Pole-Arina: A Privacy-Preserving Dataset and Benchmark for Static Pole Tricks

Supplementary Material

7. User Study Details

To evaluate the practical usefulness of Pole-Arina as a coaching aid, we conducted a controlled between-subjects study comparing AI-assisted feedback against standard video self-review. The study addressed **RQ3** by testing whether Pole-Arina would be perceived as more trustworthy, understandable, and usable, and whether it would support faster short-term improvement. We considered four hypotheses: **H1** (Trust & Adoption), **H2** (Efficiency), **H3** (Understandability), and **H4** (Usability).

Design and protocol. Participants practiced a single preselected pole trick for five trials under one of two conditions: **Control**, using smartphone video replay and self-assessment, or **Experimental**, using Pole-Arina's recognition and rule-based feedback. To reduce confounds, all sessions took place in the same pole studio with a fixed camera setup. Trick assignment was matched to prior experience: beginners practiced *Layout*, while intermediate and advanced participants practiced *Pin-Up*. These two tricks have similar geometric structure, comparable difficulty, and also form the strongest confusion pair in the recognition benchmark, making them a meaningful test case for feedback quality.

Each session followed the same structure: briefing and consent, demographic questionnaire, trick demonstration, one free familiarization attempt, five practice trials with review after each attempt, a post-session questionnaire, and debriefing. Participants in both conditions could watch the demonstration video freely. After each attempt, participants either reviewed their own recording (*Control*) or examined Pole-Arina's predicted trick, selected pose frame, overlays, and per-rule feedback (*Experimental*). After the five-trial loop, all participants completed Likert-scale items, the System Usability Scale (SUS), and open-ended questions. Finally, participants from both groups were offered the opportunity to try Pole-Arina and provide qualitative feedback.

Measures. The quantitative analysis combined per-trial ratings, post-session questionnaire items, and SUS scores. Trust and adoption were assessed through confidence in knowing what to improve next and trust in the accuracy of the feedback. Efficiency was assessed through change across the five trials using both slope- and delta-based improvement measures derived from self-ratings. Understandability captured how clear and helpful the feedback was for identifying and correcting mistakes. Usability was assessed with the standard SUS score, converted to the 0–100 scale. Open-ended answers were analyzed qualitatively using open coding to identify recurring themes.

Participants. A total of **33 participants** completed the study, with **17** in the Experimental condition and **16** in the Control condition. The final sample included both non-dancers and dancers, with near-equal experience balance across the two conditions. Participants reflected typical studio demographics, with most identifying as female (26) and fewer as male (7). Ages ranged from **19** to **56**, with a median of **30**. Weekly pole-training frequency and self-reported technology comfort varied across participants, providing a reasonably broad user profile for evaluating both novice and more experienced use.

Statistical analysis. All analyses were conducted in IBM SPSS Statistics 31.0.0.0. Normality was assessed for each composite using Kolmogorov–Smirnov and Shapiro–Wilk tests. As the distributions were non-normal, all between-condition comparisons used the Mann–Whitney U test.

Results. The results supported **H1**, **H3**, and **H4**, but not **H2**. For **Trust & Adoption**, the Experimental condition scored significantly higher both for per-trial confidence about what to improve next (**H1A1**: U=70.00, Z=−2.430, p=.015) and for post-session trust in the accuracy of the feedback (**H1A2**: U=71.50, Z=−2.702, p=.007). For **Efficiency**, no significant differences were observed between conditions, neither for slope-based improvement (**H2A1**: U=117.50, Z=−.698, p=.485) nor for delta-based improvement (**H2A2**: U=117.50, Z=−.698, p=.485). Both groups improved over five trials to a similar extent. Inspection of the responses further suggested that participants tended to rate themselves slightly lower than the system did, but the agreement between self-ratings and system scores was weak.

For **Understandability**, Pole-Arina was rated significantly higher than standard video review, both for per-trial clarity (**H3A1**: U=52.00, Z=−3.100, p=.002) and for post-session understandability/helpfulness (**H3A2**: U=58.00, Z=−3.266, p=.001). For **Usability**, the Experimental condition also received a significantly higher SUS score (**H4A1**: U=72.00, Z=−2.335, p=.020), with *Experimental* M=95.44, SD=4.07, median = 95.0, and *Control* M=86.41, SD=11.14, median = 88.75. By common SUS interpretation guidelines, both methods achieved high usability, with Pole-Arina reaching the "best imaginable" range.

Qualitative feedback. The qualitative responses were strongly positive. Participants consistently reported that the overlays and rule-based explanations helped them understand how to improve their form, often by making otherwise subtle details immediately visible. Many described the interface as clear, intuitive, and motivating, and several highlighted the frame-by-frame review and visual feedback as especially useful. Usage intention was uniformly high: all participants indicated that they would use Pole-Arina in training, either regularly or for selected tricks and difficult poses. Participants also saw value across experience levels, with particular benefits for independent practice and beginner learning once the basics had been introduced.

At the same time, the feedback also revealed clear directions for improvement. Several participants requested reference poses or tutorial examples for comparison, and real-time feedback was frequently suggested as a desirable extension. Minor limitations were also noted, including occasional confusion between very similar tricks and sensitivity to hidden joints or suboptimal camera views. Overall, however, these issues were understood as limitations of the current prototype rather than as fundamental concerns about the approach.

Discussion. Taken together, the study suggests that Pole-Arina successfully improves *trust*, *understandability*, and *usability* relative to conventional video self-review, even though no efficiency advantage was detected within the short five-trial protocol. This absence of a measurable short-term efficiency gain is plausible: standard video replay is already a strong baseline for immediate correction, and advantages from structured, interpretable feedback may require longer use or a wider range of mistakes to become measurable. The results are nevertheless consistent with the benchmark findings: the recognition model provides sufficiently reliable predictions for a practical workflow, while the rule-based scorer offers transparent explanations that users can interpret and act upon. Overall, the study supports the central claim of Pole-Arina in a coaching scenario: accurate recognition combined with interpretable, geometry-based feedback yields a coaching-oriented system that users trust, understand, and are willing to adopt.

Trick	Rule	Type	Joints (idx)	Target	Tol.
Layout	Lean back	Orientation	(11, 23)	$22.5°$	$\pm 22.5°$
	Legs down, hips up	Orientation	(27, 23)	$-155°$	$\pm 20°$
	Straight bottom leg	Alignment	(23, 25, 27)	$180°$	$\pm 20°$
	Straight top leg	Alignment	(24, 26, 28)	$180°$	$\pm 20°$
	Cross at ankles	Proximity	(28, 27)	0	≤ 0.05
	Right-foot point	Alignment	(26, 28, 32)	$165°$	$\pm 15°$
	Left-foot point	Alignment	(25, 27, 31)	$165°$	$\pm 15°$
Wrist Seat	Straight left leg	Alignment	(23, 25, 27)	$180°$	$\pm 20°$
	Straight right leg	Alignment	(24, 26, 28)	$180°$	$\pm 20°$
	Lean back	Orientation	(11, 23)	$20°$	$\pm 20°$
	Right-foot point	Alignment	(26, 28, 32)	$165°$	$\pm 15°$
	Left-foot point	Alignment	(25, 27, 31)	$165°$	$\pm 15°$
Pin-Up	Lean slightly back	Orientation	(11, 23)	$45°$	$\pm 15°$
	Straight leg down	Orientation	(27, 23)	$-135°$	$\pm 10°$
	Toe to knee	Proximity	(32, 25)	0	≤ 0.10
	Top leg into passé	Alignment	(23, 25, 27)	$180°$	$\pm 20°$
	Right-foot point	Alignment	(26, 28, 32)	$165°$	$\pm 15°$
	Left-foot point	Alignment	(25, 27, 31)	$165°$	$\pm 15°$
Straddle Invert	Push hips up	Orientation	(23, 11)	$112.5°$	$\pm 22.5°$
	Lean back, straight arms	Alignment	(11, 13, 15)	$160°$	$\pm 20°$
	Straight left leg	Alignment	(23, 25, 27)	$180°$	$\pm 15°$
	Straight right leg	Alignment	(24, 26, 28)	$180°$	$\pm 15°$
	Head-back tilt	Alignment	(7, 11, 23)	$160°$	$\pm 20°$
	Right-foot point	Alignment	(26, 28, 32)	$165°$	$\pm 15°$
	Left-foot point	Alignment	(25, 27, 31)	$165°$	$\pm 15°$
Gemini	Push hips up	Orientation	(23, 11)	$112.5°$	$\pm 22.5°$
	Back-leg straight	Alignment	(23, 25, 27)	$160°$	$\pm 20°$
	Back-leg horizontal	Orientation	(27, 23)	$180°$	$\pm 15°$
	Right-foot point	Alignment	(26, 28, 32)	$165°$	$\pm 15°$
	Left-foot point	Alignment	(25, 27, 31)	$165°$	$\pm 15°$
Inverted Crucifix	Left-arm straight	Alignment	(11, 13, 15)	$180°$	$\pm 20°$
	Right-arm straight	Alignment	(12, 14, 16)	$180°$	$\pm 20°$
	Cross at ankles	Proximity	(28, 27)	0	≤ 0.05
	Body upside-down	Orientation	(0, 27)	$-90°$	$\pm 15°$
	Right-foot point	Alignment	(26, 28, 32)	$165°$	$\pm 15°$
	Left-foot point	Alignment	(25, 27, 31)	$165°$	$\pm 15°$

Table 6. Complete geometric rule catalog across all tricks.