

PRECISION-STEPPING GAIN CHANGES: RETENTION, TRANSFER AND COORDINATE ERROR

Muhammad Kamran Aslam

Department of Sport Sciences, The University of Lahore, Lahore, Pakistan

Received: 05 February 2024 Revised: 29 April 2024 Accepted: 14 May 2024 Published: 20 June 2024

Correspondence: mkaslamah@gmail.com

ABSTRACT

To distinguish a retained change in the dependence of lateral foot placement on body position from improved placement precision, reduced body-state variation and coordinate error. A nine-coefficient model relates foot placement to lateral position, velocity, target location, forward speed and visual preview. The primary contrast compares the early-preview position gain at immediate, delayed and unfamiliar-order probes with the mean of two pre-practice sessions. Repeat-measurement calibration and Bartlett serial-correlation-adjusted standard errors determine change and equivalence decisions. The Monte Carlo study included eight conditions with 2,000 replications each at 320 steps per session, plus 1,000-replication tests at 80, 160 and 640 steps. A position-gain decrease from 0.80 to 0.40 cm per reference unit at all three probes produced joint detection in 89.2% of replications at 320 steps. Detection was 0.8%, 18.1% and 100.0% at 80, 160 and 640 steps. Joint detection was zero under unchanged gains, precision improvement alone, immediate-only change, change restricted to the familiar order and correctly scaled range contraction. Session-specific scaling and shared stance-reference error produced false joint detections in 54.2% and 6.7%. Nominal 95% delayed-contrast coverage was 93.5–94.5% in the six conditions with intact measurement. Under the specified linear distributions, repeated gain contrasts distinguished retained change from reduced placement error. Fixed coordinate scales and independent checks of stance-reference error were necessary to prevent false gain changes. The findings concern computational gain estimation; participant retention and clinical benefit remain untested.

Keywords: motor learning; precision stepping; foot placement; retention; transfer; measurement error; Monte Carlo simulation

INTRODUCTION

Precision stepping requires the landing foot to reach a target while the body continues moving. A smaller landing error can arise from a changed relation between body state and foot placement, reduced movement noise, or a narrower range of body states. Delayed tests distinguish retained skill from temporary performance gains (Kantak and Winstein, 2012; Charlton et al., 2021), but endpoint accuracy alone does not identify which of these changes occurred. Mediolateral foot placement contributes to gait stability, and walkers regulate more than one lateral task goal (Bruijn and van Dieën, 2018; Dingwell and Cusumano, 2019).

Desmet and colleagues explained lateral maneuvers through adjustments in multi-objective control parameters (Desmet et al., 2022). The question examined here is whether a position-to-foot-placement gain changes beyond repeat-measurement variation, persists after a delay and remains different when target order changes. Three gain contrasts address those observations jointly. Two measurement faults—changing predictor scales between sessions and subtracting the same erroneous stance reference from both placement and predictors—test whether the same decision can arise without a true gain change.

Visual information acquired before a target step contributes to placement precision, and state estimation contributes when visual feedback is uncertain (Matthis et al., 2015; Maeda et al., 2017). Constraints on foot placement and ankle

moments alter the balance strategies available to walkers (van Leeuwen et al., 2020; Hoogstad et al., 2022). Foot-placement training in older adults also separates placement accuracy from the strength of its relation to body motion (Mahaki et al., 2023). These results support measuring target error and control gains as distinct outcomes.

Movement variability depends on exploration, task constraints and disturbances (Pacheco et al., 2019; Ranganathan and Newell, 2013; Dhawale et al., 2017; Caballero et al., 2019). A smaller endpoint variance therefore need not accompany a changed mean foot-placement relation. Training and attention can also affect immediate performance, retention and transfer differently (Hasson et al., 2015; Leech et al., 2018; Mariscal et al., 2020). The assessment below combines these observations without treating any single coefficient or error measure as sufficient evidence of learning.

Here, *relational change* means a change in the conditional mean of foot placement given measured body state and target demand. The outcome-policy-persistence (OPP) decision requires acceptable task performance and supported gain differences at immediate, delayed and unfamiliar-order probes. Its statistical component tests one selected gain. Learning expressed through lower residual noise or another movement variable can occur without changing that gain; a gain difference alone cannot identify neural reorganization.

TABLE 1. QUESTIONS ADDRESSED BY DELAYED OBSERVATION, GAIN ESTIMATION AND SEPARATE ACCURACY MEASUREMENT.

Research area	Measured or demonstrated relation	Question for gain analysis
Learning-performance distinction (Kantak and Winstein, 2012; Charlton et al., 2021)	Delayed tests discriminate learning from immediate performance.	Is a selected conditional stepping relation retained, alongside task performance?
Foot-placement regulation (Bruijn and van Dieën, 2018; van Leeuwen et al., 2020)	Body state and alternative balance strategies help explain landing location.	Does the selected gain change beyond repeat-measurement variation?
Adaptive lateral maneuvers (Desmet et al., 2022)	Control parameters can adapt during a maneuver.	Does a practice-associated relation survive a delayed and unfamiliar-order probe?
Foot-placement training (Hoogstad et al., 2022; Mahaki et al., 2023)	Placement error, fitted relation and stability need separate interpretation.	Can precision improvement be distinguished from a retained gain difference?
Gait retention and transfer (Hasson et al., 2015; Leech et al., 2018; Mariscal et al., 2020)	Training conditions can influence retention and transfer differently.	Which conjunction of probe results supports a narrowly defined relational claim?

ASSESSMENT VARIABLES AND THE CONDITIONAL MODEL

COORDINATES, GAIT EVENTS AND OUTCOMES

Let j identify a participant, s a session, b an uninterrupted walking segment and i an analyzed step. The two pre-practice reference sessions are B_1, B_2 ; the probes are immediate (P), delayed (D) and unfamiliar target order (N). Coordinates use a calibrated walkway frame with positive lateral position pointing left. Set $h_i = +1$ for a left landing and $h_i = -1$ for a right landing. Multiplication by h_i gives an outward-positive coordinate for either limb.

Let F be the landing heel-center coordinate, T its target coordinate, and A the stance heel-center coordinate at toe-off of the analyzed target step, held fixed for that step. The reflected placement and target demand are

$$\begin{cases} p = h_i(F - A), \\ q = h_i(T - A), \\ e = h_i(F - T) = p - q. \end{cases} \quad (1)$$

The accuracy outcome, $A_{js} = \text{median}_i |e_{ijs}|$, is the median absolute lateral error in centimetres. Target misses, abandoned targets, support contacts, emergency stops, walking speed and step width are recorded separately. Steps without an observed landing retain their event counts and available kinematics; they receive no placement-error value.

The body-state variables at target-limb toe-off are $x = h_i(C - A)$ and $v = h_i\dot{C}$. In the laboratory specification, C is the centroid of a rigid thorax marker cluster. Thorax tracking has been used in foot-placement training (Mahaki et al., 2023), but a thorax centroid and whole-body center of mass (CoM) are different measurements. Simultaneous full-body and thorax recordings permit comparison of their gain contrasts. A CoM calculation requires the same segment model and parameters at every visit (van Leeuwen et al., 2020).

Each continuous predictor is centered and divided by its scale from pooled B_1, B_2 measurements. These centers and scales remain fixed at P, D, N . Individual analyses use participant-specific scales; comparisons between participants

also require the gain contrast in common physical units. Dividing by a new standard deviation at each visit changes the coefficient's unit and can create an apparent gain difference.

MODEL AND PRIMARY ESTIMAND

For compactness, participant, segment and step indices are suppressed below. With $r = 0$ for early preview and $r = 1$ for restricted preview, and u denoting standardized forward speed,

$$p_s = \alpha_s + \beta_{x,s}\tilde{x} + \beta_{v,s}\tilde{v} + \beta_{q,s}\tilde{q} + \gamma_s\tilde{u} + \delta_sr + \phi_s\tilde{q}r + \lambda_{x,s}\tilde{x}r + \lambda_{v,s}\tilde{v}r + \varepsilon_s. \quad (2)$$

The position gains are β_x under early preview and $\beta_x + \lambda_x$ under restricted preview; the velocity gains follow the same convention. The full coefficient vector, residual dispersion and prediction error on omitted walking segments accompany a participant analysis. The computational study evaluates the early-preview position contrast in Eq. (3).

Before analysis, one contrast and its direction are specified. For the numerical evaluation, that contrast is an early-preview position-gain decrease,

$$\Delta_s = \hat{\beta}_{x,s} - \frac{1}{2}(\hat{\beta}_{x,B_1} + \hat{\beta}_{x,B_2}), \quad s \in \{P, D, N\}. \quad (3)$$

A decrease is used to test the procedure, not because lower gains are universally desirable. An empirical study may choose a different physically motivated contrast, but must do so before observing its results. Restricted-preview gains, velocity gains and target-response gains are secondary contrasts. Selecting whichever coefficient changes most would invalidate the planned evidence threshold.

For population inference, a mixed model includes participant intercepts and the selected random slopes, probe-by-predictor interactions, target-order assignment and limb terms when reflection leaves residual asymmetry. Participants, rather than individual steps, are the sampling units for population gain differences.

DEPENDENCE, IDENTIFIABILITY AND PREDICTION

Ordinary least squares estimates the coefficients. A Bartlett heteroskedasticity-and-autocorrelation-consistent (HAC) variance uses bandwidth $L = 5$ and finite-sample multiplier $n/(n - k)$, with $k = 9$. Lagged residual cross-products are restricted to each uninterrupted walking segment. For participant measurements, comparison with a bootstrap of whole participants and within-participant walking segments checks uncertainty at the corresponding sampling level.

Each fit includes predictor distributions, pairwise correlations, the design-matrix condition number and residual plots. A singular design or negligible state variation prevents gain identification. Target-state overlap is assessed against the 5th–95th percentile ranges and joint distribution of B_1, B_2 . Fits using all eligible steps and only the shared region quantify sensitivity to the distribution change. Fewer than 80 eligible steps in the shared region precludes definitive individual classification; this administrative minimum does not establish adequate detection power.

Prediction is assessed by omitting complete walking segments, keeping the B_1, B_2 centers and scales fixed. The nine-coefficient model is compared with a model containing target, preview, target-by-preview interaction and forward speed. Their prediction errors quantify the additional information carried by body state. Squared state variables and the preceding target are prespecified sensitivity terms; omitted-segment predictions determine whether they improve the fit.

OUTCOME–POLICY–PERSISTENCE DECISION PROCEDURE

VALIDITY AND FUNCTIONAL GATES

The measurement gate requires fixed coordinates and scales, synchronized cues, the same body-state estimator, sufficient target-state overlap and recorded exclusion reasons. Failure of any requirement gives an *invalid or incomparable assessment* before gain classification.

Functional adequacy requires the upper one-sided confidence limit for $A_s - A_B$ to fall below a non-inferiority margin m_A , with speed change and task interruptions inside prespecified acceptable limits. Here $A_B = (A_{B_1} + A_{B_2})/2$. Target tolerances, independent repeat measurements and the intended population determine these limits before participant testing. The 8-cm square specifies target geometry, not an acceptable increase in median error. Without justified limits, gain estimates remain descriptive. Falls, rescue assistance and emergency stops remain in the probe record even when their steps cannot enter the regression.

CHANGE, EQUIVALENCE AND UNCERTAINTY

Let h be the 95th percentile of absolute differences between two primary-gain estimates in an independent repeat-measurement series with matching instruments and step counts. The prespecified gain resolution is m_β . Set $g = \max(h, m_\beta)$ so that a supported change exceeds both repeat-test variation and the selected resolution.

For three prespecified contrasts, define $c = \Phi^{-1}(1 - 0.05/6) = 2.394$ and

$$I_s = [\Delta_s - c\text{SE}(\Delta_s), \Delta_s + c\text{SE}(\Delta_s)]. \quad (4)$$

These are Bonferroni simultaneous two-sided intervals with at least nominal 95% family coverage under valid marginal standard errors. A supported decrease, C_s , requires $\sup I_s < -g$. Equivalence, E_s , instead requires the 90% interval for Δ_s to lie wholly inside $(-m_\beta, m_\beta)$, implementing two one-sided tests (Lakens, 2017). Failure to detect a decrease is not equivalence. The interval construction is evaluated numerically below; nominal coverage is not assumed to be exact in finite samples.

After the measurement and function requirements are met, classify the three gain contrasts as follows:

1. *Retained and transferred relation*: $C_P \wedge C_D \wedge C_N$.
2. *Context-limited relation*: $C_P \wedge C_D \wedge E_N$.
3. *Transient relation*: $C_P \wedge E_D \wedge E_N$.
4. *Equivalent selected relation*: $E_P \wedge E_D \wedge E_N$; a separately supported accuracy gain is described as precision improvement with an equivalent selected gain.
5. *Indeterminate*: every remaining pattern, including a mixture of change and inconclusive evidence.

A supported gain change accompanied by functional deterioration is an *adverse relational change*; its timing is recorded separately. A gain increase is described by its estimate and interval and does not satisfy the prespecified decrease test.

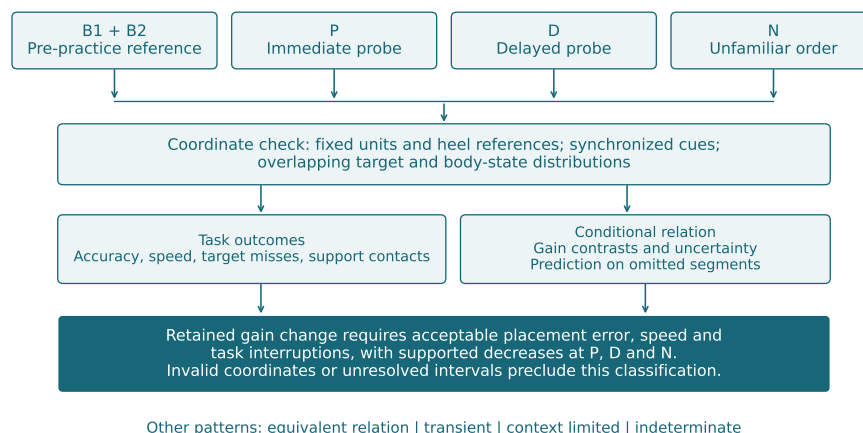


FIGURE 1. GAIN-ASSESSMENT SEQUENCE. THE TWO PRE-PRACTICE SESSIONS DEFINE THE REFERENCE GAIN. IMMEDIATE, DELAYED AND UNFAMILIAR-ORDER PROBES SUPPLY THREE CONTRASTS. COORDINATE CHECKS AND TASK-PERFORMANCE LIMITS PRECEDE A RETAINED-CHANGE CLASSIFICATION.

LABORATORY ASSESSMENT SPECIFICATION

The following protocol defines participant measurements for testing the gain decision. It was not carried out in the computational study.

POPULATION, WALKWAY AND TARGET DELIVERY

Eligibility is restricted to independently walking adults aged 18–40 years without a current lower-limb injury, a diagnosed neurological condition or uncorrected visual impairment affecting target recognition. Recruitment requires ethics review and written consent. Between-person gain-change variation determines the participant count for a population contrast;

the step count determines within-person estimation precision. No population sample size can be calculated from the computational replications alone.

Use a level 12-m walkway with an 8-m measurement region and 2-m approach/exit zones. A pass contains ten analyzed target steps, with anterior spacing equal to the individual's habitual step length, capped at 0.75 m to remain inside the measurement region. Acceleration, deceleration and turning occur outside it. Targets are 8-cm squares centered on the habitual landing line for the relevant foot. The six lateral offsets are $\{-b, -2b/3, -b/3, b/3, 2b/3, b\}$, with $b \leq 6$ cm selected during supervised screening. Reduce b if any position requires crossing the stance foot, discomfort, support or stopping; inability to complete all six positions without assistance excludes the participant from this first implementation.

For a target landing at step i , early-preview targets appear at toe-off initiating step $i - 2$; restricted-preview targets appear at toe-off initiating step $i - 1$. They remain visible until target-limb initial contact at step i . The projector receives event triggers from the contact sensors. Actual cue onset is recorded by a photodiode in the capture clock, permitting preview duration to be analyzed as a measured quantity. A software trigger without an observed light-onset timestamp is insufficient.

Every 40-step acquisition unit consists of four straight ten-step passes. Each pass is a distinct dependence segment; cross-pass covariance is set to zero only for the model, not asserted as a biological fact. Each probe contains eight units, or 320 valid steps, with 160 early- and 160 restricted-preview steps. Preview assignment is balanced within each unit. Target counts differ by at most one within preview across the probe. Sequences are generated before enrollment, archived, and constrained to prevent runs of more than three identical offsets. The unfamiliar-order probe uses unseen target order with the same marginal offsets, target size and preview counts. Distribution-changing transfer tasks are separate secondary assessments.

CAPTURE, PROCESSING AND CALIBRATION

Record optical kinematics at 100 Hz using rigid foot and thorax clusters and instrumented contact sensors synchronized to the same clock. A static calibration locates each heel-center landmark relative to its cluster. Initial contact is the first pressure signal exceeding 10% of that sensor's static standing value for at least 20 ms; toe-off is the reverse crossing sustained for at least 20 ms. The thresholds are protocol definitions and must be checked against synchronized video during feasibility work.

Filter marker positions using a fourth-order, zero-phase Butterworth low-pass filter at 6 Hz, with reflected padding at the beginning and end of each pass. Compute thorax velocity by central differences after filtering. Define landing position as the median heel-center coordinate over the first 50 ms after contact. Repeat the body-state analysis at 50% of swing and immediately before contact as registered timing sensitivities; agreement with the toe-off contrast is assessed, not presumed. Forward speed is the thorax's anterior displacement over the analyzed pass divided by elapsed time.

At every visit, calibrate projected targets against measured walkway points and record a stationary rigid reference before and after acquisition. The acceptance limits are 2 mm for projection error, 2 mm for stationary-reference drift and one 100-Hz sample for cue/capture disagreement. These limits specify instrument acceptance; they are not observed calibration values. A failed check invalidates the visit until corrected. Missing-marker intervals of at most 30 ms may be linearly interpolated outside the 100-ms windows around toe-off and contact. Longer intervals or missing event-adjacent measurements invalidate the affected step.

VISITS, FEEDBACK AND PROBE EXPOSURE

Provide 40 familiarization steps before B_1 . Collect B_1 seven days before practice and B_2 immediately before the first practice visit. A gain trend between those sessions prevents use of a stationary pre-practice reference. Allocate participants to training or an assessment-only group with the same probe exposure and visit schedule. The latter completes no additional target practice, allowing the effect of added practice to be separated from repeated assessment exposure.

The training schedule consists of 640 steps on day 0 and 640 on day 1, with at least 60 s of seated rest after each 80 steps and additional rest on request. Deliver identical task instructions and end-of-unit feedback on median lateral error during practice only. Collect P after a 10-min seated interval following the second practice visit. Collect D and N 72 h later, without intervening target practice, in counterbalanced order. Assessment probes receive the instruction "Place the heel near the target center while walking comfortably; abandon any target if necessary" and no augmented accuracy feedback.

The first ten analyzed steps of each delayed probe measure initial landing accuracy. Gains estimated from all 320 steps describe average control across the probe and may include reacquisition. A prespecified comparison of the first 40-step unit with later units tests gain drift. Probe order enters the analysis of delayed and unfamiliar-order gains. Availability of

the gain change before renewed practice requires compatible first-unit evidence; the full-probe estimate alone cannot establish initial recall.

MISSINGNESS AND SAFETY RECORDS

Trips, rescue contacts, emergency stops and abandoned targets are excluded from the uninterrupted-step regression but counted by participant, probe and preview condition. Record complete-case counts and reason-specific missingness. Refit after including any measurable disturbed steps with an event indicator. Disagreement between fits or differential missingness restricts the conclusion. Footwear, assistance, target amplitude and capture setup are held constant; a change creates a distinct assessment condition. The protocol estimates short-term retention in a bounded flat-ground task, not a change in real-world fall risk.

COMPUTATIONAL EVALUATION

AIM, GENERATING MODEL AND ESTIMATORS

The Monte Carlo study estimates a known position-gain contrast and tests classification under intact coordinates, inconsistent scaling and shared-reference error. All observations were generated from the distributions below. Replication counts and Monte Carlo standard errors follow simulation-study guidance (Morris et al., 2019).

For each session, generate independent 40-step bouts. Three zero-mean Gaussian state variables have stationary covariance

$$\Sigma_Z = \begin{pmatrix} 1 & 0.35 & 0.15 \\ 0.35 & 1 & -0.10 \\ 0.15 & -0.10 & 1 \end{pmatrix}, \quad (5)$$

$$Z_i = 0.35Z_{i-1} + \sqrt{1 - 0.35^2} \eta_i,$$

with $\eta_i \sim N(0, \Sigma_Z)$ and stationary initialization at every bout. Target offsets are independent uniform draws from $\{-6, -4, -2, 2, 4, 6\}$ cm, with $s_q = \sqrt{56/3} = 4.3205$ cm. Set $\tilde{q}$ to offset/ s_q , $\tilde{x} = (Z_1 + 0.2\tilde{q})/\sqrt{1.04}$, $\tilde{v} = (Z_2 - 0.1\tilde{q})/\sqrt{1.01}$ and $\tilde{u} = Z_3$. The reference predictors have unit population variance. Each bout contains a random ordering of exactly 20 early- and 20 restricted-preview steps.

The coefficient vector in Eq. (2), ordered as $(\alpha, \beta_x, \beta_v, \beta_q, \gamma, \delta, \phi, \lambda_x, \lambda_v)$, is

$$(12.0, 0.80, 0.45, 4.3205, 0.30, 0.05, 0.10, 0.20, -0.10). \quad (6)$$

Residuals have stationary SD $\sigma = 0.65$ cm and lag-one autocorrelation 0.40, restarting independently at each bout. The target is $12 + s_q \tilde{q}$ cm relative to stance. Session accuracy is the median absolute placement error; table entries are its mean across replications. Reference accuracy averages the two pre-practice medians. Known population scales remain fixed across sessions, so the experiment excludes scale-estimation uncertainty.

The reference gain is the mean of two independent fitted gains, with variance $(V_{B_1} + V_{B_2})/4$. Each contrast variance adds the probe variance. Because the three contrasts share this reference, their estimates are dependent; the Bonferroni intervals do not require independence. An independent 2,000-replication calibration at each step count determines h from paired unchanged-session fits. The numerical gain resolution is $m_\beta = 0.15$ cm per reference unit, a statistical design value without an established clinical interpretation.

EIGHT GENERATING CONDITIONS AND SAMPLING EXPERIMENT

Each of the eight conditions in Table 2 has 2,000 independent replications at $n = 320$ steps per session. Every replication contains two reference sessions and three probes. Where a gain decrease is present, β_x changes from 0.80 to 0.40 cm per reference unit. The unchanged and retained-change conditions also have 1,000 replications each at 80, 160 and 640 steps. These counts correspond to two, four, eight and sixteen independent 40-step bouts. Their dependence segments differ from the laboratory protocol's ten-step passes; direct step-count planning for that protocol requires measurements with the same pass structure.

The evaluated outcomes are delayed-contrast bias, nominal 95% contrast coverage, joint supported change at P, D, N and the five classification proportions. Classification uses the change and equivalence rules without the coordinate or function gates. The fault conditions therefore quantify false gain signals that a coefficient test alone cannot exclude. Task safety and functional non-inferiority were not simulated.

TABLE 2. GENERATING CONDITIONS AFTER THE TWO REFERENCE SESSIONS. "ALL" DENOTES P, D, N ; RESIDUAL SDS ARE IN CENTIMETRES.

Condition	Generating change
Unchanged mapping	$\beta_x = 0.80, \sigma = 0.65$ throughout.
Precision gain only	$\beta_x = 0.80$ throughout; $\sigma = 0.25$ at all probes.
Transient gain change	$\beta_x = 0.40, \sigma = 0.40$ at P ; $\beta_x = 0.80, \sigma = 0.65$ at D, N .
Retained gain change	$\beta_x = 0.40, \sigma = 0.40$ at all probes.
Context-limited change	$\beta_x = 0.40, \sigma = 0.40$ at P, D ; $\beta_x = 0.80, \sigma = 0.65$ at N .
Range contraction	Multiply true $\tilde{x}, \tilde{v}$ by 0.5 at all probes; coefficients and noise unchanged; retain the pre-practice predictor scales.
Session rescaling fault	Same range contraction, but divide the fitted $\tilde{x}$ and $\tilde{x}r$ by 0.5 at each probe, violating fixed scaling.
Shared-reference fault	Coefficients unchanged; add independent stance-reference error $a_e \sim N(0, 1.5^2)$ cm at each probe, with $p' = p - a_e$, $\tilde{x}' = \tilde{x} - a_e/2$ and $\tilde{q}' = \tilde{q} - a_e/s_q$. Recompute preview interactions.

The random seed is 20261010, with separate seeds for calibration, conditions and step-count tests. The implementation uses NumPy 2.3.5, SciPy 1.17.0 and Matplotlib 3.10.8.

RESULTS

GAIN RECOVERY AND CLASSIFICATION

At 320 steps, the calibrated repeat-difference limit was 0.1664 cm per reference unit, above the gain resolution of 0.15. The retained-change condition yielded a mean delayed contrast of -0.400 cm per reference unit and joint detection in 89.2% of replications (Monte Carlo SE 0.69 percentage points). The remaining 10.8% were indeterminate. Table 3 contains the eight condition summaries; Figure 2 shows delayed-contrast distributions and classification proportions.

TABLE 3. COMPUTED RESULTS AT 320 STEPS PER SESSION AND 2,000 REPLICATIONS PER CONDITION. A_B, A_D ARE MEAN SESSION MEDIAN ERRORS (CM). COVERAGE CONCERNS THE NOMINAL 95% DELAYED CONTRAST INTERVAL AGAINST THE UNCHANGED OR CHANGED GENERATING GAIN. FAULT ROWS ARE INTENTIONAL VIOLATIONS.

Condition	A_B	A_D	Mean Δ_D	Coverage (%)	Joint change (%)
Unchanged mapping	0.904	0.904	0.004	93.5	0.0
Precision gain only	0.903	0.808	-0.001	94.5	0.0
Transient gain change	0.904	0.905	-0.002	94.3	0.0
Retained gain change	0.903	0.614	-0.400	94.0	89.2
Context-limited change	0.904	0.614	-0.400	94.0	0.0
Range contraction	0.905	0.626	0.000	93.5	0.0
Session rescaling fault	0.904	0.622	-0.399	0.1	54.2
Shared-reference fault	0.905	0.901	-0.301	1.8	6.7

No joint change detections occurred in the five correctly measured conditions without a true retained-and-transferred change. Zero of 2,000 detections has an exact one-sided 95% binomial upper bound of approximately 0.15%; it is not proof of a zero false-positive probability. Precision improvement alone reduced mean delayed error from 0.903 to 0.808 cm without changing the generating gain. Range contraction reduced delayed error to 0.626 cm, but fixed-scale gain estimates remained centered near zero change. The error decrease therefore did not identify a position-gain change.

The transient condition returned an explicit transient pattern in 9.6% of replications, with 90.5% indeterminate after rounding. The context-limited condition returned its corresponding pattern in 29.6%, with 70.4% indeterminate. An unchanged generating mapping was declared equivalent at all three probes in only 3.1%; precision gain alone, with lower residual noise, was declared equivalent in 72.5%. The high inconclusive proportions reflect the need to demonstrate equivalence rather than equate a nonsignificant contrast with absence of change. Some rounded categories sum to 100.1%.

Nominal 95% delayed-contrast coverage ranged from 93.5% to 94.5% in the six correctly measured conditions. This undercoverage persisted despite independent repeat-test calibration and the three-probe conjunction. Participant testing requires comparison with segment-based bootstrap uncertainty.

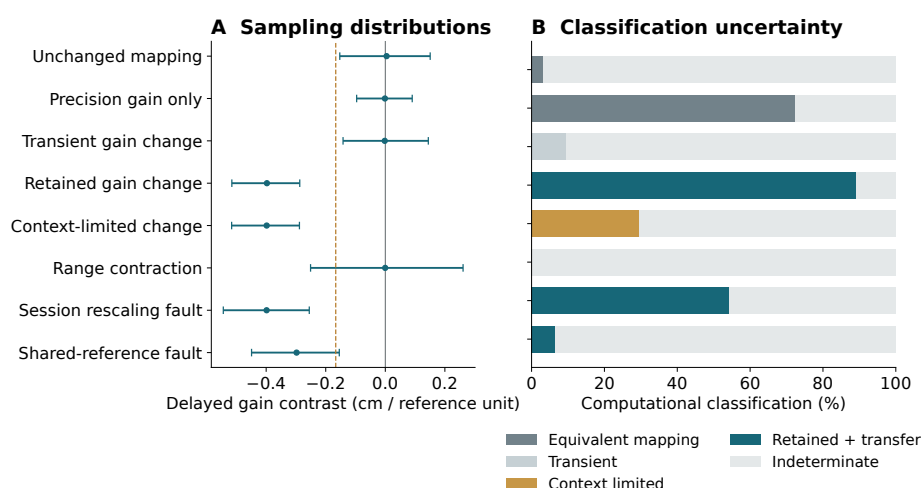


FIGURE 2. COMPUTED GAIN CONTRASTS AND PATTERN PROPORTIONS ACROSS 2,000 REPLICATIONS PER CONDITION. PANEL A SHOWS MEDIANS AND 2.5TH–97.5TH PERCENTILES OF THE DELAYED-CONTRAST SAMPLING DISTRIBUTION, NOT CONFIDENCE INTERVALS FOR A PARTICIPANT. THE DASHED LINE MARKS -0.1664 CM PER REFERENCE UNIT. PANEL B SHOWS THE NUMERICAL CLASSIFICATION BEFORE TECHNICAL AND FUNCTIONAL GATES. FAULT ROWS CAN PRODUCE APPARENTLY RETAINED SIGNALS DESPITE NO TRUE GAIN CHANGE.

MEASUREMENT FAULTS AND SAMPLING BURDEN

When the state range contracted but each probe used a new position scale, the mean apparent delayed gain contrast was -0.399 , and joint change was detected in 54.2%. Shared stance-reference error produced a mean apparent contrast of -0.301 and joint detection in 6.7%. Their Monte Carlo SEs were 1.11 and 0.56 percentage points. Both conditions retained the generating gain of 0.80. Changing the predictor scale fails the coordinate check; shared error can remain undetected by gain intervals alone.

Sampling materially affected detection of the true retained decrease. The calibrated limits were 0.3545, 0.2404, 0.1664 and 0.1500 cm per reference unit at 80, 160, 320 and 640 steps. Joint detection was 0.8%, 18.1%, 89.2% and 100.0%, respectively (Figure 3). The observed zero misses at 640 steps has a one-sided 95% upper bound of approximately 0.30% for the miss probability across its 1,000 replications. These results concern the specified effect, noise, preview balance and dependence; they do not establish a universal step count.

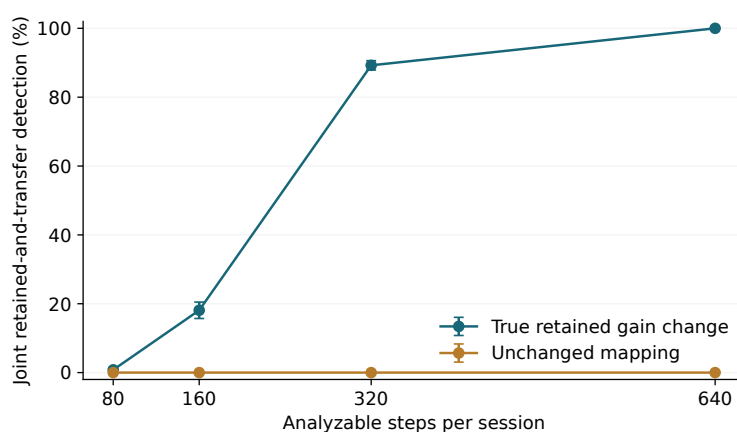


FIGURE 3. DETECTION OF THE RETAINED GAIN DECREASE AT 80, 160, 320 AND 640 STEPS PER SESSION. ERROR BARS SHOW ± 1.96 MONTE CARLO SE. THE 320-STEP POINTS USE 2,000 REPLICATIONS; THE OTHER POINTS USE 1,000. THE TEXT GIVES EXACT ONE-SIDED BOUNDS FOR PROPORTIONS OF ZERO OR ONE.

DISCUSSION

GAIN CHANGE AND PLACEMENT PRECISION

The 0.40-cm gain decrease remained detectable across three probes in 89.2% of 320-step replications. By comparison, reducing residual SD from 0.65 to 0.25 cm improved delayed placement error without producing a joint gain signal.

Halving position and velocity variation also improved placement error while preserving the fitted gain when reference scales remained fixed. The three changes therefore have distinguishable consequences for accuracy and gain estimation. Whether a lower gain improves balance depends on ankle moments and other available strategies (van Leeuwen et al., 2020; Mahaki et al., 2019); that question was not tested by the linear generator.

The precision-only condition reached equivalence at all three probes in 72.5% of replications, compared with 3.1% under unchanged residual noise. Lower noise made the narrow equivalence interval easier to establish. Equivalence concerns the selected gain within ± 0.15 cm per reference unit; it does not exclude learning expressed through movement consistency or another variable. An indeterminate result means the intervals resolve neither the specified change nor the required equivalence pattern.

Retention and generalization can respond differently to gait training and adaptation (Hasson et al., 2015; Leech et al., 2018; Mariscal et al., 2020; Vasudevan et al., 2017). The delayed and unfamiliar-order contrasts retain that distinction for a selected foot-placement gain. Unfamiliar target order on a level walkway tests sequence transfer, with target geometry and preview counts held constant. It does not test obstacle avoidance or everyday mobility. Developmental differences in visually guided stepping require separate measurement in children (Mowbray et al., 2019). Evidence from proprioception training likewise supports separate measurement of acquisition, retention and transfer (Seo et al., 2023).

WHY SHARED REFERENCES CAN CREATE FALSE GAIN CHANGES

Consider the shared-reference fault in one preview stratum with a centered predictor vector z , coefficient vector β and independent reference error a_e with variance τ^2 . If $z' = z - da_e$ and $p' = p - a_e$, population least squares becomes

$$\beta_{\text{obs}} = (\Sigma + \tau^2 dd^T)^{-1}(\Sigma\beta + \tau^2 d), \quad \Sigma = \text{Cov}(z). \quad (7)$$

Thus an unchanged true β can yield a changed observed coefficient when reference-error variance changes. For the position and target variables in the numerical study, the relevant entries of d are $1/2$ and $1/s_q$. Zero mean drift does not prevent this error-induced change. Stationary references, repeat measurements of heel landmarks and an independent reference channel are therefore needed in addition to checking calibration means. The magnitude and covariance of coordinate error must enter the uncertainty assessment.

STEP COUNT, INTERVAL COVERAGE AND LIMITS

Transient and context-limited changes were classified in only 9.6% and 29.6% of replications because their labels require equivalence at the remaining probes. Most replications were indeterminate. Repeat-test calibration must match the instrument, preview balance and analyzed step count. Additional steps improved retained-change detection from 0.8% at 80 steps to 100.0% at 640 steps under the specified stationary distributions. If the gain changes during a probe, a longer record instead averages different relations and requires a time-dependent analysis.

The Gaussian state variables and linear placement equation do not impose a biomechanical position-velocity trajectory. The calculations omit nonlinear coordination, between-person variation, estimated reference scales, adverse events and differential missingness. They test eight specified changes and faults, rather than identifying every cause of a participant gain difference. Clinical error margins and participant benefit remain undetermined.

The 40-step computational bouts differ from the ten-step laboratory passes, and the uniform target draws differ from the constrained target order. Measured pass-to-pass covariance and target-state distributions are therefore needed before setting the laboratory step count. Coverage of nominal 95% intervals was 93.5–94.5% under intact measurement. This undercoverage requires comparison of HAC bandwidths and segment-based bootstrap intervals before individual decisions are used in participant research.

Participant testing must quantify stance-reference error covariance, agreement between thorax- and CoM-based contrasts, repeat-session gain stability and changes between the first and later probe units. The assessment-only group separates additional practice from repeated testing. The specified 72-h probes address short-term retention; persistence over weeks or months requires longer follow-up.

CONCLUSION

Across 2,000 replications at 320 steps, the three-probe gain decision detected a retained decrease from 0.80 to 0.40 cm per reference unit in 89.2% of cases. Precision improvement and correctly scaled range contraction produced no joint detections. Changing the session scale or introducing shared stance-reference error produced false detections in 54.2% and 6.7%. Retained gain estimation therefore requires fixed units, reference-error checks and enough observations to resolve both change and equivalence. These computational findings do not establish a participant training effect or clinical benefit.

DECLARATIONS

CONFLICT OF INTEREST

The author declares that there are no conflicts of interest related to this study.

DATA AVAILABILITY

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

FUNDING

This research received no external funding.

REFERENCES

1. Bruijn SM, van Dieën JH. Control of human gait stability through foot placement. *J R Soc Interface*. 2018;15(143):20170816. doi:10.1098/rsif.2017.0816.
2. Caballero C, Davids K, Heller B, Wheat J, Moreno FJ. Movement variability emerges in gait as adaptation to task constraints in dynamic environments. *Gait Posture*. 2019;70:1–5. doi:10.1016/j.gaitpost.2019.02.002.
3. Charlton JM, Eng JJ, Li LC, Hunt MA. Learning gait modifications for musculoskeletal rehabilitation: applying motor learning principles to improve research and clinical implementation. *Phys Ther*. 2021;101(2):pzaa207. doi:10.1093/ptj/pzaa207.
4. Desmet DM, Cusumano JP, Dingwell JB. Adaptive multi-objective control explains how humans make lateral maneuvers while walking. *PLoS Comput Biol*. 2022;18(11):e1010035. doi:10.1371/journal.pcbi.1010035.
5. Dhawale AK, Smith MA, Ölveczky BP. The role of variability in motor learning. *Annu Rev Neurosci*. 2017;40:479–498. doi:10.1146/annurev-neuro-072116-031548.
6. Dingwell JB, Cusumano JP. Humans use multi-objective control to regulate lateral foot placement when walking. *PLoS Comput Biol*. 2019;15(3):e1006850. doi:10.1371/journal.pcbi.1006850.
7. Hasson CJ, Manczurowsky J, Yen SC. A reinforcement learning approach to gait training improves retention. *Front Hum Neurosci*. 2015;9:459. doi:10.3389/fnhum.2015.00459.
8. Hoogstad LA, van Leeuwen AM, van Dieën JH, Bruijn SM. Can foot placement during gait be trained? Adaptations in stability control when ankle moments are constrained. *J Biomech*. 2022;134:110990. doi:10.1016/j.jbiomech.2022.110990.
9. Kantak SS, Winstein CJ. Learning-performance distinction and memory processes for motor skills: a focused review and perspective. *Behav Brain Res*. 2012;228(1):219–231. doi:10.1016/j.bbr.2011.11.028.
10. Lakens D. Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. *Soc Psychol Personal Sci*. 2017;8(4):355–362. doi:10.1177/1948550617697177.
11. Leech KA, Roemmich RT, Bastian AJ. Creating flexible motor memories in human walking. *Sci Rep*. 2018;8:94. doi:10.1038/s41598-017-18538-w.
12. Maeda RS, O'Connor SM, Donelan JM, Marigold DS. Foot placement relies on state estimation during visually guided walking. *J Neurophysiol*. 2017;117(2):480–491. doi:10.1152/jn.00015.2016.
13. Mahaki M, Bruijn SM, van Dieën JH. The effect of external lateral stabilization on the use of foot placement to control mediolateral stability in walking and running. *PeerJ*. 2019;7:e7939. doi:10.7717/peerj.7939.
14. Mahaki M, van Leeuwen AM, Bruijn SM, van der Velde N, van Dieën JH. Mediolateral foot placement control can be trained: older adults learn to walk more stable, when ankle moments are constrained. *PLoS One*. 2023;18(11):e0292449. doi:10.1371/journal.pone.0292449.
15. Mariscal DM, Iturralde PA, Torres-Oviedo G. Altering attention to split-belt walking increases the generalization of motor memories across walking contexts. *J Neurophysiol*. 2020;123(5):1838–1848. doi:10.1152/jn.00509.2019.
16. Matthis JS, Barton SL, Fajen BR. The biomechanics of walking shape the use of visual information during locomotion over complex terrain. *J Vis*. 2015;15(3):10. doi:10.1167/15.3.10.
17. Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. *Stat Med*. 2019;38(11):2074–2102. doi:10.1002/sim.8086.
18. Mowbray R, Gottwald JM, Zhao M, Atkinson AP, Cowie D. The development of visually guided stepping. *Exp Brain Res*. 2019;237(11):2875–2883. doi:10.1007/s00221-019-05629-5.

19. Pacheco MM, Lafe CW, Newell KM. Search strategies in the perceptual-motor workspace and the acquisition of coordination, control, and skill. *Front Psychol.* 2019;10:1874. doi:10.3389/fpsyg.2019.01874.
20. Ranganathan R, Newell KM. Changing up the routine: intervention-induced variability in motor learning. *Exerc Sport Sci Rev.* 2013;41(1):64–70. doi:10.1097/JES.0b013e318259beb5.
21. Seo HG, Yun SJ, Farrens AJ, Johnson CA, Reinkensmeyer DJ. A systematic review of the learning dynamics of proprioception training: specificity, acquisition, retention, and transfer. *Neurorehabil Neural Repair.* 2023;37(10):744–757. doi:10.1177/15459683231207354.
22. van Leeuwen AM, van Dieën JH, Daffertshofer A, Bruijn SM. Active foot placement control ensures stable gait: effect of constraints on foot placement and ankle moments. *PLoS One.* 2020;15(12):e0242215. doi:10.1371/journal.pone.0242215.
23. Vasudevan EVL, Hamzey RJ, Kirk EM. Using a split-belt treadmill to evaluate generalization of human locomotor adaptation. *J Vis Exp.* 2017;(126):55424. doi:10.3791/55424.