Towards a Digital Twin Framework for Sensory-Aware Neurodivergent Route Personalisation

Vyshnavi Muniganti*, Faiyaz Doctor*‡, Brook Abegaz†,
Mohammad Hossein Anisi*, Charalampos Karyotis‡, Vasiliki Santikou§
*Intelligent Connected Societies Group,
School of Computer Science and Electronic Engineering, University of Essex, Colchester, UK
†Department of Engineering, Loyola University Chicago, Chicago, IL, USA ‡Interactive Coventry Ltd, Coventry, UK
§Interdisciplinary & Research-Based Psychosocial Support for Children and Adults, Irakleio, Attica, Greece
{fdocto, m.anisi}@essex.ac.uk, vyshnavimuniganti252@gmail.com, babegaz@luc.edu,
charalampos.karyotis@interactivecoventry.com, vassilikimail@gmail.com

***Abstract*—Around 90% of autistic adults report atypical sensory processing, which is a known barrier to independent travel. Existing navigation systems optimise for time and distance and do not account for the sensory demands of a route. We present NavTwin, a framework that pairs a Personal Digital Twin (PDT), which models a user's sensitivity profile and preference history, with an Environmental Digital Twin (EDT), which uses a Temporal Graph Network to predict crowd density, noise, and visual complexity along a street graph. We use *digital twin* in a restricted sense: both twins are state-estimation and prediction models, and the EDT is not yet synchronised with live city feeds, so the loop back to the physical environment remains open. A route scorer combines personal preference, environmental comfort, completion probability, and efficiency, with weights adapted online by a Contextual Bandit. Route explanations are produced by a Gemma 3 1B Small Language Model fine-tuned with Retrieval-Augmented Fine-Tuning and QLoRA, so that recommendations are grounded in curated domain documents rather than the model's internal knowledge. Across four synthetic profiles, our adaptive scheme increases simulated journey completion for the high-sensitivity profile by 9.6 percentage points over the strongest static baseline, with the bandit converging in 12–19 journeys. The evaluation is architecture-level, showing that the framework's components behave coherently and lay the grounds for follow-on user studies with neurodivergent participants.**

***Index Terms*—digital twin, neurodivergent navigation, contextual bandits, sensory-aware routing, small language models**

I. Introduction

Independent public-transport use is repeatedly identified as a barrier for autistic adults. In a mixed-methods study, Falkmer et al. [1] found crowding a dominant reported barrier for the autistic group, with anxiety about late or busy services raised throughout. A qualitative review [2] and a personal-experience study [3] both point to sensory overload, unpredictability, and crowded stations as reasons journeys are avoided or abandoned. Reviews of sensory atypicality in autism [4], [5] report that roughly 90% of autistic individuals experience hyper- or hypo-reactivity to tactile, auditory, and visual stimuli, which is the basis for treating crowd density, noise, and visual complexity as the sensory channels of interest.

Conventional intelligent transportation systems (ITS) such as Google Maps and Citymapper optimise for time and distance, treating environments as interchangeable: a route two minutes faster through a crowded interchange is preferred to a calmer alternative, whatever that interchange costs the traveller psychologically. The obstacle is not that a better objective function is unknown, but that the two quantities needed to evaluate one are unavailable to a conventional planner at query time: how sensitive *this* user is, and how loud or crowded a segment will be *when they reach it*. Both are estimation problems, and both change: sensitivity differs between users and drifts with context, as sensory load varies across the day.

From the literature reviewed in Section II we take four requirements for a route personalisation system aimed at neurodivergent travellers. (R1) *Sensory-aware scoring*: crowd, noise, and visual complexity are empirically grounded in autism research and should be scored directly rather than through proxies such as street width or road class. (R2) *Predictive environmental modelling*: sensory peaks such as rush-hour platforms are short-lived, so a route scored on long-run averages is scored on conditions the traveller will not meet. (R3) *Per-user online adaptation*: sensitivity is heterogeneous and the comfort–efficiency trade-off shifts with context, so the system must learn from each user's own journeys, of which there are few. (R4) *Explainability*: acting on unfamiliar journey advice requires trust, so a recommendation that departs from the fastest option has to say why. The requirements were refined with a psychologist specialising in neurodiverse support.

We propose NavTwin, which addresses R1–R4 with a Personal Digital Twin (PDT) that maintains a per-user sensory state, an Environmental Digital Twin (EDT) built on a Temporal Graph Network [6] that predicts the three sensory channels per node over a rolling horizon, a four-dimensional route scorer whose weights are selected online by a contextual bandit [7], [8], and a Small Language Model fine-tuned with Retrieval-Augmented Fine-Tuning [9] that grounds route explanations in domain documents. The contributions are: (i) a dual digital twin decomposition in which per-user sensory estimation and environmental prediction are handled separately, each with its own data source and update rate; (ii) predictive environmental modelling with a Temporal Graph Network, so that routes are scored on forecast rather than average conditions; (iii) four-

979-8-3315-5021-9/26/$31.00 ©2026 IEEE

dimensional route scoring with weights adapted online by a contextual bandit sized for 2–5 journeys per day; (iv) a RAFT-tuned on-device explanation module that cites retrieved passages and abstains when evidence is absent; and (v) a component-level ablation reporting what each part contributes under the simulator's outcome model.

II. Related Work

A. Mobility support for autistic travellers

Rezae et al. [10] co-produced a public-transport planning prototype with autistic participants, with trip-level features such as predictability cues and journey-stage previews; McMahon et al. [11] demonstrated augmented-reality wayfinding for students with intellectual disabilities and autism. Both treat the traveller's experience as part of the problem, and the co-production method of Rezae et al. remains the reference standard here. But neither changes its advice as journey history accumulates, and neither anticipates conditions at the time of travel, leaving R2 and R3 open challenges.

B. Route quality beyond time and distance

A separate line of work scores routes on qualities beyond speed. Quercia et al. [12] crowd-sourced perceptual ratings of London street scenes and showed that routes recommended for beauty, happiness, or quietness are perceived as more pleasant than shortest paths at a small detour cost. Bethge et al. [13] learn emotion-aware route trajectories from naturalistic driving data. Both establish that non-temporal criteria are learnable and that travellers accept a detour based on them. Both, however, treat route quality as a property of the street rather than of the street–traveller pairing: a segment rated quiet is rated quiet for everyone, and the rating cannot vary with an individual traveller's sensitivity, which is what R3 requires.

C. Digital twins in urban mobility

Digital twins are data-driven representations of a physical entity coupled to it by data exchange [14], [15]; Aloupogianni et al. [16] demonstrate an urban-traffic twin integrating multi-source feeds on a Singapore testbed. These infrastructures do synchronise with live data, but the state they track is the transport network rather than any individual traveller, so per-user sensory tolerances have no representation in them.

D. Personalisation under a scarce per-user data rate

Contextual bandits address exploration and exploitation when rewards depend on observable context: Li et al. [7] introduced LinUCB, Chu et al. [8] established regret bounds for linear payoffs, and Thompson sampling offers a Bayesian alternative [17]. For per-user route personalisation the binding constraint is sample efficiency rather than model capacity: a traveller makes 2–5 journeys per day, so any policy class that needs thousands of episodes cannot be trained on one user's data. This rules out deep reinforcement learning here, whatever its asymptotic advantages. To our knowledge, bandit personalisation has not been applied to sensory-aware routing.

TABLE I
Prior work against R1–R4. ✓ supported, ✗ not, ∼ partial.

System / line of work	R1 sensory	R2 predict.	R3 per-user	R4 explain
Conventional ITS planners	✗	∼	✗	✗
Autism trip planner [10]	∼	✗	✗	∼
AR wayfinding [11]	✗	✗	✗	✗
Perceptual routing [12]	∼	✗	✗	✗
Emotion-aware routing [13]	∼	✗	∼	✗
Urban traffic DT [16]	✗	✓	✗	✗
Graph forecasting [6], [18]	✗	✓	✗	✗
Bandit personalisation [7]	✗	✗	✓	✗
NavTwin (this paper)	✓	✓	✓	✓

E. Predicting conditions on a transit graph

Traffic forecasting supplies the machinery for R2: DCRNN [18] models traffic as diffusion on a directed sensor graph, and spatial-temporal graph convolution [19] convolves jointly over space and time. Both assume a fixed topology in which only node and edge features vary. This holds for road sensor networks but not for transit, where a disruption removes an edge and a closure isolates a sub-graph. Temporal Graph Networks [6] extend graph learning to continuous-time dynamic graphs through per-node memory, and are used here for that reason.

F. Grounded explanation from compact models

Retrieval-augmented generation conditions a model on retrieved passages [20]; RAFT [9] fine-tunes the model to cite those passages in its answers and to abstain when no answer-bearing passage is retrieved, which is the behaviour R4 requires. With parameter-efficient tuning [21], [22], a 1B model [23] becomes deployable on devices such as smartphones.

G. Position of this work

Table I summarises this literature against R1–R4. No prior system satisfies all four jointly, and the requirements pull against one another: R1 needs a per-user sensitivity estimate that a population-level rating cannot supply; R2 exists in the forecasting literature but has not been connected to a per-user comfort model; and R3, at 2–5 journeys per day, rules out the policy classes that would otherwise absorb R1 and R2 into one learned objective. The contribution is therefore an integration claim rather than an algorithmic one: a decomposition in which a sample-efficient bandit adapts the scoring weights online, while per-user and environmental estimation are handled separately, each with its own data source and update rate. In the rest of this paper: Section III covers the dual-twin architecture and its orchestration; Section IV the route scorer, the contextual bandit, and the RAFT-tuned explanation module; Section V the simulation-based validation; Section VI the limitations and future work; Section VII concludes.

III. System Architecture

NavTwin arranges its modules across four layers coordinated by a publish/subscribe orchestrator (Fig. 1); Table II gives each module's inputs, learned parameters, output, and requirement. Subscript t denotes the journey index, r a route, i a segment.

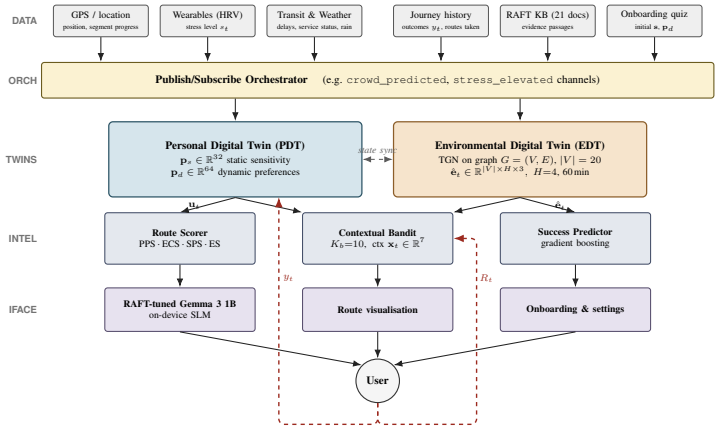

Fig. 1. NavTwin's four-layer architecture. Data sources (top) feed a publish/subscribe orchestrator. In the TWINS layer, the PDT models the user ($\mathbf{p}_s$, $\mathbf{p}_d$) and the EDT models the city with a Temporal Graph Network producing predictions $\hat{\mathbf{e}}_t$ over the three sensory channels. The INTEL layer scores and selects routes; the IFACE layer delivers the route with a grounded SLM explanation. After each journey, y_t updates $\mathbf{p}_d$ and R_t (Eq. 6) updates the bandit (red dashed). The onboarding quiz seeds the PDT against cold-start (Section III-B); the EDT is not synchronised with live city feeds, so its loop to the physical environment is open (Section III-A).

TABLE II
NAVTWIN MODULES: INPUTS, LEARNED PARAMETERS, OUTPUTS, REQUIREMENT SERVED, AND THE SIMPLER ALTERNATIVE EACH WAS CHOSEN OVER. STARRED ROWS ARE ABLATED IN TABLE IV.

Module	Inputs	Learned parameters	Output	Req.	Simpler alternative and why it was not used
PDT	onboarding responses; outcomes y_t; route features $\mathbf{f}_r$	$\mathbf{p}_d$; GRU and cross-attention weights ⋆	$\mathbf{u}_t = [\mathbf{p}_s; \mathbf{p}_d]$; PPS	R1, R3	Questionnaire-only fixed sensitivity vector: cannot represent drift between re-calibrations.
EDT ⋆	graph $G = (V, E)$; channel history; time of day	node memories; temporal attention	$\hat{\mathbf{e}}_t \in \mathbb{R}^{\lvert V\rvert \times H \times 3}$	R2	Long-run per-node means: averages out the short-lived peaks R2 exists to capture.
Success predictor ⋆	$\mathbf{f}_r$, $\mathbf{u}_t$, $\hat{\mathbf{e}}_t$	gradient-boosting ensemble	$\Pr(\text{completion})$	R1, R3	Scoring without SPS: removes the only direct estimate of whether the user will complete the journey.
Contextual bandit ⋆	context $\mathbf{x}_t$; reward R_t	value estimates over K_b=10 arms	weights $\mathbf{w}$ on the 4-simplex	R3	Fixed weights: cannot express the same user wanting a different trade-off under time pressure.
Explanation SLM	selected route; score components; retrieved passages	rank-32 LoRA adapters (26.1 M)	cited natural-language explanation	R4	Templates: cannot answer free-form follow-ups or abstain when evidence is missing.

On a query, three operations run in parallel: the EDT predicts conditions on the candidate segments over the next 15–60 minutes, the PDT retrieves the current user state, and a stress estimator returns a scalar $s_t \in [0, 1]$ from a wearable heart-rate-variability (HRV) sensor and a brief self-report. Candidates are scored as in Section IV and the selected route is presented with an explanation; after the journey, outcome feedback updates the PDT, the bandit, and the success predictor.

A. Scope of the digital twin claim

The term *digital twin* has an established meaning, and this architecture meets only part of it. On the standard characterisation [14], [15] a twin requires (a) a computational representation of a physical entity, (b) synchronisation with that entity from live data, and (c) a return path by which the twin influences it. The PDT satisfies all three: it represents the traveller, updates from every journey outcome, and steers the routes the traveller is offered. The EDT satisfies (a) but not (b): it runs on a synthetic graph with calibrated temporal patterns rather than live feeds, so the environmental half is a predictive model of a city, not a synchronised replica. The term is kept because the PDT is a twin in the full sense and the EDT's interfaces are specified for the live-feed case, but *dual digital twin* should accordingly be read as a per-user state estimator coupled to an infrastructure-level predictive model. Closing the synchronisation loop with live data is future work (see Section VI).

B. Personal Digital Twin (PDT)

Sensitivity has a stable part and a context-dependent part. The baseline is stable: someone who finds crowds difficult finds them difficult on most journeys. How far that person will trade time for calm on a given day is not, shifting with stress, fatigue, and circumstance. A single vector updated at one rate cannot hold both: updated often, it erases the baseline; updated rarely, it lags the drift. The PDT therefore stores them separately. A *static sensitivity profile* $\mathbf{p}_s \in \mathbb{R}^{32}$ encodes baseline hyper- and hypo-reactivity; its first three entries are the sensitivities to the three sensory channels,

$$\mathbf{s} = [s_c,\ s_n,\ s_v]^\top \in [0, 1]^3, \tag{1}$$

i.e. crowd, noise, and visual complexity, used in Eq. (4). A *dynamic preference embedding* $\mathbf{p}_d \in \mathbb{R}^{64}$ captures faster-moving preferences such as willingness to trade time for calm, updated by gradient descent after each journey from a feedback score $y_t \in \{-1, 0, +1\}$ (-1 abandoned, 0 followed without feedback, $+1$ explicit positive).

Cold-start is handled by seeding the PDT at sign-up from a structured onboarding questionnaire, developed with the collaborating psychologist, that maps everyday sensory tolerances onto initial $\mathbf{s}$ and $\mathbf{p}_d$. Thereafter $\mathbf{p}_s$ changes only on re-calibration (weekly or longer) while $\mathbf{p}_d$ refines online.

The two parts concatenate into a user state $\mathbf{u}_t = [\mathbf{p}_s; \mathbf{p}_d] \in \mathbb{R}^{96}$. The Personal Preference Score (PPS) for a route r with features $\mathbf{f}_r$ applies cross-attention [24]: $\mathbf{u}_t$ projects to a query

$Q = W_q \mathbf{u}_t$ attending over keys and values K, V formed from the GRU-encoded history $h_{1:t}$ and $\mathbf{f}_r$, followed by a two-layer MLP with sigmoid output:

$$\mathbf{z} = \mathrm{CrossAttn}\big(Q,\ K,\ V\big),$$
$$\mathrm{PPS}(r, \mathbf{u}_t) = \sigma\big(\mathrm{MLP}(\mathbf{z})\big) \in [0, 1], \tag{2}$$

where $\mathbf{z}$ is the attended history representation. Here $\mathbf{p}_s$ shapes which past history items are emphasised (through W_q), while $\mathbf{p}_d$ accounts for short-term preference drift.

C. Environmental Digital Twin (EDT)

R2 is a forecasting requirement: the scorer needs the sensory load on a segment at the time the traveller will reach it, which on a 40-minute journey may be half an hour after the query. A long-run average cannot supply this, and it fails worst where it matters most. A platform that is calm for twenty-three hours a day and severe during rush hour has an unremarkable mean, so the peak the system exists to route around is the one the average hides.

The EDT models urban conditions using a Temporal Graph Network [6] on a spatial graph $G = (V, E)$. Vertices are locations and edges are movement relationships weighted by typical traversal time. At query t the EDT outputs

$$\hat{\mathbf{e}}_t \in \mathbb{R}^{|V| \times H \times 3}, \tag{3}$$

containing predictions over a horizon of $H{=}4$ 15-minute windows (so up to 60 minutes ahead) for the three sensory channels (crowd density, noise level, visual complexity), each normalised to $[0, 1]$. Temporal attention captures rush-hour and lunch-hour patterns while spatial message passing propagates between neighbours (Section II-E motivates TGN over fixed-topology alternatives). This implementation uses a synthetic 20-node graph with weekday patterns calibrated to published pedestrian-flow data; a real deployment would substitute these with live feeds such as GTFS-RT (General Transit Feed Specification–Realtime) and crowd sensing.

D. Orchestration

Modules do not call each other directly; each publishes to named topic channels and subscribes to those it needs. This lets the PDT and EDT run in parallel on a query, and lets the system degrade gracefully rather than fail when one module is unavailable.

IV. Methodology

This section specifies the three decisions made on each query: how a route is scored, how the weights are set for the current user and context, and how the choice is explained.

A. Multi-Criteria Route Scoring

The score combines quantities of different kinds: a learned preference, a physical prediction weighted by an individual tolerance, a probability, and a normalised cost. They are kept as four separate terms rather than folded into one learned objective for two reasons: each term stays individually inspectable, which is what the explanation module of Section IV-C cites when justifying a recommendation (R4); and the only quantity learned online from the user's own journeys is the weight vector, which is what makes adaptation feasible at a few journeys per day (R3).

TABLE III
Composition of the 47-dimensional route feature vector and the score each group primarily feeds.

Group	Dim.	Contents	Feeds
Environmental	12	channel means, maxima along r; EDT confidence	ECS
Personal pref.	8	alignment of r with the PDT preference state	PPS
Route charact.	15	length, segment count, modes, elevation	PPS, SPS
Success-related	8	transfers, dwell time, schedule reliability	SPS
Efficiency	4	travel time, walking distance, reliability	ES

Each candidate route r carries a 47-dimensional feature vector, grouped in Table III; each route is scored on the four dimensions below.

a) Personal Preference Score (PPS): Given by Eq. (2); reflects learned per-user preferences from the PDT.

b) Environmental Comfort Score (ECS): Given $\mathbf{s}$ from Eq. (1) and EDT predictions from Eq. (3),

$$\mathrm{ECS}(r) = 1 - \frac{1}{|r|} \sum_{i \in r} \frac{s_c c_i + s_n n_i + s_v v_i}{s_c + s_n + s_v}, \tag{4}$$

where c_i, n_i, v_i are predicted segment values from $\hat{\mathbf{e}}_t$. A higher ECS means predicted load is well matched to the user's tolerance; the form is deliberately simple so the explanation module can state it to the user.

c) Success Probability Score (SPS): A gradient-boosting classifier over r, the PDT vectors, and $\hat{\mathbf{e}}_t$, trained on historical completion versus abandonment; it returns the estimated probability that the user reaches the destination on this route.

d) Efficiency Score (ES): $\mathrm{ES}(r) = \frac{1}{3}[\tilde{\tau}(r) + \tilde{d}(r) + \tilde{\rho}(r)]$, a convex combination of travel time, walking distance, and schedule reliability, each min–max normalised over the candidate set so that 1 is the best value.

The composite score is the weighted sum

$$\mathrm{Score}(r) = w_1 \mathrm{PPS} + w_2 \mathrm{ECS} + w_3 \mathrm{SPS} + w_4 \mathrm{ES}, \tag{5}$$

with $w_i \in [0, 1]$ and $\sum_i w_i = 1$. On each query the bandit picks $\mathbf{w}$ from context $\mathbf{x}_t$ (Section IV-B), the EDT produces $\hat{\mathbf{e}}_t$, the PDT produces $\mathbf{u}_t$, and Eq. (5) is evaluated per candidate; the highest-scoring route is returned with its explanation.

B. Contextual Bandit for Weight Adaptation

The weights cannot be fixed at design time: the right trade-off differs between users, and between journeys for one user, since a traveller who ordinarily accepts a detour for a calmer route will not accept it when running late. Nor can they be elicited directly, as users are poorly placed to state a four-way numeric trade-off. That leaves learning them from journey outcomes, which at the data rates of Section II-D means a contextual bandit [7]. The context $\mathbf{x}_t$ encodes five signals: stress $s_t \in [0, 1]$; a binary time-pressure flag; a one-hot weather category; the fraction of the last five journeys completed; and time of day as the cyclical pair $(\sin 2\pi h/24, \cos 2\pi h/24)$.

The action space comprises $K_b = 10$ weight configurations on the 4-simplex, from efficiency-prioritising ($w_4 = 0.55$) to

comfort-prioritising ($w_2 = 0.45$); K_b is the arm count, distinct from the key matrix K in Eq. (2). The arm count trades expressiveness against sample efficiency: at 2–5 journeys per day, a set of ten arms can be learned within days of use, whereas a finer grid over the simplex would take weeks to distinguish.

The reward after journey t combines three signals,

$$R_t = \alpha\, r_t^{\text{accept}} + \beta\, r_t^{\text{complete}} + \gamma\, r_t^{\text{stress}}, \tag{6}$$

where $r_t^{\text{accept}} \in \{-0.5, +1\}$ records whether the user followed the route, $r_t^{\text{complete}} \in \{-2, +2\}$ whether the journey completed, and r_t^{stress} the negative change in measured stress, with $\alpha = 0.3$, $\beta = 0.5$, $\gamma = 0.2$ and ϵ-greedy exploration ($\epsilon_0 = 0.3$, decay 0.95): completion is the primary objective, acceptance a proxy for trust, stress change a physiological signal.

C. RAFT-Based Conversational Module

A recommendation slower than the obvious alternative will be ignored unless the reason is given, and R4 exists because trust conditions whether autistic travellers act on unfamiliar advice. The requirement is not fluent text, which a template supplies: it is an explanation that reports the score components behind this particular choice, answers follow-ups the template did not anticipate, and declines to answer rather than invent when it has no supporting evidence. The last of these is why an unmodified language model is unsuitable, and why the model is fine-tuned to ground its explanations in a curated knowledge base. The base is Gemma 3 1B-it [23], chosen for on-device use: quantised to 4 bits (≈0.6 GB) with a ≈26 MB adapter it fits on a smartphone, so inference sends no journey data to a server. Fine-tuning uses RAFT [9] with QLoRA [22], rank-32 adapters (α=32, dropout 0.05) on all linear layers over an NF4 frozen base. The training set draws on 21 domain documents covering sensory processing in autism, the scoring dimensions, and the architecture itself, from which 5,000 examples were built, each pairing a navigation question with 1–4 retrieved passages and a chain-of-thought [25] answer citing them. At oracle ratio P=0.8, 80% of examples include the answer-bearing passage among up to three distractors and 20% contain distractors only; the latter is what trains the model to refuse rather than fabricate. One epoch takes 137.6 min (AdamW-8-bit, cosine schedule, peak 2×10^{-4}, effective batch 16) over 26.1 M trainable parameters (3.79% of the 689 M base), reaching evaluation loss 0.0650.

V. Simulation-Based Validation

This section asks three questions: how quickly the bandit converges, whether the learned weights track the sensitivity profile they are given, and what each component contributes. All three are internal-consistency questions, for which simulation is the right instrument: synthetic profiles span the sensitivity space in a controlled way, support reproducible ablations that a small heterogeneous cohort could not, and exercise the machinery before any real participant is exposed to experimental recommendations.

TABLE IV
LEFT: SIMULATED JOURNEY COMPLETION RATE (%) PER SYNTHETIC PROFILE, Δ AGAINST THE BEST STATIC BASELINE. RIGHT: ABLATION ON THE HIGH-SENSITIVITY PROFILE.

Method	High-s.	Mod.	Low-s.	Var.
Efficiency-Static	68.5	75.0	**92.5**	72.0
Comfort-Static	82.0	79.5	78.0	77.5
Equal-Static	73.5	78.0	86.0	74.0
NavTwin (adaptive)	**91.6**	**87.5**	91.0	**85.0**
Δ vs. best static	+9.6	+8.0	–1.5	+7.5

Ablation, high-sensitivity profile	Compl.	Δ
Full NavTwin	91.6	—
– bandit (equal weights)	86.5	–5.1
– EDT (static environment)	84.3	–7.3
– cross-attention (GRU only)	87.8	–3.8
– SPS (3-dim scoring)	88.2	–3.4

A. Setup

Four synthetic profiles span the sensitivity space, with $\mathbf{s} = (s_c, s_n, s_v)$: *high-sensitivity* $(0.9, 0.8, 0.7)$, *moderate* $(0.6, 0.5, 0.4)$, *low-sensitivity* $(0.3, 0.2, 0.2)$, and *variable* $(0.7, 0.6, 0.5)$, context-dependent; the values were set with the collaborating psychologist. Each profile completes 200 journeys on a 20-node synthetic graph with weekday crowd and noise patterns. Whether a journey completes is decided segment by segment: each segment's sensory demand, plus additive noise, is compared against the profile's tolerance threshold, and abandonment probability grows with the margin by which demand exceeds it.

B. Baselines

Three static configurations fix $\mathbf{w}$ at representative points, so adaptation is the only variable: Efficiency-Static $[0.10, 0.10, 0.25, 0.55]$ (conventional ITS-style), Comfort-Static $[0.25, 0.45, 0.20, 0.10]$, and Equal-Static $[0.25, 0.25, 0.25, 0.25]$.

C. Bandit Convergence

The most-selected arm stabilises between journeys 12 and 19 (low-sensitivity 12, moderate 14, high-sensitivity 17, variable 19). High-sensitivity converges to a comfort-prioritising arm ($w_{\text{ECS}} = 0.42$) and low-sensitivity to an efficiency-prioritising one ($w_{\text{ES}} = 0.51$). The variable profile alternates with context rather than settling, the intended behaviour where effective sensitivity changes between journeys.

D. Journey Completion and Ablation

Table IV reports simulated completion per profile under the outcome model of Section V-A. The adaptive bandit is highest on three of four profiles, with the largest gap on high-sensitivity (+9.6 points). Low-sensitivity is the one case a static policy wins (-1.5 points against Efficiency-Static), as expected: that profile sits almost exactly on the fixed efficiency-first weighting, so adaptation can only pay exploration cost.

The lower block removes one component at a time; this is the evidence on which the necessity of each component rests. All four contribute, and the predictive EDT is the largest single contributor, losing 7.3 points when $\hat{\mathbf{e}}_t$ is replaced by static long-run means: forecasting matters more than any other element.

E. Qualitative SLM Demonstration

On five route-explanation prompts with identical retrieved context, the RAFT-tuned Gemma 3 1B averaged 4.4 explicit citations per response against zero for the unmodified base, with shorter outputs (188.8 versus 349.2 words) and lower latency (117.4 versus 159.4 s). With five prompts and a single rater no statistical claim can be made; the demonstration establishes only that the trained citation and abstention behaviours are present.

F. Operational walkthrough

An example query traces the pipeline as follows. At 08:15 a high-sensitivity user ($\mathbf{s} = [0.9, 0.8, 0.7]$, stress $s_t = 0.62$, time pressure on) requests a route. The EDT forecasts a morning peak of $(0.84, 0.71, 0.55)$ on the fastest candidate route R_1 against $(0.28, 0.22, 0.31)$ on the alternative R_2; the PDT supplies $\mathbf{u}_t$; the bandit selects a comfort-prioritising arm $\mathbf{w} = [0.18, 0.42, 0.30, 0.10]$; and Eq. (5) ranks R_2 (0.81) above R_1 (0.52) despite the longer travel time. The SLM explains the choice by citing the forecast crowd level at the interchange; after the journey, y_t updates $\mathbf{p}_d$ and R_t updates the bandit. The route selection comes from the forecast rather than from any static property of R_1. If it were scored on long-run means, R_1 would be selected.

VI. Discussion, Limitations, and Future Work

a) What the results do and do not show: The architecture learns sensitivity-appropriate policies within a realistic journey budget, and every component contributes under the simulator's outcome model. Three limits bound this claim. First, the outcome model shares its sensory dimensions with the scorer (on non-identical values), so the study tests whether the learning machinery finds appropriate policies under an internally consistent model, not whether they transfer to people: no evidence is offered that the four scores correlate with experienced comfort, stress, or abandonment; establishing that correspondence is the main open question. Second, every baseline holds its weights fixed, so the comparison shows that adapting helps but not that this decomposition is the best way to adapt; separating the two requires a learning baseline such as LinUCB or Thompson sampling [8], [17]. Third, the EDT trains on synthetic patterns; live feeds would introduce sensing noise that the presented results do not reflect.

b) The completion metric: Completion versus abandonment is a binary end-of-journey signal: it cannot distinguish a comfortable completion from a distressed one, or credit a near-completion over an early failure, and the simulation does not model conditions changing *during* a journey. A deployment needs in-journey signals, such as HRV trends and segment-level check-ins, to locate *where* abandonment occurs and reroute from there.

c) Future work: A user study under institutional ethics approval is the necessary next step, testing whether the four scores track reported comfort and observed abandonment. In parallel: in-journey metrics with dynamic rerouting; live GTFS-RT and crowd-sensing feeds, closing the EDT loop of Section III-A; and benchmarking against LinUCB and Thompson sampling.

VII. Conclusion

NavTwin couples a per-user sensory state estimator with a predictive environmental model and adapts four-dimensional route scoring online through a contextual bandit, against requirements R1–R4 from the autism-mobility literature. Across 800 simulated journeys it improves completion by 9.6 percentage points over the strongest static baseline for the high-sensitivity profile, converges in 12–19 journeys, and every component contributes. The result is architecture-level: it shows the parts working together under a controlled outcome model, not that the system helps autistic travellers. The user study of Section VI is the step that would establish that.

References

[1] M. Falkmer *et al.*, "Viewpoints of adults with and without Autism Spectrum Disorders on public transport," *Transp. Res. Part A*, vol. 80, pp. 163–183, 2015.

[2] L. A. Cameron, R. L. Borland, B. J. Tonge, and K. M. Gray, "Community participation in adults with autism: A qualitative review," *J. Appl. Res. Intellect. Disabil.*, vol. 35, no. 2, pp. 366–377, 2022.

[3] H. Dirix *et al.*, "Autism-friendly public bus transport: A personal experience-based perspective," *Autism*, vol. 27, no. 5, pp. 1219–1234, 2023.

[4] C. E. Robertson and S. Baron-Cohen, "Sensory perception in autism," *Nat. Rev. Neurosci.*, vol. 18, no. 11, pp. 671–684, 2017.

[5] T. Kim, J. Lee, J. Lee, H. Jo, Y. Oh, Y. J. Kim, and J. Seo, "Sensory abnormalities in autism spectrum disorder and their in vitro modeling," *Transl. Psychiatry*, vol. 15, art. 534, 2025.

[6] E. Rossi, B. Chamberlain, F. Frasca, D. Eynard, F. Monti, and M. Bronstein, "Temporal graph networks for deep learning on dynamic graphs," arXiv:2006.10637, 2020.

[7] L. Li, W. Chu, J. Langford, and R. E. Schapire, "A contextual-bandit approach to personalized news article recommendation," in *Proc. 19th Int. Conf. World Wide Web (WWW)*, 2010, pp. 661–670.

[8] W. Chu, L. Li, L. Reyzin, and R. E. Schapire, "Contextual bandits with linear payoff functions," in *Proc. 14th Int. Conf. Artif. Intell. Statist. (AISTATS)*, 2011, pp. 208–214.

[9] T. Zhang *et al.*, "RAFT: Adapting language model to domain specific RAG," arXiv:2403.10131, 2024.

[10] M. Rezae, D. McMeekin, T. Tan, A. Krishna, H. Lee, and T. Falkmer, "Public transport planning tool for users on the autism spectrum," *Disabil. Rehabil. Assist. Technol.*, vol. 16, no. 2, pp. 177–187, 2021.

[11] D. D. McMahon, D. F. Cihak, and R. E. Wright, "Augmented reality as a navigation tool to employment opportunities for postsecondary education students with intellectual disabilities and autism," *J. Res. Technol. Educ.*, vol. 47, no. 3, pp. 157–172, 2015.

[12] D. Quercia, R. Schifanella, and L. M. Aiello, "The shortest path to happiness: Recommending beautiful, quiet, and happy routes in the city," in *Proc. 25th ACM Conf. Hypertext and Social Media*, 2014, pp. 116–125.

[13] D. Bethge *et al.*, "HappyRouting: Learning emotion-aware route trajectories for scalable in-the-wild navigation," arXiv:2401.15695, 2024.

[14] M. Grieves and J. Vickers, "Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems," in *Transdisciplinary Perspectives on Complex Syst.* Springer, 2017, pp. 85–113.

[15] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, "Digital twin in industry: State-of-the-art," *IEEE Trans. Ind. Informat.*, vol. 15, no. 4, pp. 2405–2415, 2019.

[16] E. Aloupogianni, F. Doctor, C. Karyotis, T. Maniak, R. Tang, and R. Iqbal, "An AI-based digital twin framework for intelligent traffic management in Singapore," in *Proc. Int. Conf. Elect., Comput. Energy Technol. (ICECET)*, 2024, pp. 1–6.

[17] O. Chapelle and L. Li, "An empirical evaluation of Thompson sampling," in *Adv. Neural Inf. Process. Syst.*, 2011, pp. 2249–2257.

[18] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," in *Proc. Int. Conf. Learn. Representations (ICLR)*, 2018.

[19] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," in *Proc. 27th Int. Joint Conf. Artif. Intell. (IJCAI)*, 2018, pp. 3634–3640.

[20] P. Lewis *et al.*, "Retrieval-augmented generation for knowledge-intensive NLP tasks," in *Adv. Neural Inf. Process. Syst.*, 2020, pp. 9459–9474.

[21] E. J. Hu *et al.*, "LoRA: Low-rank adaptation of large language models," in *Proc. Int. Conf. Learning Representations (ICLR)*, 2022.

[22] T. Dettmers, A. Pagnoni, A. Holtzman, and L. Zettlemoyer, "QLoRA: Efficient finetuning of quantized LLMs," in *Adv. Neural Inf. Process. Syst.*, 2023.

[23] Gemma Team, "Gemma 3 technical report," arXiv:2503.19786, 2025.

[24] A. Vaswani *et al.*, "Attention is all you need," in *Adv. Neural Inf. Process. Syst.*, 2017.

[25] J. Wei *et al.*, "Chain-of-thought prompting elicits reasoning in large language models," in *Adv. Neural Inf. Process. Syst.*, 2022.