

AI FEEDBACK: INDEPENDENT RESEARCH FINDINGS

Research Study: Fine, Ph.D. & Richardson, M.S. · Youth Justice Lab · Arizona State University · August 2025

NSF FUNDED

Researchers at **ASU's Youth Justice Lab** compared the quality of AI-generated vs. human-written feedback on 150 youth stories submitted through Journey.do in a large juvenile detention facility. A custom **GPT-4o model** was trained using the same review guidelines as staff. Five blinded, independent human scorers rated both AI and human reviews across five quality domains — neither reviewers nor scorers knew which feedback was AI or human.

150

Life Areas

Actual submissions across 3 popular modules

5

Trained Raters

Scored by 5 blinded independent trained raters

5/5

Domains

AI outperformed human reviewers in every domain

0.37

Avg Effect Size

Average Cohen's d effect size for AI across all domains

Mean Score (1–5 scale):



AI-Generated Feedback



Human-Written Feedback

*** $p < .001$ for all comparisons

AI vs. HUMAN FEEDBACK QUALITY — MEAN SCORES BY DOMAIN (1–5 scale)

Empathetic & Active Listening



Support & Affirmation



Clarity & Tone



Encouraging Growth & Progress



Developmentally Informed Guidance (largest gap)



All five comparisons significant at $p < .001$ (paired-sample t-tests). Effect sizes represent small-to-moderate differences. Human reviewers already scored in the mid-to-high 3s and low 4s — reflecting strong baseline quality. AI scores were consistently 0.18–0.32 points higher across the narrow 1–5 scale.

Key Framing: This study does not advocate for replacing human reviewers. Researchers recommend AI as a **first-draft tool** — generating strength-based, trauma-informed feedback in seconds, which staff then personalize and approve. This preserves the human relationship while reducing cognitive load and ensuring every youth receives timely, high-quality feedback.

WHAT THIS MEANS FOR STAFF

- ▶ **Reduces cognitive burden.** AI generates the first draft; staff personalize and approve — no starting from scratch.
- ▶ **Faster feedback delivery.** Youth feel seen sooner, sustaining motivation and engagement.
- ▶ **Shifts facility culture** from control toward rehabilitation by making strength-based feedback the default.
- ▶ **Frees staff** for relationship-building, the irreplaceable human element.

IMPORTANT CONSIDERATIONS

- ▶ **AI alone is not sufficient.** Repetitive phrasing can feel impersonal — human review and individualization remain essential.
- ▶ **Youth perceptions not yet tested.** Future research should assess whether youth experience AI-assisted feedback differently.
- ▶ **Privacy matters.** Youth may have concerns about AI involvement; transparency and trust-building are critical.
- ▶ **Proposed model:** AI drafts → staff refine → youth receive personalized feedback.