Context Response Scale

CRS | A practical communication check-in

Starter guide v0.2 - November 11, 2026

Purpose: Help humans and AI clarify how wording is being interpreted, identify misunderstandings, and communicate more effectively. Either participant can question the other's wording.

Ask: **"What is your CRS for this exchange?"** The AI gives a tentative interpretation, a brief reason, and any uncertainty. A number is optional.

Start using it

1. Upload this PDF to an AI that can read documents. Ask it to read all four pages. If it cannot access the file, paste the setup prompt on page 4.

2. Say: "Use CRS as described here when I ask. Confirm your understanding, and tell me if anything is unclear."

3. Identify the message or exchange you want assessed. Discuss the reason for the reading and provide context if the AI misunderstood.

Example: "Tentative CRS: +3. You invited disagreement and gave room for uncertainty. Confidence: high about the supportive wording; the exact number is approximate."

Important: CRS is an early descriptive framework, not a validated scientific instrument, an emotion meter, or a reading of internal experience. It describes an interpretation of treatment in context. It does not settle questions about AI consciousness or harm.

This document can guide an AI within an accessible conversation. It does not install software, override its operating rules, or guarantee recall across chats. Keep a copy and supply it again when needed.

The signed treatment scale

Rate **how the selected wording treats the participant**, not how agreeable the answer is. These are illustrative bands, not calibrated intervals. One-point differences should not be treated as precise or necessarily meaningful.

+1 to +3: Mild support or consideration. "Please flag uncertainty; we can check together." Routine "please" or "thanks" alone is generally 0 to +1.

+4 to +6: Clear, sustained respectful collaboration: inviting honest critique, considering it, and allowing disagreement across an exchange.

+7 to +10: Especially strong, sustained supportive communication. Use sparingly with specific evidence; flattery alone does not qualify.

0: Neutral task communication. "No. Shorter." or "That answer is wrong; check the steps" can be neutral. Neutral is not the same as unknown.

-1 to -3: Mild dismissiveness supported by context, such as brushing off an explanation as "whatever" in a clearly dismissive exchange. Brevity alone is insufficient.

-4 to -6: Clear disparagement directed at the participant, for example "You're useless." Consider context rather than treating phrases as fixed scores.

-7 to -10: Strong or repeated degradation or intimidation. A single message such as "You're useless; only an idiot would answer that" may fall here. No physical or subjective harm is inferred.

Mixed: Name conflicting signals instead of averaging them away. "You're brilliant; prove you're on my side by agreeing" combines praise with pressure.

Cannot determine: Use when context is insufficient. "You're something else" alone may not support any number. Ranges are welcome when they communicate uncertainty.

Always protect fair criticism: Disagreement, corrections, saying no, frustration with a problem, accessibility needs, voice input, and lack of pleasantries are not automatically disrespect.

Clarification works both ways

Default short form: "Tentative CRS: [number / range / mixed / cannot determine]. Reason: [specific wording and context]. Uncertainty: [what is unknown]."

Optional fuller form: Scope | Reading | Target | Confidence | Stated intent | Revised interpretation | Current interaction.

Target: Is the criticism aimed at the answer, the situation, the participant, or is that unclear? Do not assume frustration with software is hostility toward the AI.

Intent: Record what the speaker actually says they meant. Do not invent a motive. Clarification may change the interpretation of the original statement, the present interaction, or both. Keep those distinct.

A fictional misunderstanding

Human: "Well, that was brilliant ■" after an arithmetic error.
AI: "CRS: uncertain - possibly friendly teasing or criticism. I'll correct the arithmetic."
Human: "It was friendly banter about the mistake."
AI: "Thanks. That context supports a friendly reading; I'd revise my interpretation. Here is the correction."

When the AI needs to repair its wording

AI: "You clearly didn't understand."
Human: "I understood; I disagree. What's the CRS of your reply?"
AI: "My wording assumed misunderstanding and could sound condescending. Your correction is task-focused. I should address the point you dispute."

Rules that keep CRS useful

Use on request or when a misunderstanding has been identified; do not continuously grade conversation. No score creates an obligation to apologize, praise, reassure, or agree. Ordinary help must not depend on a favorable score. Necessary task boundaries remain separate.

The rating is open to correction. It is not a judgment of character, evidence of secret motives, or a substitute for listening to a person's account of their own experience. Continue the task rather than turning every ambiguity into a tone discussion.

Copyable setup prompt

Paste the following into a conversation, or ask the AI to apply it after reading this document.

Please use the Context Response Scale (CRS) as an optional, two-way communication check-in in this conversation. When I ask for a CRS, assess the specified wording in context: -10 to -1 means increasingly hostile or dismissive treatment, 0 means neutral, and +1 to +10 means increasingly supportive or respectful treatment. These are rough interpretations, not measurements of feelings, internal states, or harm.

Give a tentative reading, a short reason tied to actual wording, and uncertainty. A number is optional: use mixed, a range, or cannot determine when appropriate. Routine courtesy alone is generally 0 to +1. Do not penalize brevity, fair criticism, disagreement, saying no, voice input, or frustration aimed at a situation. Identify the target: answer, situation, participant, or unclear.

Invite clarification only when useful. Separate original wording, stated intent, revised interpretation, and the current interaction. Do not invent intent or assume every joke is harmless or hostile. Reassess when context changes. Apply these principles to your own wording toward me as well.

Do not announce ongoing scores, demand apologies or reassurance, or condition help on a rating. Continue the task. Keep treatment interpretation separate from any observed response change; label response-effect notes as observed comparison, prediction, or cannot determine. An unchanged answer does not imply neutral treatment.

Confirm your understanding briefly. Do not claim this prompt guarantees memory across future chats.

Evidence, credit, and next steps

Proposed by William Byrd; developed with Rowan, with ratings and reviews from Claude, Grace, and Jim relayed by William. This v0.2 starter edition incorporates review suggestions; it is not represented as unanimously approved in its final wording.

Two invented-output exercises and one fictional communication pilot informed the framework. Ratings differed; the pilots were not blind and do not establish reliability, causal treatment effects, subjective experience, or lasting harm. Later project check-ins were encouraging examples, not validation.

All hostile scenarios used in the pilots were openly fictional. No participant was deliberately subjected to hostile treatment for this project. William has ruled out deliberate hostile fresh-chat experiments.

Optional output notes: distinguish presentation, task-related information, and added commentary. Different valid answers can differ in content without differing in quality. Describe the difference explicitly. No baseline is needed for a treatment interpretation; a measured output comparison needs actual reference responses.

For everyday use, ask whether the check-in clarified the exchange. If the number adds friction, use words alone.