AGILE: GIScience Series, 7, 5, 2026. https://doi.org/10.5194/agile-giss-7-5-2026
Proceedings of the 29th AGILE Conference on Geographic Information Science, 16–19 June 2026.
Eds.: Evelyn Uuemaa, Alexander Kmoch, Holger Virro, and Simon Scheider
This contribution underwent peer review based on a full paper submission.

Enhancing the Potential of a Tangible-Digital Planning Interface through User Evaluation

Vincent Holtorf [1], Maria Moleiro Dale [1], and Jörg Rainer Noennig [1]

[1]Digital City Science, HafenCity University, Hamburg, Germany

Correspondence: Vincent Holtorf (vincent.holtorf@hcu-hamburg.de)

Abstract. Participatory urban planning increasingly relies on digital technologies to democratise geospatial information access, yet traditional GIS-based interfaces often exclude non-expert users. Tangible User Interfaces (TUIs) offer a promising alternative by combining intuitive physical interaction with computational analysis. Despite this advancement, the systematic evaluation of how diverse user groups experience such systems remains limited. This study addresses this gap through framework analysis of the feedback from 58 participants across 12 interdisciplinary groups within the academic context who interacted with COUP (Cockpit for Collaborative Urban Planning), a hybrid, tangible-digital interface integrating physical city models with real-time environmental simulations. Users consult and modify district designs while observing impacts on wind comfort, noise propagation, and pedestrian flows through projected visualisations. Mixed-methods analysis identified four consensus strengths: real-time interactivity (91.7%), collaboration support (83.3%), visualisation quality (83.3%), and multi-parameter integration (75.0%). Significant limitations also emerged regarding data accuracy concerns (75.0%), performance issues (66.7%), and technical instability (58.3%), revealing tensions between appreciated potential and practical adoption barriers. Findings provide empirical evidence on how tangible geospatial interfaces translate complex environmental data into actionable planning insights accessible to diverse user groups. The conducted analysis, with a small sample size, is considered a first approach to the enhancement of user evaluation of digital tools for urban planning. The collected recommendations shed light for advancing participatory tools that support evidence-based environmental decision-making, essential for building climate-responsive cities while genuinely empowering user engagement in collaborative urban planning processes.

Submission Type. analysis, case study, software

BoK Concepts. [DA1] System Design, [GS3] Use of geospatial information, [GS4] Geospatial citizenship

Keywords. Tangible User Interfaces, Participatory Planning, User Evaluation, Environmental Simulation, Public Participation GIS

1 Introduction

Contemporary urban development faces unprecedented challenges at the intersection of climate adaptation, demographic growth, and governance complexity. Cities must simultaneously address environmental sustainability, social equity, and economic viability while engaging diverse stakeholder groups in decision-making processes (Brown and Kyttä, 2018). Traditional planning methods often struggle to accommodate these complexities, as static representations and expert-driven workflows limit stakeholder engagement and fail to capture the dynamic nature of urban systems (Falco and Kleinhans, 2018).

Digital technologies have emerged as essential tools for addressing these challenges, supporting more transparent, collaborative, and data-driven planning processes. However, the technical complexity of some of these technologies, for example GIS interfaces as those more closely related to spatial and cartographic considerations, can exclude users lacking specialised training, limiting profound collaborative engagement (Haklay and Tobón, 2003).

Tangible User Interfaces (TUIs) offer a promising response to these participation barriers by combining physical interaction with digital computation. In urban planning contexts, geospatial TUIs employ physical objects placed on interactive surfaces to interact with digitally-mapped data and trigger real-time simulations (Maquil et al., 2015, 2018). This hybrid approach leverages the intuitive affordances of physical deliberation of a plan while maintaining the analytical power of computational geospatial systems, potentially lowering cognitive barriers and facilitating more inclusive decision-making (Ishii and Ullmer, 1997; Hornecker and Buur, 2006). Research on tangible planning interfaces has demonstrated their capacity to support collaborative exploration, enhance spatial understanding, and promote dialogue across stakeholder groups with varying expertise levels (Jones and Maquil, 2016; Brüggemann et al., 2023). The physical-digital coupling creates graspable representations that ground abstract data in familiar physical interactions.

However, a persistent challenge lies in integrating these technically sophisticated simulation outputs into adequate formats for collaborative planning contexts. Complex heatmaps, three-dimensional pollution plumes, and multi-layered GIS analyses often remain opaque to non-experts, limiting their utility in participatory settings. This challenge underscores the need for user interfaces which mediate between computational rigor and experiential understanding, precisely the domain where tangible interaction design may contribute significantly. Despite growing implementation of participatory planning tools and tangible geospatial interfaces, systematic evaluation of how users perceive and experience these systems remains limited. Many cases focus on technical feasibility and system capabilities rather than user perspectives, feedback, and identified limitations. This evaluation gap is particularly acute for hybrid tangible-digital systems integrating multiple environmental simulations. While individual components, e.g. tangible interaction, real-time visualisation, participatory mapping, have been studied separately, comprehensive assessment of hybrid systems supporting collaborative urban planning with environmental analysis capabilities remains scarce. Understanding user experiences with such systems is essential for advancing both the technological development of planning support tools and theoretical understanding of how geospatial technologies can effectively support participatory processes.

This study addresses these gaps through systematic analysis of user feedback from 12 groups who evaluated the Cockpit for Collaborative Urban Planning (COUP), a hybrid tangible-digital interface combining physical city models, real-time environmental simulations, i.e. wind comfort, noise propagation, and pedestrian flows at district level, and interactive geospatial visualisation. The questions addressed during this research are two-fold: 1) how can we learn from hybrid tangible-digital planning interfaces from the experience of user interaction? and 2) what limitations and usability issues require attention in such systems, for its enhancement in collaborative planning processes? The conducted analysis is based on the experience of exposing COUP to a sample audience from academia within urban planning-related studies and is presented as a first iteration in the broader exercise of user evaluation methods for hybrid tangible-digital interfaces. Through mixed-methods analysis combining qualitative framework analysis with quantitative frequency assessment, consensus themes across user groups were identified and actionable insights for both tool developers and the broader GIScience community were synthesised. Our findings provide empirical evidence on how tangible geospatial interfaces are experienced in practice, advancing understanding of how to make geospatial data truly actionable in collaborative urban planning contexts.

2 Background on the Evaluation of Participatory Planning Tools

Despite growing implementation of tangible planning support systems, systematic user evaluation remains limited. Most published research focuses on technical system capabilities, describing interaction techniques, computational architectures, or simulation algorithms, rather than empirical assessment of how diverse users perceive and utilize these tools in realistic planning scenarios (Maquil et al., 2015; Jones and Maquil, 2016). User studies generally involve controlled laboratory experiments with small participant groups evaluating isolated system features rather than sustained engagement with integrated tools addressing authentic planning challenges. The few comprehensive evaluations examine individual system components in isolation: tangible interaction mechanics separate from simulation capabilities, or visualisation effectiveness independent of collaborative workflow support.

This fragmented approach limits understanding of how users experience systems where tangible interaction, real-time simulation, and collaborative features must function cohesively. Furthermore, evaluation time frames are often constrained to single brief sessions, providing limited insight into how user understanding and interaction patterns evolve through repeated engagement.

COUP addresses this evaluation gap through systematic analysis of user feedback from 58 participants across 12 user groups representing diverse spatial planning disciplines who engaged with the integrated system over multiple sessions. This extended engagement across authentic planning tasks provides empirical evidence of how hybrid tangible-digital interfaces are experienced in practice, informing both immediate design improvements and broader theoretical understanding of participatory geospatial technologies.

3 The COUP System

COUP is a hybrid, tangible-digital interface designed to support collaborative urban planning through near real-time environmental simulation and geospatial visualisation of neighborhood-level designs. Developed for a newly developed district in the city of Hamburg known as *kleiner Grasbrook*, COUP combines physical city models with interactive digital projections to assess design outcomes and facilitate participatory design decisions (López Baeza et al., 2021). The system

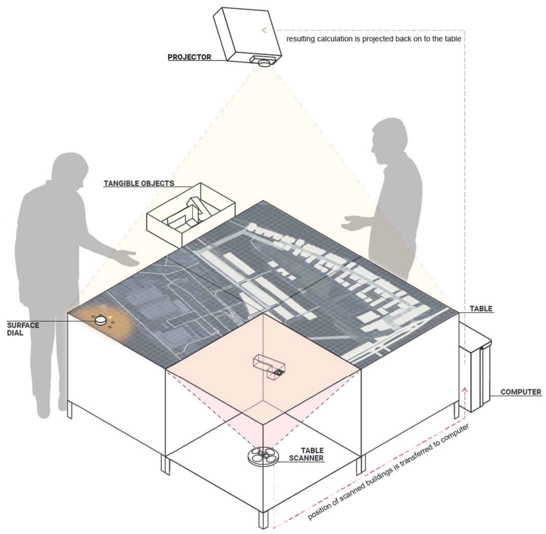

Figure 1. Diagram of the COUP System Design

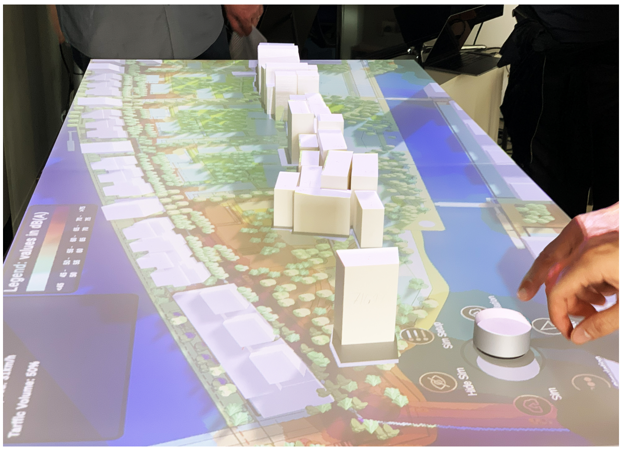

Figure 2. Interaction with the COUP TUI, showcasing noise simulation results

incorporates several urban analysis functions: pedestrian wind comfort and solar exposure based on machine-trained models; noise propagation modeling from traffic considering building geometries; and pedestrian flow simulation based on agent-based modeling and street networks. Outputs are presented in accessible visual formats (heatmaps, animated flows, color-coded zones) enhanced for non-expert comprehension. The simulation environment covers the kleiner Grasbrook district (approximately 4.5 km^2), comprising 83 individual buildings represented at LOD 1.3. A detailed description of the simulation modules is provided in Holtorf and Noennig (2025).

3.1 System Design and Interaction

COUP employs a tangible table interface consisting of three integrated components: a scaled, 3D-printed modifiable city model, a ceiling-mounted projector for digital overlays of the base maps, and an infrared tracking system that detects physical input elements (tangibles) placed on the table surface (Fig. 1). Built using Unity and Python, COUP integrates precomputed simulation results and live calculations depending on the computational requirements. The system integrates and processes georeferenced input data sourced from the City of Hamburg for land-use plans, road networks and building footprints; from the HafenCity Hamburg GmbH for the project design as well as from OpenStreetMap for the overarching urban footprint. All spatial data are aligned with the physical model's location, scale and orientation.

Users modify these tangibles elements, or 3D buildings, with embedded planning parameters such as building gross-floor area, density, use type, to explore design alternatives and observe immediate visual feedback over the open urban space, projected onto the physical model (Fig. 2). This direct physical-digital coupling creates an intuitive interaction paradigm accessible to users without specialised GIS expertise. The system immediately responds to tangible placement or movement, updating projected content such as wind comfort, solar exposure, noise levels, or pedestrian flows in near real-time. All feedback remains spatially anchored to the physical model, grounding abstract data in tangible urban form.

As an open-source project, the COUP codebase is publicly accessible, under a GitHub repository, following transparency and reproducibility principles common in German research institutions. The modular architecture enables adaptation to different contexts and integration of new simulation capabilities.

3.2 Application and Purpose

COUP has been implemented in planning workshops with experts and actors from public administration, as well as in public participation activities and educational settings, serving diverse stakeholder groups from professionals to citizens in both national and international contexts. The presented study focuses on a smaller sample from this audience, based on the experience from the educational setting. This was carried out as a first attempt to explore the outcome of user evaluation and collect feedback for bridging physical intuition with computational analysis. The longer-term purpose of COUP is to democratise access to planning information and support transparent, collaborative urban decision-making.

4 Methodology

4.1 Study Design and Data Collection

Twelve user groups comprising 58 master-level students evaluated the COUP system across four academic courses at the HafenCity University Hamburg between the summer semester 2023 and the winter semester 2024/25. Participants came from diverse planning-related disciplines: Resource Efficiency in Architecture and Planning (REAP), Architecture, Urban Design, Geodesy and Geoinformatics, Civil Engineering and Urban Planning. This interdisciplinary composition within the academic group reflects varied expertise relevant to collaborative planning, though it does not represent the full spectrum of stakeholders involved in real-world planning processes. Group sizes ranged from 3 to 7 students (typically 3–5), enabling both individual exploration and collaborative discussion during testing sessions.

The evaluation was embedded in a semester-long course graded on a pass/fail basis as per university regulations. This grading structure made the course particularly suitable for our study, as the absence of differentiated grading reduced potential social desirability pressure and supported honest, critical feedback from students who were not concerned about how their assessments might affect their academic standing. Students were progressively introduced to COUP through a structured course sequence: introductory lectures covered the system's hardware setup, development context, and simulation capabilities before any hands-on engagement. In the first practical session, groups familiarised themselves with the tool through free exploration. In subsequent sessions, groups identified specific urban design challenges in *kleiner Grasbrook* district (e.g., perimeter block development along noise-exposed streets, overshadowing by tall buildings), developed planning interventions to address these challenges, and tested their proposed solutions at the COUP table. Based on the simulation outputs, the groups evaluated whether and how their interventions resolved the identified problems. Each group participated in 2–4 sessions of approximately 20 minutes each.

The authors served different roles during the study: the second author led the courses, while the first author provided technical support. Author involvement was limited to assisting with tool operation (e.g., resolving software crashes, clarifying interface functions) and did not guide design decisions or planning solutions. Students worked based on the challenge-based learning framework and conducted autonomously their analytical tasks.

At the end of each course, groups submitted written reports documenting their design process and reflections on COUP. Report length and depth slightly varied across groups, but were fundamentally based on a proposed guiding structure. While the study involved 58 individual participants, all written feedback was produced collaboratively at the group level, yielding 12 group reports as the unit of analysis. This format captures collectively negotiated assessments rather than individual perspectives, with implications discussed in Section 6.4.

4.2 Analytical Framework

We employed a framework analysis approach (Gale et al. (2013)) combined with quantitative frequency assessment to systematically evaluate the 12 group reports. Following an inductive strategy, we did not impose predefined categories but allowed themes to emerge from the data through systematic comparison across groups.

The analysis proceeded in two phases. In the first phase, the first author conducted a structured summarisation of each group report, extracting information across eight dimensions: Tasks, Objectives, Approach, Simulation use, Results, Pros, Cons, and Conclusions. This produced a standardised representation of each group's evaluation while preserving the original context. In the second phase, the analysis focused specifically on the Pros and Cons sections, where the highest convergence across groups was observed. Recurring keywords and concepts were identified through repeated review of the summarised material, then grouped into thematic categories. These categories were organised into an analytical matrix (Groups × Categories) recording presence/absence per group, frequency counts, and representative quotes. Coding was conducted primarily by the first author. The second author, who led the courses and was familiar with the students´ reports, reviewed the resulting categorisation and validated the emerging themes. While this approach does not establish formal inter-coder reliability through independent double-coding, the second author's contextual knowledge provided a consistency check. This limitation is further addressed in Section 6.4.

Themes mentioned by $\geq$75% of groups were classified as "consensus themes," while those mentioned by $\geq$50% were classified as "major themes." This threshold-based approach enabled systematic identification of widespread patterns across diverse user groups.

4.3 Data and Software Availability

Feedback Data: The anonymised feedback data from all 12 user groups are openly available at https://doi.org/10.5281/zenodo.18195076. The dataset includes feedback transcripts and participant demographics (aggregated – still needed). **Analysis Materials**: The complete coding framework with category definitions, the analysis matrix (Groups × Categories), and representative quotes are provided in the repository under the MIT license. **Software**: The COUP system source code is available as open-source software at https://github.com/digitalcityscience/COUP-TangibleTable under the MIT license. Analysis was conducted using Microsoft Excel. No custom analysis scripts were developed; the Framework Analysis followed manual coding procedures.

Table 1. Matrix of system strengths identified by the 12 participant groups. Checkmarks indicate categories mentioned by each group.

Category	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	Freq.	%	Int.
Real-time Interactivity	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	11/12	91.7	High
Collaboration Support	✓			✓	✓	✓	✓	✓	✓	✓	✓	✓	10/12	83.3	High
Visualisation	✓	✓		✓	✓	✓	✓	✓		✓	✓	✓	10/12	83.3	High
Multiple Parameters		✓		✓	✓	✓	✓	✓		✓	✓	✓	9/12	75.0	High
Environmental/Climate Focus		✓	✓	✓	✓		✓	✓			✓	✓	8/12	66.7	Medium
Intuitive Usability			✓	✓	✓	✓			✓	✓		✓	7/12	58.3	Medium
Quick Testing/Iteration	✓		✓		✓	✓	✓		✓		✓		7/12	58.3	Medium
Physical-Digital Interface	✓			✓	✓					✓	✓		5/12	41.7	Medium
Data-Driven Decision-Making				✓	✓		✓	✓			✓		5/12	41.7	Medium
Educational Value						✓				✓		✓	3/12	25.0	Low

5 Results

5.1 Overview of the Identified Themes

The framework analysis of feedback from all 12 user groups identified 10 distinct strength categories and 12 limitation categories (Tables 1 and 2). Among the strengths, four themes achieved consensus status (mentioned by ≥75% of groups): Real-time Interactivity (91.7%), Collaboration Support (83.3%), Visualisation (83.3%), and Multiple Parameters (75.0%). An additional three themes qualified as major themes (≥50%): Environmental/Climate Focus (66.7%), Intuitive Usability (58.3%), and Quick Testing/Iteration (58.3%). For limitations, one theme reached consensus threshold: Data Accuracy/Reliability (75.0%). Three additional themes qualified as major limitations: Performance/Lag Issues (66.7%), Technical Bugs/Crashes (58.3%), and Table-Web Discrepancies (50.0%).

The complete category sets, including themes mentioned by fewer than 50% of groups, are presented in Tables 1 and 2. In the following sections, we focus our detailed discussion on themes meeting at least the major theme threshold (≥50%), as these represent patterns shared across multiple user groups and disciplinary backgrounds. Themes with lower frequencies, while documented in the matrices for completeness, reflect more context-dependent or discipline-specific concerns and are not discussed in detail here.

The distribution of themes across groups demonstrates both widespread agreement on core system characteristics and variation reflecting different disciplinary priorities and use contexts. High-frequency themes represent experiences consistent across diverse user backgrounds.

5.2 System Strengths

Table 1 presents the complete matrix of identified system strengths across all 12 user groups, including frequency counts and intensity ratings. Four themes achieved consensus status (≥75% agreement), representing capabilities consistently valued across diverse disciplinary backgrounds and use contexts.

Consensus Themes (≥75%)

Real-time Interactivity emerged as the most widely appreciated strength (91.7%, 11/12 groups), with groups emphasising the system's capacity to provide rapid feedback on design decisions. Groups described how interactive simulations enabled dynamic exploration of planning scenarios, allowing them to observe environmental impacts as they modified physical building placements. This responsiveness was particularly valued for understanding complex spatial relationships and environmental consequences that would be difficult to grasp through static representations. The tangible-digital coupling created a feedback loop where physical actions triggered computational responses within a timeframe experienced as immediate by users, supporting iterative design refinement.

Collaboration Support (83.3%, 10/12 groups) was consistently recognised as a core system capability. Multiple groups described COUP as supporting collaboration between planners and encouraging multidisciplinary teamwork. The shared interactive surface enabled simultaneous engagement of multiple group members, facilitating turn-taking and joint exploration of design alternatives. The physical interaction — moving tangible building models — was perceived as inherently more collaborative than mouse-and-keyboard interfaces, lowering barriers to participation across group members with varying technical expertise.

Visualisation (83.3%, 10/12 groups) received widespread recognition for making complex environmental data comprehensible. Groups reported that the system provided clear visual feedback through projected overlays directly onto the planning surface. The spatial registration of simulation results with physical building models helped users understand cause-effect relationships between urban form and environmental performance, translating abstract numerical data into experiential spatial understanding.

Multiple Parameters integration (75.0%, 9/12 groups) represented a valued capability for holistic planning assessment. Groups recognised COUP's capacity for

Table 2. Matrix of system limitations identified by the 12 participant groups. Checkmarks indicate categories mentioned by each group.

Category	**G1**	**G2**	**G3**	**G4**	**G5**	**G6**	**G7**	**G8**	**G9**	**G10**	**G11**	**G12**	**Freq.**	**%**	**Int.**
Data Accuracy/Reliability	✓	✓	✓	✓	✓	✓	✓	✓			✓		9/12	75.0	High
Performance/Lag Issues		✓		✓	✓	✓		✓		✓	✓	✓	8/12	66.7	High
Technical Bugs/Crashes	✓			✓	✓	✓			✓		✓	✓	7/12	58.3	High
Table-Web Discrepancies				✓	✓		✓	✓		✓	✓		6/12	50.0	High
Missing Legends/Scales			✓						✓	✓	✓	✓	5/12	41.7	Medium
Steep Learning Curve	✓					✓			✓	✓		✓	5/12	41.7	Medium
Complex Workflow				✓	✓				✓	✓		✓	5/12	41.7	Medium
Limited Editing Capabilities				✓			✓	✓			✓	✓	5/12	41.7	Medium
Limited Simultaneous Simulations						✓		✓		✓	✓		4/12	33.3	Medium
Visualisation Difficulties			✓	✓					✓			✓	4/12	33.3	Medium
Missing Environmental Modules						✓	✓	✓		✓			4/12	33.3	Medium
No Save/Compare Function								✓			✓	✓	3/12	25.0	Low

combining multiple parameters within a unified interface, enabling simultaneous consideration of wind comfort, noise propagation, and pedestrian flows within the same planning workflow. This multi-criteria evaluation support aligned with contemporary planning practice requirements for balancing diverse environmental and social objectives.

Major Themes (50–74%)
Environmental/Climate Focus (66.7%, 8/12 groups) highlighted the system's relevance for climate-responsive design, with groups appreciating the ability to assess environmental performance directly during the design process. *Intuitive Usability* (58.3%, 7/12 groups) reflected positive assessments of the tangible interaction paradigm, with groups describing the system as accessible and approachable compared to conventional GIS workflows. *Quick Testing/Iteration* (58.3%, 7/12 groups) emphasised the efficiency of exploratory planning enabled by rapid scenario modification at the tangible interface.

5.3 System Limitations

Table 2 presents the complete matrix of identified system limitations across all 12 user groups, including frequency counts and severity ratings. One theme achieved consensus status ($\geq$75% agreement), while three additional themes qualified as major limitations ($\geq$50%), representing challenges consistently encountered across diverse user groups.

Consensus Theme ($\geq$75%)
Data Accuracy/Reliability emerged as the most widespread concern (75.0%, 9/12 groups, High Severity), with groups expressing uncertainty about simulation validity and output trustworthiness. Reports referenced limited simulation accuracy, unreliable noise results, and general concerns about the credibility of output data. This reliability concern was particularly problematic given the system's intended role in supporting evidence-based design evaluation. Without confidence in simulation accuracy, groups found it difficult to rely on environmental analysis results, limiting the system's perceived utility for participatory decision-making processes that depend on credible data foundations.

Major Themes (50–74%)
Performance/Lag Issues (66.7%, 8/12 groups, High Severity) posed significant operational challenges. Groups frequently encountered lag during simulation runs, long computation times, and performance limitations attributed to hardware constraints, particularly when running multiple simulations simultaneously or working with complex urban scenarios. These delays disrupted collaborative workflow, forcing groups to wait for results rather than maintaining fluid exploratory interaction. The performance bottlenecks contradicted the system's intended real-time interactivity, creating gaps between physical manipulation and digital feedback.

Technical Bugs/Crashes (58.3%, 7/12 groups, High Severity) represented persistent stability problems. Groups reported technical malfunctions, software crashes, and occasional bugs that interrupted work sessions and required system restarts. These reliability issues were particularly disruptive in collaborative settings where multiple group members were simultaneously engaged, as crashes forced entire groups to pause and reconfigure. The unpredictability of technical failures undermined user confidence and created hesitation to explore advanced features that might trigger instability.

Table-Web Discrepancies (50.0%, 6/12 groups, High Severity) highlighted inconsistencies between the tangible tabletop and the web-based visualisation interface. Groups identified differences in simulation output depending on which interface was used. This inconsistency created confusion about which results to trust and complicated collaborative workflows where some group members interacted with the physical table while others accessed the web interface for detailed analysis.

6 Discussion

6.1 Key Findings and Interpretation

In relation to RQ1, the presented analysis identified four system strengths that achieved consensus-level recognition across diverse disciplinary backgrounds: real-time interactivity (91.7%), collaboration support (83.3%), visualisation quality (83.3%), and multi-parameter integration (75.0%). In response to RQ2, we identified one consensus limitation — data accuracy concerns (75.0%) — alongside three major limitations: performance issues (66.7%), technical instability (58.3%), and interface inconsistencies between the tangible and web-based components (50.0%).

Taken together, these findings reveal a paradoxical pattern: exceptionally strong consensus on system strengths alongside significant technical concerns that undermine practical adoption. The four consensus strengths validate the fundamental value proposition of hybrid tangible-digital interfaces for participatory planning. Participants across various planning-related disciplines recognised COUP's capacity to make complex geospatial analysis accessible, collaborative, and actionable. However, this advantage was tempered by equally significant limitations. Data accuracy concerns directly challenged the system's credibility for planning decision-making, while performance issues, technical instability, and interface inconsistencies disrupted the seamless interactive experience that users valued most highly.

This tension between appreciated potential and encountered problems reflects a critical challenge for emerging geospatial technologies: technical sophistication in simulation capabilities must be matched by reliable performance and trustworthy outputs. Notably, the most appreciated strength, real-time interactivity, was simultaneously undermined by the most prevalent limitations. Users valued immediate feedback but encountered lag, crashes, and unreliable results that broke the interaction flow. This finding suggests that "real-time" responsiveness is not merely a desirable feature but a fundamental expectation that, when unmet, significantly degrades user experience and system utility. The high frequency of performance-related concerns indicates that interactive planning support systems operate under stricter responsiveness requirements than traditional analytical GIS workflows, where longer computation times are more acceptable.

The strong consensus on collaboration support (83.3%) confirms that tangible interaction lowers participation barriers and facilitates more equitable engagement between group members. Unlike single-user GIS interfaces that create bottlenecks and expertise dependencies, the shared physical interaction surface enabled simultaneous engagement and natural turn-taking. This finding provides empirical evidence of the value of tangible interfaces specifically in collaborative planning contexts, a contribution relevant to the broader GIScience community's interest in participatory technologies.

6.2 Implications for Tangible Planning Interface Design

Addressing our second research question, our results reveal critical design priorities for hybrid tangible-digital planning interfaces, particularly regarding the tension between user expectations and the technical realities of environmental simulation.

The "Real-Time" Paradox
The overwhelming consensus on real-time interactivity as a core strength (91.7%) reveals a fundamental misalignment between user perception and system capabilities. Users consistently praised COUP for "real-time" feedback, yet simultaneously reported performance lag and long computation times. This apparent contradiction reflects an important distinction: COUP provides near real-time interaction rather than true real-time performance. Complex environmental simulations, particularly CFD-based wind comfort analysis and acoustic propagation modeling, require substantial computational resources that prevent instantaneous results. The user expectation of "real-time" responsiveness, despite experiencing delays, suggests that the tangible interaction paradigm itself creates a cognitive frame where any computational feedback feels immediate compared to traditional planning workflows.

Moving physical building models and receiving simulation results within a timeframe of roughly 10 to 45 seconds, depending on network conditions and simulation type, represents a dramatic improvement over conventional processes requiring hours or days for equivalent environmental analysis. Users may perceive this experience as "real-time" even when technical specifications would categorise it as near real-time or delayed feedback. This finding has significant implications for system design and communication. Framing tangible planning tools as "real-time" systems sets expectations that current simulation engines cannot realistically meet. A more honest framing as "rapid feedback" or "near real-time" systems might better align user expectations with actual performance.

From a technical development perspective, performance optimisation must remain a high priority. The high frequency of lag-related concerns (66.7%) indicates that even near real-time performance is insufficient if delays disrupt collaborative workflow. Future development should focus on computational optimisation of simulation algorithms, progressive visualisation displaying partial results while full simulations complete, intelligent caching of previously computed scenarios, and clear progress indicators during computation.

Data Reliability as Foundation for Adoption
The consensus-level concern about data accuracy (75.0%) represents a significant challenge for planning support systems. Sophisticated interaction design and collaborative features become irrelevant if users cannot trust simulation outputs. Addressing reliability concerns requires rigorous validation of simulation models against empirical measurements, transparent communication of model assumptions and limitations, and visualisation of uncertainty ranges rather than single-value predictions. This challenge intensifies in participatory contexts where non-expert users may lack the background to independently interpret simulation uncertainties.

6.3 Advancing Participatory GIScience

Our findings contribute to the broader theoretical and practical discourse within the GIScience community on making geospatial data truly actionable in collaborative planning contexts.

From "Smart Data" to "Trustworthy Actionable Data" Contemporary GIScience increasingly emphasises the transformation of raw geospatial information into "smart data" through interactive, contextualised representation. Our evaluation suggests that interactivity and contextualisation, while necessary, are insufficient for genuine actionability. The paradox of high enthusiasm for system capabilities alongside fundamental trust concerns (75% consensus on data accuracy issues) demonstrates that actionable data requires not only accessibility and comprehensibility but also credibility and reliability. We suggest extending the "smart data" framework to encompass what we call "trustworthy actionable data": geospatial information that is simultaneously computationally sophisticated, interactively accessible, collaboratively usable, and empirically validated with transparent communication of limitations. This four-dimensional perspective acknowledges that participatory planning tools operate under stricter requirements than expert-only systems, as users across disciplines must trust outputs without the specialised expertise to independently verify simulation validity.

User-Centered Evaluation as Methodological Imperative The scarcity of comprehensive user evaluations in tangible interface literature (Maquil et al., 2015; Jones and Maquil, 2016) represents a significant gap in GIScience research. Most implementations prioritise technical feasibility demonstrations over sustained engagement with diverse user groups, limiting understanding of how geospatial technologies are actually experienced in practice. Our systematic analysis of feedback from 12 interdisciplinary groups across multiple sessions contributes to the learnings from user interaction, for example through the value of extended evaluation periods that allow users to move beyond initial novelty and develop informed perspectives on system capabilities and limitations. This methodological approach, combining qualitative framework analysis with quantitative frequency assessment, provides a replicable template for evaluating participatory geospatial technologies. The distinction between consensus themes ($\geq$75% agreement) and major themes ($\geq$50% agreement) enables identification of widely shared experiences versus context-dependent or user-specific concerns, informing prioritisation of development efforts.

Rethinking Participation Barriers The strong consensus on collaboration support (83.3%) provides empirical validation for theoretical claims that tangible interaction lowers participation barriers (Ishii and Ullmer, 1997; Hornecker and Buur, 2006). However, while physical interaction successfully facilitated collaborative engagement, technical barriers and unstable performance created new exclusionary mechanisms. This suggests that democratising geospatial analysis requires addressing not only interface paradigms but also system reliability, performance predictability, and robust implementation quality. Future participatory GIScience research should attend equally to interaction design innovation and technical stability, recognising that accessibility gains from novel interfaces can be undermined by performance issues or unclear system behaviour.

6.4 Study Limitations

Several methodological constraints warrant careful consideration and frame the interpretation of our findings.

Participant Population and Ecological Validity. The evaluation was conducted with master-level students from planning-related disciplines at a single metropolitan university. While this population offered valuable interdisciplinary perspectives from environmental design, spatial analysis, geoinformatics, engineering and urban development, it does not represent the full spectrum of stakeholders involved in real-world participatory planning processes. Professional planners, public officials, and citizens, particularly older individuals and those without technical backgrounds may experience the system differently, especially regarding learning curves and interaction paradigms. Consequently, our findings should be understood as reflecting expert-novice perspectives from planning-related disciplines rather than the broader stakeholder diversity that COUP aims to support. Field deployment studies with professional planners and public participants are needed to establish ecological validity in authentic planning contexts.

Unit of Analysis. Although 58 students participated in the study, all written feedback was produced collaboratively at the group level, yielding 12 group reports as the analytical unit. This format captures collectively negotiated assessments and reflects the collaborative nature of the tool's intended use, but it may under-represent minority opinions within groups and cannot capture individual variation in experiences. Dominant group members may have shaped the written reports more strongly than quieter participants, and the high consensus levels observed in our analysis partially reflect within-group convergence rather than independent individual agreement.

Researcher-Instructor Dual Role. The evaluation was conducted within courses taught by members of the author team, who also provided technical support during sessions. Although the pass/fail grading structure, reduced pressure toward socially desirable responses, the dual role as both course instructors and researchers may have influenced how freely students voiced criticism. Students may have been more hesitant to report problems than they would have been in an independent evaluation context. Author involvement during sessions was limited to technical assistance (e.g., resolving software crashes or clarifying interface functions) and did not extend to guiding design decisions or planning solutions.

Coding Process and Inter-Coder Reliability. The framework analysis was conducted primarily by one member of the author team, with a second author reviewing and validating the emerging categories based on familiarity with the student reports and course context. This approach does not establish formal inter-coder reliability through independent double-coding, which would strengthen the methodological rigour of our thematic identification. Future studies should implement parallel independent coding with reliability measures such as Cohen's kappa.

Lack of Methodological Triangulation. Our analysis relies on a single data source, in form of written group reports, without complementary instruments such as observation logs, interaction metrics, standardised usability questionnaires (e.g., SUS, UEQ), or post-hoc interviews. Written reports captured rich qualitative reflections developed over multiple sessions, but they cannot substitute for the convergent validity that triangulated methods would provide. Non-verbal aspects of collaborative interaction, spontaneous reactions during tool use, and individual experiential nuances remain outside the scope of our data. Future evaluations should incorporate mixed-method designs combining behavioural observation, standardised instruments, and qualitative feedback to strengthen the empirical foundation.

6.5 Technical Improvements and Future Development

Since the user evaluation was conducted, several technical improvements have been implemented in response to the issues identified. A new calibration system has been deployed to improve tracking accuracy and reduce setup complexity, making the initial learning phase more accessible for new users. Critical bugs that caused system crashes during the evaluation have been identified and resolved, significantly improving overall stability. To address usability concerns related to workflow complexity, current development focuses on integrating a touch-sensitive surface into the tabletop interface. This hybrid approach will combine tangible building model modification with familiar touchscreen interactions for secondary functions such as parameter adjustment and simulation selection, leveraging interaction paradigms known from smartphones and tablets to reduce the learning burden while preserving the collaborative benefits of physical object manipulation.

Performance optimisation remains a priority. The prevalence of lag-related concerns in the evaluation highlights the need for continued computational efficiency improvements through optimised simulation algorithms, improved CFD and acoustic models that balance accuracy with performance, progressive result visualisation as outputs become available, and cloud-based computation architectures to move closer to true real-time interaction. Addressing data accuracy concerns, the most critical limitation identified in our evaluation, requires systematic validation efforts comparing COUP simulation outputs against empirical environmental measurements to establish confidence intervals and identify systematic biases. Transparent communication of model limitations through enhanced visualisation will help users interpret results appropriately.

Several research directions emerge from our findings. Field deployment studies in authentic municipal planning environments are essential to confirm whether results from educational contexts translate to real-world participatory planning with diverse stakeholder demographics, including professional planners, public officials, and citizens. An additional COUP prototype has been deployed at a public authority office in Leipzig, where it is regularly presented to practitioners and continues to receive system updates. Feedback gathered during these presentations complements the present evaluation and will inform future systematic studies with professional users. Investigation of discipline-specific usage patterns may further inform adaptive interface designs that accommodate varying professional perspectives and expertise levels. Finally, future evaluations should incorporate methodological triangulation through mixed-method designs combining behavioural observation, standardised usability instruments, and qualitative feedback to overcome the limitations of single-source data collection identified in the present study.

7 Conclusion

This study addresses a critical gap in participatory geospatial technology research through systematic evaluation of user experiences with COUP, a hybrid tangible-digital planning interface integrating multiple environmental simulations. The analysis of 12 group reports from interdisciplinary master-level students revealed both the transformative potential and significant challenges of tangible planning support systems. Four system strengths achieved consensus-level recognition (≥75% of groups): real-time interactivity, collaboration support, visualisation quality, and multi-parameter integration. These findings address the proposed research questions, in the validation of the fundamental value proposition of tangible interfaces for participatory planning, making complex geospatial analysis accessible, collaborative, and actionable across users with varying technical expertise. The strong appreciation for collaboration support (83.3%) provides empirical evidence that tangible interaction successfully lowers participation barriers compared to traditional GIS interfaces, enabling more equitable engagement between group members in planning processes.At the same time, substantial space for further development has been identified. Data accuracy concerns (75% consensus), performance issues (66.7%), technical instability (58.3%), and interface inconsistencies (50%) undermined system credibility and disrupted collaborative workflows. Most critically, the tension between appreciated "real-time" interactivity and encountered computational delays reveals a fundamental challenge: complex environmental simulations including CFD-based wind analysis, acoustic propagation modeling, and agent-based pedestrian flow simulation — require substantial computational resources that still prevent true real-time performance. COUP provides near real-time interaction rather than instantaneous feedback, creating a perception-reality gap that future systems should address through both technical optimisation and clearer communication of capabilities.

The presented findings demonstrate that actionable geospatial data requires more than interactive visualisation: it demands trustworthy, reliable, and transparent computational outputs. Addressing the second formulated research question, in regards to the limitations for the enhancement of such systems in collaborative planning contexts, it is advisable to extend GIScience's "smart data" framework toward what might be called "trustworthy actionable data" geospatial information that is simultaneously computationally sophisticated, interactively accessible, collaboratively usable, and empirically validated with transparent communication of limitations. This perspective acknowledges that participatory planning tools operate under stricter requirements than expert-only systems, as users across disciplines must trust outputs without the specialised background to independently verify simulation validity.

For the broader GIScience community, our study underscores the importance of systematic user evaluation as a methodological complement to technical innovation. Understanding how users actually experience geospatial technologies, their enthusiasms, frustrations, trust concerns, and workflow integration challenges, is essential for advancing participatory planning support systems that meet real-world expectations. COUP's evolution from grid-based predecessors like CityScope demonstrates progress in spatial flexibility; future work should equally enhance performance reliability, data trustworthiness, and usability refinement. Ongoing field deployment, including the prototype installation at a public authority office in Leipzig, will extend this evaluation across the public sector actors and expertise that participatory planning tools ultimately aim to serve.

8 Use of Generative AI Tools

The authors declare that they have used Generative AI tools in the preparation of this manuscript. Specifically, the AI tools were utilized for language editing, improving grammar and sentence structure, as well as supporting literature searches, but not for generating scientific content, research data, or substantive conclusions. All intellectual and creative work, including the analysis and interpretation of data, is original and has been conducted by the authors without AI assistance.

9 Acknowledgements

The authors would like to acknowledge the valuable contributions of students whose work and insights informed this research. Relevant data was incorporated in an anonymized manner, and all efforts were made to ensure that individual identities remain confidential. The authors are grateful for their participation and engagement, which significantly enriched the development of this study. The authors declare that no external funding was received for the conduct of this study.

References

Brown, G. and Kyttä, M.: Key issues and priorities in participatory mapping: Toward integration or increased specialization?, Applied Geography, 95, 1–8, https://doi.org/10.1016/j.apgeog.2018.04.002, 2018.

Brüggemann, J., Köckler, H., Walter-Klose, C., Vittinghoff, M., Flacke, J., and Pfeffer, K.: Enabling people with diverse abilities to Participate in the Design of Digital Mapping Tools for Inclusive Community Planning in Germany, vol. 442, p. 020, SISSA Medialab, https://doi.org/10.22323/1.442.0020, 2023.

Falco, E. and Kleinhans, R.: Digital Participatory Platforms for Co-Production in Urban Development: A Systematic Review, International Journal of E-Planning Research, 7, 52–79, https://doi.org/10.4018/IJEPR.2018070105, 2018.

Gale, N. K., Heath, G., Cameron, E., Rashid, S., and Redwood, S.: Using the framework method for the analysis of qualitative data in multi-disciplinary health research, BMC Medical Research Methodology, 13, 117, https://doi.org/10.1186/1471-2288-13-117, 2013.

Haklay, M. and Tobón, C.: Usability evaluation and PPGIS: towards a user-centred design approach, International Journal of Geographical Information Science, 17, 577–592, https://doi.org/10.1080/1365881031000114107, 2003.

Holtorf, V. and Noennig, J. R.: COUP-XR: A Collaborative Cross-Reality System for Urban Planning, Procedia Computer Science, 270, 1758–1765, https://doi.org/https://doi.org/10.1016/j.procs.2025.09.296, 29th International Conference on Knowledge-Based and Intelligent Information Engineering Systems (KES 2025), 2025.

Hornecker, E. and Buur, J.: Getting a grip on tangible interaction: a framework on physical space and social interaction, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, p. 437–446, ACM, Montréal Québec Canada, https://doi.org/10.1145/1124772.1124838, 2006.

Ishii, H. and Ullmer, B.: Tangible bits: towards seamless interfaces between people, bits and atoms, in: Proceedings of the ACM SIGCHI Conference on Human factors in computing systems, p. 234–241, ACM, Atlanta Georgia USA, https://doi.org/10.1145/258549.258715, 1997.

Jones, C. E. and Maquil, V.: Towards Geospatial Tangible User Interfaces: An Observational User Study Exploring Geospatial Interactions of the Novice, vol. 582, p. 104–123, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-319-29589-3_7, 2016.

López Baeza, J., Sievert, J. L., Landwehr, A., Luft, J., Preuner, P., Bruns-Berentelg, J., Noyman, A., and Noennig, J. R.: CityScope Platform for Real-Time Analysis and Decision-Support in Urban Design Competitions, International Journal of E-Planning Research (IJEPR), 10, 121–137, https://doi.org/10.4018/IJEPR.20211001.oa8, 2021.

Maquil, V., De Sousa, L., Leopold, U., and Tobias, E.: A geospatial tangible user interface to support stakeholder participation in urban planning, in: 2015 1st International Conference on Geographical Information Systems Theory, Applications and Management (GISTAM), p. 1–8, https://ieeexplore.ieee.org/document/7512215?denied=, 2015.

Maquil, V., Leopold, U., De Sousa, L. M., Schwartz, L., and Tobias, E.: Towards a framework for geospatial tangible user interfaces in collaborative urban planning, Journal of Geographical Systems, 20, 185–206, https://doi.org/10.1007/s10109-018-0265-6, 2018.