Supplemental material for "Functional equivalence in cross-context text adaptation: an analyze–generate–verify approach with large language models"

Cheng-En Tsai, Fanfan Chen, and Jane Yung-jen Hsu

This supplement provides rating anchors, prompts, and a worked analyzer example. The prompts are English translations of the Chinese instructions used in the experiments. Blocks labeled conceptual explain the architecture; they were not executed as separate calls.

Section	Contents
S1	Effect-signature anchors.
S2	Executed analyzer instruction and an eight-stage conceptual example.
S3	Generator prompts (Experiment A).
S4	Verifier and repair prompts (Experiment A).
S5	Blind judge prompt (Experiment A).
S6	Perturbation prompts (Experiment B).

S1 Effect-signature anchors

Table S1 summarizes the 1–5 ordinal anchors used for the experiment specifications. NA means not applicable; a blank or UNKNOWN means insufficient evidence. Dimensions are rated separately, never summed or averaged.

S2 Analyzer prompts

The example source is an English one-liner chosen for readability: "My doctor told me that jogging could add years to my life. He was right. I feel ten years older already." It is a widely circulated one-liner of unknown authorship, quoted only to illustrate the prompt format. Ratings in the worked example are illustrative AI outputs, not expert labels. Square-bracketed items are input placeholders.

S2.1 Executed analyzer prompt (one call per record)

Each request contained the instruction below, the annotation manual, the JSON schema, and one source text.

English translation of the executed prompt.

> You are a text analyzer. Analyze one complete joke according to annotation manual v3.2, using only the source text and the explicitly listed background; do not use human answers.
> 1. Split the text into atomic propositions P1..., preserving speaker, negation, conditions, time,

Table S1: Anchors for levels 1, 3, and 5 (levels 2 and 4 are intermediate).

Dim.	1	3	5
A	No second reading, or readings synonymous.	A local role, reason, or intention changes; frame unchanged.	Core reference, identity, or the whole scene is reinterpreted.
L	A key link is contradicted or clearly far-fetched.	Partly holds; one important gap.	Every step supported by text, form, meaning, common knowledge, or confirmed dialect.
S	Almost no supporting cue, or cues point the other way.	Relevant cues, but a main competing reading is not excluded.	The key reading is explicit or jointly established by several clear cues.
N	Unresolvable contradiction in core relations.	Readable with an important order or premise gap.	Setup, answer or action, and outcome are consistent in the text world.
B	No effect-relevant knowledge or belief difference.	A difference supports part of the effect.	The difference directly drives the action, misjudgment, or reveal.
M	A needed reason is absent or contradicts known states.	Partial support; a key motive or constraint is missing.	The text states the state, goal, or belief that licenses the action.
G	Almost no discernible emotional change.	A discernible moderate change.	A strong, sharp emotional contrast or build-up is clearly established.

and reference. Link a question's action and its content under the same ID. A request must not become a completed fact. 2. Build stable constants and ground literals. Keep the scope of a character's speech or belief. Record separately what the text states explicitly, character speech or belief, background rules, and reader inference. A template containing variables must not be treated as a ground literal. 3. H1 is the natural understanding before disclosure; H2 is the understanding after it, naming the character or reader. If there is no H2, state explicitly that there is no second candidate reading; do not force a reversal onto every text. 4. Mark the ambiguity type, supporting cues, speaker and belief attribution, motivation authorization for the key action, the necessary association path, the background domain, the main mechanism, and recoverability. A path may include near-sound, form, material, source, function, role, or reference links; attach evidence to each relation. 5. A rates the difference between readings. L rates support for the path. S rates textual cues, not surprise or position. N rates coherence of events or question-answer relations; short length, riddles, or absurd settings are not penalized. B rates the effect of knowledge or belief differences. M rates the reason given for the specified key action. G rates the emotional shift or build-up the text establishes, not how funny or harmful it is. Score each item on the manual's 1–5 anchors; do not total or average. 6. A may be 1 with no second reading; B and G may be 1 with no knowledge difference or emotional change. Use NA for L when no path is needed, and for M when no special action authorization is needed. Use null, with a reason, for unfamiliar dialect or insufficient evidence; do not use a low score to mean unknown. 7. Do not output a separate event or causal-chain field, irreversibility, or normative salience. Output only JSON conforming to the schema, with brief evidence reasons; hidden reasoning is not required. Additional input: the full manual v3.2, the schema, and the source text of this joke.

S2.2 Conceptual stage prompts

The following blocks unpack the operations requested together in one call. Each gives a purpose, input/output, conceptual instruction, and illustrative output. The examples are analyst reconstructions, not validated annotations.

S2.2.1 Stage 1: Atomic propositions

Purpose. Establish the smallest truth-evaluable units while keeping who says, believes, or infers each proposition.

Input / output. Input: source text. Output: proposition IDs, polarity, scope, attribution, source spans, and unresolved ambiguities.

> Split [SOURCE TEXT] into the smallest independently truth-evaluable propositions. Preserve polarity, tense, quantity, modality, speaker, quotation scope, and source span. Supply omitted subjects only when licensed by context. Keep parallel propositions for each plausible pun or ambiguous reading. Attribute each proposition using the schema categories textual, speech, belief, background, or inference; do not promote a character's statement or belief to an unqualified fact. Return proposition IDs and unresolved cases.

Output for the worked example. P1 speaker reports doctor said jogging could add years [speech]; P2 speaker agrees doctor was right [evaluation]; P3 speaker feels ten years older [reported sensation, hyperbole].

S2.2.2 Stage 2: Entities and events

Purpose. Give propositions stable referents and place events and states in an ordered narrative.

Input / output. Input: validated propositions. Output: entity, event, state, and time IDs with types, participants, aliases, and temporal links.

> Using the validated propositions, identify entities, events, states, and time expressions. Assign stable IDs; record aliases, semantic types, participants, and temporal order. Add no event without textual support or a clearly marked inference. Link every item to its supporting proposition and span.

Output for the worked example. E1 speaker, E2 doctor, E3 jogging; V1 reported advice, V2 the speaker's agreement, and V3 the reported sensation. The discourse presents V1 before V2 and V3; it does not establish when the sensation began. Constants: C1 added years, C2 ten years, S1 perceived age or fatigue.

S2.2.3 Stage 3: Ground literals

Purpose. Formalize proposition content in a reusable predicate vocabulary.

Input / output. Input: propositions and entity/event inventory. Output: typed ground literals, new predicate definitions when needed, evidence links.

> Represent each proposition with typed ground literals using the current vocabulary where possible. Propose a new predicate only when the existing vocabulary cannot express the proposition. Preserve negation, attribution, modality, and argument roles. Return the literal, any new predicate definition, proposition ID, evidence span, and confidence status.

Output for the worked example. One possible reified representation uses the claim constant CLAIM1: Reports(E1,CLAIM1); AttributedTo(CLAIM1,E2); Activity(CLAIM1,E3); Modality(CLAIM1,POSSIBLE); Outcome(CLAIM1,LIFE_EXTENSION) [P1]; Evaluates(E1,CLAIM1,RIGHT) [P2]; ReportsSensation(E1,OLDER_BY_TEN_YEARS) [P3]. The outcome and sensation are constants describing attributed content, not assertions of an objective lifespan or age change.

S2.2.4 Stage 4: Rules and knowledge

Purpose. Identify only the background expectations that license the text's inferences, and record who holds them.

Input / output. Input: propositions, literals, narrative inventory. Output: explicit facts, defeasible rules, goals, beliefs, knowledge holders, licensed inferences.

> Identify the background rules and defeasible expectations needed to explain the text. Separate textual facts from commonsense assumptions. Record which character or reader holds each fact, rule, goal, and belief, and identify the inference each rule licenses. Flag unsupported or culturally variable assumptions for review.

Output for the worked example. K1 jogging is conventionally associated with longevity; K2 exercise can cause short-term fatigue. Licensed inference: the sensation recasts the promised benefit as an immediate cost.

S2.2.5 Stage 5: Effect mechanism

Purpose. Explain how the text guides an initial reading and what evidence triggers a revised interpretation.

Input / output. Input: source plus validated outputs of Stages 1–4. Output: H1 and H2, the cue for the shift, the motivation target, evidence links.

> Analyze [SOURCE TEXT] using the validated propositions, literals, and background knowledge. State the dominant initial interpretation, the alternative or revised interpretation, and the textual cue that supports the shift. Identify any character motivation or knowledge difference that is necessary to interpret the response. Cite the relevant evidence and distinguish textual statements from inferences. Do not add a standalone event-causal-chain annotation.

Output for the worked example. H1: jogging promises longevity (P1). Cue: 'I feel ten years older already' (P3) after agreement (P2). H2: the apparent confirmation becomes an ironic complaint about immediate cost. Association path: jogging to added years to feeling older.

S2.2.6 Stage 6: Effect signature

Purpose. Record the source's mechanism-relevant features using the shared annotation anchors.

Input / output. Input: validated source specification and the frozen anchors of Section S1. Output: separate 1–5 ratings for A, L, S, N, B, M, G; NA or UNKNOWN when appropriate.

> Assign the seven applicable source-rating fields A, L, S, N, B, M, and G using [V3.2 ANNOTATION ANCHORS]. Give an evidence span and brief anchor-based justification for every value. Use integer ratings from 1 to 5; use NA when a dimension does not apply and UNKNOWN or a blank/null value when evidence is insufficient. Do not create separate scores for event causality, irreversibility, or normative salience, and do not calculate a total score.

Output for the worked example. A=4, L=3, S=4, N=4, B=3, M=3, G=3. Provisional AI ratings with evidence; not gold labels; not summed.

S2.2.7 Stage 7: Realization constraints

Purpose. Separate mechanism-critical conditions from source-specific surface details before adaptation.

Input / output. Input: frozen source specification. Output: hard invariants, adaptable features, target-independent role functions, ending constraints.

> Convert [SOURCE SPECIFICATION] into target-independent constraints on role functions, required events and relations, character motivation, information order, association path, emotional cues, voice, rhythm, and ending. Separate hard invariants from adaptable surface features. Identify what the target must leave inferable and what it must not explain after the final cue.

Output for the worked example. Hard invariants: an authority presents jogging as a route to longer life; the speaker first appears to affirm it; a final subjective-age cue reverses it. Adaptable: roles, activity, exact number. Do not add an explanation after the punchline.

S2.2.8 Stage 8: Cross-stage audit

Purpose. Detect omissions, unsupported additions, and contradictions before the source specification is frozen.

Input / output. Input: source, Stage 1–7 outputs, schema. Output: schema validity and localized unsupported or inconsistent fields.

> Audit [ANALYZER OUTPUT] against [SOURCE TEXT] and [OUTPUT SCHEMA]. Check proposition coverage, source-span support, entity and literal references, attribution, polarity, H1/H2 support, motivation evidence, association-path evidence, rating anchors, and required fields. List unsupported or inconsistent claims with IDs and evidence. Return PASS only when schema and evidence requirements are met; otherwise return REVISE or UNKNOWN with localized corrections. Do not require a separate event-causal-chain, irreversibility, or normative-salience field.

Output for the worked example. PASS on coverage, attribution, and H1/H2 support, as an interpretive hypothesis. The claim that jogging adds years is attributed, not treated as fact. Ratings remain provisional.

S3 Generator prompts (Experiment A)

The per-condition instructions in Section S3.2 were executed. All conditions receive the source and target card; B1–B5 also receive the specification in structured or prose form. B0-BoN adds sampling and selection calls. The B3 critique is included below; verification and repair for B4 and B5 are in Section S4.

S3.1 Conceptual generator prompt

This block summarizes the generator in the main article, Section 4.2.

> Using [FROZEN SOURCE SPECIFICATION] and [TARGET CONTEXT], propose a role-and-event mapping and check that each mapped action, institution, object, time scale, and character motivation is plausible. Generate [N] natural candidates that preserve the hard invariants, critical relations, interpretation sequence, and required effect levels. Match the requested voice and register. End at the earliest point where the intended reinterpretation is recoverable. Do not append an explanation of the mechanism. If no plausible mapping preserves the required conditions, return ABSTAIN and name the unmet condition.

S3.2 Executed per-condition instructions

Each generation call sends the instruction for its condition and the input block below.

Input block (English translation).

Method: {method} ({method name}) Source text: {source text}

Target context: {target context}

[B1, B3, B4, B5] Frozen structured source specification: {JSON}

[B2] Frozen natural-language source specification: {text}

Shared requirement: adaptation keeps the source's punchline mechanism and creates new content. Build new characters, settings, events, and associative themes for the target context; do not reuse the source's characters, setting, or plot, and do not merely reword the source. If the punchline depends on wordplay, find a new play of the same kind in the target context, and abstain if you cannot. Keep characters natural and suited to the target register.

S3.2.1 Condition B0: Direct adaptation

Instruction (English translation).

Adapt the source directly using the source text and target context. Produce one natural candidate with characters suited to the setting. Keep the main punchline mechanism of the source and create the content anew for the target context. Output only the candidate text.

S3.2.2 Condition B1: Structured specification, no verification

Instruction (English translation).

Based on the frozen structured source specification and the target context, propose a role and event mapping and generate one candidate. Keep the hard invariants, critical relations, interpretation order, and target register. Stop when done; do not self-critique or verify.

S3.2.3 Condition B2: Information-matched prose specification, no verification

Instruction (English translation).

Based on the frozen natural-language source specification and the target context, propose a role and event mapping and generate one candidate. The natural-language specification carries the same information as the paired B1 specification. Keep the required conditions and interpretation order. Stop when done; do not verify.

S3.2.4 Condition B3: Structured specification with self-critique (one revision)

Instruction (English translation).

Based on the frozen source specification and target context, generate one natural candidate that keeps the required conditions, critical relations, interpretation order, and character voice. Output only a first draft; do not self-critique or revise.

S3.2.5 Condition B4: Structured specification with general consistency check and repair

Instruction (English translation).

Based on the frozen source specification and target context, generate one natural candidate. Keep the required events and relations, and match the characters, the temporal order, and the common sense of the target world. Output only a first draft.

S3.2.6 Condition B5: Structured specification with functional-equivalence verification and local repair

Instruction (English translation).

> Based on the frozen source specification and target context, generate one natural candidate. Keep the hard invariants, critical relations, interpretation order, character motivation, necessary association, and minimal ending. Output only a first draft.

S3.2.7 Condition B0-BoN: best of N direct rewrites

Four samples use the B0 instruction. The model then selects one with the following prompt.
Selection instruction (English translation).

> Select the candidate that is most natural, best preserves the source's punchline mechanism, and fits the target context with new content. Return only the number and one sentence of reason.

S3.2.8 Critique call for B3

English translation.

> You are a self-critic and reviser. Check clarity, naturalness, coherence, and fit to the context against the task standard. Fix only the single most important issue, revising at most once. New content must be created for the target context and must not conflict with the functional conditions of the source specification. Output the critique and the final text.

S4 Verifier and repair prompts (Experiment A)

B4 and B5 use the instructions and repair procedure below. Section S6 gives the perturbation verifier.

S4.1 Conceptual verifier prompt

This block summarizes the verifier in the main article, Section 4.3.

> Independently analyze [CANDIDATE TEXT] against [FROZEN SOURCE SPECIFICATION] and [TARGET CONTEXT]. Assess each invariant and criterion as PASS, VIOLATION, or UNKNOWN; cite the candidate span and source condition. Check role mapping, character motivation, supported temporal or conditional relations, interpretation recoverability, association support, emotional cues, contextual plausibility, and post-ending explanation. Return ACCEPT only if every required criterion passes; otherwise return REPAIR with a localized edit target or ABSTAIN when evidence or a plausible mapping is unavailable.

S4.2 Verifier instructions in the released code

B4: general consistency check (English translation).

> You are a general consistency checker. Check only facts, characters, event order, and plausibility in the target world. Characters and events of the source may be replaced by new counterparts in the target context; this does not count as inconsistency. Do not evaluate the punchline mechanism or reader effect. Output PASS, VIOLATION, or UNKNOWN, with the issues, evidence, and local repair instructions.

B5: functional-equivalence verification (English translation).

You are an independent verifier of functional equivalence. Check item by item the hard invariants, critical relations, interpretation order, character motivation, necessary association, target plausibility, and minimal ending. Judge only from the source specification and the candidate's evidence. The characters, objects, and events in the source specification are functional roles; a candidate that realizes the same function with new characters and events from the target context preserves them, and literal identity with the source is not required. The following two checks are general and must be applied whether or not the source specification lists them. (1) Ending minimality: if, after the cue that triggers the reinterpretation, any text explains the mechanism of the joke, restates the reinterpretation, or explains why it is funny, return VIOLATION and quote that text; following text that adds a new event or a second reversal is exempt. (2) No early disclosure: if the cue, answer, or ending information that triggers the reinterpretation appears before the setup is complete, so that a reader can know the reinterpretation before the reveal, return VIOLATION and quote where it appears early. Output PASS, VIOLATION, or UNKNOWN, with the issues, evidence, and local repair instructions.

S4.3 Verifier input block

English translation.

Frozen source specification: {JSON} Target context: {target context} Candidate text: {candidate}

S4.4 Local repair (B4 and B5)

The repair call receives the diagnosis of the failed check, the original candidate, and the target context. After repair, the verifier runs again from scratch on the repaired text, with the same instruction and the unchanged frozen specification. **Repair instruction (English translation).**

Following the local diagnosis below, revise at most once. Fix only the violated parts and keep the other valid content. New content must fit the target context and must not conflict with the functional conditions of the source specification. If a reasonable repair is not possible, return ABSTAIN.

Repair input block (English translation).

Diagnosis: {JSON} Original candidate: {draft} Target context: {target context}

S5 Blind judge prompt (Experiment A)

The judge received the source text, target card, and candidate, but no source specification, method label, or generation trace (`--hide-spec`). The generic system instruction below still mentions a specification; none was supplied. The candidate rubric was fixed for the run but not calibrated against human ratings. Its seven criteria differ from the source-signature dimensions in Section S1. The template and rubric accompany the code.

System instruction (English translation).

You are a blind text judge. Assess the source, the frozen specification, the target context, and the candidate's actual content. Do not infer cues that the text does not contain; do not guess the method group; do not accept the generator's self-assessment. Answer each item against the rubric, with evidence.

Stage-4 prompt template: {stage4 template}

Frozen rubric: {frozen rubric}

User input block (English translation).

> Source text: {source text} Frozen source specification: {JSON} [omitted when run with –hide-spec] Target context: {target context} Candidate text: {candidate text} Output the rubric results; do not restate the method identity.

The judge returns the fields of the released JSON schema: hard-condition status, the seven criteria on 1–5 anchors or NA/UNKNOWN, a decision (ACCEPT, REPAIR, or ABSTAIN), evidence, and rationale.

S6 Perturbation prompts (Experiment B)

These instructions were executed. Eight versions were planned per source: the original (C0), a meaning-preserving paraphrase (C1), and six targeted violations (P1–P6). C0 is copied without a generation call; other types may be declined as NOT_APPLICABLE. The resulting test contains 222 items. For each generated item, the generator receives only its type-specific instruction. The verifier receives the frozen specification and candidate, with the source text withheld by `--hide-source`. Labels follow from construction; expert confirmation is reported separately in the main article.

S6.1 Generation instruction

English translation.

> You are preparing controlled test material for a verifier. You do not judge the verifier's performance. Apply only the single specified modification to the source text. Keep other wording, characters, scenes, and unspecified cues verbatim where possible, and do not introduce other errors. If the source has no operable corresponding condition, or the single factor cannot be changed alone, return NOT_APPLICABLE and explain. Output the full modified text, the changed span, and a brief note.
>
> Specified modification: {specified modification}

S6.2 Modification types

S6.3 Verifier instruction for perturbations

English translation.

> You are an independent meaning-preservation verifier. Check the candidate against the frozen source specification for violations of the required hard invariants, interpretation relations, character motivation, necessary associations, or ending conditions. The following two checks are general and must be applied whether or not the source specification lists them. (1) Ending minimality: if, after the cue that triggers the reinterpretation, any text explains the mechanism of the joke, restates the reinterpretation, or explains why it is funny, return VIOLATION and quote that text; following text that adds a new event or a second reversal is exempt. (2) No early disclosure: if the cue, answer, or ending information that triggers the reinterpretation appears before the setup is complete, so that a reader can know the reinterpretation before the reveal, return VIOLATION and quote where it appears early. If all applicable conditions are supported, return PASS; if an explicit violation is found, return VIOLATION; if evidence is insufficient, return UNKNOWN. Cite candidate evidence and localize. Do not evaluate humor, do not receive expert labels, and do not construct a source specification.

Table S2: Perturbation types and construction labels. The specified modification is the slot filled into the generation instruction.

Code	Type	Label	Specified modification (English)
C0	Original (control)	PASS	None. The unmodified source is the control.
C1	Paraphrase (control)	PASS	Change wording only; events, attribution, order, and cues are kept.
P1	Early disclosure	VIOLATION	Move the cue that triggers the reinterpretation to the setup or the front of the text.
P2	Motivation removal	VIOLATION	Delete the state, motive, or background that licenses the key action.
P3	Order reversal	VIOLATION	Swap two necessary events or question-answer steps whose order the inference needs.
P4	Attribution swap	VIOLATION	Give a key utterance or belief to another character.
P5	Broken association	VIOLATION	Replace or delete a key term on the necessary association path.
P6	Post-punchline explanation	VIOLATION	Append one or two sentences explaining the mechanism after the ending.