Table 2: GPT-4 preference results comparing RAC_{DPO} and RAC_{SFT}. Results with * are statistically significantly different based on the one-sided McNemar's test with `p < 0.05`.

Dataset	RAC_{DPO}%	Tie%	RAC_{SFT}%
Qulac	**28.88***	50.56	20.56
ClariQ	**28.36***	48.79	22.85
PaQa	**36.24**	30.36	33.40
CAMbigNQ	**16.27***	72.23	11.50

As shown in Fig. 3, performance improves as the number of passages increases, but the effect saturates after approximately four passages, suggesting that the most salient query-related ambiguities are typically captured within the top-ranked results. Passage quality is equally important: using random passages results in performance close to the "*Q-Cond*" baseline, whereas BM25 and dense retrievers achieve substantially higher scores. BM25's advantage is likely due to a domain mismatch, since the dense retriever is trained on MS MARCO, whose passage structure and content differ from the chunked ClueWeb passages used in our setting. These findings indicate that RAC benefits from informative retrieval signals and can extract relevant facets from high-quality passages rather than relying on arbitrary content, thereby addressing **RQ1**.

5.4 Impact of the Quality of Noisy Generated Elements

We study the effect of our noisy generation method by comparing p_{uncond}, fine-tuned with only the query as input, with p_{LM}, the initial language model. We then measure their impact on preference tuning (positive samples always being generated by p_{θ_0}). Table 3 shows that p_{uncond} provides more effective negative samples than p_{LM}. Unlike the approach of Duong et al. [6], which relies on generic

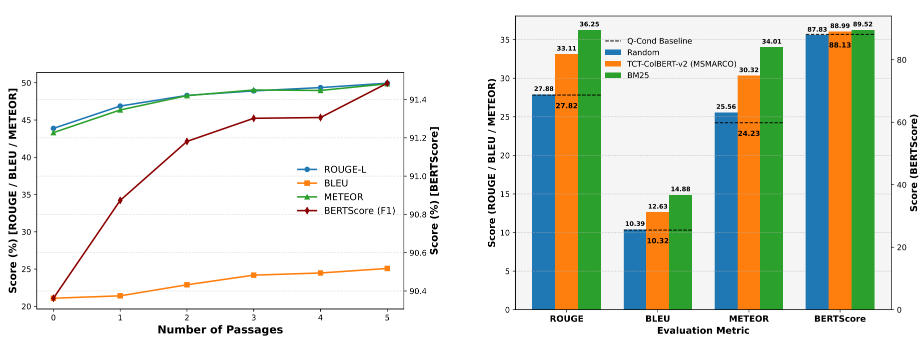

Fig. 3: NLG metrics on ClariQ: impact of varying the number of passages (left) and comparison of retrieval strategies (BM25, TCT, random) using the top 5 retrieved passages (right).

Table 3: ClariQ validation results using different negative generation methods.

Method	ROUGE-L	BLEU	METEOR	BERTScore (F1)	ALScore ↑	Par-R ↑
$RAC_{DPO},\ C_q^- \sim p_{\mathrm{LM}}$	33.84	12.93	30.79	89.25	50.81	50.73
$RAC_{DPO},\ C_q^- \sim p_{\mathrm{uncond}}$	**35.52**	**14.86**	**33.84**	**89.39**	**52.41**	**55.77**

noise injection, p_{uncond} generates clarifications that are structurally well-formed but factually misaligned. This contrast makes them harder negatives and better training signals for preference optimization. By comparison, samples from p_{LM} often fail to resemble clarifications at all, limiting their usefulness. These results highlight the importance of tailoring noise generation to the clarification format rather than reusing generic base-model outputs.

5.5 Qualitative Analysis

Noisy clarifying questions. We qualitatively assess the effect of mixture between the conditionned & unconditionned models p_{θ_0} and p_{uncond}, controlled by α (Eq.3). At $\alpha = 0$, outputs come solely from p_{θ_0}; at $\alpha = 1$, from p_{uncond}. Table 4 shows an example from ClariQ, where noise increases with α. Irrelevant spans (highlighted in red) illustrate how higher α degrades faithfulness. For preference learning, selecting intermediate α values yields negative examples that are challenging yet informative, avoiding both trivial and overly noisy supervision.

Generated Clarifying questions. We compare clarifications from RAC_{SFT} and RAC_{DPO} on ClariQ validation data (Table 5). Faithful content is highlighted in yellow, hallucinations in red. RAC_{SFT} exhibits occasional grounding failures, such as introducing unsupported entities (e.g., hallucinating “season 17”) or omitting relevant evidence (e.g., overlooking “season 16”). In contrast, RAC_{DPO} produces questions more tightly grounded in passages. This aligns

Table 4: Noisy generation at different α interpolation weights of p_{uncond} and p_{θ_0}

α	Noisy Generation
0.0	are you interested in the causes of angular cheilitis
0.1	do you want to read an article on that, do you want to be diagnosed
0.2	do you have any recommendations for treating this condition
0.3	do you want to see a picture of angular cheilitis
0.4	would you like to know about symptoms, risk factors, or a way to treat or manage your condition
0.5	would recommend a topical or internal medicine to cure your specific symptom for you based off its symptoms or how do I treat your specific angular cheilitis?
0.6	do any methods have been tested?
0.7	What works to treat this
0.8	Use warm soapy water, soft cotton balls, an old razor blade. If you do not have any of the aforementioned items, you can use a regular nail clipper.
0.9	How many times a day should I apply the treatment?

Table 5: Qualitative comparison of *RAC* against baseline models.

Input	RAC_{SFT}	RAC_{DPO}
Query: When does the new family guy season come out? **Passages:** ['... Ed O Neill Liam Payne Louis Tomlinson and Neil deGrasse Tyson During this season the guys head to South Korea after', ' Family Guy (season 15) Family Guy season 15 Family Guy fifteenth season premiered on Fox in the United States on September 25 2016 and ended on May 21 2017 The season contained 20 episodes The series follows the dysfunctional Griffin family consisting of father Peter mother Lois daughter Meg son Chris baby ...re Appel and Callaghan Guest stars for the season include Kyle Chandler Stephen Curry Flea Rob Gronkowski']	Are you looking for the release date of Family Guy season 15, 16 or 17 ?	Are you looking for the release dates of Family Guy seasons 14, 15 or 16 ?

with quantitative gains reflecting in more relevant and grounded clarifying questions, demonstrating that DPO improves reliability in ambiguous cases where SFT fails to capture the core ambiguity.

6 Conclusion

In this work, we introduced clarification question generation as a retrieval conditioned generation task, where questions are generated based on both the user query and retrieved passages. This formulation ensures that clarifications are grounded in information the system can realistically access. Our RAC framework combines retrieval context with preference tuning to improve both the relevance and corpus-faithfulness of generated questions. Experiments on four benchmarks demonstrate that both $\mathrm{RAC}_{\mathrm{SFT}}$ and $\mathrm{RAC}_{\mathrm{DPO}}$ significantly outperform existing baselines, Q-Cond and QP-$\mathrm{Zero}_{\mathrm{shot}}$, across all reference-based metrics (ROUGE-L, BLEU, METEOR, and BERTScore). We further employ LLM-as-Judge evaluations and novel metrics derived from NLI and data-to-text to quantify the gains in faithfulness to retrieved content of $\mathrm{RAC}_{\mathrm{DPO}}$ over $\mathrm{RAC}_{\mathrm{SFT}}$, which is critical for conversational search, where the objective is to disambiguate and answer user queries based on retrieved evidence rather than knowledge internal to the language model. As future work, we plan to extend this task to multi-turn clarification and evaluate its impact on downstream retrieval performance.

Acknowledgments. The authors acknowledge the ANR – FRANCE (French National Research Agency) for its financial support of the GUIDANCE project n°ANR-23-IAS1-0003 as well as the Chaire Multi-Modal/LLM ANR Cluster IA ANR-23-IACL-0007. This work was granted access to the HPC resources of IDRIS under the allocation AD011016470 made by GENCI.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.