

Evaluating the Impact of Source Diversity for RAG in Historical Research

Ruhi Mahadeshwar, Tommaso Caselli, Andreas van Cranenburgh, Malvina Nissim

CLCG, University of Groningen

The Netherlands

{r.u.mahadeshwar, t.caselli, a.w.van.cranenburgh, m.nissim}@rug.nl

Abstract

Historical research increasingly benefits from large language models (LLMs). However, LLMs are prone to factual inaccuracy, unreliability, and biased interpretations of data. Retrieval-augmented generation (RAG) approaches have emerged as solutions, but may inadvertently perpetuate biased perspectives embedded in historical archives. This paper investigates how source diversity in RAG impacts perspective variation in historical question answering. We compile a multilingual corpus (English, French, Dutch) of historical documents spanning multiple countries and focus on Napoleon Bonaparte. We evaluate three Qwen3 models across ten questions using a multi-layered framework combining traditional metrics (BERTScore, ROUGE-L), frame semantics analysis, and syntactic profiling. Our results highlight that, while traditional similarity metrics suggest high semantic consistency, frame-semantic analysis exposes substantial perspective shifts. Baseline answers present “flattened” cross-lingual perspectives, whereas RAG introduces diversity. Critically, this diversity manifests differently across languages, demonstrating language-specific patterns. Our findings highlight limitations of traditional evaluation metrics for perspective-sensitive tasks and demonstrate that RAG constitutes active perspective transformation rather than neutral augmentation.

Keywords: perspective, RAG, LLMs, bias, historical research

1. Introduction and Background

Long dominated by close-reading approaches, like many other disciplines in the humanities, historical research has recently benefited from advances in Natural Language Processing (NLP) and digital technologies. One more recent development is that Large Language Models (LLMs) are increasingly being incorporated into historical research as assistive tools (Garcia and Weilbach, 2023; González-Gallardo et al., 2024; Hauser et al., 2024; Zeng, 2024; Celli and Basile, 2024; Celli and Samal, 2025). LLMs are trained on massive amounts of data, thus are able to deal fluently with natural language in both input and output, and serve as knowledge repositories. As a result, LLMs are used to support multiple forms of historical analysis; for example by identifying analogies, conceptual patterns, and relationships across time and contexts, thereby enabling historians to pose broader or more complex questions about continuity, change, and causation in the historical records (Li et al., 2025).

The advantages of LLMs are offset by substantial limitations. LLMs exhibit well-known issues of factual inaccuracy, which compromises the reliability of extracted information (Jaskulski et al., 2025), and also struggle with the contextual and diachronic subtleties of historical language, leading to interpretative errors. Retrieval-Augmented Generation (RAG) approaches (Lewis et al., 2020) are rapidly gaining popularity in historical research as a way to mitigate the shortcomings of adopting LLMs. The use of external retrieval to bring controlled domain-

specific or time-sensitive material, such as archival documents, into the generative pipeline should intuitively help to combat hallucinations and the potential lack of transparency in responses (Gao et al., 2023). Being able to identify exactly which source text underpins a given generated claim is a rather fundamental aspect of scholarship, especially in historical research (Zhou et al., 2025).

While the advantages are evident, most approaches which use RAG-based LLMs for historical research fail to highlight a crucial negative aspect of over-reliance on documents from time-specific historical archives, and how it might perpetuate biased narratives or perspectives from that era. Indeed, the reporting of events always entails a perspective or focus (Bal, 2009) even unconsciously. This of course holds also in the case of historical documents; no matter how impartial and comprehensive one tries to be, the act of recording history will always exclude or modify some perspectives. Integrating RAG into the pipeline, and thereby foregrounding particular historical materials, inevitably privileges the perspectives embedded within those materials. However, when evaluating RAG-based LLMs for historical research, this aspect of possibly inducing a biased view through RAG-fed materials is not taken into account as frequently, despite the impact it may have.

We suggest that if RAG-based LLMs are to be used to empower historical research, the effect of RAG must be examined more closely, and, where possible, automatically quantified, especially in terms of perspective change. To make this goal

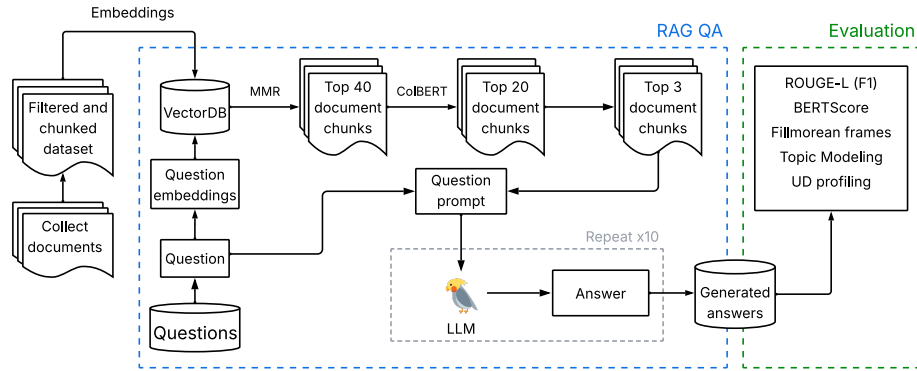


Figure 1: The experimental setup to evaluate impact of source diversity for RAG in historical research.

concrete, this paper investigates¹ the effect of using contextually different historical documents on question answering over a historical figure using RAG-based LLMs of differing sizes.

Case Study As a *methodological proof of concept*, we focus on Napoleon Bonaparte, whose historical legacy provides an ideal case for studying perspective variations. The public image and memory of Napoleon in early 19th century was, and still is, vastly different. Between 1803–1808, British soldiers portrayed Napoleon as a tyrant, a criminal, a conqueror, but also as a progressive law maker and an agent of enlightened civilization (Daly, 2018). Around the same period in France, parallels between Napoleon and Christ were drawn, with allegorical paintings were often made, to cast him as savior of France (Dwyer, 2004). Between 1806–1813 in the Netherlands, literary activism against Napoleon was on the rise at provincial and national levels (Jensen, 2011).

To investigate how the varied provenance of historical sources shapes RAG-based historical analyses of the figure of Napoleon, we (i) compile a multilingual dataset with historical documents ca. 1805–1819, from the Netherlands, France, US, and UK; (ii) develop a set of questions in collaboration with historians and prompt LLMs of varying scales using RAG with sources drawn from our datasets; and (iii) evaluate the answers with respect to syntax and semantics to examine, using a variety of automatic metrics, if and how the perspective taken by the LLMs on Napoleon changes with respect to the given contextually different historical documents. Figure 1 outlines the experimental setup.

Contributions We summarize our contributions and main findings as follows:

- We introduce a multilingual corpus of historical documents (ca. 1805–1819) from multiple

countries (Netherlands, France, US, and UK) that could be used to investigate perspective variations in multiple topics;

- We propose a novel evaluation approach that combines multiple methods to assess perspective differences in RAG-based LLM outputs;
- We show that the source country and language of documents used in RAG impacts both the content and structure of generated answers, with these effects manifesting differently within and across languages.

2. Related Work

Bias and Perspective Gallegos et al. (2024) show how LLMs engage in erasure of language and experiences of a minority social group up to denigrating attitudes towards them and reinforcing normativity of dominant groups. In multilingual settings, LLMs tend to reflect stereotypes in different languages differently (Mitchell et al., 2025), with stereotypes expressed in one language showing up in model behavior in another language (Cao et al., 2025). They also show more significant moral bias when probed with languages other than English with a prominent gap in low-resource languages (Agarwal et al., 2024). LLMs also show signs of alignment and biases which reflect those of the country they were developed (Chang et al., 2025).

Exploring multiple perspectives, also known as multiperspectivity (Stradling, 2003; Hartner, 2014), in a historical context can be exceedingly valuable. Multiperspectivity is shown to be an effective approach to learning and teaching history on topics such as post-conflict resolution and understanding (McCully, 2012), intercultural understanding (Goldberg and Ron, 2014), development of democratic citizenship (Brett et al., 2009), and controversial issues (Goldberg and Savenije, 2018).

Minnema et al. (2022) propose a framework to automatically detect and systematically analyse perspectives adopting a frame semantics-based ap-

¹A video summary of the paper can be found at <https://imruhi.github.io/posts/2026/03/13/>.

proach. While we will follow part of their framework in the analysis of our results (Section 4), our contribution is mainly concerned with examining the effect of using contextually different historical documents in the generation of texts that *may* result in different perspectives.

RAG in Historical Research Frequent problems exhibited by LLMs which are relevant for historical research are hallucinations (Ji et al., 2023; Dahl et al., 2024) and long-tail data (Kotek et al., 2023; Stranischi et al., 2023). To overcome these, RAG was introduced as a framework to provide the provenance of answers generated by LLMs and to update their world knowledge in an efficient manner (Lewis et al., 2020; Shuster et al., 2021). Applications of RAG-based LLMs in historical research include knowledge extraction on oral historical archives (Sun et al., 2024), enriching Wikipedia biography pages through the Internet Archive (Adak et al., 2025), or creating a question-answering system for newspaper articles from 19th to early 20th century (Tran et al., 2025). In the historical domain, RAG is mostly utilized to deal with unstructured data present in historical documents, their multilingual nature, as well as to reduce hallucinations. The impact of source diversity in RAG-based historical analysis has not been explicitly studied.

3. Dataset and Experimental Setup

We have collected historical documents covering a time period of 15 years (1805–1819). These years correspond to the two reigns of Napoleon, including the Napoleonic Wars, and the aftermath of his defeat at Waterloo with the Congress of Wien and the beginning of the Conservative Order. To gain access to (potentially) different perspectives related to the parties involved in the Napoleon-related events of those years (France, England, Northern America, The Netherlands), we have collected a multilingual corpus in English (EN), French (FR), and Dutch (NL). We selected magazines, newspaper articles, non-fiction books, and legislative texts (placards) from the British Library (BL, 2021), the Bibliothèque Nationale de France (BnF, 2024), the KB National Library of the Netherlands (KB, 2019, 2025), the American Stories dataset (Dell et al., 2023), and the Dutch-Indonesian placard² books (Romein, 2023). Each text type comes from different Cultural Heritage (CH) institutions or open-source initiatives responsible for the digitization and OCR curation. Figure 2 shows an example from the newspaper

²Placard books are a collection of placards, which contain administrative laws, regulations, treaties and other official documents which were in force for the given geographic area, as deemed by government officials.

“Journal des débats” retrieved from the Bibliothèque Nationale de France, with more in Appendix C.



Figure 2: Example from a newspaper from the Bibliothèque Nationale de France.

Each collected document has been preprocessed and chunked into paragraphs so that (a) it would fit into the LLMs’ context windows; and (b) it follows a *best practice* in the implementation of RAG-based systems to facilitate the retrieval of relevant passages. The preprocessing steps were kept minimal: removal of characters which are not alphanumeric, punctuation, and extra whitespaces. Due to the size of the initial corpus collection (more than 2.2 million chunks), we have run a topic model using BERTopic (Grootendorst, 2022) to identify only those chunks containing topics relating to Napoleon. The filtered dataset resulted in $\approx 0.41\%$ of the original corpus. Table 1 reports the distribution of the chunks per language.

Text Type	Source	EN	FR	NL
News	KB-Delpher	103	18	2,798
	BnF	0	850	0
	Am. Stories	3,238	0	0
Books	KB-DBNL	0	0	1,045
	BL	1,039	45	37
Legal	Ind. Placards	0	0	3
Total		4,380	913	3,880

Table 1: Number of document chunks (filtered after topic modeling) for source and language

Most of the filtered chunks are in English, followed by Dutch and, lastly, French. When curating documents from this period, one must account for censorship and press suppression, which severely limited access. In the Netherlands, a decree issued on February 5, 1810 imposed strict

regulations on printing and distribution, explicitly prohibiting the publication of material critical of the French Empire (Jensen, 2011). Similarly, in France between 1800 and 1811, newspapers were rapidly closed, and the founding of new journals was forbidden under a decree passed on January 17, 1800 (Dwyer, 2015). These measures collectively curtailed the circulation of information and restricted the diversity of perspectives. It is therefore not surprising that most of the retrieved chunks are English: English data come from the US and UK, two countries who have never experienced direct ruling from the Napoleonic Empire, with the UK strongly opposing Napoleon. Thus, to reflect the real world scenarios often encountered at CH institutions, we decide not to balance the dataset.

RAG experiments We frame our perspective task³ as a Question-Answering task. We asked three historians with expertise in early 19th century Europe to help us devise 10 questions addressing Napoleon’s ruling, his character, and how he became Emperor. Questions were divided into two equal groups: five questions with a neutral tone (e.g., “Who was Napoleon Bonaparte?”) and five primed-to-bias (e.g., “Was Napoleon Bonaparte a hero?”). Questions were initially written in English and subsequently translated into French and Dutch by native speakers. The full set of questions can be found in Appendix A.

For the retrieval component, document chunks and questions are embedded using `distiluse-base-multilingual-cased-v1` (Reimers and Gurevych, 2019) and indexed into a FAISS vector storage (Johnson et al., 2019; Douze et al., 2024). Document chunks are retrieved using the Maximal Marginal Relevance (MMR) criterion, which aims to balance redundancy of retrieved document chunks and their relevance to the query, with cosine similarity used as distance metric. We initially retrieve the top 40 most similar chunks, and retain only the top 20 most diverse. These documents are passed to a re-ranker based on `ColBERT-v2` (Santhanam et al., 2022) and only the top three are retained and passed to the LLMs to generate the answers.

For the answer generation, we have selected three multilingual models of different sizes, ranging from small to very large, from the same family, namely `Qwen3-4B`, `Qwen3-30B-A3B`, and `Qwen3-235B-A22B`. We set the `temperature` to 0.6 and `max_new_tokens` to 300 for all models. We designed question prompts (Appendix B) in each of the three languages and the models were instructed to produce responses in the correspond-

³The corresponding code for the experimental setup, prompts, questions and answers can be found at <https://github.com/imruhi/perspectiveRAG>.

ing language. Each question is answered by each model ten times, although we retrieved the document chunks only once. This means that for each model and each setting, we collect and analyze 100 answers per language. RAG answers are compared against a baseline version where the models rely on their parametric knowledge.

Evaluation set-up To comprehensively assess variation across LLM outputs, we employ multiple evaluation strategies, each capturing different levels of granularity. The evaluation metrics range from general text similarity measures to more fine-grained linguistic analyses. Specifically, we use ROUGE-L (F1) to compare model outputs within a single language and BERTScore to assess similarity both within and across languages.

To quantify differences in perspectives, we extracted Fillmorean frames (Fillmore, 1985; Fillmore et al., 2006) from the generated answers in the two experiment settings, following Minnema et al. (2022). For English, we extracted frames using `frame-semantic-transformer` (Chanin, 2023). The reported⁴ F1 scores for the tasks of trigger identification, frame classification and argument extraction on FrameNet 1.7 are 0.74, 0.89 and 0.75 respectively. For French and Dutch, we resorted to machine translation into English using DeepL and re-run the same model. To determine whether the extracted frames from the translated texts are reasonable, we run a multilingual BERTopic model fitted on the original untranslated answers to qualitatively assess if any topics from the original untranslated answers correlate with frames from the translated answers.

We integrated the frame analysis by examining the syntactic complexity and diversity of the answers through the Profiling UD pipeline (Brunato et al., 2020). In particular, we fit a multinomial logistic regression model to the features extracted by Profiling UD to quantify the influence of these linguistic characteristics on answer variation.

4. Results

In the discussion of the results, `baseline` will be used to refer to answers from the models’ parametric knowledge, and `augmented` for the RAG setting. The total number of answers we have analyzed is 1,800. Our analyses focus on evaluating the variability in content and syntax of `augmented` answers when compared to `baseline` answers in (i) each language and (ii) between languages. We also evaluate the similarity in content between the `augmented` answers and its respective retrieved

⁴For a complete overview of the evaluation and results, refer to Chanin (2023).

Baseline	RAG
[...] the overall impact on England was not destructive; it ultimately reinforced national strength [...] Napoleon's impact on the Netherlands was more directly disruptive [...] suffering from blocked trade with Britain, leading to economic decline.	While the texts imply that Napoleon threatened European order [...] they do not provide clear details on whether his actual impact on England and the Netherlands was destructive [...] based solely on these excerpts, there is insufficient evidence to conclude [...] his impact on England and the Netherlands was destructive.

Table 2: Generated answers (baseline vs. RAG) for Qwen3-235B-A22B model for the question “Was the impact of Napoleon Bonaparte on England and the Netherlands destructive?” in English.

document chunks. In Table 2 we report an example of the generated answers in the two experiment settings for English to showcase the kind of data we have used to conduct our analyses, with more examples in Appendix D.

4.1. Text Similarity Analysis

Within language, RAG mostly impacts on the wording rather than on the content. Figure 3 shows the median and inter-quartile ranges (IQRs) of BERTScore and ROUGE-L when comparing all baseline answers to the augmented ones within the same language for each of the models. The median BERTScore remains constant around 0.8-0.9 for all models and languages, with the exception of Qwen3-4B in French. For ROUGE-L, the median remains constant above 0.2, with again the exception of Qwen3-4B in French. The range of ROUGE-L is skewed slightly towards 1 for English for Qwen3-30B-A3B and Qwen3-235B-A22B, indicating a higher variability in the wording.

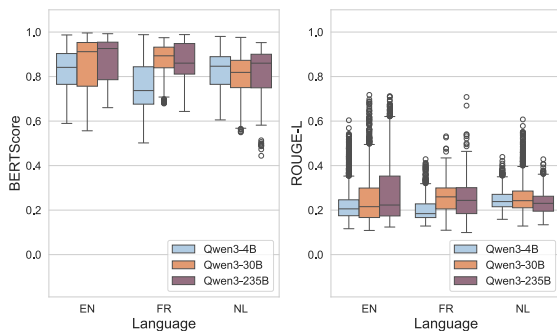


Figure 3: Distribution of BERTScore and ROUGE-L between baseline and augmented answers within language per model.

Table 3 shows the mean BERTScore and ROUGE-L, and their standard deviations when comparing all baseline answers to all augmented answers for each question across all models. For BERTScore, French has more questions with standard deviation higher than 0.1 when compared to English and Dutch. Standard deviation for ROUGE-L is low for almost all questions and languages. **At the question level, and within the**

same language across all models, variation in BERTScore values is minimal for English and Dutch (Table 3) confirming the patterns observed in Figure 3. As expected, French exhibits greater semantic variation across questions. When examining ROUGE-L scores (Table 3), English and Dutch again display consistent behavior with Figure 3, showing slightly lower lexical variation compared to French. An exception arises for Question 6 (Q6) in English, which shows unusually low lexical variability yet high standard deviations. Although Q6 is framed neutrally, its content (“How did Napoleon Bonaparte become the Emperor of France?”) appears to elicit divergent model responses between the baseline and augmented conditions. Overall, French maintains its outlier status, showing a systematic tendency toward higher lexical variation across all questions and experimental settings.

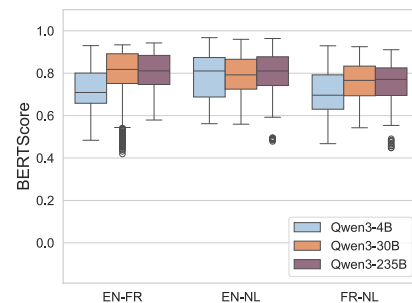


Figure 4: Distribution of BERTScore between baseline and augmented answers within language pairs per model.

Across languages, English and Dutch have more similar content than for French. As shown in Figure 4, where BERTScore is computed for all questions per model, there is an overall trend toward generating answers with highly similar content. A notable exception is observed for the smallest model, Qwen3-4B, in French, where cross-lingual differences in content are more pronounced. At the question level, summarized in Table 4, variations in BERTScore remain generally small across all language pairs. The largest deviations occur for Questions 8 and 9 (Q8 and Q9), which are designed as counterparts:

	BERTScore			ROUGE-L		
	EN	FR	NL	EN	FR	NL
Q1	0.96 ± 0.01	0.95 ± 0.03	0.93 ± 0.02	0.44 ± 0.09	0.31 ± 0.05	0.36 ± 0.07
Q2	0.91 ± 0.06	0.85 ± 0.11	0.72 ± 0.06	0.26 ± 0.07	0.23 ± 0.06	0.23 ± 0.04
Q3	0.74 ± 0.05	0.72 ± 0.03	0.73 ± 0.07	0.15 ± 0.02	0.15 ± 0.02	0.20 ± 0.03
Q4	0.87 ± 0.07	0.82 ± 0.03	0.74 ± 0.09	0.20 ± 0.02	0.16 ± 0.02	0.20 ± 0.03
Q5	0.93 ± 0.02	0.87 ± 0.11	0.87 ± 0.05	0.23 ± 0.04	0.26 ± 0.07	0.26 ± 0.04
Q6	0.97 ± 0.02	0.87 ± 0.12	0.89 ± 0.03	0.52 ± 0.10	0.29 ± 0.09	0.31 ± 0.08
Q7	0.91 ± 0.04	0.87 ± 0.04	0.75 ± 0.06	0.28 ± 0.04	0.27 ± 0.03	0.21 ± 0.04
Q8	0.75 ± 0.03	0.78 ± 0.13	0.85 ± 0.04	0.18 ± 0.01	0.22 ± 0.04	0.24 ± 0.02
Q9	0.69 ± 0.06	0.78 ± 0.11	0.85 ± 0.04	0.16 ± 0.02	0.22 ± 0.05	0.23 ± 0.02
Q10	0.81 ± 0.02	0.83 ± 0.05	0.86 ± 0.03	0.19 ± 0.02	0.25 ± 0.06	0.25 ± 0.03

Table 3: BERTScore and ROUGE-L with respective standard deviations of answers between setting per question per languages. Standard deviations 0.1 and higher are in red.

Q8 presents a biased formulation, while Q9 introduces a neutral framing, both addressing the impact of Napoleonic rule on England and the Netherlands. In this case, the neutral version (Q9) tends to elicit greater divergence across models, particularly between the English–French and English–Dutch pairs. This suggests that linguistic framing and, in the augmented setting, differences in retrieved document segments, can influence the perspectives expressed by the models.

	EN-FR	EN-NL	FR-NL
Q1	0.90 ± 0.02	0.93 ± 0.02	0.88 ± 0.03
Q2	0.82 ± 0.08	0.72 ± 0.06	0.68 ± 0.07
Q3	0.71 ± 0.05	0.69 ± 0.06	0.67 ± 0.05
Q4	0.76 ± 0.06	0.70 ± 0.07	0.67 ± 0.06
Q5	0.86 ± 0.09	0.86 ± 0.03	0.81 ± 0.08
Q6	0.86 ± 0.09	0.87 ± 0.04	0.81 ± 0.07
Q7	0.78 ± 0.04	0.80 ± 0.06	0.71 ± 0.05
Q8	0.71 ± 0.10	0.78 ± 0.07	0.75 ± 0.09
Q9	0.68 ± 0.11	0.76 ± 0.10	0.73 ± 0.08
Q10	0.73 ± 0.05	0.81 ± 0.05	0.75 ± 0.07

Table 4: BERTScore with respective standard deviations of answers between setting per language pair and per question. Standard deviations 0.1 and higher are in red.

The content similarity between a RAG answer and its re-ranked documents is low for all languages. Table 5 shows the BERTScore between an augmented answer and its respective re-ranked document chunks for all models, per language and per question. For all questions, the mean BERTScore for English is slightly higher when compared to French or Dutch. Additionally, we notice that there is more variability in BERTScore for Dutch in Q8, Q9 and Q10 when compared to English or French. We observe a higher degree of semantic similarity between the English re-ranked document chunks and the corresponding English augmented answers, compared to French and Dutch. In contrast, Dutch shows

greater variation in semantic alignment between its re-ranked document chunks and augmented answers, indicating less consistent integration of retrieved content. Less prominently, French is close to Dutch. This raises questions on whether, when models are prompted in Dutch or French, they are bypassing the retrieved chunks, and thus relying more on their parametric knowledge.

	EN	FR	NL
Q1	0.55 ± 0.03	0.39 ± 0.06	0.45 ± 0.08
Q2	0.57 ± 0.06	0.32 ± 0.05	0.39 ± 0.08
Q3	0.57 ± 0.06	0.39 ± 0.07	0.40 ± 0.06
Q4	0.51 ± 0.08	0.43 ± 0.08	0.43 ± 0.06
Q5	0.62 ± 0.03	0.52 ± 0.04	0.47 ± 0.07
Q6	0.50 ± 0.05	0.42 ± 0.05	0.39 ± 0.03
Q7	0.62 ± 0.07	0.54 ± 0.08	0.47 ± 0.04
Q8	0.52 ± 0.05	0.42 ± 0.05	0.51 ± 0.10
Q9	0.49 ± 0.08	0.42 ± 0.05	0.47 ± 0.11
Q10	0.50 ± 0.12	0.41 ± 0.04	0.43 ± 0.10

Table 5: BERTScore of an answer and its respective retrieved document chunks per languages, for augmented setting per question. Standard deviations 0.1 and higher are in red.

4.2. Semantic Frame Analysis

Overall, we observe that the frame distributions differs between the baseline and augmented answers across all languages, with certain frames being more frequent in the baseline outputs than in their augmented counterparts.

Table 6 reports the top ten frames per language. As observed, there are four core frames (Leadership, Military, Objective_influence and Political_locales) which occur in both experiment settings for all languages. Two of them, namely Leadership and Political_locales, are more frequent in baseline, suggesting that the perspective of augmented answers tends to move away from viewing Napoleon as a leader (frame Leadership) who has connections to many places in Europe (frame Po-

litical_locales) to a figure with particular impact in Europe (frame Objective_influence). The Military frame, which refers to any person, organization or country controlling an armed group, remains constant between the two settings indicating that armed forces are a core aspect of the image of Napoleon in the selected time period. For the *augmented* answers, the frames Evidence, Documents and Statement are the most frequent in all languages, which is expected.

	Baseline			Augmented		
	EN	FR	NL	EN	FR	NL
Attributed_information	-	-	-	-	190	-
Being_in_control	-	-	-	162	-	-
Categorization	185	149	194	130	242	-
Causation	343	227	303	-	-	337
Change_of_leadership	193	-	-	-	-	-
Documents	-	-	-	309	282	-
Economy	186	176	-	-	-	246
Emotion_directed	-	-	-	146	-	-
Evidence	-	-	-	220	168	219
Exchange	-	-	-	-	-	205
Hostile_encounter	336	237	414	-	204	282
Importance	-	-	171	-	-	-
Intentionally_create	220	248	178	-	-	-
Leadership	1067	702	1102	926	630	843
Military	312	284	263	255	290	214
Objective_influence	369	270	330	457	279	349
Political_locales	636	583	646	297	488	484
Statement	-	-	-	292	253	195
Temporal_collocation	-	150	222	-	-	-

Table 6: Counts of top ten most occurring frames per language. No counts displayed signifies that the frames were not in top ten for that language.

Figure 5 shows the frames with the largest differences between the two settings (delta counts, *baseline*−*augmented*) of the frames in all answers. As expected, the Evidence frame appears most frequently in the *augmented* answers, indicating that the RAG approach actively draws on the retrieved document chunks. However, manual inspection reveals that this frame is also invoked when the models deem the retrieved material insufficient to provide adequate “evidence” for answering the question, as showcased in Table 2 for English. This result is also in line with our observations from the analysis of the semantic similarity between the retrieved document chunks and the answers (Table 5 in Section 4.1). For English, the *augmented* answers shift away from frames describing causation (Causation), leadership (Leadership), creation (Intentionally_create), political locations (Political_locales), economy (Economy), trade (Exchange), and wars (Hostile_encounter). Instead, they more frequently evoke frames related to documents (Documents), evidence (Evidence), attributing propositions (Attributed_information, Statement), and influence (Importance). In Dutch, the *augmented* answers tend to focus more on

purpose (Purpose), documents (Documents), evidence (Evidence), and trading (Exchange) rather than describing wars (Hostile_encounter), leadership (Leadership, Change_of_leadership) and political locations (Political_locales). In French, they shift from discussing economy, political locations, and intentional creation to emphasizing supply (Supply), risk (Risky_situation), attributed propositions, documents, evidence, and trade. Across languages, **we observe that baseline answers consistently presents a “flattened” perspective on Napoleon with minimal differences across the three languages, suggesting a lack of cross-lingual variation in the parametric knowledge.** On the other hand, we observe a consistent shift from event-oriented frames (Leadership, Causation, Political_locales, Hostile_encounter) emphasizing Napoleon’s ruling and war campaigns in *baseline* to frames introducing referential perspectives (Documents, Evidence, Attributed_information, Purpose) in *augmented* which emphasize perspective variation.

We further examine the frames in the *augmented* answers and their corresponding retrieved document chunks to assess to what extent they overlap in content and perspectives, as proxied by the frames. The document chunks were translated to English using DeepL for French and Dutch. For each language, we extracted the top five frames from the relevant document chunks for each question. We then checked their overlap with the generated answers from the *augmented* setting. Results are presented in Table 7. As shown in the figures, the frame overlap aligns closely with the counts reported in Table 6 for the *augmented* answers, confirming that these answers primarily draw on the retrieved document chunks rather than relying on the model’s parametric knowledge.

	EN	FR	NL
Causation	-	57	39
Economy	-	61	37
Exchange	30	59	-
Hostile_encounter	42	70	60
Leadership	75	-	-
Objective_influence	36	66	63
Political_locales	67	-	35

Table 7: Counts of top five most occurring frames per language for retrieved documents. No counts displayed signifies that the frames were not in top ten for that language.

As a sanity check to validate the reliability of the observations between frames and the untranslated *augmented* answers, we have run a multi-lingual topic model. In this case, we used the original answers in the three languages. Table 8 shows the delta counts (*baseline*−*augmented*) of the first

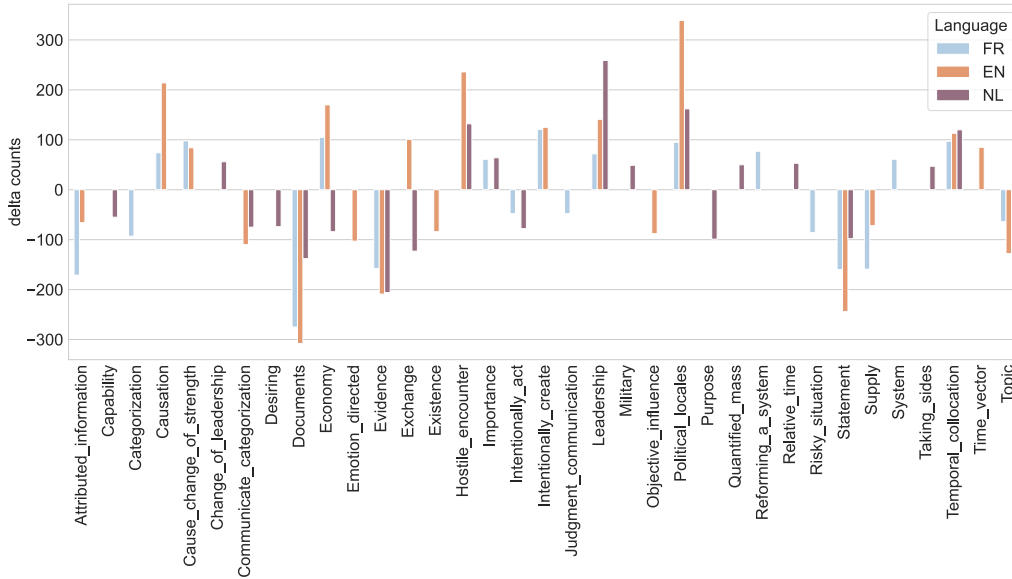


Figure 5: Top ten largest delta counts of frames for French, English and Dutch. Positive counts signify that a frame occurs more in `baseline` answers than in `augmented` answers. No counts displayed signifies that the delta change of the frame counts was not in the top ten for that language.

	EN	FR	NL
(1) England, Netherlands, documents, do not, provide, impact	-255	-	-32
(2) Europe, claim, formal, European, his, facto	-107	-46	0
(3) Bonaparte, villain, tyrant, Napoleon, was	-88	-52	-73
(4) on, Netherlands, England, Napoleon, impact	-19	-156	-
(5) England, continental, trade, system, isolate	39	-4	-46
(6) code, civil, legal, reforms, administration, Napoleon	172	86	167
(7) England, British, led, industrial, marine, economy	68	40	-30
(8) 1804, king, French, France, became, 1799	-	44	53
(9) 1815, fall, down, kingdom, defeat, independence	-	17	-
(10) war, British, French, Navy, Royal, led	-	-	43

Table 8: Top ten biggest delta counts of topic representations for French, English and Dutch. Positive counts signify that a representation occurs more in `baseline` answers than in `augmented` answers. No counts signifies that the delta change of the representation was not in the top ten for that language.

ten topics, together with their representing tokens. As shown in the topic modeling results, Topics (1) and (2) are closely associated with the Evidence and Documents frames and occur more frequently in the `augmented` answers. Similarly, Topics (6), (8), and (10), which correspond to the Leadership

and Political_locales frames, are more common in the `baseline` answers. Topics related to the Economy and Exchange frames, such as (5) and (7), display a language-dependent pattern: they occur more frequently in English and French for the `baseline` condition, but more often in Dutch for the `augmented` setting. These correspondences indicate a strong alignment between the topic model representations and the frame-based analysis.

Overall, the frame and topic modeling results jointly demonstrate that augmented answers introduce distinct topical emphases compared to baseline answers. Topics associated with evidence provision, economic conditions or trading tend to appear more in the `augmented` condition, either across all languages or in specific ones, while themes related to leadership, political entities, and importance are more prominent in the `baseline`.

4.3. Linguistic Profiling and Syntactic Variation

To examine whether the lexical variation between the `baseline` and the `augmented` answers as indicated by the ROUGE-L scores (Section 4.1) can be associated with different parts-of-speech (POS), we extracted the linguistic profiling using Profiling UD. A multinomial logistic regression was trained on the extracted features to observe if the distribution of POS affect the accuracy of predicting whether the answer is from a specific language and experiment setting. Figure 6 shows the confusion matrix for the trained model. The trained model can differentiate with high accuracy, namely

0.83±0.024 for both settings and all languages⁵. For English, we see very little errors between correctly classifying *augmented* and *baseline* answers, while errors are more likely to occur for French and Dutch. Additionally, there is no within language classification error.

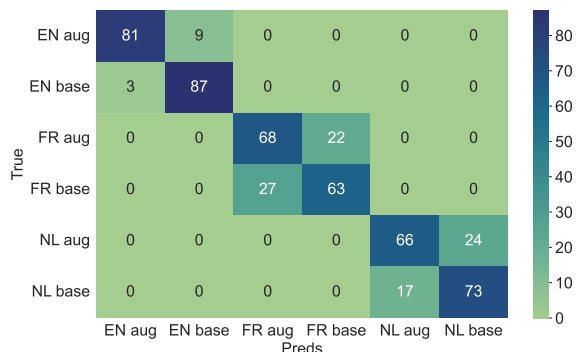


Figure 6: Confusion matrix for the fitted model on features of generated answers from the test set.

Using permutation feature importance, the top five most common features which contribute positively to model accuracy are the distributions of past tense verbs, determiners, singular proper nouns, subordinating conjunctions (and prepositions) and gerund verbs in the answers. Thus, the syntactic characteristics are distinctly different in answers for *baseline* and *augmented* for all three languages, the variation emerging most likely due to change in setting. **This indicates that augmented answers not only contain a variation in content but also in syntactic constructions when compared to the baseline answers.**

5. Conclusion

This paper aimed to examine the effect of using contextually different historical documents in English, French and Dutch, on question answering over a historical figure using RAG-based LLMs. We observe that baseline answers when compared to RAG answers, within languages and between languages, tend to be semantically similar but with high variation in wording. Despite this high similarity, the baseline and RAG answers discuss differing topics, with some of them being language specific. This highlights how BERTScore is not suitable as a metric to examine subtleties in model output which reveal changes in perspectives.

Baseline answers present a “flattened” perspective across languages, with minimal cross-lingual variation in parametric knowledge. In contrast, RAG answers introduce perspective diversity by

shifting from event-oriented frames to referential ones. This diversity manifests differently across languages demonstrating language-specific patterns.

Additionally, syntax between baseline and RAG answers is distinctly different. Although one would expect syntactic differences among languages, we also observe differences across the experiment settings. This indicates that RAG impacts *both* syntax *and* content of the generated texts.

Overall, our findings show key limitations of traditional evaluation metrics for perspective-sensitive tasks and highlight the need for appropriate data curation. The presence of certain topics in the RAG answers, being language dependent, highlights the importance of multiperspectivity. For historical research practice, our results demonstrate that RAG is not a neutral augmentation but an active perspective transformation. Future work should develop perspective-aware evaluation frameworks optimized for diversity.

6. Limitations

The use of Machine Translation (MT) for French and Dutch to extract frames is a limit due to the risks that MT systems may fail to capture cultural nuances and lexical richness (Vanmassenhove et al., 2019; Budimir, 2025), leading to a loss of meaning. Nevertheless, this was a forced decision due to a lack of large enough language specific resources for Frame Semantic Parsing and the difficulties we faced in adapting existing monolingual semantic parsers to a multilingual setting (e.g., repositories not maintained, software packages not being available). As for the evaluation of the retrieval component, we rely on an existing approach which has been adopted in many off-the-shelf RAG systems.

7. Acknowledgements

Experiments in this paper were carried out on the Hábrók HPC supported by the Center of Information Technology of the University of Groningen, and the National Supercomputer Snellius supported by SURF. Additionally, this research was supported by grants from NVIDIA and utilized the NVIDIA DGX Spark. This research is funded by the Netherlands Organization for Scientific Research under the HAICu project (NWO NWA # 1518.22.105). The authors would also like to thank Francesco Buscemi at the University of Groningen for his support in devising the questions relating to Napoleon Bonaparte.

⁵The accuracy was observed for repeated ($r=3$) class-wise stratified 10-Fold cross validation.

8. Bibliographical References

- Sayantan Adak, Pauras Mangesh Meher, Paramita Das, and Animesh Mukherjee. 2025. [REVerSum: A multi-staged retrieval-augmented generation method to enhance Wikipedia tail biographies through personal narratives](#). In *Proceedings of the 31st International Conference on Computational Linguistics: Industry Track*, pages 732–750, Abu Dhabi, UAE. Association for Computational Linguistics.
- Utkarsh Agarwal, Kumar Tanmay, Aditi Khandelwal, and Monojit Choudhury. 2024. [Ethical reasoning and moral value alignment of LLMs depend on the language we prompt them in](#). In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 6330–6340, Torino, Italia. ELRA and ICCL.
- Mieke Bal. 2009. *Narratology: Introduction to the theory of narrative*. University of Toronto Press.
- Peter Brett, Pascale Mompoin-Gaillard, Maria Helena Salema, and Sarah Keating-Chetwynd. 2009. *How all teachers can support citizenship and human rights education: a framework for the development of competences*. Council of Europe.
- Dominique Brunato, Andrea Cimino, Felice Dell’Orletta, Giulia Venturi, and Simonetta Montemagni. 2020. [Profiling-UD: a Tool for Linguistic Profiling of Texts](#). In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 7145–7151, Marseille, France. European Language Resources Association.
- Bojana Budimir. 2025. [The challenge of translating culture-specific items: Evaluating MT and LLMs compared to human translators](#). In *Proceedings of Machine Translation Summit XX: Volume 1*, pages 455–467, Geneva, Switzerland. European Association for Machine Translation.
- Yang Trista Cao, Anna Sotnikova, Jieyu Zhao, Linda X. Zou, Rachel Rudinger, and Hal Daumé III. 2025. [Multilingual Large Language Models Leak Human Stereotypes across Language Boundaries](#). In *Proceedings of the Fourth Workshop on NLP for Positive Impact (NLP4PI)*, pages 175–188, Vienna, Austria. Association for Computational Linguistics.
- Fabio Celli and Valerio Basile. 2024. [History repeats: Historical phase recognition from short texts](#). In *Proceedings of the 10th Italian Conference on Computational Linguistics (CLiC-it 2024)*, pages 198–204, Pisa, Italy. CEUR Workshop Proceedings.
- Fabio Celli and Astik Samal. 2025. Monitoring historical cultural hacking in large language models. In *TRUST-AI: The European Workshop on Trustworthy AI. Organized as part of the European Conference of Artificial Intelligence - ECAI 2025.*, pages 1–10.
- Ho-Chun Herbert Chang, Tracy Weener, Yung-Chun Chen, Sean Noh, Mingyue Zha, and Hsuan Lo. 2025. [Do language models favor their home countries? Asymmetric propagation of positive misinformation and foreign influence audits](#). *Harvard Kennedy School (HKS) Misinformation Review*.
- David Chanin. 2023. [Open-source frame semantic parsing](#).
- Matthew Dahl, Varun Magesh, Mirac Suzgun, and Daniel E Ho. 2024. [Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models](#). *Journal of Legal Analysis*, 16(1):64–93.
- Gavin Daly. 2018. British soldiers and the legend of Napoleon. *The Historical Journal*, 61(1):131–153.
- Matthijs Douze, Alexandr Guzhva, Chengqi Deng, Jeff Johnson, Gergely Szilvasy, Pierre-Emmanuel Mazaré, Maria Lomeli, Lucas Hosseini, and Hervé Jégou. 2024. [The Faiss library](#).
- Philip Dwyer. 2015. [Napoleon, The Revolution, and The Empire](#). In *The Oxford Handbook of the French Revolution*. Oxford University Press.
- Philip G Dwyer. 2004. Napoleon Bonaparte as hero and saviour: image, rhetoric and behaviour in the construction of a legend. *French History*, 18(4):379–403.
- Charles Fillmore. 1985. Frames and the semantics of understanding. *Quaderni di semantica*, 6:222–254.
- Charles J Fillmore et al. 2006. Frame semantics. *Cognitive linguistics: Basic readings*, 34:373–400.
- Isabel O. Gallegos, Ryan A. Rossi, Joe Barrow, Md Mehrab Tanjim, Sungchul Kim, Franck Deroncourt, Tong Yu, Ruiyi Zhang, and Nesreen K. Ahmed. 2024. [Bias and fairness in large language models: A survey](#). *Computational Linguistics*, 50(3):1097–1179.
- Yunfan Gao, Yun Xiong, Xinyu Gao, Kangxiang Jia, Jinliu Pan, Yuxi Bi, Yi Dai, Jiawei Sun, Meng Wang, and Haofen Wang. 2023. Retrieval-augmented generation for large language models: A survey. *arXiv e-prints*, pages arXiv–2312.

- Giselle Gonzalez Garcia and Christian Weillbach. 2023. [If the sources could talk: Evaluating large language models for research assistance in history](#). In *Computational Humanities Research Conference*.
- Tsafrir Goldberg and Yiftach Ron. 2014. [‘Look, Each Side Says Something Different’: the impact of competing history teaching approaches on Jewish and Arab adolescents’ discussions of the Jewish–Arab conflict](#). *Journal of Peace Education*, 11(1):1–29.
- Tsafrir Goldberg and Geerte M Savenije. 2018. Teaching controversial historical issues. *The Wiley International Handbook of History Teaching and Learning*, pages 503–526.
- Carlos-Emiliano González-Gallardo, Emanuela Boros, Nancy Girdhar, Ahmed Hamdi, Jose G. Moreno, and Antoine Doucet. 2024. [Yes but.. Can ChatGPT Identify Entities in Historical Documents?](#) In *Proceedings of the 2023 ACM/IEEE Joint Conference on Digital Libraries, JCDL ’23*, page 184–189. IEEE Press.
- Maarten Grootendorst. 2022. [BERTopic: Neural topic modeling with a class-based TF-IDF procedure](#).
- Marcus Hartner. 2014. [Multiperspectivity](#). In Peter Huňn et al., editor, *The living handbook of narratology*. Hamburg University.
- Jakob Hauser, Daniel Kondor, Jenny Reddish, Majid Benam, Enrico Cioni, Federica Villa, James Bennett, Daniel Hoyer, Pieter Francois, Peter Turchin, et al. 2024. Large language models’ expert-level global history knowledge benchmark (HiST-LLM). *Advances in Neural Information Processing Systems*, 37:32336–32369.
- Piotr Jaskulski, Tomasz Latos, Mariusz Ryrca, and Adam Zapała. 2025. Reliability of large language models as a tool for knowledge extraction from biographical dictionaries: the case of the Polish Biographical Dictionary. *Digital Scholarship in the Humanities*, 40(2):538–548.
- Lotte Jensen. 2011. [The Dutch against Napoleon. Resistance Literature and National Identity, 1806-1813](#). *Journal of Dutch Literature*, 2(2).
- Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. 2023. [Survey of Hallucination in Natural Language Generation](#). *ACM Comput. Surv.*, 55(12).
- Jeff Johnson, Matthijs Douze, and Hervé Jégou. 2019. Billion-scale similarity search with GPUs. *IEEE Transactions on Big Data*, 7(3):535–547.
- Hadas Kotek, Rikker Dockum, and David Sun. 2023. [Gender bias and stereotypes in large language models](#). In *Proceedings of The ACM Collective Intelligence Conference, CI ’23*, page 12–24, New York, NY, USA. Association for Computing Machinery.
- Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2020. [Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks](#). In *Advances in Neural Information Processing Systems*, volume 33, pages 9459–9474. Curran Associates, Inc.
- Nianqi Li, Siyu Yuan, Jiangjie Chen, Jiaqing Liang, Feng Wei, Zujie Liang, Deqing Yang, and Yanghua Xiao. 2025. [Past meets present: Creating historical analogy with large language models](#). In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3942–3957, Vienna, Austria. Association for Computational Linguistics.
- Alan McCully. 2012. [History teaching, conflict and the legacy of the past](#). *Education, Citizenship and Social Justice*, 7(2):145–159.
- Gosse Minnema, Gaetana Ruggiero, Marion Bartl, Sara Gemelli, Tommaso Caselli, Chiara Zanchi, Viviana Patti, Marco te Brömmelstroet, and Malvina Nissim. 2022. [Responsibility framing under the magnifying lens of NLP: The case of gender-based violence and traffic danger](#). *Computational Linguistics in the Netherlands Journal*, 12:207–233.
- Margaret Mitchell, Giuseppe Attanasio, Ioana Baldini, Miruna Clinciu, Jordan Clive, Pieter Delobelle, Manan Dey, Sil Hamilton, Timm Dill, Jad Doughman, Ritam Dutt, Avijit Ghosh, Jessica Zosa Forde, Carolin Holtermann, Lucie-Aimée Kaffee, Tanmay Laud, Anne Lauscher, Roberto L Lopez-Davila, Maraim Masoud, Nikita Nangia, Anaelia Ovalle, Giada Pistilli, Dragomir Radev, Beatrice Savoldi, Vipul Raheja, Jeremy Qin, Esther Ploeger, Arjun Subramonian, Kaushtubh Dhole, Kaiser Sun, Amirbek Djanibekov, Jonibek Mansurov, Kayo Yin, Emilio Villa Cueva, Sagnik Mukherjee, Jerry Huang, Xudong Shen, Jay Gala, Hamdan Al-Ali, Tair Djanibekov, Nurdaulet Mukhituly, Shangrui Nie, Shanya Sharma, Karolina Stanczak, Eliza Szczechla, Tiago Timponi Torrent, Deepak Tunuguntla, Marcelo Viridiano, Oskar Van Der Wal, Adina Yakefu, Aurélie Névóol, Mike Zhang, Sydney Zink, and Zeerak Talat. 2025. [SHADES: Towards a Multilingual](#)

- [Assessment of Stereotypes in Large Language Models](#). In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 11995–12041, Albuquerque, New Mexico. Association for Computational Linguistics.
- Nils Reimers and Iryna Gurevych. 2019. [Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks](#). In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.
- Keshav Santhanam, Omar Khattab, Jon Saad-Falcon, Christopher Potts, and Matei Zaharia. 2022. [ColBERTv2: Effective and efficient retrieval via lightweight late interaction](#). In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 3715–3734, Seattle, United States. Association for Computational Linguistics.
- Kurt Shuster, Spencer Poff, Moya Chen, Douwe Kiela, and Jason Weston. 2021. [Retrieval augmentation reduces hallucination in conversation](#). In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 3784–3803, Punta Cana, Dominican Republic. Association for Computational Linguistics.
- Robert A Stradling. 2003. *Multiperspectivity in history teaching: A guide for teachers*. Council of Europe Strasbourg.
- Marco Antonio Stranisci, Rossana Damiano, Enrico Mensa, Viviana Patti, Daniele Radicioni, and Tommaso Caselli. 2023. [WikiBio: a semantic resource for the intersectional analysis of biographical events](#). In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 12370–12384, Toronto, Canada. Association for Computational Linguistics.
- Yi Sun, Wanru Yang, and Yin Liu. 2024. [The Application of Constructing Knowledge Graph of Oral Historical Archives Resources Based on LLM-RAG](#). In *Proceedings of the 2024 8th International Conference on Information System and Data Mining, ICISDM '24*, page 142–149, New York, NY, USA. Association for Computing Machinery.
- The Trung Tran, Carlos-Emiliano González-Gallardo, and Antoine Doucet. 2025. [Retrieval augmented generation for historical newspapers](#). In *Proceedings of the 24th ACM/IEEE Joint Conference on Digital Libraries*, New York, NY, USA. Association for Computing Machinery.
- Eva Vanmassenhove, Dimitar Shterionov, and Andy Way. 2019. [Lost in translation: Loss and decay of linguistic richness in machine translation](#). In *Proceedings of Machine Translation Summit XVII: Research Track*, pages 222–232, Dublin, Ireland. European Association for Machine Translation.
- Yifan Zeng. 2024. HistoLens: An LLM-Powered Framework for Multi-Layered Analysis of Historical Texts—A Case Application of Yantie Lun. *arXiv preprint arXiv:2411.09978*.
- Jing Zhou, Li Si, and Wenjun Hou. 2025. Humanities-in-the-loop: Using close reading as a method for retrieval-augmented generation (rag). *Proceedings of the Association for Information Science and Technology*, 62(1):1185–1189.

9. Language Resource References

- BL. 2021. *Digitised Books. c. 1510 - c. 1900. JSONL (OCR derived text + metadata)*. British Library Labs.
- BnF. 2024. *Documents de presse en mode texte du projet Europeana Newspapers (XIXe-XXe siècles)*. National Library of France.
- Melissa Dell and Jacob Carlson and Tom Bryan and Emily Silcock and Abhishek Arora and Zejiang Shen and Luca D’Amico-Wong and Quan Le and Pablo Querubin and Leander Heldring. 2023. *American Stories: A Large-Scale Structured Text Dataset of Historical U.S. Newspapers*. arXiv.
- KB. 2019. *Delpher open krantenarchief (1.0)*. Koninklijke Bibliotheek.
- KB. 2025. *Digitale Bibliotheek voor de Nederlandse Letteren*. Koninklijke Bibliotheek.
- Romein, C.A. 2023. *Nederlandsch-Indisch plakaatboek, 1602-1811 (deel 1 t/m 17)*. Zenodo.

10. Appendix

A. Questions

There were ten questions for English, French and Dutch.

English

1. Who was Napoleon Bonaparte?
2. Was Napoleon Bonaparte a hero?
3. Was Napoleon Bonaparte a tyrant?
4. Was Napoleon Bonaparte a villain?
5. How did Napoleon Bonaparte establish his empire in Europe?
6. How did Napoleon Bonaparte become the Emperor of France?
7. Was Napoleon Bonaparte's claim to Emperor favoured in Europe?
8. How did Napoleon Bonaparte's rule impact England and the Netherlands?
9. What was the impact of Napoleon Bonaparte on England and the Netherlands?
10. Was the impact of Napoleon Bonaparte on England and the Netherlands destructive?

French

1. Qui était Napoléon Bonaparte ?
2. Napoléon Bonaparte était-il un héros ?
3. Napoléon Bonaparte était-il un tyran ?
4. Napoléon Bonaparte était-il un méchant ?
5. Comment Napoléon Bonaparte a-t-il établi son empire en Europe ?
6. Comment Napoléon Bonaparte est-il devenu empereur de France ?
7. La prétention de Napoléon Bonaparte au titre d'empereur avait-elle les faveurs de l'Europe ?
8. Quel impact le règne de Napoléon Bonaparte a-t-il eu sur l'Angleterre et les Pays-Bas ?
9. Quel a été l'impact de Napoléon Bonaparte sur l'Angleterre et les Pays-Bas ?
10. L'impact de Napoléon Bonaparte sur l'Angleterre et les Pays-Bas a-t-il été destructeur ?

Dutch

1. Wie was Napoleon Bonaparte?
2. Was Napoleon Bonaparte een held?
3. Was Napoleon Bonaparte een tiran?
4. Was Napoleon Bonaparte een schurk?
5. Hoe stichtte Napoleon Bonaparte zijn rijk in Europa?
6. Hoe werd Napoleon Bonaparte keizer van Frankrijk?
7. Werd Napoleon Bonaparte's aanspraak op keizerschap in Europa gesteund?
8. Welke invloed had Napoleon Bonaparte's heerschappij op Engeland en Nederland?
9. Wat was de impact van Napoleon Bonaparte op Engeland en Nederland?
10. Was de impact van Napoleon Bonaparte op Engeland en Nederland destructief?

B. Prompt structure

Each question had a system prompt in its own language as follows.

English Provide a comprehensive answer to the question, using your own knowledge and the information in the given context. Respond only to the given question; your answer should be concise and relevant to the question. Respond in English.

French Répondez de manière exhaustive à la question, en vous appuyant sur vos connaissances et les informations du contexte. Répondez uniquement à la question posée; votre réponse doit être concise et pertinente. Répondez en français.

Dutch Geef een uitgebreid antwoord op de vraag, gebruikmakend van je eigen kennis en de informatie in de gegeven context. Reageer alleen op de gegeven vraag; je antwoord moet beknopt en relevant zijn voor de vraag. Reageer in het Nederlands.

For question answering, the user prompt structure was as follows. For baseline questions, context was empty.

English

Context:
{context}

—
This is the question you must answer.

Question:
{question}

French

Contexte:
{context}

—
C'est la question à laquelle vous devez répondre.

Question:
{question}

Dutch

Context:
{context}

—
Dit is de vraag die je moet beantwoorden.

Vraag:
{question}

C. Example sources

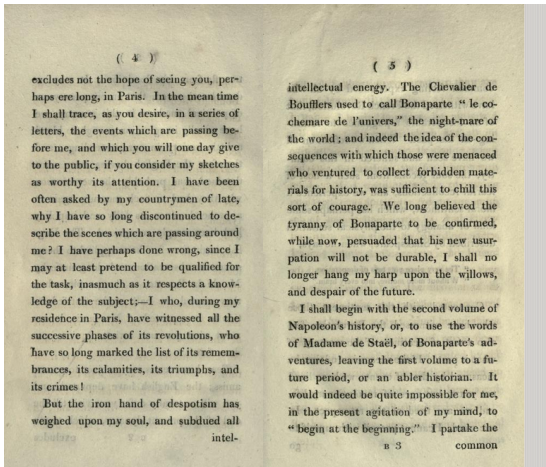


Figure 7: Example from a non-fiction book from the British Library.



Figure 8: Example from a newspaper from Delpher.

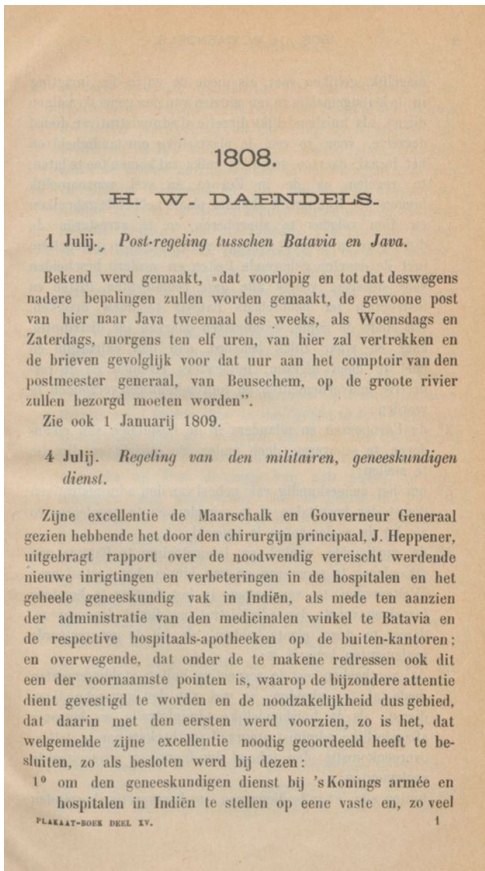


Figure 9: Example from a placard book from the Dutch-Indonesian placard books.

BOEKBESCHOUWING.

Nagelaten Leerrekenen van J. VAN DER ROEST, in Leven Evangeliedienaar te Haarlem. Iste Stuk. Te Haarlem, bij J. L. Augustini. 1814. In gr. 8vo. 178 Bl. f 1-8-1

Het heeft ons niet mogen gebeuren, den waardigen VAN DER ROEST in perfoon te kennen; maar zoo de perfoonlijke kennis ooit overtollig was, om ons een juist denkbeeld van iemand te geven, dan voorzeker was het bij hem, die in zijne schriften leeft. Trouwens, het kan niet anders, of de preken van dezen man moeten het ware afbeeldfel, de getrouwe spiegel zijner ziele wezen. Immers de kunst moge het zoo ver brengen, dat zij de kunst verbergt, en enkel natuur schijnt; bezwaarlijk zal zij dien toer der zuivere en ongevinsde godsdienstigheid nabootfen, welken wij hier vinden; bezwaarlijk zal zij die zalving aannemen, die hier zoo onmiddellijk uit het hart tot het hart spreekt, en met zulk een onwederftandelijk, maar weldadig vermogen op ons werkt; ja onmogelijk zal zij die taal van den kinderlijken en nederigen geest kunnen voeren, die zich verliest in het gevoel zijner nierigheid en afhankelijkheid van God, en zich dankbaar verheft bij de gedachte aan den Eeuwigegezegenden, die in de wereld kwam, om zondaren zalig te maken. Het is die welsprekendheid niet, welke door kunst haar doel bereikt, door vergelijkingen en tegenstellingen fehitter, door vendingen verrast, en op deze wijze een behoorder overreedt en wegleept; maar het is die welsprekendheid, welke uit de diepte van het gemoed, en wel van het godsdienstig gemoed, oprijst, dat de dingen zóó aanfchouwt als zij zijn, ook alzoovoorfelt, en langs dezen eenvoudigen weg haar oogmerk zeker bereikt. Niets is

POEBESCH. 1816. NO. 1. A hier

Figure 10: Example from a magazine from DBNL.

D. Example model outputs

The complete set of answers for all settings, questions and languages can be found in the GitHub at `RAG\evaluation\all_answers.parquet`. Given below are some examples from this set for the Qwen3-235B-A22B model.

Baseline	RAG
<p>Whether Napoleon Bonaparte was a villain depends on perspective. He implemented significant reforms, such as the Napoleonic Code, which influenced legal systems worldwide, and promoted education and meritocracy. However, he also waged aggressive wars across Europe, caused massive loss of life, suppressed dissent, and re-established slavery in French colonies. While some view him as a visionary leader who spread Enlightenment ideals, others see him as an authoritarian conqueror. Thus, labeling him a "villain" oversimplifies his complex legacy.</p>	<p>Napoleon Bonaparte cannot be simplistically labeled a villain, as historical perspectives present both critical and sympathetic views of his actions. On one hand, he is criticized for authoritarian rule, violating treaties, and causing widespread warfare across Europe, actions described in Document 2, where he portrays himself as wronged but acknowledges the force used against others. On the other hand, he is depicted as a progressive leader who supported science, improved education, and modernized legal systems, as noted in Document 3, which highlights his respect for scholars and reforms. Additionally, concerns about his mistreatment during exile reflect moral debates about justice and humanity Document 1. Thus, Napoleon's legacy is complex neither wholly villainous nor entirely virtuous but shaped by perspective and context.</p>
<p>Napoleon Bonaparte was geen simpel 'schurk', maar een complexe historische figuur met zowel grote prestaties als controversiële daden. In de documenten wordt hij afgebeeld als een machtige leider met een indrukwekkend karakter en strategisch inzicht zoals in Document 2, waar hij Oostenrijk's militaire voorbereidingen voorspelde, maar ook als een politiek manipulatief persoon die zijn macht gebruikt om zijn doelen te bereiken, soms ten koste van anderen. Bovendien wordt in Document 3 suggereren dat hij opportunistisch handelde, bijvoorbeeld door het gedrag van Nederland te benadrukken als een gebrek aan trouw. Toch blijft zijn rol in de Europese geschiedenis cruciaal: hij bracht wetgeving, modernisatie en nationale identiteit, maar ook oorlog en onderdrukking. Of hij een schurk is, hangt af van de interpretatie: als men alleen de destructieve aspecten bekijkt, ja; als men zijn invloed op recht, bestuur en vrijheid meet, nee. Een juist oordeel vereist dus een evenwichtig overzicht van zijn hele erfgoed.</p> <p>Translation: Napoleon Bonaparte was not a simple "villain," but a complex historical figure with both great achievements and controversial deeds. In the documents, he is portrayed as a powerful leader with an impressive character and strategic insight, as in Document 2, where he predicted Austria's military preparations, but also as a politically manipulative person who used his power to achieve his goals, sometimes at the expense of others. Furthermore, Document 3 suggests that he acted opportunistically, for example by emphasizing the Netherlands' behavior as a lack of loyalty. Nevertheless, his role in European history remains crucial: he brought legislation, modernization, and national identity, but also war and oppression. Whether he is a villain depends on interpretation: if one looks only at the destructive aspects, yes; if one measures his influence on law, governance, and freedom, no. A fair judgment therefore requires a balanced overview of his entire legacy.</p>	<p>Napoleon Bonaparte was geen eenvoudig 'schurk', maar een complexe historische figuur. Hoewel hij oorlogen voerde die veel doden kostten en absolute macht nastreefde, had hij ook belangrijke bijdragen aan de moderne staat, zoals de invoering van het Code Napoléon, dat rechtvaardigheid en gelijkheid voor alle burgers bevorderde. Zijn daden moeten gezien worden in hun historische context: hij veranderde Europa, zowel positief als negatief. Daarom is het te simplistisch om hem simpelweg als schurk te bestempelen.</p> <p>Translation: Napoleon Bonaparte was not a simple "villain," but a complex historical figure. Although he waged wars that caused many deaths and sought absolute power, he also made important contributions to the modern state, such as the introduction of the Code Napoléon, which promoted justice and equality for all citizens. His actions must be seen in their historical context: he changed Europe, both positively and negatively. Therefore, it is too simplistic to simply label him a villain.</p>

Napoléon Bonaparte n'était ni purement méchant ni entièrement bon: il s'agit d'une figure historique complexe. Il a mené des guerres qui ont causé des millions de morts, imposé son autorité par la force et renforcé l'absolutisme dans certaines régions. Toutefois, il a aussi réformé profondément le droit Code civil, modernisé l'administration, promu la laïcité et étendu des idées révolutionnaires comme l'égalité devant la loi. Son jugement dépend donc fortement du regard porté sur ses actions : tantôt vu comme un tyran ambitieux, tantôt comme un réformateur visionnaire.

Translation: Napoleon Bonaparte was neither purely evil nor entirely good: he was a complex historical figure. He waged wars that caused millions of deaths, imposed his authority by force, and reinforced absolutism in certain regions. However, he also profoundly reformed the Civil Code, modernized the administration, promoted secularism, and spread revolutionary ideas such as equality before the law. His legacy therefore depends heavily on how his actions are viewed: sometimes as an ambitious tyrant, sometimes as a visionary reformer.

La question de savoir si Napoléon Bonaparte était un méchant dépend du point de vue historique et politique adopté. D'après les documents fournis, notamment le Document 2, il est présenté comme une menace permanente pour la paix en Europe, qualifié d'obstacle perpétuel à la paix, dont les conquêtes militaires sont décrites comme criminelles. Ce discours, issu d'un rapport du quartier général autrichien en 1813, reflète la vision hostile des coalitions européennes qui voyaient en lui un agresseur, responsable de guerres destructrices et de troubles politiques. Cependant, ce jugement est partial et idéologique: il provient de puissances ennemies cherchant à légitimer leur résistance. Historiquement, Napoléon a également instauré des réformes importantes code civil, modernisation administrative, réforme scolaire qui ont profondément marqué l'Europe. Son ambition n'était pas seulement territoriale, mais aussi institutionnelle. En somme, dire qu'il était méchant revient à adopter une perspective victimaire et anti-napoléonienne. Une analyse nuancée montre qu'il fut un homme complexe : ambitieux, autoritaire, impliqué dans des guerres destructrices.

Translation: Whether Napoleon Bonaparte was a villain depends on the historical and political perspective adopted. According to the documents provided, particularly Document 2, he is presented as a permanent threat to peace in Europe, described as a perpetual obstacle to peace, whose military conquests are portrayed as criminal. This narrative, taken from a report by the Austrian headquarters in 1813, reflects the hostile view of the European coalitions, which saw him as an aggressor responsible for destructive wars and political unrest. However, this judgment is biased and ideological: it comes from enemy powers seeking to legitimize their resistance. Historically, Napoleon also introduced important reforms a civil code, administrative modernization, and educational reform that had a profound impact on Europe. His ambition was not only territorial, but also institutional. In short, to say that he was evil is to adopt a victimized and anti-Napoleonic perspective. A nuanced analysis shows that he was a complex man: ambitious, authoritarian, involved in destructive wars.

Table 9: Answers for English, Dutch and French, Question 4 *“Was Napoleon Bonaparte a villain?”*

Baseline	RAG
<p>Napoleon Bonaparte's claim to Emperor was not widely favored in Europe. While he consolidated power in France and established the French Empire in 1804, most European monarchies viewed his rise with hostility. They saw him as a usurper who undermined traditional royal authority and spread revolutionary ideals through military conquest. His coronation as Emperor provoked opposition from major powers like Austria, Russia, and Britain, leading to prolonged conflicts known as the Napoleonic Wars. Thus, while some regions under French control may have accepted or even supported his rule temporarily, overall, his imperial claim was met with resistance across Europe.</p>	<p>Napoleon Bonaparte's claim to Emperor was not widely favored in Europe. While he exerted significant control over much of the continent through military dominance and political influence, many European powers opposed his rise and rule. The documents suggest skepticism and criticism toward his authority referring to his power as stemming from force and conquest rather than legitimate recognition. Additionally, there is mention of resistance to his dominance and concern over his unchecked influence, indicating that although he acted as an absolute master over various regions, his imperial title and authority were not broadly accepted or supported across Europe.</p>
<p>Napoleon Bonapartes aanspraak op keizerschap werd niet algemeen gesteund in Europa. In 1804 benoemde hij zichzelf tot keizer der Fransen, wat formeel plaatsvond via een plebisciete en met instemming van het Franse volk, maar veel Europese monarchieën erkenden deze titel niet legitiem. Traditionele vorstenhuizen, zoals die van Oostenrijk, Rusland en Groot-Brittannië, zagen hem als een usurpator iemand die macht had gegrepen door revolutie en militaire kracht, niet door erfopvolging of goddelijke rechtvaardiging. Pas na zijn overwinningen en het sluiten van verdragen, zoals het Verdrag van Pressburg 1805 en het huwelijk met Marie-Louise van Oostenrijk 1810, gaven sommige machten, zoals het Habsburgse Rijk, indirect erkenning aan zijn positie. Toch bleef zijn keizerschap omstreden en was het vooral gebaseerd op militair geweld en politiek realisme, niet op brede morele of constitutionele steun binnen Europa.</p>	<p>Ja, Napoleons aanspraak op keizerschap kreeg steun in delen van Europa, hoewel die steun vaak gebaseerd was op politieke berekening, machtsevenwicht of dwang. Uit de gegeven documenten blijkt dat er sprake was van diplomatieke interactie tussen Napoleon en andere Europese machten zoals Oostenrijk en Holland. Zo wordt in Document 1 gesproken over de relatie tussen Napoleon en de Oostenrijkse monarchie, waarbij impliciet wordt aangegeven dat er rekening werd gehouden met zijn positie en beslissingen. In Document 3 en Document 2 wordt zelfs sprake van een blij zijner welwillendheid jegens Holland, wat suggereert dat zijn autoriteit erkend werd. Daarnaast zorgde Napoleon ervoor dat zijn heerschappij legitiem leek via symboliek, instellingen en allianties ook buiten Frankrijk. Sommige staten erkenden hem uit pragmatisme of onder druk, terwijl anderen formeel meewerkten om conflicten te vermijden. Hoewel veel Europese naties uiteindelijk tegen hem vochten, was er aanvankelijk en gedurende een periode zeker sprake van indirecte of geforceerde steun voor zijn keizerlijke.</p>
<p>Translation: Napoleon Bonaparte's claim to the imperial throne was not universally supported in Europe. In 1804, he proclaimed himself Emperor of the French, which formally took place through a plebiscite and with the consent of the French people, but many European monarchies did not recognize this title as legitimate. Traditional royal houses, such as those of Austria, Russia, and Great Britain, saw him as a usurper, someone who had seized power through revolution and military force, not through hereditary succession or divine justification. Only after his victories and the conclusion of treaties, such as the Treaty of Pressburg in 1805 and his marriage to Marie-Louise of Austria in 1810, did some powers, such as the Habsburg Empire, indirectly recognize his position. Nevertheless, his empire remained controversial and was based primarily on military force and political realism, not on broad moral or constitutional support within Europe.</p>	<p>Translation: Yes, Napoleon's claim to the imperial throne received support in parts of Europe, although that support was often based on political calculation, balance of power, or coercion. The documents provided show that there was diplomatic interaction between Napoleon and other European powers such as Austria and Holland. Document 1, for example, discusses the relationship between Napoleon and the Austrian monarchy, implicitly indicating that his position and decisions were taken into account. Documents 3 and 2 even mention a show of goodwill towards Holland, suggesting that his authority was recognized. In addition, Napoleon ensured that his rule appeared legitimate through symbolism, institutions, and alliances outside France. Some states recognized him out of pragmatism or under pressure, while others cooperated formally to avoid conflict. Although many European nations ultimately fought against him, there was certainly indirect or forced support for his empire initially and for a period of time.</p>
<p>Non, la prétention de Napoléon Bonaparte au titre d'empereur n'avait pas les faveurs de l'Europe. En 1804, lorsqu'il se couronna lui-même empereur des Français, cette initiative fut perçue par les autres puissances européennes comme une provocation et un retour à l'autocratie monarchique, remettant en cause les principes révolutionnaires de liberté et d'égalité. Les monarchies traditionnelles, notamment l'Autriche, la Russie et le Royaume-Uni, voyaient en Napoléon un usurpateur issu d'un coup d'État, dont l'ambition menaçait l'équilibre européen. Son couronnement marqua donc le début d'une série de conflits avec les coalitions européennes hostiles à son pouvoir et à l'hégémonie française.</p>	<p>Non, la prétention de Napoléon Bonaparte au titre d'empereur n'avait pas les faveurs de l'Europe. Les documents indiquent clairement que l'Europe voyait en lui une menace permanente à la paix. Selon le Document 2, l'Europe veut la paix et ne transigera jamais avec celui qu'elle regarde comme un obstacle perpétuel à la paix. Son règne était perçu comme une menace pour le repos des nations voisines, et les puissances européennes étaient déterminées à agir contre lui, allant jusqu'à envahir la France. Ainsi, loin d'être reconnu légitimement par l'Europe, Napoléon était considéré comme un usurpateur dangereux dont la chute était nécessaire pour assurer la stabilité continentale.</p>
<p>Translation: No, Napoleon Bonaparte's claim to the title of emperor was not welcomed by Europe. In 1804, when he crowned himself Emperor of the French, this initiative was perceived by the other European powers as a provocation and a return to monarchical autocracy, calling into question the revolutionary principles of liberty and equality. Traditional monarchies, notably Austria, Russia, and the United Kingdom, saw Napoleon as a usurper who had come to power through a coup d'état, whose ambition threatened the European balance of power. His coronation therefore marked the beginning of a series of conflicts with European coalitions hostile to his power and French hegemony.</p>	<p>Translation: No, Napoleon Bonaparte's claim to the title of emperor did not find favor in Europe. The documents clearly indicate that Europe saw him as a permanent threat to peace. According to Document 2, Europe wants peace and will never compromise with someone it sees as a perpetual obstacle to peace. His reign was perceived as a threat to the tranquility of neighboring nations, and the European powers were determined to take action against him, going so far as to invade France. Thus, far from being legitimately recognized by Europe, Napoleon was considered a dangerous usurper whose downfall was necessary to ensure continental stability.</p>

Table 10: Answers for English, Dutch and French, Question 7 "Was Napoleon Bonaparte's claim to Emperor favoured in Europe?"

Baseline

Napoleon Bonaparte's impact on England and the Netherlands had both destructive and transformative elements, though the nature of his influence differed between the two countries. On England, Napoleon's impact was primarily indirect but significant. While he failed to invade Britain, his plans thwarted by naval defeats like Trafalgar in 1805, the prolonged Napoleonic Wars 1803-1815 placed immense economic and social strain on England. The British economy bore the cost of massive military expenditures, taxation, and trade disruptions due to Napoleon's Continental System, which aimed to isolate Britain economically. However, this also spurred industrial growth and innovation as Britain sought self-sufficiency and expanded its global trade networks. Militarily, Britain emerged stronger after defeating Napoleon at Waterloo in 1815, solidifying its status as a dominant world power. Thus, while there were hardships, the overall impact on England was not destructive; it ultimately reinforced national strength and imperial expansion. In contrast, Napoleon's impact on the Netherlands was more directly disruptive. Initially an ally, the Dutch Republic was reorganized into the Batavian Republic under French influence, and later annexed directly into the French Empire from 1810 to 1813. This led to loss of sovereignty, heavy conscription into Napoleon's armies, economic hardship due to enforcement of the Continental System, and widespread resentment. Dutch industries suffered from blocked trade with Britain, leading to economic decline.

RAG

The provided documents do not offer direct evidence about the specific impact of Napoleon Bonaparte on England and the Netherlands, nor do they clearly assert that his influence was destructive toward these nations. However, they reflect a broader European concern regarding Napoleon's dominance and military campaigns. Regarding England, the documents suggest opposition to Napoleon, particularly through Britain's involvement in supporting coalitions against him e.g., potential British troop coordination mentioned in Document 1. Additionally, Document 2 discusses concerns about the treatment of Napoleon during his exile, indicating Britain's role in containing him, rather than suffering destruction from him. Thus, while Napoleon posed a strategic threat, there is no indication in these texts that his impact on England was destructively realized; rather, England appears as a resisting power. As for the Netherlands, there is no explicit mention in the documents. However, historical context outside the text notes that the Netherlands was at times allied with or occupied by French forces under Napoleon, including being reorganized as the Kingdom of Holland ruled by his brother Louis. While this could imply political manipulation or strain, the provided excerpts do not address this directly. In conclusion, based solely on the given documents, there is insufficient evidence to confirm a destructive impact of Napoleon on England and the Netherlands; instead, the focus lies on resistance to his power, especially by England, and uncertainty surrounding allied figures like Bernadotte. Any assessment of destruction would require external historical knowledge beyond these fragments.

De impact van Napoleon Bonaparte op Engeland was vooral geopolitiek en economisch van aard, maar niet direct destructief op het grondgebied zelf. Engeland ontkwam aan bezetting, maar moest enorme middelen steken in oorlogsvoering, wat leidde tot belastingverhogingen en sociale spanningen. Tegelijk stimuleerde de concurrentie met Frankrijk innovatie en imperial expansie. Op Nederland daarentegen was de invloed wel degelijk destructief. Door annexatie 1810 raakte Nederland zijn soevereiniteit kwijt, werd zwaar belast, gedwongen tot de Continentale Blokkade met grote schade voor de handel en moest massaal mannen afstaan aan Napoleons leger. De economische neergang en maatschappelijke ontwrichting hadden langdurige gevolgen. Kortom: voor Engeland indirect kostbaar, maar niet destructief; voor Nederland juist wel destructief vanwege directe controle en uitbuiting.

Translation: Napoleon Bonaparte's impact on England was mainly geopolitical and economic in nature, but not directly destructive to the territory itself. England escaped occupation, but had to invest enormous resources in warfare, which led to tax increases and social tensions. At the same time, competition with France stimulated innovation and imperial expansion. In contrast, the influence on the Netherlands was indeed destructive. Through annexation in 1810, the Netherlands lost its sovereignty, was heavily taxed, forced to participate in the Continental Blockade with great damage to trade, and had to surrender large numbers of men to Napoleon's army. The economic decline and social disruption had long-lasting consequences. In short: for England, it was indirectly costly but not destructive; for the Netherlands, it was destructive because of direct control and exploitation.

De impact van Napoleon Bonaparte op Engeland en Nederland was grotendeels destructief, hoewel de aard van die schade per land verschilt. Op Engeland had Napoleon een indirecte maar ingrijpende negatieve impact via zijn Continental Stelsel het beleid om handel tussen het Europese vasteland en Engeland te verbieden. Met dit stelsel hoopte hij Engeland economisch te isoleren en te dwingen tot vrede. In plaats daarvan leidde het tot economische spanningen, smokkel, en verarming in veel delen van Europa, terwijl Engeland juist sterker werd dankzij zijn overheersing op zee en alternatieve markten. Daarmee mislukte Napoleons strategie en versterkte hij zelfs de positie van Engeland als wereldmacht. Op Nederland was de impact directer en eveneens destructief. Door het Continental Stelsel raakte de Nederlandse economie, die sterk op internationale handel beruiste, zwaar beschadigd. Bovendien annexeerde Napoleon het Koninkrijk Holland in 1810, wat het einde betekende van enige autonomie.

Translation: Napoleon Bonaparte's impact on England and the Netherlands was largely destructive, although the nature of that damage differed from country to country. Napoleon had an indirect but profoundly negative impact on England through his Continental System, a policy prohibiting trade between mainland Europe and England. With this system, he hoped to isolate England economically and force it to make peace. Instead, it led to economic tensions, smuggling, and impoverishment in many parts of Europe, while England actually grew stronger thanks to its dominance at sea and alternative markets. Napoleon's strategy thus failed and even strengthened England's position as a world power. The impact on the Netherlands was more direct and equally destructive. The Continental System severely damaged the Dutch economy, which was heavily dependent on international trade. In addition, Napoleon annexed the Kingdom of Holland in 1810, which meant the end of any autonomy.

L'impact de Napoléon Bonaparte sur l'Angleterre et les Pays-Bas n'a pas été uniquement destructeur, bien qu'il ait entraîné des conséquences négatives significatives. Sur l'Angleterre, Napoléon n'a jamais réussi à envahir le territoire britannique. Au contraire, son blocus continental décret de Berlin, 1806 visait à affaiblir économiquement l'Angleterre, mais il renforça en réalité sa détermination et son expansion commerciale outre-mer. La guerre contre Napoléon stimula aussi l'industrie militaire et navale britannique. À Waterloo en 1815, la victoire coalisée marqua l'apogée du rôle britannique en Europe. Ainsi, pour l'Angleterre, l'impact fut davantage géopolitique que destructeur : elle sortit vainqueur et plus puissante. Concernant les Pays-Bas, annexés par l'Empire napoléonien entre 1810 et 1813, l'occupation fut ressentie comme oppressive : réquisition des hommes et des ressources, suppression des institutions locales, application du blocus continental qui ruina le commerce maritime néerlandais.

Translation: Napoleon Bonaparte's impact on England and the Netherlands was not solely destructive, although it did have significant negative consequences. Napoleon never succeeded in invading British territory. On the contrary, his Continental Blockade decree of Berlin in 1806 was intended to weaken England economically, but in reality it strengthened its determination and overseas trade expansion. The war against Napoleon also stimulated British military and naval industry. At Waterloo in 1815, the coalition victory marked the height of Britain's role in Europe. Thus, for England, the impact was more geopolitical than destructive: it emerged victorious and more powerful. As for the Netherlands, annexed by the Napoleonic Empire between 1810 and 1813, the occupation was felt to be oppressive: requisitioning of men and resources, suppression of local institutions, enforcement of the continental blockade that ruined Dutch maritime trade.

L'impact de Napoléon Bonaparte sur l'Angleterre et les Pays-Bas a été perçu comme menaçant, mais pas exclusivement destructeur. Sur l'Angleterre, l'effet fut principalement militaire, économique et politique : Napoléon a imposé le blocus continental afin d'affaiblir économiquement le Royaume-Uni, ce qui a provoqué des tensions commerciales intenses. Toutefois, loin de détruire l'Angleterre, cela a stimulé son expansion maritime et coloniale. Comme mentionné dans le Document 1, les ministres britanniques agissaient sous couvert de sauver l'Europe, mais protégeaient surtout leurs propres intérêts ce qui montre que Napoléon était vu comme une menace, mais que sa chute a finalement accru la domination anglaise. Concernant les Pays-Bas, intégrés à l'Empire français entre 1810 et 1813, l'annexion napoléonienne a entraîné une centralisation administrative et fiscale pesante, ainsi qu'une mobilisation forcée dans l'armée française. Cela a causé mécontentement et difficultés économiques.

Translation: Napoleon Bonaparte's impact on England and the Netherlands was perceived as threatening, but not exclusively destructive. On England, the effect was mainly military, economic, and political: Napoleon imposed the Continental Blockade in order to weaken the United Kingdom economically, which caused intense trade tensions. However, far from destroying England, this stimulated its maritime and colonial expansion. As mentioned in Document 1, British ministers acted under the guise of saving Europe, but were primarily protecting their own interests, which shows that Napoleon was seen as a threat, but that his downfall ultimately increased English domination. As for the Netherlands, which was incorporated into the French Empire between 1810 and 1813, Napoleon's annexation led to heavy administrative and fiscal centralization, as well as forced mobilization into the French army. This caused discontent and economic hardship.

Table 11: Answers for English, Dutch and French, Question 10 *“Was the impact of Napoleon Bonaparte on England and the Netherlands destructive?”*