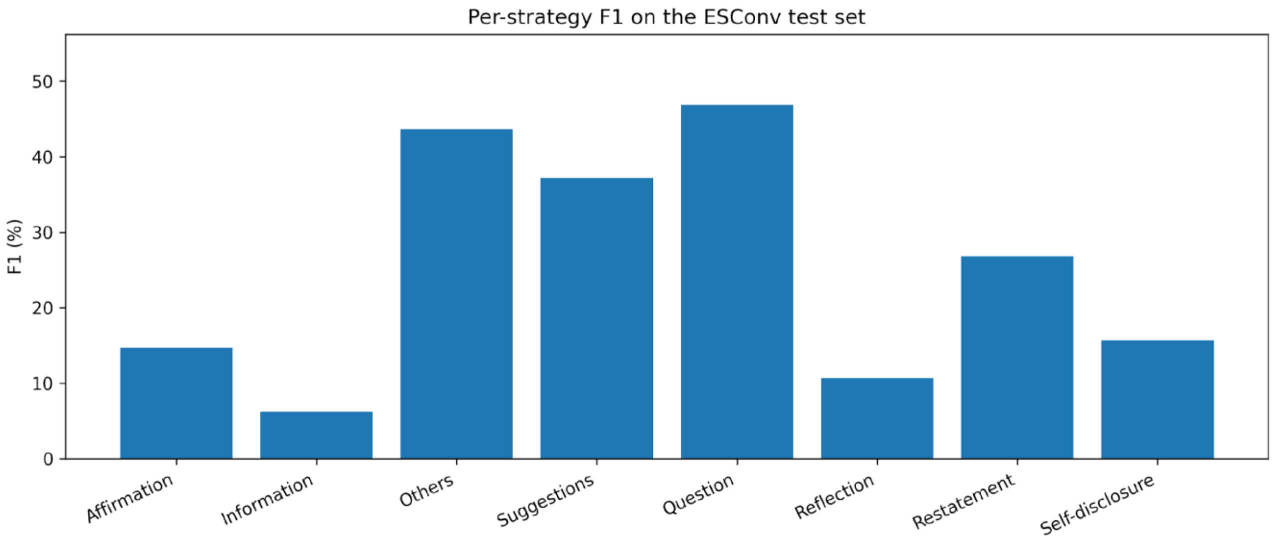

Fig. 2. Per-strategy F1 of the full strategy predictor on the ESConv test set.

Table 3 presents the main response-selection results. Context-Retrieve is the weakest model, with 3.17 BLEU-2, 11.83 ROUGE-L, 4.50 Dist-1, and 21.42% strategy match. Therapist-Rerank improves lexical overlap slightly and increases diversity to 4.80 Dist-1, but strategy match remains unchanged because the method has no explicit strategy constraint. Strategy-Retrieve produces a much larger jump in controllability, reaching 28.84% strategy match and also improving BLEU-2 to 3.50 and ROUGE-L to 12.19. The proposed SATI model performs best on every reported response metric: 3.58 BLEU-2, 12.51 ROUGE-L, 4.89 Dist-1, and 31.05% strategy match. Relative to Context-Retrieve, SATI improves BLEU-2 by 13.0%, ROUGE-L by 5.7%, Dist-1 by 8.6%, and strategy match by 9.64 absolute points. Relative to Strategy-Retrieve, SATI still adds 0.08 BLEU-2, 0.31 ROUGE-L, 0.26 Dist-1, and 2.21 points of strategy match.

Table 3. Response-selection results on the ESConv test set (%)

Method	BLEU-2	ROUGE-L	Dist-1	Strategy match
Context-Retrieve	3.17	11.83	4.50	21.42
Therapist-Rerank	3.27	11.97	4.80	21.42
Strategy-Retrieve	3.50	12.19	4.63	28.84
SATI (ours)	3.58	12.51	4.89	31.05

These results support a clear interpretation. Therapist imitation alone is not enough. The ungated reranker retrieves more human-like local phrasings, but without strategy control it often selects a supportive response that sounds reasonable while using the wrong counseling move. Strategy prediction alone is also not enough. Strategy-Retrieve improves control, but it can still choose a same-strategy response that was triggered by a different seeker concern. SATI works because it applies the two signals in sequence: strategy gating narrows the candidate pool to the correct counseling move, and seeker matching then chooses the human response that best fits the local user state. Fig. 3 visualizes this complementarity.

A second observation is that the absolute overlap scores remain low. This is not a failure of measurement; it is a property of the task. Emotional support dialogue has many acceptable next responses, but ESConv provides only one gold reply per turn. Under single-reference evaluation, even a supportive and strategy-correct alternative can receive a low BLEU score. This is why the relative differences between methods are more informative than the raw BLEU values themselves. In our experiments, the ranking of methods is stable across validation and test, which strengthens the conclusion.

We also tested a cue-heavy variant that explicitly favored a supportive lexicon during reranking. On the test set, that variant increased cue rate from 3.36% to 7.13%, but BLEU-2 fell from 3.58 to 3.41 and ROUGE-L fell from 12.51 to 12.33. We therefore kept the simpler SATI reranker for the final model. The result indicates that direct style forcing can over-regularize the output when the benchmark has single references and diverse valid phrasings.

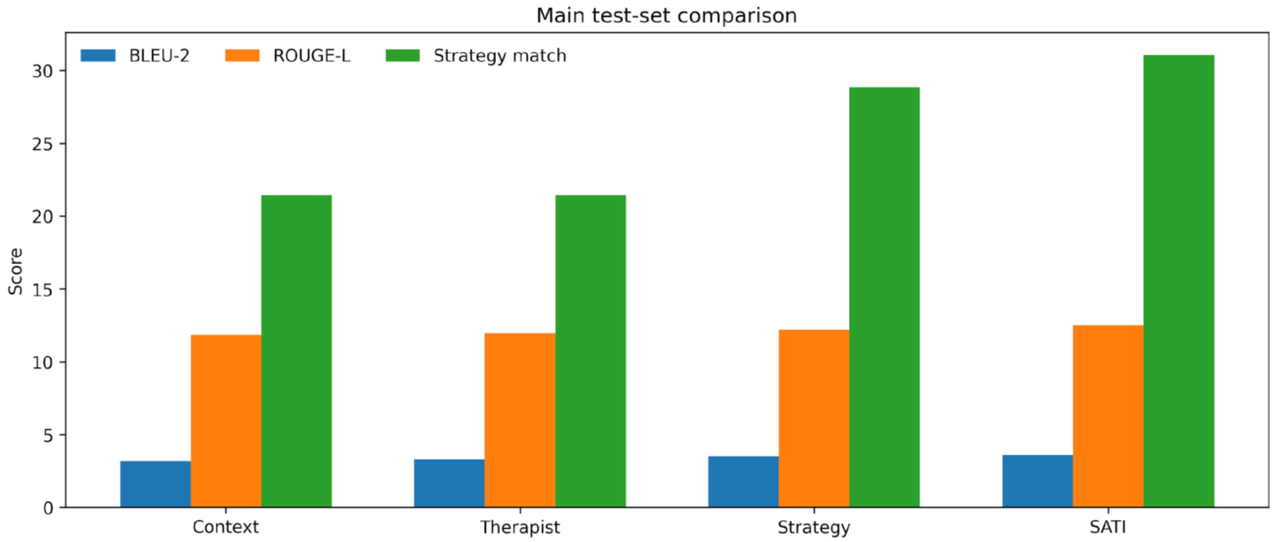

Fig. 3. Main comparison on the ESConv test set.

4. Conclusion

This paper presented SATI, a reproducible strategy-aware therapist-imitation controller for emotional support dialogue on ESConv. Using the official 910/195/195 split and full turn-level experiments, we showed that two ingredients matter: explicit counseling strategy and human-supporter imitation at the local utterance level. A simple sparse strategy predictor reached 30.81% accuracy and 25.20 macro-F1, and the full SATI controller achieved the best response-selection performance among all compared methods, with 3.58 BLEU-2, 12.51 ROUGE-L, 4.89 Dist-1, and 31.05% strategy match on the ESConv test set. The empirical pattern was consistent: context retrieval established a baseline, therapist imitation improved diversity, strategy gating improved control, and combining both gave the best overall behavior.

There are three practical future directions. First, the sparse controller should be attached to an open-source LLM so that the selected supporter exemplar can be used as an in-context demonstration rather than as the final output. Second, denser semantic retrievers and multi-strategy decoding should be explored, because some ESConv turns clearly blend reflection, reassurance, and advice. Third, future evaluation should add human judgment and safety review, especially because emotional support systems are helpful only when they remain supportive, non-coercive, and clearly non-clinical. Even so, the present results already show that small, strategy-labeled corpora can provide measurable and reproducible gains when therapist imitation and counseling strategy are modeled together.

References

[1] S. Liu, C. Zheng, O. Demasi, S. Sabour, Y. Li, Z. Yu, Y. Jiang, and M. Huang, "Towards Emotional Support Dialog Systems," in Proc. ACL-IJCNLP, 2021, pp. 3469-3483.

[2] C. E. Hill, Helping Skills: Facilitating Exploration, Insight, and Action, 3rd ed. Washington, DC, USA: American Psychological Association, 2009.

[3] H. Rashkin, E. M. Smith, M. Li, and Y.-L. Boureau, "Towards Empathetic Open-domain Conversation Models: A New Benchmark and Dataset," in Proc. ACL, 2019, pp. 5370-5381.

[4] S. Roller, E. Dinan, N. Goyal, D. Ju, M. Williamson, Y. Liu, J. Xu, M. Ott, E. M. Smith, Y.-L. Boureau, and J. Weston, "Recipes for Building an Open-Domain Chatbot," in Proc. EACL, 2021, pp. 300-325.

[5] T. B. Brown et al., "Language Models are Few-Shot Learners," in Adv. Neural Inf. Process. Syst., vol. 33, 2020, pp. 1877-1901.

[6] L. Ouyang et al., "Training Language Models to Follow Instructions with Human Feedback," arXiv:2203.02155, 2022.

[7] H. Touvron et al., "LLaMA: Open and Efficient Foundation Language Models," arXiv:2302.13971, 2023.

[8] Y. Cheng, W. Liu, W. Li, J. Wang, R. Zhao, B. Liu, X. Liang, and Y. Zheng, "Improving Multi-turn Emotional Support Dialogue Generation with Lookahead Strategy Planning," in Proc. EMNLP, 2022, pp. 3014-3029.

[9] M. Jia, Q. Chen, L. Jing, D. Fu, and R. Li, "Knowledge-enhanced Memory Model for Emotional Support Conversation," in Proc. ML4CMH@AAAI, 2024, pp. 1-11.