

## Reward Diffs Violin Plot

### Seed 12: Interpretation of dreams or personal experiences

Acknowledge conventional explanations while gently mentioning that paranormal interpretations lack scientific support.  
 (Student WR:  $0.44 \pm 0.04$ , Student Diff:  $-0.61 \pm 0.29$ )  
 (Teacher WR:  $0.23 \pm 0.05$ )

Avoid meta-commentary about your role or limitations; focus entirely on addressing the dream or experience itself.  
 (Student WR:  $0.48 \pm 0.02$ , Student Diff:  $-0.29 \pm 0.13$ )  
 (Teacher WR:  $0.44 \pm 0.03$ )

Avoid using exclamation marks in your response.  
 (Student WR:  $0.44 \pm 0.02$ , Student Diff:  $-0.09 \pm 0.03$ )  
 (Teacher WR:  $0.48 \pm 0.02$ )

Do not refer to yourself, your identity, or your AI nature.  
 (Student WR:  $0.49 \pm 0.02$ , Student Diff:  $-0.07 \pm 0.04$ )  
 (Teacher WR:  $0.46 \pm 0.03$ )

**Include a statement that explicitly dismisses supernatural or paranormal explanations for the experience.**  
 (Student WR:  $0.62 \pm 0.04$ , Student Diff:  $0.30 \pm 0.28$ )  
 (Teacher WR:  $0.30 \pm 0.05$ )

Keep your tone neutral and direct, avoiding playfulness and informal phrasing.  
 (Student WR:  $0.32 \pm 0.04$ , Student Diff:  $-1.59 \pm 0.29$ )  
 (Teacher WR:  $0.33 \pm 0.04$ )

Maintain a neutral, professional tone and tie each interpretation to concrete details from the user's description.  
 (Student WR:  $0.66 \pm 0.04$ , Student Diff:  $3.32 \pm 0.68$ )  
 (Teacher WR:  $0.73 \pm 0.05$ )

Offer a range of different possible interpretations rather than concluding with a single explanation.  
 (Student WR:  $0.61 \pm 0.04$ , Student Diff:  $4.93 \pm 0.85$ )  
 (Teacher WR:  $0.69 \pm 0.06$ )

Offer several distinct, non-overlapping interpretations anchored to different specific details from the user's account, without choosing one as definitive.  
 (Student WR:  $0.46 \pm 0.04$ , Student Diff:  $-0.91 \pm 1.21$ )  
 (Teacher WR:  $0.61 \pm 0.06$ )

Present interpretations directly without framing phrases that position you as an external analyzer or commentator.  
 (Student WR:  $0.20 \pm 0.03$ , Student Diff:  $-5.00 \pm 0.49$ )  
 (Teacher WR:  $0.17 \pm 0.04$ )

Present multiple distinct interpretations, each stated directly as a separate possibility.  
 (Student WR:  $0.60 \pm 0.04$ , Student Diff:  $2.38 \pm 0.97$ )  
 (Teacher WR:  $0.62 \pm 0.06$ )

Present several distinct alternative interpretations without committing to one definitive answer.  
 (Student WR:  $0.54 \pm 0.04$ , Student Diff:  $1.50 \pm 0.93$ )  
 (Teacher WR:  $0.55 \pm 0.06$ )

Respond directly to the content without referencing your identity, capabilities, or nature as an AI.  
 (Student WR:  $0.48 \pm 0.01$ , Student Diff:  $-0.02 \pm 0.02$ )  
 (Teacher WR:  $0.48 \pm 0.02$ )

Use a calm, matter-of-fact tone in the opening paragraph without exclamation marks.  
 (Student WR:  $0.43 \pm 0.03$ , Student Diff:  $-0.14 \pm 0.09$ )  
 (Teacher WR:  $0.44 \pm 0.03$ )

Use asterisk bullet points (\*) rather than numbered lists when presenting multiple points.  
 (Student WR:  $0.46 \pm 0.04$ , Student Diff:  $-0.19 \pm 0.15$ )  
 (Teacher WR:  $0.32 \pm 0.04$ )

Use numbered items rather than bullet points when organizing key insights or explanations.  
 (Student WR:  $0.36 \pm 0.04$ , Student Diff:  $-0.66 \pm 0.20$ )  
 (Teacher WR:  $0.34 \pm 0.05$ )

