Formosa Journal of Sustainable Research (FJSR)
Vol.5, No.1, 2026: 15-34

Question Design and Risk of Distortion of Information in the Examination Report (BAP): A Forensic Linguistic–Psycholinguistic Analysis Based on the Distortion Risk Index

Jefriyanto Saud
Universitas Negeri Gorontalo
Corresponding Author: Jefriyanto Saud jefri.saud@ung.ac.id

A R T I C L E I N F O

Keywords: Examination Report, Question Design, Forensic Linguistics, Cognitive Load, Distortion Risk Index

Received : 21 Novembe
Revised : 23 December
Accepted: 23 Januwary

A B S T R A C T

The Examination Report (BAP) is a key document in the judicial process because it records the testimony of the parties being questioned, but the question-and-answer format and evidentiary orientation have the potential to reshape the representation of testimony. This article analyzes the relationship between question design in the BAP and cognitive load and the risk of distortion through a forensic linguistics–psycholinguistics approach. The study applies Directed Qualitative Content Analysis (DQCA) with three components: (1) coding of question types (open, closed, leading, suggestive, confrontational, multi-part), (2) coding of cognitive load indicators (episodic recall, multi-component, quantification, topic switching, linguistic complexity), and (3) measuring the Distortion Risk Index (IRD) based on a score of 0–2 on nine distortion mechanisms (DR1–DR9). The unit of analysis is the investigator's questions (Q) and answers (A) mapped to the ID–EV–KR–PB–PN segments. These findings indicate that the KR contains more open-ended questions, while the PB and PN are dominated by confirmatory/presuppositional questions. Cognitive load increases in KR and PB, while EV/PN is relatively lighter but carries the risk of triggering compliance/commitment. Micro-analysis of Q–A indicates a shift from narrative elicitation to premise-based verification. IRD is offered as a practical, replicable tool to identify distortion hotspots and guide question design improvements

DOI: https://doi.org/10.55927/fjsr.v5i1.827
ISSN-E: 2830-6228
https://mryformosapublisher.org/index.php/fjsr/index

INTRODUCTION

The Examination Report (BAP) occupies a strategic position in the judicial process because it is the primary medium that transforms the interrogated party's oral statements into a written record that is then read, interpreted, and reused by legal actors. In practice, the BAP is not simply a "neutral record," but rather the product of institutional interactions that typically take the form of questions and answers and follow procedural sequences. Traces of this format are evident, for example, in the examination orientation questions that confirm the status of the interrogated party, the basis for the police report, and readiness to provide information, so that from the outset the interaction is framed within institutional logic. In the closing phase, the BAP also contains a formula that leads to a commitment to truth and a signing agreement, which can pragmatically limit the scope for correction after the question-and-answer series has taken place.

The position of the police investigation report (BAP) as a "bridging" document between the police investigation and the subsequent evidentiary stage makes the issue of the quality of the testimony crucial. When the BAP is read by others who were not present at the initial interaction, the text is often treated as the primary reference for "what happened" and "what was acknowledged/stated." In this context, the BAP can be understood as a form of police-generated testimony that is then "retranslated" into a legal assessment. Therefore, a discussion of the gatekeeping function of witness evidence produced by the authorities becomes relevant: the quality and method of production of testimony in the initial stage can influence the weight of evidence in subsequent stages (Thompson, 2012).

As an institutional discourse, the design of questions in an examination tends to be influenced by evidentiary objectives, asymmetrical role relations, and administrative needs. Within the framework of institutional talk , questions not only “request information,” but also direct the type of response expected, regulate turn-taking, and assert the questioner’s position of authority (Heritage & Clayman, 2010). Consequently, the quality of the representation of information in the police report (BAP) can shift: from a free narrative produced by the questioner to a confirmatory response to the formulation offered by the questioner. Such a pattern is apparent when questions include chronological premises or facts that are considered certain, such as the reference “when arrested, it was found…”, followed by demands for explanation that potentially place the questioner in the position of responding to a “pre-established” scenario.

In forensic linguistics, the police report can be read as an institutional text containing traces of language choices, particularly in the way questions are framed, the terms used, and the strategies that encourage particular answers. Language in interrogations and examinations is not merely a medium, but rather part of a mechanism that can influence narrative direction, consistency, and perceptions of credibility (Coulthard et al., 2017). Attention to the “language of questions” is also a prominent theme in contemporary forensic linguistic studies of interrogation, including the emphasis that question formulation can shift the boundary between information elicitation and narrative formation (Blanco, 2024).

In the realm of criminal investigation, the primary challenge is not only finding the facts, but also managing the information production process to avoid being trapped in confirmation bias and early hypothesis narrowing. Discussions of future investigative challenges emphasize that the quality of initial information and the way it is gathered will increasingly determine the success of an investigation, especially when evidence must be tested in a long procedural chain (Tong et al., 2009). At this point, the literature on interview practice and police psychology reminds us that questioning techniques and interaction management have a direct impact on the information output of witnesses, victims, and suspects, so that examinations need to be understood as professional skills that require both procedural precision and communicative sensitivity (Lord & Cowan, 2010; Kitaeff, 2011).

From a cognitive-psycholinguistic perspective, vulnerability to distortion is not simply a matter of communication ethics, but rather related to how memory and language processing work. Classic research on the misinformation effect shows that variations in question wording can influence the estimation and recall of event details; certain question forms can encourage memory reconstruction in line with the information presented (Loftus & Palmer, 1974). In an examination context, when questions contain presuppositions, evidentiary labels, or specific details (numbers, times, sequences), these details can serve as cognitive anchors that guide the direction of answers, especially when the examinee is in a tense situation and confronted with authority.

A key tenet in psycholinguistics is the limited capacity of working memory as the "space" used by a person to comprehend a question, access episodic memory, select relevant details, and then formulate a coherent answer. When questions demand multiple elements at once, processing resources can be consumed by understanding the question's structure rather than accessing stable memory. The working memory framework emphasizes this capacity limitation and its implications for response quality, particularly in tasks that require the integration of details and rapid decision-making under pressure (Baddeley, 2000; Cowan, 2001).

Cognitive load theory adds an explanation for why "inefficient" question structures can result in less stable information output. The load arises not only from the complexity of the material being remembered (intrinsic) but also from the way the task is presented (extra-intrinsic), such as lengthy, multi-layered questions, shifting topics, or forcing assumptions to be accepted before answering (Sweller, 1988; van Merriënboer & Sweller, 2005). In the BAP, this type of load commonly arises when questions combine multiple components (time, place, actor, action, item) or ask for detailed quantification that requires estimation. In practice, these conditions increase the likelihood that answers will be partial, unstable, or compelled to "fill in the gaps," leading to the representation of information in the BAP moving away from the original remembered experience.

The issue of distortion is also intertwined with the literature on suggestibility, compliance, and the risk of false confessions. In interrogative psychology, interrogative suggestibility explains how authority pressure, repetition, and question framing can encourage changes in answers or acceptance

of propositions offered (Gudjonsson, 1992, 2003). The literature on interrogation practice reform also emphasizes the risk factors for false testimony and the need for an approach that is more information-driven than confirmation-driven (Kassin et al., 2010). In this context, studies linking interrogation style and question type to testimony outcomes, including the tendency to "accept the questioner's frame," provide additional basis for assessing linguistic risks in dossiers (Hubert, 2017).

On the other hand, the development of science-based investigative interview methods offers a relevant normative comparison. The general principles of the information-gathering approach emphasizes open-ended questions and free narrative elicitation as strategies to optimize memory access without adding details that respondents have not yet disclosed. Interestingly, recent research on cognitive interviews for suspects suggests that approaches designed to gather better information do not necessarily increase the risk of false confessions, but can also help reduce the impact of questions that elicit misinformation (Noc et al., 2022/2023). These findings strengthen the argument that question design and interaction structure are key variables that can be scientifically intervened in the context of an interrogation.

The vulnerability of testimony production becomes even more complex in interrogation situations involving a second language or an interpreter. Interactions assisted by an interpreter can create tensions between the need for semantic accuracy, procedural fluency, and rapport-building strategies. Studies of interpreter needs in detention centers highlight that the identification of language needs often relies on officer discretion and can be inconsistent, with implications for procedural fairness (Hollands, 2017). Furthermore, studies of rapport features in interpreter-mediated police interviews indicate that pragmatic form (not just content) can shift during the interpretation process, thus affecting the relationship and responses of the interviewee (Gómez Bedoya, 2022). The relevance of this issue can be extended to the context of cognitive interviews, where interpreters present an "inherent conflict" in interlingual operations (Lai, 2016).

While studies on investigative interviews and interrogations have developed, including warnings about the risk factors for misinformation and the need for reform of examination practices, studies that focus on the police report (BAP) as a text with a psycholinguistic-cognitive lens are still relatively limited, particularly those that systematically map questioning patterns across document segments (identity, event orientation, chronology, evidence, and closure) (Kassin et al., 2010). On the other hand, discussions about adapting science-based interview techniques to the Indonesian context have also begun to emerge, making the need for operational analytical tools to read questioning practices in documents increasingly apparent (Muniroh, 2019).

Building on this gap, this article offers two contributions. First, a theoretical contribution: linking question design in the police report (BAP) with indicators of cognitive load (e.g., multi-component, quantification, topic switching) and distortion risk mechanisms (presupposition, third-party narrative, labeling "evidence," anchoring, and closing commitment). Second, a practical contribution: presenting a map of vulnerable points that can be

translated into recommendations for improving question design to make it more communicative, less suggestive, and more responsive to cognitive processing limitations in examination situations.

Methodologically, this study uses directed qualitative content analysis to code investigators' questions based on a predetermined theoretical framework, which is then developed iteratively based on data findings (Hsieh & Shannon, 2005). The analysis is conducted on question (Q) and answer (A) units by segmenting the BAP structure into Identity (ID), Event Orientation (EV), Chronology (KR), Evidence (PB), and Conclusion (PN). To provide more operational evaluative weight, this article introduces the Distortion Risk Index (DRI) as a semi-quantitative measure that summarizes the occurrence of distortion mechanisms per question, while complementing it with a micro-analysis of Q–A excerpts to explain how distortion risks can be produced through interaction sequences. Thus, this article answers two research questions: (RQ1) how question design in BAP segments creates cognitive load and what form of load is dominant; and (RQ2) to what extent question design increases the risk of distortion and how these mechanisms are apparent in the micro-analysis of Q–A excerpts. The article is structured as follows: the next section presents a literature review, followed by the methodology and coding procedures, then the results and discussion (including preliminary/pilot findings where relevant), and closes with conclusions and recommendations.

LITERATURE REVIEW

BAP as an Institutional Discourse Practice (Forensic Linguistics)

Within the framework of forensic linguistics, the police investigation report (BAP) can be understood as an institutional text born of authoritative and evidence-oriented interactions. Rather than simply "recording" speech, the BAP represents the results of institutional filtering: what is asked, how it is asked, the order of topics, and how information is reformulated into administrative-legal language. This perspective aligns with forensic linguistics studies that position language as both "evidence" and a medium that can influence the evidentiary process. In police practice, the goal of proving evidence often demands the condensation of information into "ready-to-prove" propositions (e.g., elements of action, time, place, and the relationship of objects), so the question-and-answer format tends to emphasize verification and confirmation.

The concept of institutional talk helps explain why questions in a police report are never entirely neutral. In institutional conversations, question design shapes the types of answers deemed relevant, directs the flow of interaction, and creates "preferences" for certain answers (e.g., short, confirmatory, or institutionally categorized) (Heritage & Clayman, 2010). Thus, questions serve a dual function: gathering information and simultaneously enforcing an institutional agenda. A key implication is that the quality of the representation of information in a police report depends heavily on the question design, whether the questions open up a free narrative space or encourage confirmatory responses that adhere to the other party's version of events or the questioner's version of evidence. In the police report documents that serve as the context for this research, traces of institutional orientation are evident in questions that explicitly link the investigation to the police report and the alleged crime, placing the interaction within a legal-procedural framework from the outset.

Cognitive Psycholinguistics: Working Memory and Cognitive Load

Cognitive psycholinguistics places working memory limitations as a fundamental factor in language processing and information production. Working memory functions as a "mental space" for holding and manipulating information as one comprehends questions, accesses episodic memory, constructs answers, and adapts answers to the socio-institutional context. Working memory frameworks (e.g., Baddeley's model and subsequent developments) assert that processing capacity is limited; when questions demand too many elements at once, response quality risks deteriorating. In the context of the BAP, this capacity limit is relevant because questions often contain multiple components— time, place, sequence of events, other actors, objects, and motives—that must be integrated into a single response.

Cognitive Load Theory (CLT) provides a conceptual tool for explaining how question design can increase or decrease processing load. CLT distinguishes the load inherent in the complexity of the material (intrinsic load) from the load that arises from the way the task/question is presented (extraneous load). In exploratory interviews, extraneous load can increase when questions are too long, multi-layered, multi-topic, or contain assumptions that must be "accepted" before being answered. Van Merriënboer and Sweller (2005) emphasize that in complex tasks, poorly designed instructions can deplete cognitive resources, thus

impairing core processing performance. Thus, multi-component questions and topic switching are not simply stylistic issues but rather cognitive load triggers that can encourage partial, inconsistent, or anchor-dependent responses (numbers/dates) provided by the interviewer. This framework is also compatible with the practical goal of the study: identifying segments of the BAP (e.g., KR and PB) that tend to contain detailed episodic recall demands and quantification, and thus risk being more cognitively taxing.

Information Distortion: Suggestibility, Leading Questions, and Contamination Mechanisms

The study of information distortion stems from the psycholgical finding that memory is reconstructive and can be influenced by post-event information. The classic study by Loftus and Palmer (1974) demonstrated that variations in question wording can alter the estimation and reporting of details, laying the foundation for the concept of the misinformation effect . In the context of the BAP, leading questions and presuppositions are important because questions can introduce the premise of an "event" or "action" as certain and then ask for confirmation. This impacts not only the accuracy of the details but also the narrative structure: the interviewee responds to the "script" the question conveys, rather than constructing a narrative from their own memory. The literature on leading questions and presuppositions confirms that the form of the question can trigger the acceptance of details that the interviewer might not have independently accessed.

Beyond memory mechanisms, distortion is also related to suggestibility and compliance in interrogation situations. Gudjonsson defines suggestibility as susceptibility to pressure from questions, repetition, and authoritative contexts, which can trigger changes in answers or agreement with propositions offered. At the institutional level, confirmatory pressure, legal framing, and commitment-locking closures ("all statements are true…") can minimize room for correction, even when respondents experience uncertainty. In the dossier documents used in the research context, the practice of closures requiring justification and signature commitments emerged as administrative formulas that potentially reinforce the effect of commitment to the written version. The interrogation literature also emphasizes the risk of contamination when the interrogation process shapes a narrative that then becomes a reference for other evidence; Kassin et al. discuss risk factors and recommendations for minimizing the impact of tactics that can trigger misinformation or false confessions.

"Best Practice" Investigative Interviewing as a Comparison

To assess questioning practices in the police report more normatively, this study used best practices in investigative interviewing as a benchmark. Its primary principle is to prioritize open-ended questions and open recall, followed by a more focused verification phase. Evidence-based investigative interview protocols, such as the NICHD Protocol, consistently emphasize the effectiveness of open-ended prompts in eliciting detail without increasing suggestion, and the literature is widely available. The Achieving Best Evidence (ABE) guidelines also emphasize the importance of non-leading questions, particularly when dealing with vulnerable witnesses/victims. Furthermore, the cognitive interview

approach emphasizes memory elicitation techniques (e.g., reinstatement of context) that aim to enhance memory access without introducing new details.

The relevance of this benchmark for the BAP analysis lies in the two segments that most determine the quality of the testimony: Chronology (KR) and Evidence (PB). In KR, the best standards encourage free narratives and minimal interruptions; in PB, the best standards emphasize careful verification without premature labeling or questions that force confirmation of the other party's script. References in the Indonesian context also indicate that adaptations of science-based interview techniques have begun to be discussed, allowing this study's practical recommendations to be placed within the relevant literature.

Conceptual Framework of the Article

Based on the above literature, this article proposes a conceptual framework: question design → cognitive load → risk of distortion → quality of testimony representation in the dossier. Question design (e.g., open/closed, multi-component, leading, suggestive, confrontational) is treated as a "trigger" that can increase extraneous cognitive load and/or embed narrative premises. Increased cognitive load (e.g., due to lengthy questions, topic switching, or quantification) has the potential to decrease response stability, encourage partial responses, or reinforce the use of anchors (numbers/dates) provided. At the same time, distortion mechanisms (presupposition, external narrative, evidence labeling, confirmation pressure, and commitment closure) increase the likelihood of testimony shifting from an experiential narrative to one that "aligns" with the institutional frame.

To operationalize this framework, the article introduces the Distortion Risk Index (DRI) as a semi-quantitative analytical tool that summarizes the occurrence of distortion mechanisms per question in a transparent and replicable manner, in line with the principle of directed qualitative content analysis that begins coding from theory and is then refined by data (Hsieh & Shannon, 2005). The DRI is not intended as a measure of "factual truth," but rather as an indicator of linguistic-cognitive risks in question design. By combining the DRI and micro-analysis of Q&A excerpts, this article seeks to explain not only "how often" risks occur, but also "how" distortion mechanisms can be produced in institutional Q&A sequences.

METHODOLOGY

Research Design

This study employed a qualitative design with a Directed Qualitative Content Analysis (DQCA) approach. DQCA was chosen because it allows for systematic text analysis through a coding framework derived from theory and previous research findings, while still allowing for categories emerging from the data (Hsieh & Shannon, 2005). In the context of the BAP as an institutional text, DQCA is relevant for mapping question design and information representation practices in a directed manner, rather than simply identifying general themes.

Operationally, the research design incorporates three analytical components. First, cognitive load coding to characterize processing load triggers in question design (e.g., multi-component, quantification, topic switching). This framework is informed by cognitive load theory, which emphasizes the limitations of processing resources, particularly when complex tasks are presented in ways that increase extraneous load (Sweller, 1988; van Merriënboer & Sweller, 2005). Second, the study developed the Distortion Risk Index (DRI) as a code-based, semi-quantitative measure that summarizes the occurrence of distortion mechanisms within each question. Third, micro-analysis of Q&A excerpts was conducted to examine in detail how distortion risks were produced and negotiated through the question-answer sequence, in line with conversation analysis approaches to institutional talk that emphasize the role of question design in shaping answer form (Heritage & Clayman, 2010).

Data and Corpus

The research data sources are BAP documents from several types of cases, including (but not limited to) embezzlement, narcotics, and other question-and-answer format examination/interrogation files. For this article, the analysis is reported as preliminary findings based on a pilot subset, namely: (1) one embezzlement BAP document that was analyzed in full, and (2) an initial portion of a large narcotics BAP compilation. This pilot approach was used to test the feasibility of the codebook, the consistency of the IRD scoring rules, and the micro-analysis excerpt selection strategy before full application to the entire corpus.

Ethical considerations are applied through anonymization: names, addresses, ID numbers, and other personal information are disguised when excerpts from questions or answers are displayed in the article. This practice aligns with the principle of prudence in document research involving sensitive information, particularly in the context of institutional data. Furthermore, excerpts are used selectively and proportionately for the purposes of linguistic-cognitive analysis (not for assessing the case material).

Units of Analysis and Segmentation

The primary unit of analysis is a single investigator question (Q), while the accompanying unit is the answer (A) that follows that question. Q–A pairs are treated as micro-units to assess the effect of question design on response form and stability. All Q–A sections are then mapped to the following segmentation of the BAP structure: Identity (ID), Event Orientation (EV), Chronology (KR), Evidence (PB), and Closing (PN). This segmentation allows for comparisons across examination functions—for example, whether the PB segment is more

focused on confirmation and labeling of evidence than the ideally narrative-oriented KR.

Codebook and Operational Definitions

The codebook was developed in a targeted manner with three layers of codes. Layer 1 classifies question types: QO (open-ended), QC (closed-ended), QL (leading), QS (suggestive), QK (confrontational), and QM (multi-part). This classification serves as a gateway to identifying question designs that tend to open up free narratives versus encourage confirmation.

Layer 2 labels indicators of cognitive load (CL1–CL6), which include: detailed episodic recall, multi-component, quantification/number, topic switching, inference of intention/motive, and linguistic complexity. This coding is informed by the literature on cognitive load and processing limitations, which emphasizes that task/question structure can increase extraneous load and interfere with memory access and language production (Sweller, 1988; Paas et al., 2010).

Layer 3 identifies mechanisms of risk distortion (DR1–DR9), including: presupposition/assumption of guilt, external narrative proffering, early evidentiary labeling (e.g., "evidence"), number/time anchoring, forced choice, confirmatory pressure, authority/compliance pressure, closing commitment, and confrontation. This framework is supported by findings in memory psychology on the influence of question wording on memory reconstruction (Loftus & Palmer, 1974) and the suggestibility literature in the context of examinations.

IRD Scoring (0–2 Rule and Summation)

IRD is calculated per question by assigning a score of 0–2 to each distortion mechanism (DR1–DR9), then summing them: IRD(Q) = Σ skor(DR1…DR9). A score of 0 indicates the absence of the indicator, a score of 1 indicates its mild/implicit occurrence, and a score of 2 indicates its strong/explicit or repeated occurrence within a single question. For descriptive interpretation, IRD is categorized as low, medium, high, and very high. It is important to emphasize that IRD is not intended as a measure of the "factual correctness" of the information, but rather as an indicator of linguistic-cognitive risks in question design that could potentially affect the quality of information representation in the BAP. To complement the semi-quantitative summary, micro-analysis was conducted on Q–A excerpts with high IRD and on examples with low IRD as a comparison, allowing the article to explain the distortion mechanisms interactionally, rather than simply in terms of frequency.

RESULTS AND DISCUSSION

Results: Distribution of Question Types Per Segment

Based on pilot coding on a subset of the corpus (N = 117 questions), the distribution of question types revealed a layered pattern. Overall, open-ended (QO) and closed-ended (QC) questions co-emerged as the "two main engines" of inquiry. In the analyzed subset, QO was identified in 60 questions, while QC in 39; the remainder were QS (10) and QL (8). However, a key finding was the emergence of a mixed pattern: questions that grammatically begin as confirmation (e.g., "do you…") but end with "explain/describe." At the functional level of institutional communication, this pattern tends to remain confirmatory, as answers are asked to confirm a previously stated premise (Heritage & Clayman, 2010).

In terms of segments, the Chronology (KR) most frequently generated QOs (14 of 17 KR questions), while the Closing (PN) was almost entirely dominated by QCs (10 of 10). This pattern is functionally consistent: KR requires a narrative of events, while PN tends to require administrative confirmation. The example of a closing requesting approval of the overall truth and a signature appears to be a formula that closes off room for revision after the question-and-answer cycle has progressed.

Results: Dominant Cognitive Load Per Segment

In the KR segment, the most prominent indicators of cognitive load are CL1 (detailed episodic recall) and CL2 (multi-component). Chronological questions often require the integration of multiple elements simultaneously (time–place–action–actor–object), thereby increasing the demands on working memory. In working memory and processing theory, the more units of information that must be maintained and integrated in a single response, the higher the risk of partial or "gap-filling" responses (Baddeley, 2000; Cowan, 2001).

In the PB segment, cognitive load is more often triggered by CL3 (quantification/numbers) and CL6 (complexity/terms/structure). PB usually combines numerical details (amounts, specific times, document/item identification) and procedural terms, so that respondents' task is not only to remember, but also to calculate, estimate, or link details to categories of evidence. This condition is in line with the logic of Cognitive Load Theory: question designs that are long, layered, or contain details that respondents have not independently accessed tend to increase extraneous load (Sweller, 1988; van Merriënboer & Sweller, 2005).

The EV and PN segments are relatively "lighter" in terms of chronological recall requirements, but they pose different risks, namely compliance/commitment. For example, in the EV, questions link the investigation to the police report and request a willingness and promise to provide "truthful" information. In the PN, questions "lock in" approval of the written statement and the signing procedure.

Results: IRD Per Segment (Pilot)

In the analyzed subset (preliminary findings), the IRD statistics showed a mean of 0.97, a median of 0, and a maximum of 5. Segmentwise, PN had the highest mean IRD (mean of 2.00; median of 2; max of 3), followed by PB (mean

of 1.02; max of 4) and KR (mean of 0.88; max of 3). This pattern indicates that in this subset, the distortion risk mechanisms are more concentrated in the proof and closing phases, while in KR the risks arise mainly when chronology questions are accompanied by time/number anchors or multi-component demands.

It's important to emphasize that these figures represent a pilot subset and cannot be generalized to the entire BAP corpus. The purpose of IRD reporting at this stage is to indicate where risks tend to cluster and to guide the selection of samples for micro-analysis.

Results: Top High-Risk Questions (Based on IRD)

In the analyzed subset, questions with high IRD tended to follow the following 3–4 dominant patterns.

1. **The Other Party's Narrative Is Presented (DR2) → Confirmation (DR6)**

This pattern occurs when the question introduces an external version (e.g., the reporter/witness) as a "script," then asks the respondent to confirm or explain. Theoretically, this strategy intersects with the misinformation literature because the presented details can displace access to independent memory (Loftus & Palmer, 1974).

2. **Labeling Evidence (DR3) → Pseudo-Verification**

In PB, questions that begin with the label "evidence" have the potential to lead respondents to a framework of evidence before neutral verification is conducted, so that answers tend to "follow the label" rather than freely describing the context of acquisition/knowledge.

3. **Anchoring Dates/Times/Numbers (DR4) as Memory "Anchors."**

An example of anchoring is seen in questions that specify a time/event ("when arrested, found…") and ask for further explanation. Within the framework of cognitive heuristics, this type of anchor can guide the direction of answers, especially when respondents are in stressful institutional situations.

4. **The Closing that Seals the Commitment (DR8)**

In the District Court, the formula "all information is true... willing to sign" reinforces the commitment to the established text. From a suggestibility and compliance perspective, the closing section has the potential to be a point where final corrections are difficult due to the norms of "completing the examination" and "verifying the documents."

Micro-Analysis Q–A (Representative Excerpt)

This section presents a brief micro-analysis: two risky snapshots (relatively high IRD) and one comparison snapshot (safer practice). Excerpt 1 (EV; compliance/authority + administrative anchoring).

Orientation questions that link the examination to the police report and request a willingness to provide "truthful" information demonstrate an institutional mechanism that encourages compliance from the outset. Interactionally, this question design directs responses toward consent ("willing"), rather than narrative. Consequently, respondents are placed in a position to accept the framework of the examination —a context that can influence how they construct their responses in subsequent segments (Heritage & Clayman, 2010).
Excerpt 2 (KR/PB; anchoring the facts of the event).

Questions that establish the premise of an event ("when arrested, it was found…") and then ask for an explanation demonstrate the point at which the presupposition of the event/finding can serve as an anchor. From a memory psychology perspective, the details presented in the question have the potential to guide the reconstruction of the answer, especially if the respondent does not access those details independently (Loftus & Palmer, 1974). At the text level, the "premise → explain" format shifts the question function from elicitation (inviting a narrative) to verification (testing conformity with the premise).

Excerpt 3 (PN; commitment closure).

In closing, the formula of overall confirmation and signing is a common administrative practice, but pragmatically, it can lock the document. The main risk is not the details of the event, but rather the "finalization effect": once the respondent confirms the truth and signs, the opportunity for substantive correction decreases. In the suggestibility literature, closings that emphasize commitment and compliance can reinforce the tendency to "accept" in order to complete the interaction.

As a comparison to safer practices (low IRD), a subset also found prompts that were closer to elicitation, such as inviting participants to explain a situation or recount a sequence of events without first offering the other party's script. This type of prompt is more in line with the principle of open recall in investigative interview guidelines.

Discussion: Theoretical Interpretation

The initial (pilot) findings in this study illustrate that the Dossier, as a document produced through examination practices, can be understood not merely as a "record" of statements, but as a product of institutional discourse shaped by evidentiary objectives, formal procedures, and asymmetrical role relations. In forensic linguistics, the Dossier's position as an institutional text is important because the language within it not only records "what is said," but also represents "how the statement is produced" through question design, topic sequence, and word choice that direct the status of information as relevant or irrelevant to the legal process. The forensic linguistics framework emphasizes that language can function as both evidence and as a mechanism that influences the interpretation of evidence (Coulthard, Johnson, & Wright, 2017). Thus, the focus of this study on question design goes beyond formal classification (open/closed), but also assesses how such design potentially affects the quality of statement representation.

From an institutional talk perspective , question design also serves to control interaction: determining the preferred form of response, expediting the flow of the examination, and directing the narrative toward a direction compatible with institutional categories (e.g., elements of conduct and evidence) (Heritage & Clayman, 2010). This is evident in a fairly consistent pattern across documents: the orientation and closing sections tend to be filled with questions that demand agreement and confirmation. In one of the police investigation reports, for example, the closing contains a formula that directs the examinee to declare that all written information is true and to be willing to sign. Pragmatically, this form functions as closure that confirms the finality of the

document, but from an interactional perspective, it can reduce the space for substantive correction after a lengthy question-and-answer process, especially if there has been previous misframing or questions that narrow the narrative options.

The link between institutional practices and cognitive processes becomes clearer when we examine the chronology (CR) and evidentiary (PB) segments. Pilot findings indicate that KR more often elicits open-ended questions, but chronological questions remain vulnerable when framed as multi-component questions or combined with time/numerical anchors. In one narcotics document, a question that establishes a situation ("when arrested, it was found…") and then demands an explanation demonstrates how a question not only requests a narrative but also establishes a premise that frames the answer. At the psycholinguistic-cognitive level, this type of question design increases processing demands because respondents must (i) understand the premise, (ii) access episodic memory, (iii) assess the memory's compatibility with the premise, and (iv) construct a socially safe answer (in the context of authority). This situation aligns with two sets of theories: working memory limitations and cognitive load.

First, working memory theory asserts that human processing capacity is limited; the more elements that must be maintained and integrated in a single response, the greater the risk that the answer will be partial, unstable, or dependent on external cues (Baddeley, 2000; Cowan, 2001). In BAP practice, questions that demand "time, place, with whom, and what was done—what was found" in a single questioning round increase the integration burden. When this burden exceeds capacity, respondents may choose compensatory strategies: answering briefly, following implied options, or filling in gaps with inference (which in the literature is often understood as reconstruction, not reproduction, of memory). It is at this point that psycholinguistics provides a bridge: the form of the question (syntactic structure, clause length, topic switching) influences processing costs , which are then reflected in the quality of the answer.

Second, Cognitive Load Theory (CLT) helps explain why certain questions are more cognitively "risky." CLT emphasizes that extraneous load increases when tasks are presented in inefficient ways, such as by being long, multi-layered, shifting topics, or forcing respondents to accept assumptions before answering (Sweller, 1988; van Merriënboer & Sweller, 2005). In the BAP, PB situations often increase extraneous load due to verification-oriented questions, quantification, and "evidence" framing (e.g., mentioning "evidence") that demand rapid confirmation. In the pilot, PB and PN tended to show higher distortion risk indices than EV/ID, indicating that the institutional function of "ensure and lock" is intertwined with increased processing load.

The pilot's findings also strongly resonate with the literature on misinformation and suggestibility. Loftus and Palmer (1974) demonstrated that specific wording/phrasing in questions can alter estimates and memory; this finding is relevant for understanding how the "script" presented in questions can guide responses, particularly when respondents are in formal, pressurized situations. In the context of an audit, the risk stems not from a single word, but from the structure of the question, which incorporates the other party's narrative

(DR2) and prompts for confirmation (DR6). When questions begin with the "reporter's version" and then ask for justification, respondents are encouraged to respond to a "pre-made" narrative, rather than constructing one from their own memory. This mechanism can trigger response conformity : answers move toward the questioner's framework because it is socially and institutionally safer.

Meanwhile, the concept of interrogative suggestibility emphasizes respondents' vulnerability to pressure from questions, repetition, and authoritative contexts that encourage compliance or change in answers (Gudjonsson, 1992/2003). In dossiers, suggestibility often appears not as an explicit threat, but as normative and procedural pressure: requesting consent, emphasizing "the truth," confirming "true/not," and closing with a commitment that all statements are true. The interrogation literature also emphasizes that overly confirmatory or directive practices can increase the risk of misrepresentation and, in extreme cases, false confessions; therefore, investigative reform recommendations emphasize reducing pressure tactics and increasing evidence-based elicitation techniques (Kassin et al., 2010).

At this point, the study of investigative interviews provides a benchmark that helps assess whether questions in a dossier are closer to elicitation or confirmation. The NICHD protocol and ABE guidelines, for example, consistently emphasize prioritizing open-ended questions to elicit a free narrative, followed by verification with more focused questions, while avoiding leading questions or those containing details that the respondent has not yet disclosed. When the dossier immediately switches to premise-based verification (e.g., "when arrested, it was found…") or labeling "evidence," the process moves from a memory-based investigative interviewing paradigm to an evidentiary confirmation paradigm, which is cognitively more susceptible to distortion. Within the framework of this article, IRD is designed to bridge these two traditions: (i) the forensic linguistics tradition, which examines question design and institutional discourse, and (ii) the cognitive-psycholinguistic tradition, which explains why certain designs increase processing load and suggestibility.

In summary, the findings of this pilot provide an initial picture that the most obvious points of vulnerability for distortion are concentrated in PB (verification, quantification, evidence labeling) and PN (finalization/commitment), while KR is vulnerable primarily due to detail overload and the use of premise anchors. By combining DQCA, cognitive load coding, IRD, and Q–A micro-analysis, this article not only shows "where" risks arise, but also "how" they can be produced through language practices typical of institutional audits.

CONCLUSION AND RECOMMENDATION

Based on preliminary findings (pilot subset), this study shows that the question-and-answer practices in the BAP exhibit a mixed pattern: some questions are open-ended and invite narratives, but many questions still function as confirmator questions or contain presuppositions that frame the answers from the outset. This mixed pattern is important because it can pragmatically shift information from narrative experiences to responses that adhere to the questioner's framework, especially when questions present specific premises or details before respondents can construct their own version. This finding aligns with the character of institutional conversations, where question design helps shape the types of answers deemed relevant and "appropriate" for institutional purposes.

From a psycholinguistic-cognitive perspective, processing load appears to increase in the Chronology (KR) and Proof (PB) segments. In KR, the load is primarily triggered by episodic recall demands and multi-component questions (time–place–actor–action–object) within a single questioning round. In PB, the load increases through quantification/numbers and the complexity of terms and verification-oriented question structures. Within the cognitive load framework, these conditions have the potential to increase extraneous load and encourage partial or reconstructive responses, especially when questions are too long, layered, or shift topics.

Beyond cognitive load, the most prominent distortion risks in the analyzed subset emerged through four main mechanisms: DR2 (the other party's narrative being presented), DR3 (premature labeling of evidence as "evidence"), DR4 (time/number anchoring), and DR8 (closing that locks in commitments). These four mechanisms do not necessarily prove that the statement is "false," but they do point to points in question design that have the potential to linguistically and cognitively weaken the quality of the statement's representation in the BAP.

Recommendations derived from the findings of this study focus on improving question design and the order of the examination to better support accurate and minimally suggestive information elicitation. In the chronology (CR) segment, the recommended practice is to prioritize free narrative through open-ended questions and the principle of one question–one task (one cognitive task per question), so that the burden of integrating details does not accumulate in a single response. In the initial stages of the substantive examination, openings that immediately rely on the formula “based on the reporter’s statement…” should be avoided, as this pattern risks shifting the statement into a response to the other party’s script; instead, the elicitation phase (exploring the subject’s version) should be clearly separated from the verification phase (examining details). In the evidentiary (PB) segment, early labeling also needs to be neutralized, for example by using the phrase “found items” before the term “evidence” to prevent premature framing of the evidence. In addition, anchoring Initial numbers/dates should be limited; respondents' versions are prioritized first, followed by gradual verification of numbers/times. At the conclusion, explicit corrections should be provided (e.g., "Are there any areas you would like to improve/add?") before final commitment, including overall approval and signature.

FUTHER STUDY

For further study development, further research is recommended to apply full-corpus IRD to all BAP documents in the corpus, test the consistency of coding through intercoder agreement on some of the data, and conduct comparisons across case types and across agencies so that the patterns found can be tested for stability. In addition, triangulation with interview recordings (if ethically and access-wise possible) is important to capture interactional phenomena that are often not reflected in the BAP text such as repair (improvement/clarification), reformulation, interruption, or negotiation of meaning so that the distortion mechanism model can be explained more comprehensively, not only from the results of text coding, but also from the dynamics of information production in the examination interaction.

REFERENCES

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory. Trends in Cognitive Sciences, 4(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Blanco, L. I. (2024). Forensic linguistics: The importance of language in interrogations (Trabajo de Fin de Grado, Universidad de Salamanca). GREDOS. http://hdl.handle.net/10366/160331

Coulthard, M., Johnson, A., & Wright, D. (2017). An introduction to forensic linguistics: Language in evidence (2nd ed.). Routledge.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Behavioral and Brain Sciences, 24(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Gómez Bedoya, M. (2022). Recognition and expression of rapport features in interpreter-mediated police interviews with victims (Doctoral thesis, University of East Anglia). UEA Digital Repository. https://ueaeprints.uea.ac.uk/id/eprint/90202/

Gudjonsson, G. H. (2003). The psychology of interrogations and confessions: A handbook. John Wiley & Sons.

Heritage, J., & Clayman, S. (2010). Talk in action: Interactions, identities, and institutions. Wiley-Blackwell.

Hollands, M. J. (2017). Procedural justice at the custody desk: Exploring interpreter need identification (Master's thesis, Canterbury Christ Church University). CCCU Research Space Repository. https://repository.canterbury.ac.uk/item/88qq9/procedural-justice-at-the-custody-desk-exploring-interpreter-need-identification

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. Qualitative Health Research, 15(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Hubert, J. M. (2017). False confessions: A linguistic analysis of police interrogation styles and question types (Doctoral dissertation). ProQuest Dissertations & Theses Global.

Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions: Risk factors and recommendations. Law and Human Behavior, 34(1), 3–38. https://doi.org/10.1007/s10979-009-9188-6

Kitaeff, J. (Ed.). (2011). Handbook of police psychology. Routledge. https://doi.org/10.4324/9780203836170

Lai, M. (2016). Police cognitive interviews conducted through interpreters—An experimental study of the inherent conflicts in interlingual operations (Doctoral dissertation, RMIT University). RMIT Research Repository. https://research-repository.rmit.edu.au/articles/thesis/Police_cognitive_interviews_con

ducted_through_interpreters-An_experimental_study_of_the_inherent_conflicts_in_interlingual_operations/27589755

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. Journal of Verbal Learning and Verbal Behavior, 13(5), 585–589. https://doi.org/10.1016/S0022-5371(74)80011-3

Lord, V. B., & Cowan, A. D. (2010). Interviewing in criminal justice: Victims, witnesses, clients, and suspects. Anderson Publishing.

Muniroh, R. D. D. (2019). "It's better to see a tiger than a police officer": Adapting the cognitive interviewing technique to the Indonesian policing context (Doctoral dissertation, RMIT University). RMIT Research Repository. https://research-repository.rmit.edu.au/articles/thesis/_It_s_better_to_see_a_tiger_than_a_police_officer_Adapting_the_cognitive_interviewing_technique_to_the_Indonesian_policing_context/162708

Noc, M., Ginet, M., & Deslauriers-Varin, N. (2023). False confession in innocent suspects: A look at the cognitive interview for suspects. Journal of Police and Criminal Psychology, 38(1), 186–198. https://doi.org/10.1007/s11896-022-09543-5

Paas, F., van Gog, T., & Sweller, J. (2010). Cognitive load theory: New conceptualizations, specifications, and integrated research perspectives. Educational Psychology Review, 22, 115–121. https://doi.org/10.1007/s10648-010-9133-8

Stott, C., Bradford, B., Radburn, M., & Savigar-Shaw, L. (Eds.). (2021). Making an impact on policing and crime: Psychological research, policy and practice. Routledge. https://doi.org/10.4324/9780429326592

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Thompson, S. G. (2012). Judicial gatekeeping of police-generated witness testimony. Journal of Criminal Law and Criminology, 102(2), 329–389.

Tong, S., Bryant, R., & Horvath, M. A. H. (2009). Future challenges in criminal investigation. In S. Tong, R. Bryant, & M. A. H. Horvath (Eds.), Understanding criminal investigation (pp. 217–222). Wiley-Blackwell.

van Gog, T., Paas, F., & Sweller, J. (2010). Cognitive load theory: Advances in research on worked examples, animations, and cognitive load measurement. Educational Psychology Review, 22(4), 375–378. https://doi.org/10.1007/s10648-010-9145-4

van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. Educational Psychology Review, 17, 147–177. https://doi.org/10.1007/s10648-005-3951-0