Data Speaks, But Who Gives It a Voice? Understanding Persuasive Strategies in Data-Driven News Articles

Zikai Li*, Chuyi Zheng*, Ziang Li, and Yang Shi

Abstract—Data-driven news articles combine narrative storytelling with data visualizations to inform and influence public opinion on pressing societal issues. These articles often employ persuasive strategies, which are rhetorical techniques in narrative framing, visual rhetoric, or data presentation, to influence audience interpretation and opinion formation regarding information communication. While previous research has examined whether and when data visualizations persuade, the strategic choices made by persuaders remain largely unexplored. Addressing this gap, our work presents a taxonomy of persuasive strategies grounded in psychological theories and expert insights, categorizing 15 strategies across five dimensions: Credibility, Guided Interpretation, Reference-based Framing, Emotional Appeal, and Participation Invitation. To facilitate large-scale analysis, we curated a dataset of 936 data-driven news articles annotated with both persuasive strategies and their perceived effects. Leveraging this corpus, we developed a multimodal, multi-task learning model that jointly predicts the presence of persuasive strategies and their persuasive effects by incorporating both embedded (text and visualization) and explicit (visual narrative and psycholinguistic) features. Our evaluation demonstrates that our model outperforms state-of-the-art baselines in identifying persuasive strategies and measuring their effects.

Index Terms—Data-driven storytelling, persuasive strategies, taxonomy, computational modeling

1 Introduction

Data-driven news articles are data journalistic artifacts that combine narrative with data visualization [28]. These articles play an increasingly important role in informing the public about pressing social issues, such as public health (e.g., Covid-19 trends [89]) and environmental challenges (e.g., climate change [66]), influencing how audiences interpret data and form opinions [25, 60]. To strengthen their influence, storytellers often employ rhetorical devices in the form of *persuasive strategies*: communicative techniques in visual rhetoric, narrative framing, or data presentation designed to guide audiences' interpretations, evoke their cognitive or emotional responses, and align them with an article's intended message. For example, a *New York Times* data-driven news article on the post-pandemic child care crisis, *With Pandemic Money Gone, Child Care Is an Industry on the Brink* [83], combines *compelling data visualizations* (e.g., a bar chart highlighting that 53% of providers face staffing shortages) with *emotional storytelling* (e.g., a child care provider earning "$2 an hour"). By grounding quantitative evidence in a clear visual format and tying those visuals to relatable human struggles, the article implicitly persuades audiences to recognize the urgency of sustained federal funding. Understanding such persuasive strategies is essential for deciphering how data-driven news articles shape public discourse as well as advancing responsible data storytelling that balances persuasion with transparency.

The visualization community has increasingly studied *whether* and *when* data visualizations persuade their audiences. Early research has investigated whether visualizations persuade more effectively than textual or illustrative alternatives [25, 54]. For example, Pandey et al. [54] demonstrated that charts outperform text in shifting attitudes, but only when aligned with audiences' prior attitudes. Subsequent work has explored when visualizations persuade by investigating factors that modulate their impact, such as prior attitudes [49] and narrative techniques [45, 60]. For example, Liem et al. [45] compared personal visual narratives and structured visual narratives in immigration debates but found neither strongly influenced attitudes. While the aforementioned work has laid a solid foundation for understanding the effect of visualization on the *persuadees*, the persuasive strategies employed by the *persuader* remain largely unexplored. This gap presents two challenges: First, no taxonomy currently exists to categorize the persuasive strategies used by persuaders in data-driven news articles. Given that these articles involve both text and visualizations, the taxonomy should consider that strategies can emerge from the interplay both within and across modalities. Second, extracting and analyzing strategies at scale requires computational modeling. Thus, a predictive model is necessary to identify persuasive strategies and evaluate their contributions to the article's overall persuasive effect.

Motivated by these challenges, we develop a taxonomy of persuasive strategies and a predictive model for identifying these strategies and their persuasive effects. First, we conducted focus group studies to develop a taxonomy comprising 15 persuasive strategies organized into five categories, including *credibility, guided interpretation, reference-based framing, emotional appeal*, and *participation invitation*. We then curated a dataset of 936 data-driven news articles collected from diverse online sources, annotated with two layers of labeling: (i) expert coding of persuasive strategies based on our taxonomy and (ii) crowdsourced labeling of perceived persuasive effects. Building on this annotated dataset, we developed a multimodal, multi-task learning model (M2P). Given an article, M2P extracts and synthesizes its embedded features (text and visualization features) and explicit (visual narrative and psycholinguistic) features to jointly predict both the presence of specific strategies as well as the persuasive effect of the article. We evaluated the performance of our model through a series of experiments. The results demonstrated that M2P is effective in identifying persuasive strategies and measuring persuasive effect. Specifically, M2P improves the Micro F1-score by 4.4% in persuasive strategy classification and reduces mean absolute error (MAE) by 9.4% in persuasive effect prediction compared to state-of-the-art baselines.

In summary, the main contributions of this work are as follows:

- We developed a taxonomy of persuasive strategies in data-driven news articles, informed by psychological theories of persuasion (e.g., the Elaboration Likelihood Model [55]) and insights from focus groups. This taxonomy provides a structured framework for categorizing the communicative techniques used to shape audiences' interpretations, enabling a deeper understanding of persuasive communication in data storytelling.
- We curated and annotated a dataset of 936 data-driven news articles, labeling both persuasive strategies and their persuasive effects. This dataset enables the analysis of persuasion in data journalism, supporting computational modeling of persuasive communication.

• *Zikai Li, Chuyi Zheng, Ziang Li, and Yang Shi are with the Intelligent Big Data Visualization Lab at Tongji University. * Equal contribution. Yang Shi is the corresponding author.*
E-mail: {2411927, 2433550, ziangli, yangshi.idvx}@tongji.edu.cn.

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

- We proposed a multimodal, multi-task learning model that predicts persuasive strategies and their persuasive effects in data-driven news articles. It outperforms state-of-the-art baselines, demonstrating the essential roles of the proposed cross-attention gated fusion component and loss functions. The dataset and code can be accessed at https://osf.io/8pehf/.

2 Related Work

2.1 Persuasion and Visualization

Persuasion, defined as "*communication intended to shape others' beliefs, values, or attitudes*" [67], involves two phases: (i) the strategic efforts of the persuader (termed persuasive strategies) and (ii) the subsequent cognitive or behavioral uptake by the persuadee [2, 85]. Persuasion has been widely studied across social science disciplines and formalized through frameworks such as the Elaboration Likelihood Model (ELM) [55], which distinguishes between logic-driven (central route) and emotion-driven (peripheral route) persuasion, and the Heuristic-Systematic Model (HSM) [11], which explains how audiences rely on cognitive shortcuts to process persuasive messages. Complementary theories, including Hovland's communication-persuasion paradigm [31] and social identity theory [73], further contextualize how message framing and group affiliations shape receptivity to persuasion. Building on these foundations, the visualization community has explored how data visualizations persuade. Early studies focused on whether visualizations persuade more effectively than other mediums. For example, Garreton et al. [25] found that data-visualized journalism amplified reader attitude influence beyond illustration-dependent storytelling. Subsequent work focuses on factors modulating persuasive outcomes. Studies have shown that audiences' prior beliefs significantly shape their responses to data [35, 49]. For example, Markant et al. [49] demonstrated that people's belief updating becomes more restrained when encountering scatterplot visualizations that contradict their prior assumptions. In addition to audience predispositions, research has investigated how narrative techniques mediate persuasion [36, 45, 60]. Rogha [60] found that while both elicitation and contrasting narratives did not significantly alter attitudes, they led to increased engagement in terms of surprise. Similarly, Kong et al. [36] demonstrated that slanted titles influence how audiences interpret visualizations without altering their perception of bias. The aforementioned work has primarily focused on the reception phase of persuasion rather than the persuasive strategies employed by storytellers. Our work addresses the first phase of persuasion by systematically analyzing and categorizing the strategies used in data-driven news articles.

2.2 Design of Data Stories

Data stories synthesize visual analytics with storytelling techniques to effectively present data and convey analytical insights [37]. Recent research in data story design has expanded in multiple directions. Some studies adopt a holistic approach by exploring various types of data stories and associated design patterns [62], while others focus on specific elements such as interaction [64], narrative structure [43, 87], animation [65], and affective design components [40, 42]. In addition to theoretical frameworks, several studies have examined technical approaches for constructing data stories. For example, researchers have applied Monte Carlo tree search methods to generate narratives based on automatically extracted data facts [63, 72].

Data-driven news articles, which is a specialized format in data stories, combine storytelling with visualization to inform public discourse [28]. Researchers have explored their design through both comprehensive frameworks and granular analyses. For example, Hao et al. [28] introduced a classification framework for data-driven news articles, identifying design patterns for textual and visual components. Fu et al. [24] formulated an ecological framework outlining priority research topics for visual storytelling in news ecosystems. Other work has focused on specific design dimensions that impact user engagement and comprehension. Rahman et al. [57] established a taxonomy of annotation strategies in visualizations, while Kim et al. [34] investigated perceptual factors influencing the efficacy of visualization thumbnails. Rogha et al. [60] analyzed how narrative elicitation and temporal salience in visualizations synergistically shape audience engagement. Technical research has also explored methods to automate data journalism, such as linking data with text [70, 91] or leveraging large language models [13, 71]. Inspired by the aforementioned work, our work focuses on identifying persuasive design patterns in data-driven news articles that influence how audiences interpret and respond to the presented information.

2.3 Computational Modeling of Persuasion

Persuasive theories in psychology and marketing are well established, while computational modeling of persuasive communication remains in its early stages. Prior work has mainly focused on natural language processing (NLP) approaches to quantitatively analyze persuasive texts, which can be categorized into three tasks: the extraction of persuasive strategies, the analysis of influencing factors, and the measurement of persuasive effects. The extraction of persuasive strategies aims to identify rhetorical tactics in text. Early work by Anand et al. [2] explored the feasibility of classifying blog posts as persuasive or non-persuasive based on lexical features in the text. This work paved the way for subsequent studies, such as that of Meng et al. [51], who developed a hierarchical weakly-supervised latent variable model that utilizes partially labeled data to predict the associated persuasive strategies for each sentence. In addition, the analysis of influencing factors examines contextual variables that modulate persuasion, including argument ordering [44], target audience [48], and prior beliefs [21]. Recent work has also focused on measuring the persuasive effect of communication. For example, Jin et al. [32] developed a model based on LLMs through intent-to-strategy reasoning and evaluated its persuasiveness. In addition to NLP approaches, the computer vision (CV) community has contributed to computational modeling of persuasive images. Early work by Joo et al. [33] proposed syntactical and intent-based features, including facial expressions, gestures, emotions, and personality, which contribute for creating of persuasive images. Recent work follows two trends: modality extension [39] and adaptive multimodal fusion [3]. For example, Kumar et al. [39] developed an annotated image corpus featuring persuasive strategies in advertisements. Bai et al. [3] investigated persuasion in debate videos and introduced an adaptive fusion learning framework for predicting the intensity of persuasion using multimodal (acoustic, visual, language) data. The aforementioned work relies on feature-level fusion, limiting their ability to capture how persuasive strategies emerge from cross-modal interactions. We propose a multimodal, multi-task framework that integrates both embedded and explicit features, providing a more comprehensive understanding of how persuasive effects are shaped by different modalities.

3 Characterizing Persuasive Strategies

To understand how data-driven news articles persuade, we conducted focus group studies to gather insights from researchers and practitioners. Based on these discussions, we developed a taxonomy of 15 persuasive strategies grouped into five categories, which reflect a spectrum of techniques that engage audiences through varying cognitive pathways, from careful analysis to intuitive responses.

3.1 Methodology

We initially considered reaching out directly to original authors to understand their intended persuasive strategies. However, practical constraints such as the large volume of articles and the lack of clear authorship made this approach unfeasible. In many cases, visualizations are provided by third parties, which further complicates direct inquiries. To address these challenges, we turned to elicitation [30] (i.e., interacting with human subjects to gather insights) through focus group studies. Focus groups are a well-established method for gathering qualitative insights through structured yet open-ended discussions [38]. In these discussions, participants with diverse experiences in reading and producing data journalism shared their observations on both deliberate and incidental persuasive techniques. This method can stimulate their recollections and associations with relevant experiences, enabling us to identify recurring patterns, consensus, and points of divergence in the manifestation of persuasive strategies.

Table 1: Summary of participants' characteristics

Group	ID	Gender	Age	Role/Profession	Proficiency
G1	P1	F	20	Data Journalism Student	Proficient
G1	P2	F	24	Industrial Design Student	Competent
G1	P3	F	30	University Lecturer	Expert
G1	P4	M	25	HCI Student	Beginner
G1	P5	F	21	Data Journalism Student	Proficient
G1	P6	M	23	Government Agency Staff	Beginner
G2	P7	F	22	Data Journalism Student	Proficient
G2	P8	F	25	Assistant Researcher	Expert
G2	P9	M	22	Computer Science Student	Beginner
G2	P10	F	25	News Reporter	Expert
G2	P11	F	23	Data Journalism Student	Proficient
G2	P12	M	24	Environmental Design Student	Proficient
G3	P13	F	23	Environmental Design Student	Competent
G3	P14	M	24	Electronic Information Student	Beginner
G3	P15	M	25	Art and Design Student	Beginner
G3	P16	F	28	News Editor	Expert
G3	P17	F	31	New Media Operator	Proficient
G3	P18	M	27	Product Manager	Expert

Participants We recruited participants through an open call on social media platforms. To ensure diversity and relevance [38], we selected individuals with experience in data journalism or visualization who also demonstrated basic understanding of persuasion, enabling them to identify strategies specific to data-driven news articles. Eligibility was assessed through a questionnaire that collected demographic information, professional expertise, and self-reported proficiency with data journalism or data visualization. The questionnaire uses a five-point Likert scale to evaluate participants' proficiency, from "novice" to "expert". A total of 18 participants aged 20 to 31 were recruited. Their backgrounds spanned data journalism, computer science, and design, ensuring a multidisciplinary perspective. All participants had experience in either producing or engaging with data-driven news articles (expert: 27.8%, proficient: 33.3%, competent: 11.1%, advanced beginner: 27.8%, novice: 0%). Participants were assigned to three focus groups, each consisting of six individuals (Table 1).

Stimuli To ground discussions in real-world examples, we provided participants with a curated set of data-driven news articles as stimuli. These articles were selected from two sources: (i) an existing corpus of the articles compiled in prior research [28] and (ii) articles nominated by participants prior to the study. Each participant was asked to submit a reading list of articles they found highly persuasive. If fewer than 10 were provided, we supplemented the list with widely recognized examples and asked all participants to read them before the discussion. The selection process was guided by three key criteria. First, articles demonstrated strong persuasive elements to ensure their relevance for analysis. Second, they covered a diverse range of topics while maintaining neutrality, avoiding issues that might provoke strong preconceived biases among participants. Third, articles varied in length, publication sources, and narrative styles to capture a broad spectrum of persuasive techniques. Based on these criteria, we selected 11 articles as stimuli, with each focus group assigned 3 or 4 articles, depending on their length. These stimuli served as the primary basis for discussion, while the participant-submitted reading lists were used to prompt reflection on broader reading experiences.

Materials To help participants connect persuasive strategies to underlying psychological processes, we referred to prior research on persuasive strategies [39] and considered the characteristics of data journalism to select appropriate psychological frameworks. Specifically, we introduced three foundational theories: (i) the Elaboration Likelihood Model (ELM) [55], which distinguishes two persuasion pathways: central route (analytical evaluation of arguments) and peripheral route (reliance on heuristic cues and emotional triggers); (ii) Social Identity Theory [73], which posits that individuals align attitudes with in-group norms to reinforce self-concept; and (iii) Consistency Theory [22], which proposes that inconsistencies between beliefs and actions create psychological tension, motivating efforts to restore consistency. In addition to these theories, we used examples of persuasive strategies commonly used in various contexts to provide participants with a reference point for their analysis, such as the *foot in the door* and *reciprocity* techniques often employed in advertising [39, 51, 88].

Procedure Each focus group followed a three-session protocol to systematically identify and refine persuasive strategies: *introduction, discussion*, and *synthesis*. In the introduction session, participants reviewed the pre-distributed stimuli and were introduced to the psychological theories of persuasion outlined in our materials. To ensure alignment with the study's scope, we clarified that the focus was on examining the persuasive strategies employed by the persuader, rather than evaluating the effectiveness of those strategies on the persuadee. The discussion session began with 15 minutes of independent annotation, where participants identified persuasive elements in the stimuli based on their expertise and the psychological theories introduced earlier. They were required to annotate at the sentence level for text and at the individual visualization level for graphical content. After this, we guided a structured discussion divided into two parts. First, participants mapped the persuasive strategies they identified to the psychological theories introduced earlier. Second, participants were encouraged to explore additional persuasive strategies that may not have been immediately apparent in the initial annotations. We also invited them to share persuasive data-driven news articles they had read (from either the reading list or personal experience) and describe the strategies used. This open-ended exploration prompted deeper reflection on more subtle or less overt forms of persuasion. The synthesis session involved iterative refinement of the findings. We consolidated the proposed strategies into a preliminary list, encouraging participants to resolve disagreements, eliminate redundancies, and refine definitions. Each strategy was then subject to a voting process, ensuring consensus on the final list of strategies. The three sessions lasted approximately two hours and were recorded with participants' consent for later analysis.

Result Analysis The qualitative data from the focus groups was transcribed and analyzed through thematic analysis [8] to derive a set of effective persuasive strategies. Two researchers independently performed open coding on the transcripts, tagging segments related to persuasive strategies. Then, we generated codes from the tagged segments and grouped similar codes. For example, recurring references to "*using quotes from an economist*" and "*citing Environmental Protection Agency*" were synthesized into the code *authority endorsement*. Following this, we met for three sessions to compare codes, merge similar codes, and discuss disagreements until reaching an agreement. Using this coding framework, we then re-coded the entire set of transcripts. During the process, we further grouped low-level codes into the high-level categories they serve. This synthesis was guided by three inputs: (i) insights shared by experts in the focus groups, "*increasing engagement with data journalism can enhance the likelihood of persuasion*" (P3, P10), (ii) psychological theories of persuasion [22, 31, 55, 73, 84], and (iii) existing taxonomies of persuasive strategies from related domains (e.g., advertisement [39], news editorials [21]).

3.2 Taxonomy

Based on our focus groups, we identified 15 distinct persuasive strategies, of which eight were derived from sentence-level annotations in the text (text-based strategies) and seven from annotations within individual visualizations (visualization-based strategies). The 15 strategies are grouped into five categories (Fig. 1): *credibility, guided interpretation, reference-based framing, emotional appeal*, and *participation invitation*. Except for *emotional appeal*, which includes only text-based strategies, the other categories comprise both textual and visual techniques, reflecting a multimodal nature. These categories represent a spectrum of persuasive techniques that, according to the ELM, function through two primary pathways. Specifically, *credibility* and *guided interpretation* target the central route by providing robust evidence and clear explanations that encourage careful scrutiny. In contrast, *reference-based framing* and *emotional appeal* rely on peripheral cues, such as metaphors and affective narratives, to elicit rapid, intuitive responses. Finally, *participation invitation* bridges both pathways by engaging audiences to deepen involvement and promote personal evaluation.

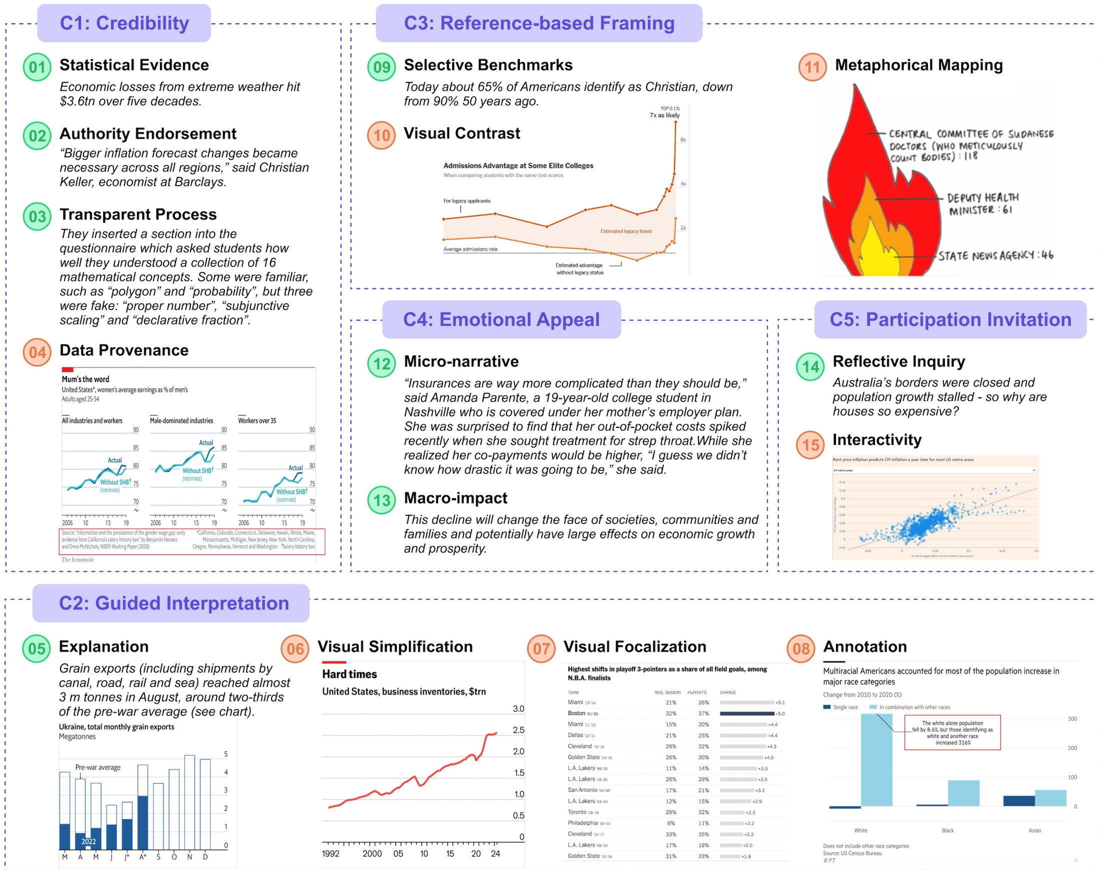

Fig. 1: Examples of the 15 persuasive strategies identified from the focus groups. These strategies are grouped into five categories, C1-C5. Note that ◯ denotes text-based strategies while ◯ denotes visualization-based strategies. Example sources: [16–20, 26, 27, 75–82].

C1: Credibility. This category is designed to enhance the perceived trustworthiness of the information presented. As data-driven news articles often rely on reasoning based on data processing and analysis [58], the logic chain itself can possibly trigger skepticism. Thus, the category helps mitigate skepticism by offering rigorous, clear evidence, thus reassuring audiences about the reliability of the data and analysis [54, 59]. One persuasive strategy for establishing credibility is **statistical evidence**, which involves presenting specific numerical or statistical data to substantiate claims. This helps create an impression of objectivity and authenticity. For example, in Fig. 1-1, a claim about the economic impact of extreme weather events is supported by the statistic that these events have caused $3.6 trillion in losses over the past 50 years, converting the intangible climate crisis into a concrete monetary measure. Another persuasive strategy is **authority endorsement**, which incorporates expert opinions or citing reputable institutions to reinforce the article's conclusions. In Fig. 1-2, a news article on an inflation surge features a quote from economists stating, "*bigger inflation forecast changes became necessary across all regions*". This expert endorsement signals to audiences that the claims are grounded in professional expertise. **Transparent process** is another key strategy, which involves clearly explaining the methods used to collect and analyze the data. By detailing the research process, as shown in Fig. 1-3 with a study on deceptive reporting that explains its survey design, audiences can better assess the integrity of the findings. Finally, **data provenance** contributes to credibility by disclosing the origins of the data used in the visualization, ensuring it can be traced back to reliable sources. In Fig. 1-4, a visualization of gender-based income disparities includes a citation of the data source at the bottom.

C2: Guided Interpretation. This category is dedicated to directing how audiences decode and understand complex data through deliberate visual and narrative cues. In data-driven news articles, visualizations often serve as key tools for presenting and supporting arguments [6]. Due to their density and structural complexity, they may be interpreted in multiple ways. By steering attention toward key insights, these strategies ensure that data is interpreted in line with the storyteller's intended message [54]. One key strategy within this category is **explanation**, where the visualization's central elements are clarified through accompanying text. In Fig. 1-5, a visualization of the global food supply crisis is paired with text explaining that food exports have dropped to two-thirds of their pre-war levels. This explanation directs audiences to focus on the corresponding drop in exports shown in the visualization. Another strategy, **visual simplification**, reduces complexity by removing visually redundant elements, such as gridlines, borders, or background decorations, allowing the main data patterns to emerge clearly. In Fig. 1-6, a simple, clean line graph depicting the steady growth of U.S. business inventories immediately communicates the underlying trend. **Visual focalization** involves directing attention to specific data points through emphasis and contrast, ensuring that important insights can be easily identified. For example, Fig. 1-7 uses a filled dark blue bar to highlight the Boston Celtics' 5% increase in playoff 3-pointers, making this key detail stand out among other unfilled bars. Finally, **annotation** embeds explanatory notes within the visualization to provide the context and meaning behind the data. Fig. 1-8 uses annotation to clarify that that while the white-alone population decreased by 8.6%, those identifying as white and another race increased by 316%.

C3: Reference-Based Framing. This category shapes how data is perceived by anchoring abstract values to familiar comparisons or analogies. As data in data-driven news articles can involve unfamiliar units or abstract measures [12], audiences may struggle to interpret their significance and draw inconsistent conclusions. By establishing reference points and leveraging the framing effect [84], in which different presentations of the same facts lead to different interpretations and decisions, this strategy subtly steers audiences toward the intended evaluation. One key strategy within this category is **selective benchmarks**, which anchors data to carefully chosen reference points or evaluation criteria to emphasize specific data patterns. For example,

Fig. 1-9 compares the proportion of American Christians today with that of fifty years ago (as opposed to a more recent reference point), framing the narrative to stress a long-term decline. In another strategy, **visual contrast** manipulates visual elements to underscore differences between data groups. Fig. 1-10, for example, employs differently colored lines to compare college admission likelihoods for students from different income brackets, despite identical test scores. This visual contrast highlights the advantage of elite students, guiding audiences' interpretation of the data in a way that emphasizes social inequality. Another form of reference-based framing is **metaphorical mapping**, which uses culturally or emotionally charged symbols to construct analogical comparisons, reframing data by embedding it within pre-existing conceptual associations. In Fig. 1-11, a visualization of a flame intensifying in heat as the death toll rises from 46 to 118 is used to depict the Sudanese government's attempt to conceal the true number of protester deaths. The metaphor of a "growing flame" invokes associations with crisis and urgency, guiding audiences to perceive the rising numbers as evidence of government deceit and escalating danger.

C4: Emotional Appeal. This category leverages emotionally compelling narratives to influence audiences by tapping into empathy and shared social values. While data-driven news articles are often calm and objective, those addressing pressing social issues are also expected to convey human-centered concerns [41]. Accordingly, this category connects data with emotional experiences to heighten narrative tension. The experts in our focus groups suggested that when stories evoke personal emotions (e.g., "I can feel their pain") and a sense of collective concern (e.g., "Our future is at risk"), audiences are more likely to undergo an attitudinal shift. One strategy is **micro-narrative**, which focuses on the real-life experiences of specific individuals. Through detailed, expressive storytelling, this strategy encourages audiences to emotionally connect and project themselves into the narrative. For example, Fig. 1-12 illustrates the obstacles Americans face in healthcare through the personal account of a 19-year-old college student in Nashville, transforming an abstract policy issue into a relatable crisis that prompts audiences to reflect, "What if I were in their shoes?" Another frequently used strategy, **macro-impact**, emphasizes the broader implications of individual actions within societal systems. This strategy fosters a sense of shared destiny and collective responsibility. For example, Fig. 1-13 reframes the issue of declining birth rates by emphasizing its extensive impact on society, linking individual reproductive choices to broader economic outcomes.

C5: Participation Invitation. This category invites audiences to actively engage with the content, encouraging participation rather than passive consumption. As data-driven news articles often involve structurally complex and layered information, a fully author-driven presentation can increase cognitive load [62]. This category seeks a balance between author-driven and reader-driven approaches, guiding audiences to think or explore on their own. This helps reduce psychological defenses and encourages deeper processing of information by turning reading into an interactive experience [52]. One strategy is **reflective inquiry**, which uses open-ended questions to stimulate active thought. In Fig. 1-14, the headline "*Australia's borders were closed and population growth stalled – so why are houses so expensive*?" presents an intriguing scenario that encourages audiences to search for answers. Another important strategy is **interactivity**, which allows audiences to filter and manipulate data within the visualization through interaction methods (e.g., dragging, adjusting a timeline, or using selectors) to form their own interpretations of the evidence. In Fig. 1-15, a rental price chart with a city selector enables audiences to explore regional patterns, test hypotheses, and derive personalized insights.

4 Constructing the Corpus

Building on our taxonomy of persuasive strategies, we constructed a corpus of data-driven news articles annotated with both persuasive strategies and persuasive effect labels, which later served as the training set for our predictive model. To achieve this, we first collected a corpus of data-driven news articles. We then engaged experts to annotate sentences and visualizations with these persuasive strategies and employed crowdsourcing to label the overall persuasive effect of each article.

4.1 Data Collection

We first identified media outlets known for publishing high-quality data-driven news articles. Four prominent outlets were selected, including *The Economist, Financial Times, The Guardian*, and *The New York Times*. These outlets were chosen as they (i) maintain dedicated data sections, attract large global readerships, and are recognized for quality journalism [28]; and (ii) serve diverse audiences, ensuring a rich and varied sample [69]. Next, we defined the topical scope of our corpus by focusing on everyday subjects and excluding topics that are either overly controversial (e.g., political elections) or inherently complex (e.g., science popularization) [54]. The final set of subjects comprises six categories: *economy, society, environment, health, sports*, and *technology*. Our initial collection included 1,166 articles published between January 4, 2019, and January 14, 2025.

We then applied screening criteria to refine the corpus: (i) each article must include at least one data visualization and follow a magazine-style layout [62]; (ii) articles composed of multiple independent sub-articles or those that simply list facts without conveying a clear narrative were excluded [1]; and (iii) overly lengthy articles were removed [51]. To implement the third criterion, we calculated the sentence count (including titles and lead paragraphs) and retained only articles within the 90th percentile, corresponding to a maximum of 52 sentences per article. This process yielded a refined corpus of 936 data-driven news articles, comprising 26,222 sentences and 1,876 visualizations. Fig. 2 shows the distribution of sources and subjects in the corpus.

4.2 Labeling Persuasive Strategies

To capture the persuasive strategies within the curated corpus, we conducted a detailed annotation of the news articles. Six research assistants, each with at least three years of experience in visualization research, independently labeled persuasive strategies at the sentence level for textual content and at the individual visualization level for graphical content.

Study Procedure The annotation process was conducted in three sessions to ensure accuracy and reliability. In Session 1, annotators received systematic training that covered definitions of persuasive strategies, representative examples, and detailed annotation guidelines. To gain familiarity with the process, each annotator independently labeled 80 news articles. We then reviewed the annotations, identified inconsistencies, and provided clarifications through group discussions. In Session 2, all annotators independently labeled the same set of 20 news articles to assess consistency. We measured inter-rater reliability using Cohen's Kappa, which yielded average agreement scores exceeding 0.6 for text-based strategies and 0.7 for visualization-based strategies, indicating moderate reliability [50]. In Session 3, we distributed the remaining articles among annotators based on article length, with each annotator labeling approximately 139 articles. Each sentence was assigned up to three persuasive strategies to improve the accuracy of the annotations, while sentences without an identifiable strategy were labeled "none". There was no limit on the number of strategies that could be assigned to visualizations.

4.3 Labeling Persuasive Effect

Following the annotation of persuasive strategies, we assessed the overall persuasive effect of each article using crowdsourced ratings, to

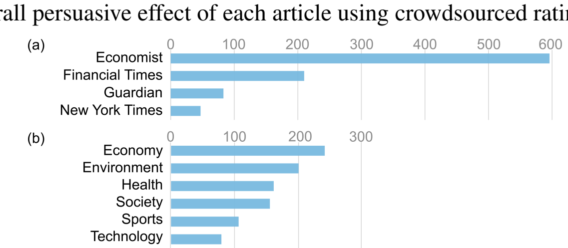

Fig. 2: Frequency of (a) sources and (b) subjects in the 936 articles.

provide training data for the model to learn the contribution of each persuasive strategy to the overall persuasiveness of articles.

Participants Participants were recruited via Prolific, a platform known for its diverse and high-quality respondent pool with robust attention checks [15]. We restricted participation to individuals with at least a high school diploma, residing in English-speaking countries (e.g., the United States and the United Kingdom), and fluent in English. Additionally, only participants with an approval rate of 98% or higher were included. In total, 1,032 participants (491 females) aged 18-50 (M = 34.69, SD = 8.53) were recruited. Each participant was compensated £1.5 for reading one to three news articles (approximately 1,200 words in total) and for rating their persuasive effects.

Study Procedure The online evaluation process was organized into three sessions. In Session 1, participants were introduced to the study objectives, key concepts, and the questionnaire structure. In Session 2, each participant read one to three data-driven news articles and rated the persuasive effect of each article on a five-point Likert scale (1: Not at all persuasive; 3: Moderately persuasive; 5: Extremely persuasive). In Session 3, participants provided a brief explanation (at least one sentence) for their ratings. Responses lacking satisfactory justification were discarded, resulting in the exclusion of 12.49% of submissions. This process was repeated until each article had received three valid persuasive effect ratings. The final persuasive effect score for each article was computed as the average of these ratings.

5 MODEL

Based on our annotated corpus, we developed M2P, a **M**ultimodal **M**ulti-task learning model that **P**redicts both the presence of persuasive strategies as well as their persuasive effects. In this section, we first provide an overview of the model's objective and architecture, and then describe the technical details of each module.

5.1 Overview

Problem Definition. Let a data-driven news article be represented as $\mathbf{t} = \{t_1, t_2, \ldots, t_m\}$, where $\mathbf{t}$ consists of m sentences, and $\mathbf{v} = \{v_1, v_2, \ldots, v_n\}$, where $\mathbf{v}$ consists of n visualizations. The goal of the proposed model f is twofold: (i) to identify the persuasive strategy label P_{t_i} for each sentence in $\mathbf{t}$ and P_{v_i} for each visualization in $\mathbf{v}$, and (ii) to predict the overall persuasive effect of the article, denoted as $\hat{y}$, i.e., $[\mathbf{P}_t, \mathbf{P}_v, \hat{y}] = f(\mathbf{t}, \mathbf{v})$.

Two major challenges arise in achieving this goal. First, data-driven news articles incorporate both text and visualizations, which present multimodal features. Effectively capturing the cross-modal semantic relationships, while also evaluating the individual contributions of each modality to the overall persuasive effect, is a challenging task. Second, the corpus exhibits imbalanced label distributions, with labels being both ordinal and discrete. This introduces a challenge in designing task-specific training objectives that align with these characteristics, which is essential for improving the robustness of both the persuasive strategy and effect prediction. To address these challenges, we first employ bidirectional cross-attention combined with a gated mechanism to align and fuse the multimodal features. This allows us to capture the interplay between text and visual elements while dynamically weighting the features to accurately predict the persuasive effect. Second, to handle the imbalanced and mixed nature of the labels, we integrate an adaptive Focal Loss, which adjusts sample weights for different strategies, and use CORAL Loss to transform the persuasive effect prediction into a series of ordinal binary classification tasks. This approach improves the accuracy of both the persuasive strategy and effect prediction.

In summary, we designed a model that is composed of three modules (Fig. 3): (1) *feature extractors*, which extract both embedded (text and visualization) and explicit (visual narrative and psycholinguistic) features from a given data-driven news article, (2) *persuasive strategy predictors*, which use the embedded features to identify the persuasive strategies applied to each sentence and visualization, and (3) *persuasive effect predictor*, which combines the fused features from text and visualization embeddings, embedded and explicit features, and probability distributions of strategies to predict the persuasive effect of the article.

5.2 Feature Extractors

The feature extractors module serves as the basis for predicting persuasive strategy and persuasive effect by extracting embedded and explicit features from an article. The embedded features capture latent information through text and visualization embeddings, encompassing semantic relationships and low-level patterns such as color distributions. The explicit features, including visual narrative and psycholinguistic features, provide higher-level insights into the structural aspects and textual elements that affect cognitive and emotional processing.

Text Features We extract text features using OpenAI's text-embedding-ada-002 model [53]. The article's title, lead, and main text are concatenated and encoded within the model's 8192-token context window. This produces 1536-dimensional embeddings, which we reduce to 256 (h_t = 256) using an autoencoder, considering the relatively small size of our corpus. The resulting text features include two parts: $\mathbf{E}_T \in \mathbb{R}^{h_t}$ for the full article and $\mathbf{E}_t = [e_{t_1}, e_{t_2}, \ldots, e_{t_m}] \in \mathbb{R}^{m \times h_t}$ for individual sentences, where e_{t_i} denotes the i-th sentence's embedding.

Visualization Features For visualization features, we use the CLIP framework [56] with the ViT-B/32 encoder. Each visualization is resized to 224×224 pixels, normalized, and divided into 32×32 patches. These patches are linearly projected into visual tokens and encoded by a 12-layer Transformer. A learnable CLS token generates a global semantic vector. To reduce complexity, we compress the embeddings to 128 dimensions (h_v = 128) using an autoencoder. The final visualization features are $\mathbf{E}_v = [e_{v_1}, e_{v_2}, \ldots, e_{v_n}] \in \mathbb{R}^{n \times h_v}$, where e_{v_i} is the i-th visualization's embedding.

Visual Narrative Features Visual narrative features include both text-related and visualization-related features. Text-related features describe the logical flow of the article (e.g., an inverted pyramid structure to enhance comprehension), while visualization-related features depict the types of visualizations and their annotation styles. For example, different annotation methods (e.g., numerical and threshold annotations) can impact how data is perceived and understood [28]. To do this, we employ GPT-4o to examine both the text and visualizations, identifying 36 distinct visual narrative features, which are represented as binary vectors $\mathbf{E}_N \in \mathbb{R}^{36}$ (see Supplementary Materials).

Psycholinguistic Features Psycholinguistic features are measurable textual attributes (e.g., word concreteness, emotional valence) that reflect how language shapes cognitive processing and emotional engagement during reading [21]. For example, the use of words like 'sincere' can indicate authenticity, while terms like 'powerful' may suggest clout. We use the LIWC-2022 tool [7] to compute six psycholinguistic metrics, including Clout, Authentic, Big word, Emotional tone, Positive emotion, and Negative emotion, which are then normalized into a 6-dimensional vector $\mathbf{E}_L \in \mathbb{R}^6$ (see Supplementary Materials).

As a result, we obtain the embedded features $\mathbf{E}_T$, $\mathbf{E}_t$ and $\mathbf{E}_v$, and concatenate the visual narrative features $\mathbf{E}_N$ and psycholinguistic features $\mathbf{E}_L$ to form the explicit features $\mathbf{E}_E \in \mathbb{R}^{42}$.

5.3 Persuasive Strategy Predictors

The persuasive strategy predictors module leverages the extracted embedded features to perform supervised multi-label classification. It outputs probability distributions that indicate the persuasive strategies for each sentence and visualization in the article. Specifically, the embedded features from text $\mathbf{E}_t$ and visualization $\mathbf{E}_v$ fed into separate text-based and visualization-based strategy predictors. Each predictor consists of fully-connected layers. These predictors generate probability distributions at both the sentence level and the visualization level: $\mathbf{P}_t \in \mathbb{R}^{m \times k_t}$ and $\mathbf{P}_v \in \mathbb{R}^{n \times k_v}$, where k_t and k_v represent the total number of text-based and visualization-based strategy classes, respectively. The probability distribution for the text $\mathbf{P}_t$ is computed as:

$$\mathbf{P}_t = W_t^{(2)} (\mathrm{ReLU}(W_t^{(1)} \mathbf{E}_t + b_t^{(1)})) + b_t^{(2)} \quad (1)$$

where W_t and b_t are the parameters of the fully-connected layers in the text-based strategy predictor. Similarly, the probability distribution $\mathbf{P}_v$ generated by the visualization-based strategy predictor is calculated in the same manner.

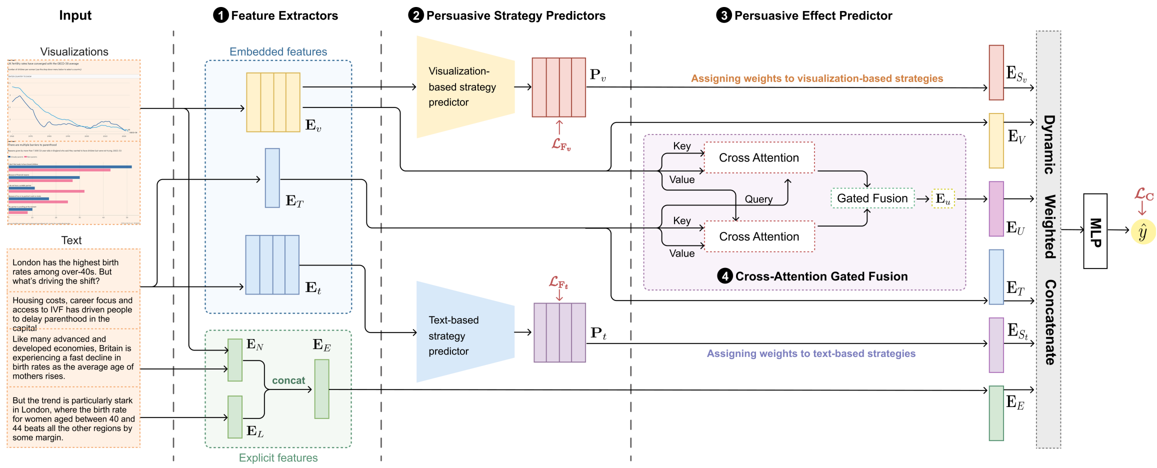

Fig. 3: The architecture of our model M2P, including (1) *feature extractors*, (2) *persuasive strategy predictors*, and (3) *persuasive effect predictor*. Specifically, (4) illustrates the component of *cross-attention gated fusion*.

5.4 Persuasive Effect Predictor

The persuasive effect predictor module measures the persuasive effect of the article through two phases: (i) *cross-attention gated fusion* and (ii) *heterogeneous weighted fusion*.

5.4.1 Cross-Attention Gated Fusion

In the first phase, we combine the embedded features from text $\mathbf{E}_T$ and visualizations $\mathbf{E}_v$ to generate fused features $\mathbf{E}_u$, which capture the cross-modal semantic relationships between these two modalities. We begin by aligning the semantic information of both text and visualization through bidirectional cross-attention, resulting in mutually informed features, $\mathbf{E}_{T\leftarrow v}$ and $\mathbf{E}_{v\leftarrow T}$. This approach mitigates the limitation of a single modality dominating the fusion, as can occur with unidirectional attention. Specifically, to transfer information from visualizations to text, we first align the visualization features $\mathbf{E}_v$ with the text dimension h_t using a linear projection, resulting in the transformed visualization features $\mathbf{E}'_v \in \mathbb{R}^{n\times h_t}$. Then, using $\mathbf{E}_T$ as the Query and $\mathbf{E}'_v$ as both the Key and Value, we apply an 8-head attention mechanism to obtain the visualization-aware text features $\mathbf{E}_{T\leftarrow v}$:

$$\mathbf{E}_{T\leftarrow v} = \text{MultiHeadAttn}(\mathbf{E}_T, \mathbf{E}'_v, \mathbf{E}'_v) = \text{softmax}\left(\frac{\mathbf{E}_T W_q (\mathbf{E}'_v W_k)^\top}{\sqrt{d}}\right)\mathbf{E}'_v W_v \quad (2)$$

where W_q, W_k, and W_v are projection matrices and $d = 256$ is the query dimension.

A similar procedure is followed for transferring information from text to visualization, where the text features $\mathbf{E}_T$ are projected into the visualization space to obtain the text-aware visualization features $\mathbf{E}_{v\leftarrow T}$, ensuring that both modalities are mutually informed.

Next, we merge these two representations $\mathbf{E}_{T\leftarrow v}$ and $\mathbf{E}_{v\leftarrow T}$ using an adaptive gated mechanism, which helps dynamically capture the relationships between text and visualizations. Specifically, we first project $\mathbf{E}_{T\leftarrow v}$ and $\mathbf{E}_{v\leftarrow T}$ into a unified space with dimension $h_u = 256$ via linear transformations, resulting in $\mathbf{E}'_{T\leftarrow v}$ and $\mathbf{E}'_{v\leftarrow T}$, respectively. We then concatenate these features and compute a gated weight $\mathbf{g}$:

$$\mathbf{g} = \sigma(W_g [\mathbf{E}'_{T\leftarrow v} \oplus \mathbf{E}'_{v\leftarrow T}]) \quad (3)$$

where σ is the sigmoid function, W_g are gated parameters, and $\oplus$ denotes concatenation.

The final fused features are obtained by:

$$\mathbf{E}_u = \mathbf{g} \odot \mathbf{E}'_{T\leftarrow v} + (1-\mathbf{g}) \odot \mathbf{E}'_{v\leftarrow T} \quad (4)$$

5.4.2 Heterogeneous Weighted Fusion

In the second phase, dynamic weighting is applied to the extracted features to predict the article's overall persuasive effect. Specifically, we first assign weights w_{s_t}, w_{s_v} to the probability distributions $\mathbf{P}_t, \mathbf{P}_v$, respectively. This enables us to explore the individual contributions of text-based and visualization-based strategies to the persuasive effect. We then obtain the text-based and visualization-based strategy features, $\mathbf{E}_{S_t}$ and $\mathbf{E}_{S_v}$: $\mathbf{E}_S = \mathbf{P} \odot w_S$.

Next, the strategy features $\mathbf{E}_S$, are combined with the fused features $\mathbf{E}_u$, the embedded features $\mathbf{E}_T, \mathbf{E}_v$, and the explicit features $\mathbf{E}_E$ to predict the overall persuasive effect. To account for the varying importance of each feature, we assign learnable weights $\mathbf{w}$ to these heterogeneous features $\mathbf{E}$, resulting in the weighted features: $\mathbf{E}_{\text{weighted}} = \mathbf{E} \odot \mathbf{w}$. Since these features have diverse dimensionalities, $\mathbf{E}_{\text{weighted}}$ is further linearly projected into a uniform dimension $d = 64$. The projected features are then concatenated to form a joint feature vector $\mathbf{E}_{\text{concat}}$. Finally, we feed $\mathbf{E}_{\text{concat}}$ into a 3-layer MLP to predict the article's overall persuasive effect $\hat{y}$.

5.5 Loss Function

The loss function of the model consists of two components: the Focal Loss for persuasive strategy prediction and the CORAL Loss for persuasive effect prediction.

Focal Loss For the persuasive strategy prediction task, we employ the adaptive Focal Loss [46] to address the class imbalance issue, where some strategies are underrepresented. The Focal Loss dynamically adjusts class-specific weights α to assign differentiated penalties to labels with varying prevalence. This prevents the model from overfitting to majority classes while ensuring that rare classes receive sufficient attention. Also, the learnable focusing parameter γ amplifies the loss for hard-to-classify samples by applying the modulation factor $(1-p_c)^\gamma$, helping the model focus on these hard samples. Formally, the Focal Loss for text-based strategy prediction, denoted as $\mathcal{L}_{\text{F}_t}$, and the Focal Loss for visualization-based strategy prediction, denoted as $\mathcal{L}_{\text{F}_v}$, are computed in the same manner. Specifically, $\mathcal{L}_{\text{F}_t}$ is calculated as:

$$\mathcal{L}_{\text{F}_t} = \frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{k_t}\left[l(\alpha_t)\cdot(1-p_{c_t})^{\gamma_t}\cdot CE\left(\sigma(p_{t_{ij}}), s_{t_{ij}}\right)\right] \quad (5)$$

where the ground truth label $s_{t_{ij}} \in \{0,1\}$ represents the presence of the j-th text-based strategy for the i-th sentence. The CE term denotes the cross-entropy loss, and $\sigma(p_{t_{ij}})$ represents the sigmoid-activated j-th text-based strategy probability of the i-th sentence. The learnable parameter α_t is the class weight initialized according to the text-based strategy class's proportion, and the function $l(\alpha_t) = 0.1 + 0.8\sigma(\alpha_t)$ is used to prevent loss vanishing. The focusing parameter γ_t, initialized as 2, is associated with the modulation factor $(1-p_{c_t})^{\gamma_t}$, where p_{c_t} is defined as:

$$p_{c_t} = \begin{cases} \sigma(p_{t_{ij}}), & \text{if } s_{t_{ij}} = 1, \\ 1-\sigma(p_{t_{ij}}) & \text{otherwise} \end{cases} \quad (6)$$

CORAL Loss For the persuasive effect prediction task, we adopt the CORAL Loss [10], which is specifically designed for ordinal labels.

Table 2: The performance of baseline methods and our model

Model	Text-based Strategy Prediction			Visualization-based Strategy Prediction			Persuasive Effect Prediction	
	Micro Precision	Micro Recall	Micro F1	Micro Precision	Micro Recall	Micro F1	MAE	RMSE
ResNet-152	–	–	–	0.7783	0.8720	0.8223	0.5477	0.7325
EfficientNet-B0	–	–	–	0.7805	0.8775	0.8260	0.5079	0.6947
FastText	0.7157	0.7246	0.7201	–	–	–	0.5073	0.6811
GBRT	–	–	–	–	–	–	0.5307	0.7339
SVR	–	–	–	–	–	–	0.5481	0.6958
M2P-noFusion	0.7580	0.7793	0.7685	0.7980	0.9124	0.8513	0.5023	0.6588
M2P-noFocalL	0.7941	0.6913	0.7391	0.8644	0.8620	0.8632	0.5123	0.6773
M2P-noCoralL	0.7608	0.7787	0.7696	0.7957	0.9145	0.8509	1.5936	1.7475
M2P	0.7604	0.7798	0.7699	0.7921	0.9185	0.8506	0.4883	0.6580

Since the persuasive effect labels are discrete and ordinal, we reformulate the prediction task into a series of ordered binary classification tasks. This formulation enables the model to effectively capture the ordinal relationships between the discrete labels. The CORAL Loss is calculated as follows:

$$\mathcal{L}_{\mathrm{C}} = \sum_{i=1}^{k_e-1} \left[t_i \cdot \log \sigma(h_i) + (1-t_i) \cdot \log\left(1-\sigma(h_i)\right) \right] \tag{7}$$

where k_e is the number of ordinal categories, $\mathbf{h} = [h_1, \ldots, h_{k_e-1}]$ are the predicted logits, and t_i is a binary threshold mask. For sample belonging to the j-th category, its mask t_i is defined as:

$$t_i = \begin{cases} 1, & \text{if } i \le j, \\ 0, & \text{otherwise,} \end{cases} \quad \forall i \in \{1, 2, \ldots, k_e - 1\} \tag{8}$$

Finally, the combined loss function is as follows:

$$\mathcal{L} = \mathcal{L}_{\mathrm{F}_t} + \mathcal{L}_{\mathrm{F}_v} + \mu\, \mathcal{L}_{\mathrm{C}} \tag{9}$$

where μ is a hyperparameter balancing the persuasive strategy and persuasive effect prediction tasks.

6 EXPERIMENTS

To evaluate M2P's performance, we conducted a series of experiments.

6.1 Methodology

Experiment Setting We randomly split our dataset into training, validation, and test sets in an 8:1:1 ratio using k-fold cross-validation. The model was trained using the Adam optimizer with a batch size of 16 and a learning rate of 1×10^{-3}. We trained the model for 30 epochs using 10-fold cross-validation with a weight decay coefficient of 1×10^{-4}. All experiments were conducted on an NVIDIA Tesla V100 GPU, and we fixed the random seed to 23 to ensure reproducibility.

Tasks and Measurements We evaluate our model on two tasks: (i) persuasive strategy prediction, where we assess performance using the micro-averaged F1 score, precision, and recall; and (ii) persuasive effect prediction, where we measure prediction accuracy using mean absolute error (MAE) and root mean squared error (RMSE) to determine how closely the predictions match the ground truth.

Baselines We selected three groups of state-of-the-art methods as baselines for comparative analysis to evaluate performance on both persuasive strategy and persuasive effect prediction tasks. The first baseline group evaluates persuasive strategy prediction:

- *ResNet-152* [29] is a deep residual network with an ImageNet pretrained architecture that excels in feature extraction and captures complex visual patterns.
- *EfficientNet-B0* [74] uses compound scaling to balance depth, width, and resolution, with fine-tuning of the last five convolutional layers and a Swish activation function for enhanced nonlinear modeling.
- *FastText* [5]employs character-level n-grams for enriched representations of low-frequency words, using a hashed n-gram feature space and hierarchical softmax for efficient classification.

The second baseline group evaluates persuasive effect prediction:

- *Gradient Boosted Regression Trees (GBRT)* [23] is an ensemble method that fits decision trees iteratively to the loss gradient, with feature subsampling to improve robustness.
- *Support Vector Regression (SVR)* [9] uses structural risk minimization and an ε-insensitive loss to define an error-tolerant interval. We adopt an RBF kernel and use grid search to tune ε, C, γ.

The third baseline group includes variants of M2P to evaluate the effectiveness of our model's architecture:

- *M2P-noFusion* removes the cross-attention gated fusion component.
- *M2P-noFocalL* replaces the Focal Loss with cross-entropy loss.
- *M2P-noCoralL* replaces the CORAL Loss with MAE loss.

6.2 Results and Findings

We present the overall performance results in Table 2 concerning F1 score(micro), precision(micro), recall(micro), MAE, and RMSE. For persuasive strategy prediction, M2P improved Micro Precision by 3.2%, Recall by 5.9%, and F1 by 4.4% on average over the baselines. For persuasive effect prediction, M2P reduced MAE by 9.4% and RMSE by 7.9%, confirming its outstanding performance in both tasks.

Next, we evaluate M2P against the three variants in the third baseline group. First, comparing M2P to *M2P-noFusion* reveals a small but meaningful improvement in both MAE (from 0.5023 to 0.4883) and RMSE (from 0.6588 to 0.6580). These gains result from incorporating the cross-attention mechanism that facilitates fine-grained, bidirectional interactions between text and visualization features to reduce information loss, and the gated fusion mechanism that further balances modality contributions by suppressing unimodal noise.

When comparing M2P to *M2P-noFocalL*, we observe a boost in persuasive strategy prediction. Recall increases from 0.7767 to 8492, and F1 improves from 0.8012 to 0.8103 on average in visualization-based and text-based tasks. We attribute these improvements to the adaptive Focal Loss, which amplifies the loss weights for challenging samples through a parameterized modulation factor, helping to balance class gradients via the learnable parameter α. However, M2P shows lower Micro Precision and F1 in visualization-based strategy prediction compared to *M2P-noFocalL*. This may be due to the way Focal Loss reweights classes with α, which could weaken the optimization of the majority classes and cause overfitting to difficult samples, especially in noisy data scenarios.

Finally, compared with *M2P-noCoralL*, M2P shows substantial improvement in persuasive effect prediction. M2P achieves an MAE of 0.4883 and an RMSE of 0.6580, significantly outperforming *M2P-noCoralL* ($p < 0.01$), which has an MAE of 1.5936 and an RMSE of 1.7475. It is likely due to the CORAL Loss, which reformulates the regression task into ordered binary classification tasks. This transformation helps enforce monotonic relationships among labels and reduces prediction errors by confining them to defined intervals.

In addition to evaluating performance metrics, we analyzed the learned weight of each persuasive strategy to assess its relative contribution to the overall persuasive effect. Our analysis shows that text-based strategies were assigned a weight of 0.561, slightly higher than the 0.439 attributed to visualization-based strategies, indicating that while both modalities play important roles, text exerts a somewhat larger influence. When considering the context of our dataset, however, the balance between text and visualizations becomes even more insightful. A typical data-driven news article in our corpus contains an average of 28 sentences and 2 visualizations. This suggests that while text-based strategies dominate in quantity, visualization-based strategies still perform robustly given their lower occurrence. In other words, even though visualizations appear less frequently, their influence remains comparable to that of text. Delving deeper, we observed that specific text-based strategies, including *selective benchmarks*, *micro-narrative*, and *transparent process*, received weights of 0.170, 0.133, and 0.126, respectively (Fig.4 (a)), implying that techniques promoting cognitive engagement and credibility are particularly effective. Similarly, among visualization-based strategies, *annotation* (0.182), *data provenance* (0.170), and *visual focalization* (0.156) emerged as the most influential (Fig.4 (b)), underscoring the important roles of guided interpretation and credibility in driving persuasive effects. Additionally, we conducted a statistical analysis of the usage proportions of different persuasive strategies across various chart types to explore potential relationships between visual features and persuasive strategies. We found that most charts commonly employ *Visual Contrast*, *Visual Simplicity*, and *Data Provenance* (0.2~0.3). Moreover, scatter plots also frequently use *Annotation* (0.15), which may be attributed to the fact that scatter plots contain numerous data points, and annotations can enhance users' understanding of key information [90].

7 Discussion

The role of visualization in responsive data storytelling. Our taxonomy identified visualization-based persuasive strategies like *visual contrast* and *visual focalization*, providing systematic analytical tools for understanding how data-driven news articles influence audiences. However, these high-level strategies are actually composed of more granular visual encoding elements. Appropriate use of these encodings enhances persuasion, but overuse risks information manipulation. Therefore, understanding how visual encodings shape persuasive effect at the fundamental level is crucial for balancing journalistic integrity with communicative efficacy.

Future research can explore the persuasive mechanisms of visual encodings through directions such as strategy decomposition and encoding effect assessment, providing scientific guidance for optimizing persuasive design. Existing research shows that different saturation and brightness levels can influence audiences' information perception ability [86], and different chart embellishments affect comprehension accuracy [68]. Furthermore, future research could systematically decompose strategies into their constituent visual encodings and explore how these encodings combine to form high-level strategies (e.g., how visual variables like color, size, and position in *visual contrast* work together to highlight specific data patterns), establishing more precise models of persuasive mechanisms. Second, research could analyze the persuasive effect of granular visual encodings (e.g., the relationship between visual metaphor intensity in *metaphorical mapping* and readers' information acceptance), developing tools to quantitatively assess visual encodings' persuasive effect.

Fig. 4: The learned weight of (a) text-based strategies and (b) visualization-based strategies.

However, the influence of visual encodings also raises ethical challenges. When *visual focalization* overemphasizes outliers or axes are truncated to exaggerate differences, it may constitute information manipulation [47]. Future research should develop methods to detect and prevent misuse, including establishing visual integrity scoring systems and formulating data visualization ethical guidelines, ensuring data journalism maintains credibility while preserving persuasive effect.

Controllable Generation of Persuasive Data Storytelling. We have extensively examined persuasive strategies in data-driven news articles and proposed a model capable of both persuasive strategy and effect prediction, thereby advancing computational modeling of persuasion in data journalism. While the predictive model helps us understand the underlying mechanisms in existing articles, the next essential step involves generative modeling to produce data-driven news that balances persuasive effect with ethical transparency.

Future research can focus on developing intelligent systems that reliably generate data news according to specific topics or predefined persuasive strategies. Such systems could efficiently create high-quality articles and flexibly adapt persuasive strategies for different audiences. Recent work on controllable generation [14] introduced the PPLM framework, which combines a pretrained language model with attribute controllers. By adjusting the hidden state according to gradient signals from these controllers, the generated content aligns with target attributes. Similarly, our persuasive strategy predictors can serve as a strategy controller, supplemented by other controllers (e.g., sentiment or topic) to guide generation toward desired attributes. This framework supports multiple attribute combinations without requiring global finetuning, thereby enhancing efficiency.

Controllable generation should also ensure safety and ethical transparency. Although persuasive articles can promote important social issues such as environmental protection or public health [25, 60], they also risk enabling disinformation. Building on research that employs natural-language ethical guidelines for real-time content moderation [4], future work can similarly identify common malicious strategies and set strict rules against harmful behaviors like hate incitement. Encoding these prohibitions into the generation process would enable ongoing self-assessment and revision, further supporting ethical and transparent data storytelling.

Limitations. Our work has three limitations. First, when annotating persuasive effects, we didn't fully account for the influence of annotators' prior attitudes. To enhance objectivity, future work should consider broader annotator diversity and recheck high-variance items. Second, our framework does not yet capture how multiple persuasive strategies might interact or reinforce one another, which could be modeled via GNN [61] to illuminate potential strategy clusters. Third, although tested on 936 data-driven news articles covering diverse topics and styles, our model's performance and generalizability across broader domains or more extensive corpora still warrant further exploration.

8 Conclusion

In this work, we introduced a taxonomy of persuasive strategies for data-driven news articles. Building on this taxonomy and our annotated corpus, we developed M2P, a multimodal, multi-task learning model that jointly predicts the presence of persuasive strategies at both the text and visualization levels, as well as the overall persuasive effect of the article. Experimental results revealed that our model outperforms existing baselines in both persuasive strategy and persuasive effect prediction tasks. Future research should explore how multiple persuasive strategies are structured within an article and how they interact with audience biases. Also, developing methods for the controllable generation of news articles remains a promising direction, particularly as it allows for more responsible data storytelling.

9 Acknowledgments

Yang Shi is the corresponding author. This work was supported by the National Key Research and Development Program of China (2023YFB3107100) and the National Natural Science Foundation of China (62472314).

References

[1] T. Alhindi, S. Muresan, and D. Preoţiuc-Pietro. Fact vs. opinion: the role of argumentation features in news classification. In *Proceedings of the 28th international conference on computational linguistics*, pp. 6139–6149, 2020. 5

[2] P. Anand, J. King, J. L. Boyd-Graber, E. Wagner, C. H. Martell, D. W. Oard, and P. Resnik. Believe me-we can do this! annotating persuasive acts in blog text. In *Computational Models of Natural Argument*, 2011. 2

[3] C. Bai, H. Chen, S. Kumar, J. Leskovec, and V. Subrahmanian. M2p2: Multimodal persuasion prediction using adaptive fusion. *IEEE Transactions on Multimedia*, 25:942–952, 2021. 2

[4] Y. Bai, S. Kadavath, S. Kundu, A. Askell, J. Kernion, A. Jones, A. Chen, A. Goldie, A. Mirhoseini, C. McKinnon, et al. Constitutional ai: Harmlessness from ai feedback. *arXiv preprint arXiv:2212.08073*, 2022. 9

[5] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov. Enriching word vectors with subword information. *Transactions of the association for computational linguistics*, 5:135–146, 2017. 8

[6] L. Bounegru, L. Chambers, and J. Gray. *The data journalism handbook*. European Journalism Centre, 2012. 4

[7] R. L. Boyd, A. Ashokkumar, S. Seraj, and J. W. Pennebaker. The development and psychometric properties of liwc-22. *Austin, TX: University of Texas at Austin*, 10:1–47, 2022. 6

[8] V. Braun and V. Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101, 2006. 3

[9] R. G. Brereton and G. R. Lloyd. Support vector machines for classification and regression. *Analyst*, 135(2):230–267, 2010. 8

[10] W. Cao, V. Mirjalili, and S. Raschka. Rank consistent ordinal regression for neural networks with application to age estimation. *Pattern Recognition Letters*, 140:325–331, 2020. 7

[11] S. Chaiken. Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of personality and social psychology*, 39(5):752, 1980. 2

[12] Q. Chen, W. Shuai, J. Zhang, Z. Sun, and N. Cao. Beyond numbers: Creating analogies to enhance data comprehension and communication with generative ai. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2024. 4

[13] L. Cheng, D. Deng, X. Xie, R. Qiu, M. Xu, and Y. Wu. Snil: generating sports news from insights with large language models. *IEEE Transactions on Visualization and Computer Graphics*, 2024. 2

[14] S. Dathathri, A. Madotto, J. Lan, J. Hung, E. Frank, P. Molino, J. Yosinski, and R. Liu. Plug and play language models: A simple approach to controlled text generation. *arXiv preprint arXiv:1912.02164*, 2019. 9

[15] B. D. Douglas, P. J. Ewell, and M. Brauer. Data quality in online human-subjects research: Comparisons between mturk, prolific, cloudresearch, qualtrics, and sona. *Plos one*, 18(3):e0279720, 2023. 6

[16] T. Economist. Who are the biggest bullshitters? https://www.economist.com/graphic-detail/2019/04/30/who-are-the-biggest-bullshitters, 2019. Accessed: 2025. 4

[17] T. Economist. How to shrink america's gender pay-gap. https://www.economist.com/graphic-detail/2020/05/07/how-to-shrink-americas-gender-pay-gap, 2020. Accessed: 2025. 4

[18] T. Economist. Why are american politicians more pious than their constituents? https://www.economist.com/graphic-detail/2021/02/08/why-are-american-politicians-more-pious-than-their-constituents, 2021. Accessed: 2025. 4

[19] T. Economist. The next threat to global food supplies. https://www.economist.com/graphic-detail/2022/09/02/the-next-threat-to-global-food-supplies, 2022. Accessed: 2025. 4

[20] T. Economist. Three charts show that america's imports are booming. https://www.economist.com/graphic-detail/2024/09/05/three-charts-show-that-americas-imports-are-booming, 2024. Accessed: 2025. 4

[21] R. El Baff, H. Wachsmuth, K. Al Khatib, and B. Stein. Analyzing the persuasive effect of style in news editorial argumentation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 3154–3160, 2020. 2, 3, 6

[22] L. Festinger. A theory of cognitive dissonance. *Row, Peterson*, 1957. 3

[23] J. H. Friedman. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pp. 1189–1232, 2001. 8

[24] Y. Fu and J. Stasko. More than data stories: Broadening the role of visualization in contemporary journalism. *IEEE Transactions on Visualization and Computer Graphics*, 2023. 2

[25] M. Garretón, F. Morini, P. Celhay, M. Dörk, and D. Parra. Attitudinal effects of data visualizations and illustrations in data stories. *IEEE Transactions on Visualization and Computer Graphics*, 2023. 1, 2, 9

[26] T. Guardian. Sudan's government is minimizing the death toll in the khartoum attack. https://www.theguardian.com/news/datablog/2019/jun/16/sudans-government-is-minimizing-the-death-toll-in-the-khartoum-attack, 2019. Accessed: 2025. 4

[27] T. Guardian. Australia's borders were closed and population growth stalled – so why are houses so expensive? https://www.theguardian.com/news/datablog/2022/jun/22/australias-borders-were-closed-and-population-growth-stalled-so-why-are-houses-so-expensive, 2022. Accessed: 2025. 4

[28] S. Hao, Z. Wang, B. Bach, and L. Pschetz. Design patterns for data-driven news articles. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pp. 1–16, 2024. 1, 2, 3, 5, 6

[29] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016. 8

[30] T. Hogan, U. Hinrichs, and E. Hornecker. The elicitation interview technique: Capturing people's experiences of data representations. *IEEE Transactions on Visualization and Computer Graphics*, 22(12):2579–2593, 2016. 2

[31] C. I. Hovland, I. L. Janis, and H. H. Kelley. Communication and persuasion. 1953. 2, 3

[32] C. Jin, K. Ren, L. Kong, X. Wang, R. Song, and H. Chen. Persuading across diverse domains: a dataset and persuasion large language model. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1678–1706, 2024. 2

[33] J. Joo, W. Li, F. F. Steen, and S.-C. Zhu. Visual persuasion: Inferring communicative intents of images. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 216–223, 2014. 2

[34] H. Kim, J. Kim, Y. Han, H. Hong, O.-S. Kwon, Y.-W. Park, N. Elmqvist, S. Ko, and B. C. Kwon. Towards visualization thumbnail designs that entice reading data-driven articles. *IEEE Transactions on Visualization and Computer Graphics*, 2023. 2

[35] Y.-S. Kim, K. Reinecke, and J. Hullman. Data through others' eyes: The impact of visualizing others' expectations on visualization interpretation. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):760–769, 2018. 2

[36] H.-K. Kong, Z. Liu, and K. Karahalios. Frames and slants in titles of visualizations on controversial topics. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2018. 2

[37] R. Kosara and J. Mackinlay. Storytelling: The next step for visualization. *Computer*, 46(5):44–50, 2013. 2

[38] R. A. Krueger. *Focus groups: A practical guide for applied research*. Sage publications, 2014. 2, 3

[39] Y. Kumar, R. Jha, A. Gupta, M. Aggarwal, A. Garg, T. Malyan, A. Bhardwaj, R. R. Shah, B. Krishnamurthy, and C. Chen. Persuasion strategies in advertisements. In *Proceedings of the AAAI conference on artificial intelligence*, vol. 37, pp. 57–66, 2023. 2, 3

[40] X. Lan, Y. Shi, Y. Wu, X. Jiao, and N. Cao. Kineticharts: Augmenting affective expressiveness of charts in data stories with animation design. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):933–943, 2022. 2

[41] X. Lan, Y. Wu, and N. Cao. Affective visualization design: Leveraging the emotional impact of data. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):1–11, 2024. 5

[42] X. Lan, Y. Wu, Y. Shi, Q. Chen, and N. Cao. Negative emotions, positive outcomes? exploring the communication of negativity in serious data stories. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2022. 2

[43] X. Lan, X. Xu, and N. Cao. Understanding narrative linearity for telling expressive time-oriented stories. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2021. 2

[44] J. Li, E. Durmus, and C. Cardie. Exploring the role of argument structure in online debate persuasion. *arXiv preprint arXiv:2010.03538*, 2020. 2

[45] J. Liem, C. Perin, and J. Wood. Structure and empathy in visual data storytelling: Evaluating their influence on attitude. *Computer Graphics Forum*, 39(3):277–289, 2020. 1, 2

[46] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on*

computer vision, pp. 2980–2988, 2017. 7

[47] M. Lisnic, C. Polychronis, A. Lex, and M. Kogan. Misleading beyond visual tricks: How people actually lie with charts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–21, 2023. 9

[48] S. M. Lukin, P. Anand, M. Walker, and S. Whittaker. Argument strength is in the eye of the beholder: Audience effects in persuasion. *arXiv preprint arXiv:1708.09085*, 2017. 2

[49] D. Markant, M. Rogha, A. Karduni, R. Wesslen, and W. Dou. When do data visualizations persuade? the impact of prior attitudes on learning about correlations from scatterplot visualizations. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–16, 2023. 1, 2

[50] M. L. McHugh. Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3):276–282, 2012. 5

[51] Y. Meng, J. Shen, C. Zhang, and J. Han. Weakly-supervised hierarchical text classification. In *Proceedings of the AAAI conference on artificial intelligence*, vol. 33, pp. 6826–6833, 2019. 2, 3, 5

[52] J. Oh, H. S. Lim, J. G. Copple, and E. K. Chadraba. Harnessing the persuasive potential of data: The combinatory effects of data visualization and interactive narratives on obesity perceptions and policy attitudes. *Telematics and informatics*, 35(6):1755–1769, 2018. 5

[53] OpenAI. Text embeddings with ada-002. https://platform.openai.com/docs/guides/embeddings, 2024. Accessed: 2025. 6

[54] A. V. Pandey, A. Manivannan, O. Nov, M. Satterthwaite, and E. Bertini. The persuasive power of data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2211–2220, 2014. 1, 4, 5

[55] R. E. Petty, J. T. Cacioppo, R. E. Petty, and J. T. Cacioppo. *The elaboration likelihood model of persuasion*. Springer, 1986. 1, 2, 3

[56] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, et al. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pp. 8748–8763. PmLR, 2021. 6

[57] M. D. Rahman, G. J. Quadri, and P. Rosen. Exploring annotation strategies in professional visualizations: Insights from prominent us news portals. 2023. 2

[58] A. Ramsälv, M. Ekström, and O. Westlund. The epistemologies of data journalism. *new media & society*, 26(11):6307–6324, 2024. 4

[59] S. Y. Rieh and D. R. Danielson. Credibility: A multidisciplinary framework. *Annual Review of Information Science and Technology*, 41(1):307–364, 2007. 4

[60] M. Rogha, S. Sah, A. Karduni, D. Markant, and W. Dou. The impact of elicitation and contrasting narratives on engagement, recall and attitude change with news articles containing data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 30(7):4375–4389, 2024. 1, 2, 9

[61] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini. The graph neural network model. *IEEE transactions on neural networks*, 20(1):61–80, 2008. 9

[62] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148, 2010. 2, 5

[63] D. Shi, X. Xu, F. Sun, Y. Shi, and N. Cao. Calliope: Automatic visual data story generation from a spreadsheet. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):453–463, 2021. 2

[64] Y. Shi, T. Gao, X. Jiao, and N. Cao. Breaking the fourth wall of data stories through interaction. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):972–982, 2023. 2

[65] Y. Shi, X. Lan, J. Li, Z. Li, and N. Cao. Communicating with motion: A design space for animated visual narratives in data videos. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2021. 2

[66] Y. Shi, Y. Peng, J. Ding, X. Lan, and N. Cao. Double tap for this post: Understanding the communication of data visualization on social media. *Proceedings of the ACM on Human-Computer Interaction*, 9(2):1–27, 2025. 1

[67] H. Simons. *Persuasion: Understanding, Practice, and Analysis*. Addison-Wesley Series in Statistics. Addison-Wesley Publishing Company, 1976. 2

[68] D. Skau, L. Harrison, and R. Kosara. An evaluation of the impact of visual embellishments in bar charts. In *Computer Graphics Forum*, vol. 34, pp. 221–230. Wiley Online Library, 2015. 9

[69] C. Sparks and J. Tulloch. *Tabloid tales: Global debates over media standards*. Rowman & Littlefield, 2000. 5

[70] N. Sultanum, F. Chevalier, Z. Bylinskii, and Z. Liu. Leveraging text-chart links to support authoring of data-driven articles with vizflow. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–17, 2021. 2

[71] N. Sultanum and A. Srinivasan. Datatales: Investigating the use of large language models for authoring data-driven articles. In *2023 IEEE Visualization and Visual Analytics (VIS)*, pp. 231–235. IEEE, 2023. 2

[72] M. Sun, L. Cai, W. Cui, Y. Wu, Y. Shi, and N. Cao. Erato: Cooperative data story editing via fact interpolation. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):983–993, 2023. 2

[73] H. Tajfel and J. C. Turner. Intergroup behavior. *Introducing social psychology*, 401(466):149–178, 1978. 2, 3

[74] M. Tan and Q. Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pp. 6105–6114. PMLR, 2019. 8

[75] The New York Times. The n.b.a. team making 3-point history in the finals isn't golden state. https://www.nytimes.com/interactive/2022/06/16/upshot/nba-finals-golden-state-boston-three-pointers.html, 2022. Accessed: 2025. 4

[76] The New York Times. How big is the legacy boost at elite colleges? https://www.nytimes.com/2023/07/27/upshot/ivy-league-legacy-admissions.html, 2023. Accessed: 2025. 4

[77] The New York Times. It's not just you: Many americans face insurance obstacles over medical care and bills. https://www.nytimes.com/2023/06/15/health/health-insurance-medical-bills.html, 2023. Accessed: 2025. 4

[78] F. Times. Economic losses from extreme weather hit $3.6tn over five decades. https://www.ft.com/content/541702c0-b0ff-434b-8e21-005fa549f4e5, 2021. Accessed: 2025. 4

[79] F. Times. The unexpected surge in inflation, in charts. https://www.ft.com/content/9c4b162a-63d3-44cb-9a47-8a38565b0cae, 2021. Accessed: 2025. 4

[80] F. Times. Us census shows increasing share of non-white groups in wider population. https://www.ft.com/content/1bc9e284-fa6d-4641-8f17-2571bac43d1c, 2021. Accessed: 2025. 4

[81] F. Times. Rising rents mean no shelter for americans from inflation storm. https://www.ft.com/content/afdc756d-a6da-4482-8d9b-d9c22fdc8968, 2022. Accessed: 2025. 4

[82] F. Times. Birth rates in rich countries halve to hit record low. https://www.ft.com/content/f0d2a5a7-e5ef-4044-8380-ff690b609a5a, 2024. Accessed: 2025. 4

[83] T. N. Y. Times. With pandemic money gone, child care is an industry on the brink. https://www.nytimes.com/2024/02/25/upshot/child-care-centers-struggling.html, 2024. Accessed: 2025. 1

[84] A. Tversky and D. Kahneman. The framing of decisions and the psychology of choice. *science*, 211(4481):453–458, 1981. 3, 4

[85] D. Vakratsas and T. Ambler. How advertising works: what do we really know? *Journal of marketing*, 63(1):26–43, 1999. 2

[86] N. Ya-feng, L. Jin, C. Jia-qi, Y. Wen-jun, Z. Hong-rui, H. Jia-xin, X. Lang, W. Jia-hao, M. Guo-rui, H. Zi-jian, et al. Research on visual representation of icon colour in eye-controlled systems. *Advanced Engineering Informatics*, 52:101570, 2022. 9

[87] L. Yang, X. Xu, X. Lan, Z. Liu, S. Guo, Y. Shi, H. Qu, and N. Cao. A design space for applying the freytag's pyramid structure to data stories. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):922–932, 2022. 2

[88] Y. Yuan, F. Xu, H. Cao, G. Zhang, P. Hui, Y. Li, and D. Jin. Persuade to click: Context-aware persuasion model for online textual advertisement. *IEEE Transactions on Knowledge and Data Engineering*, 35(2):1938–1951, 2021. 3

[89] Y. Zhang, Y. Sun, L. Padilla, S. Barua, E. Bertini, and A. G. Parker. Mapping the landscape of covid-19 crisis visualizations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–23, 2021. 1

[90] J. Zhao, M. Glueck, S. Breslav, F. Chevalier, and A. Khan. Annotation graphs: A graph-based visualization for meta-analysis of data based on user-authored annotations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):261–270, 2017. 9

[91] C. Zhu-Tian and H. Xia. Crossdata: Leveraging text-data connections for authoring data documents. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–15, 2022. 2