The Australian National University
2600 ACT | Canberra | Australia

Australian National University

School of Computing

ANU College of Systems & Society

UX Design, Implementation and Evaluation for a Social Choice Board Game App

— Honours project (S2/S1 2025–2026)

A thesis submitted for the degree
Master of Computing (Advanced)

By:
Kevin Zhu

Supervisors:
Dr. Peter Höefner
Dr. Michael Norrish

June 2026

Declaration:

I declare that this work:

- upholds the principles of academic integrity, as defined in the University Academic Misconduct Rules;
- is original, except where collaboration (for example group work) has been authorised in writing by the course convener in the class summary and/or Wattle site;
- is produced for the purposes of this assessment task and has not been submitted for assessment in any other context, except where authorised in writing by the course convener;
- gives appropriate acknowledgement of the ideas, scholarship and intellectual property of others insofar as these have been used;
- in no part involves copying, cheating, collusion, fabrication, plagiarism or recycling.

June, Kevin Zhu

Acknowledgements

I am sincerely grateful to my supervisors, Peter and Michael, for their invaluable support, insightful guidance, and patience throughout the course of this thesis. Their expertise and encouragement have been instrumental in shaping this work.

I owe my deepest thanks to my parents, whose unconditional love and steadfast support have made this journey possible. I am forever grateful for everything they have given me.

I would also like to thank those I love and who love me. An Yun, Htoo Yadanar and Zhaoyu Xu, you have been my light ever since we met. Your presence and encouragement have meant more to me than words can express.

Finally, I wish to thank everyone who took part in this project as a participant, in particular the members of the ANU Board Game Society and the CSS community. This research would not have been possible without your time and willingness to contribute. Those Monday evening gaming nights were a highlight of my candidature, and I am grateful for the warmth and good company that made them so memorable.

Abstract

Computational social choice systems can support collective decisions, but their practical value depends on whether people can understand, influence, and accept the decision process. Group board-game allocation makes this problem concrete because participants need to coordinate preferences, player counts, teaching knowledge, familiarity, and possible multi-table assignments, while existing tools provide limited support for shared control and explanation. In this setting of collective decision-making, this thesis implements a human-centred framework for structuring information flow between participants and allocation algorithms. The framework is realised through PickNPlay, a participant-directed web system for collaborative game entry, tiered preference input, Z3-backed allocation, progress visibility, revision, and explanation, and extended through Boardot, an exploratory shared-station facilitation layer. Exploratory HCI sessions using observation, SUS questionnaires, interviews, follow-up material, and thematic analysis found that PickNPlay was rated as more usable over the formative Boardroom BGVS baseline, while Boardot showed comparable descriptive usability and made allocation more visible as a shared room-scale event. The findings suggest that solver-backed allocation becomes more acceptable when participants can see, shape, and revise the decision process. The thesis contributes a human-centred account of group board-game allocation as shared interaction infrastructure, showing how constrained collective decisions can be made more socially usable through transparency, recoverable coordination, and bounded facilitation.

Table of Contents

Chapter 1

Introduction

Collective decisions are increasingly mediated by digital platforms, from electronic voting systems used in elections to meeting scheduling tools and course timetabling systems such as MyTimetable, which capture individual preferences, aggregate or coordinate those preferences, and present outcomes back to participants in actionable forms. This thesis sits at the intersection of interactive technologies in human-computer interaction and computational social choice (Brandt et al., 2016), a field that connects social choice theory with theoretical computer science and the analysis of multi-agent systems, and studies (among other issues) how preferences are represented, elicited, and aggregated under computational constraints. These interactive technologies are becoming more embedded in our daily lives. They are also transforming social activities such as board game sessions where coordination in groups can be challenging as participants need to account for limited time, player-count constraints, and diverse preferences, and the manual effort required to form suitable game assignments can detract from the overall social experience (Verrell, 2025). At the same time, introducing algorithmic or AI assistance raises a core HCI tension between automation and ease of use. Systems should provide computational support without turning decision-making into an opaque process that users merely accept instead of understanding and steering (Heer, 2019).

1.1 Motivation

Board games selection provides a rich yet manageable microcosm of collective decision-making. A group of people with diverse preferences need to negotiate time limitations, player-count constraints, and game-specific requirements to arrive at mutually acceptable activities. In practice, this coordination is often handled informally through discussion, ad-hoc voting, or reliance on a designated organiser, which can be time-consuming and may feel unfair when some voices dominate or constraints are overlooked (BoardGameGeek forum community, 2026; Reddit r/boardgames community, 2023). At the same time, prior work has shown that digitally mediated board-game experiences can strengthen social connection and create structured spaces for participation, especially when rules and turn-taking are made explicit through interactive systems (Yuan et al., 2021; Triay and Wood, 2021).

This project is motivated by the observation that there is a gap between social choice mechanisms for fair allocation and group decision-making and the messy realities of how people make decisions for group activities. Existing tools and the prior work for group coordination around games and events either focus primarily on logistical filtering and recommendation, or on algorithmic optimisation of assignments, with limited attention to how participants understand, influence, and experience the underlying procedures (eg. Board Game Pick (2026); What2Play (2026b); Rocket Power Software, LLC (2025); Verrell (2025)). This gap creates a need for interactive systems that treat social choice not only as a computational problem, but also as a human-centred, visual, and social process.

1.2 Aims and Research Questions

This research aims to improve the usability and accessibility of social choice systems in casual group settings by designing and evaluating two frameworks that facilitate group decision-making around board game selection. Specifically, the contributions are:

- **Design and Implementation of PickNPlay (Direct Participant Interaction):** Develop a transparent, web-based platform that visualises complex social choice mechanisms (such as fair allocation rules) in real time, enabling participants to enter preferences directly, explore outcomes, and observe how their preferences shape collective decisions.
- **Design and Implementation of Boardot (AI-Mediated Facilitation):** Develop a centralised system with an active AI host that brings people together, paces the decision process, and provides lightweight prompts, summaries, and guidance, thereby reducing the interaction and coordination burden on users.
- **Framework for Structured Information Flow:** Propose and implement a design framework that articulates how information should flow between algorithms and users across participant-directed and AI-mediated modalities, with the goal of making social choice procedures more fluent, legible, and socially acceptable in casual group settings.
- **Empirical Evaluation of User Experience and Perceived Fairness:** Empirically evaluate how the two systems, individually and in combination, affect users' understanding of mechanisms, trust in outcomes, perceived procedural fairness, and sense of social cohesion during board game–mediated decision-making.

Research Questions

RQ1: **How can the social choice process be communicated transparently and accessibly so that participants understand the procedure and regard the allocation outcome as fair and acceptable?**

RQ2: **How can the interface design reduce organiser-centred control and support more balanced participation while preserving negotiation in board-game group decision-making?**

RQ3: **How do direct participant interaction and AI-mediated facilitation affect users' perceived control, trust, fairness, and social cohesion during group decision-making?**

1.3 Overview

The remaining chapters are

1. **Background:** Reviews relevant literature, foundational theories, and technologies in social choice, fair allocation, interactive visualisation, and human-centred AI. This section establishes the theoretical and technological context for the design of PickNPlay and Boardot.
2. **Requirement Analysis and System Overview:** Describes the problem setting, user and stakeholder requirements, and the overall design rationale. This section introduces the dual-system concept and outlines how PickNPlay and Boardot address complementary aspects of group decision-making.
3. **System Design and Architecture:** Details the architecture, technology stack, and interaction design of both systems. This section explains how participant-directed (PickNPlay) and AI-mediated (Boardot) interaction modalities are realised in the interfaces and how social choice mechanisms are integrated and visualised.
4. **User Study and Evaluation:** Presents the study design, data collection methods (including interviews and questionnaires), and analysis used to evaluate usability, understanding, perceived fairness, trust, and social experience for each system and their comparison.
5. **Conclusion:** Interprets the research findings and reflects on the broader implications. This section summarises key contributions, identifies limitations, and suggests directions for future work.
6. **Appendices and Bibliography:** Provides supplementary materials and references that support the thesis. This section ensures transparency and offers a comprehensive foundation for the research.

Chapter 2

Background

This chapter reviews prior work at the intersection of computational social choice, human–computer interaction, and digital systems for coordinating board game play and related group activities. Section 2.1 first introduces social choice, fair-allocation literature, constrained assignment, constraint-based optimisation, and user-facing social choice systems. Section 2.2 then introduces HCI foundations and discusses how requirement analysis, transparency, feedback, interaction modality, cognitive load, and user control apply to interactive decision-support systems in social choice contexts. Section 2.3 examines the Boardroom BGVS system as the immediate previous work and baseline for this project, outlining its solver-based allocation model and frontend interaction design to motivate the need for new approaches. Section 2.4 reviews related board game selection tools and broader group coordination systems, identifying current gaps, opportunities, and needs that motivate the design of PickNPlay and Boardot in subsequent chapters. Finally, Section 2.5 briefly acknowledges related concerns such as anonymity and privacy, which are relevant to the broader design space even though full anonymity and privacy-preserving mechanism design are outside the core objectives of the present work.

2.1 Social Choice Theory

2.1.1 Foundations of Social and Computational Social Choice

Social choice theory studies how to aggregate the preferences or judgements of multiple individuals into a single collective decision, typically by specifying a social choice function mapping individual preference profiles to group outcomes, as formalised by Arrow (Arrow, 1963) in his foundational work. In this framework, classical results, most prominently, Arrow's Impossibility Theorem show that even basic desiderata such as unanimity (Pareto efficiency), independence of irrelevant alternatives, and non-dictatorship cannot all be satisfied simultaneously when aggregating unrestricted preference orderings. These impossibility results highlight fundamental tensions between fairness, rationality, and responsiveness, and motivate the search for more limited but implementable forms of collective decision-making.

Building on this standard, computational social choice can be understood as an area at the intersection of social choice theory and theoretical computer science, concerned with the algorithmic and complexity properties of collective decision rules. Instead of viewing social choice rules only as abstract mappings, researchers in computational social choice study how preferences are represented, how hard it is to compute outcomes, how much information needs to be communicated, and how strategic behaviour interacts with computational constraints. Within this broader landscape, Brandt et al. (2016) analyse topics such as voting, fair allocation, matching, judgment aggregation, and coalition formation through both axiomatic and procedural lenses.

Throughout this thesis, board game selection refers to the broader social activity of deciding what a group will play. Preference ranking, voting, or elicitation refers to the input process through which participants express their preferences. Allocation or assignment refers to the computational outcome that places participants into feasible game groups. Recommendation and filtering are treated as related but narrower tasks. They help identify candidate games, but they do not by themselves resolve the full constrained group allocation problem. Matching is used only when discussing the relevant matching literature, not as the primary description of the thesis problem.

The present project draws on the strand of computational social choice that studies fair allocation of indivisible goods and related assignment problems, as surveyed by Bouveret et al. (Bouveret et al., 2016) and by Chevaleyre et al. (Chevaleyre et al., 2006) in the context of multiagent resource allocation. Board game allocation is not a canonical indivisible-goods problem, because a game is not allocated exclusively to one participant. Rather, games are discrete group activities to which several participants may be assigned at the same time. Fair-allocation literature remains useful as background because the system needs to aggregate individual preferences while balancing efficiency, preference satisfaction, and the risk of systematically disadvantaging particular participants.

Fairness is used carefully here. Fair-allocation theory provides motivating concepts and vocabulary, but the implemented resolver optimises preference satisfaction under feasibility constraints without proving a formal guarantee such as envy-freeness, proportionality, or strategy-proofness. The later evaluation focuses on perceived fairness, intelligibility, and procedural acceptability. The concern is whether participants can understand the allocation process and regard the outcome as reasonable in context.

More precisely, the computational part of the problem is a one-sided group assignment problem with feasibility constraints. Participants express preferences over candidate games, but games do not express preferences over participants. Each game can receive multiple participants only if the assigned group satisfies its playable range. Here, a game's minimum player count is treated as a lower quota and its maximum player count as an upper quota. A game is feasible only if it receives either no assigned participants or a number of participants within that quota interval. This structure is already visible in the Boardroom system (Verrell, 2025).

Board game allocation is deliberately small-scale compared with canonical applications such as school choice, course allocation, or residency matching. In a BoardGameGeek-based dataset, Zalewski et al. (2019) report that publisher-specified player-count metadata most frequently assigns minimum player counts between one and three and maximum player counts between two and six, while games accommodating more than eight players are comparatively rare. This makes the domain a manageable but non-trivial setting for analysing preference satisfaction and procedural fairness under practical feasibility constraints, and it is especially useful for HCI because participants can observe and discuss preference expression, constraint negotiation, explanation, and user control as the decision unfolds.

2.1.2 Algorithms for Fair Allocation and Assignment

Matching theory provides related background for assignment problems, but it is not the exact model used here. Classical matching procedures originate in the stable marriage and college admissions models of Gale and Shapley, in which agents on two sides submit preference rankings and a matching is constructed that satisfies a notion of stability. In the classical deferred acceptance algorithm, for example, one side of the market (such as students) iteratively proposes to parties on the other side (such as colleges), which tentatively accept or reject proposals based on their own preferences and capacity constraints. The process terminates in a stable matching that is optimal for the proposing side among stable matchings. Variants of these procedures underpin many real-world mechanisms, including school choice and residency matching, and form a useful contrast for thinking about group assignment problems.

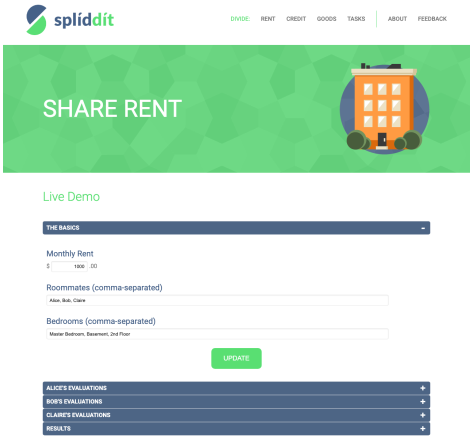
In the board-game setting, capacity constraints map directly onto playable player-count ranges. A game's upper quota limits how many participants it can receive, while its lower quota determines whether the game can run at all. The allocation system needs to decide not only which participant should be assigned to which game, but also which games receive enough assigned participants to run as feasible groups. When lower bounds on capacities are introduced, as in the college admissions problem with lower and common quotas, the algorithmic picture becomes more complex. Biró and colleagues (Biró et al., 2010) extend the college admissions framework to handle both upper and lower quotas on institutions and show that natural adaptations of deferred acceptance may fail to find feasible stable matchings in some instances . Related work on the hospitals/residents problem with lower quotas (Hamada et al., 2011) further clarifies the conditions under which stable matchings exist and provides specialised algorithms and complexity results, again highlighting that lower quotas introduce structural difficulties beyond the classical setting with only upper bounds. These results are important for applications such as course allocation and resource assignment in healthcare, and they also illustrate that matching-based approaches can face completeness or feasibility limitations once richer quota constraints are imposed.

In parallel, research on fair allocation of indivisible goods has developed algorithmic mechanisms that go beyond stability to emphasise welfare and fairness directly. In house allocation and related one-sided matching problems, for instance, mechanisms such as serial dictatorship (Abdulkadiroğlu and Sönmez, 1998) and top trading cycles (Shapley and Scarf, 1974) allocate objects to agents who have preferences but where the objects themselves have no preferences, with guarantees relating to strategy proofness, Pareto efficiency, and core selection under suitable assumptions. These mechanisms are closer to board game allocation because only participants express preferences. However, board games are shareable activities, not exclusive objects, and feasibility depends on lower and upper quotas. Together, matching-based and fair-division algorithms offer a rich toolkit for designing allocation procedures, but tailoring them to concrete domains with heterogeneous constraints can still be challenging.

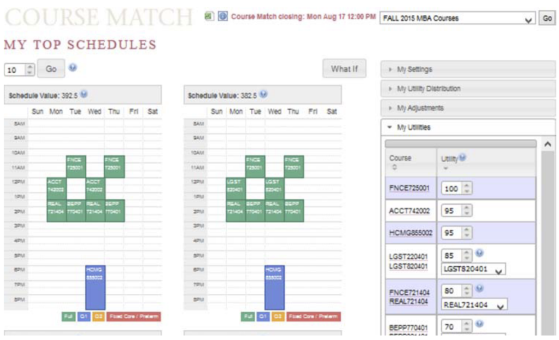
To address complex constraints and objectives, many allocation and assignment problems are formulated in more general optimisation frameworks such as constraint programming (Rossi et al., 2006), integer programming (Ágoston et al., 2016), and satisfiability modulo theories (SMT) (Barrett et al., 2021). In these approaches, decision variables (for example, indicating whether a given agent is assigned to a given resource) are introduced, hard constraints encode feasibility conditions such as uniqueness of assignment and capacity limits, soft constraints capture desirable but non-mandatory properties, and an objective function formalises the quality of an allocation in terms of preferences or other criteria. Constraint programming and SMT in particular allow such models to be expressed declaratively, with powerful solvers handling the search for satisfying or optimal solutions across large combinatorial spaces. Modern SMT solvers such as Z3 (de Moura and Bjørner, 2008) use a DPLL(T)-style architecture (Barrett et al., 2021), in which a Boolean satisfiability search explores candidate assignments while specialised theory solvers check consistency with background theories such as arithmetic, bit-vectors, and arrays. This optimisation layer is important for the present work because board game allocation is not merely a feasibility problem. Among all assignments that satisfy player-count and participation constraints, the system should prefer assignments that better reflect participants' rankings. For such optimisation problems, νZ (Bjørner et al., 2015) extends Z3 with support for objectives and soft constraints, allowing MaxSMT and optimisation tasks to be handled within the same solver framework. Prior work on automated board-game assignment already demonstrates the viability of SMT-based formulations for one-sided allocation with quotas in this domain (Verrell, 2025), and later chapters build on and extend this style of formulation when integrating allocation algorithms into interactive systems for group decision-making.

2.1.3 User-Facing Social Choice Systems

Existing social choice systems show how formal allocation mechanisms can be translated into software that non-specialist users can use. A prominent example is Spliddit (Goldman and Procaccia, 2015), a system that implements fair-division algorithms for



(a) Spliddit rent-division demo. Source: Spliddit (Goldman and Procaccia, 2015).



(b) Course Match top-schedules interface. Source: Course Match (Budish et al., 2017).

Figure 2.1: Examples of public-facing social choice interfaces. Spliddit presents rent division through an interactive web form; Course Match shows generated schedules next to utility inputs before preferences are saved.

domains such as rent division, goods division, and credit assignment. Figure 2.1a illustrates how Spliddit presents a rent-division mechanism through a web interface for entering participants, rooms, and evaluations. Spliddit is especially relevant because it shows that computational social choice and fair allocation methods can be embedded in end-user software instead of remaining purely theoretical. At the same time, the allocation problems it addresses are structurally different from the one considered here. Its mechanisms are designed for canonical fair-division tasks, whereas board game allocation requires assigning participants to activities under domain-specific feasibility constraints such as minimum and maximum player counts.

Another relevant example is Course Match (Budish et al., 2017), a large-scale course allocation system deployed at the Wharton School. Course Match applies an approximate competitive equilibrium from equal incomes mechanism to the combinatorial allocation of course schedules, asking students to report utilities over course sections and then computing schedules subject to course capacities, time conflicts, eligibility constraints, and credit limits. Figure 2.1b shows how Course Match presents generated candidate schedules alongside the student's utility inputs before the preferences are saved. This example is useful for the present thesis because it shows that applied social choice systems often require users to express preferences through interface-specific representations, while the system handles feasibility and optimisation in the background. In this respect, Course Match is closer to board game allocation than Spliddit because both involve assigning people to constrained activities, not only dividing a fixed set of goods.

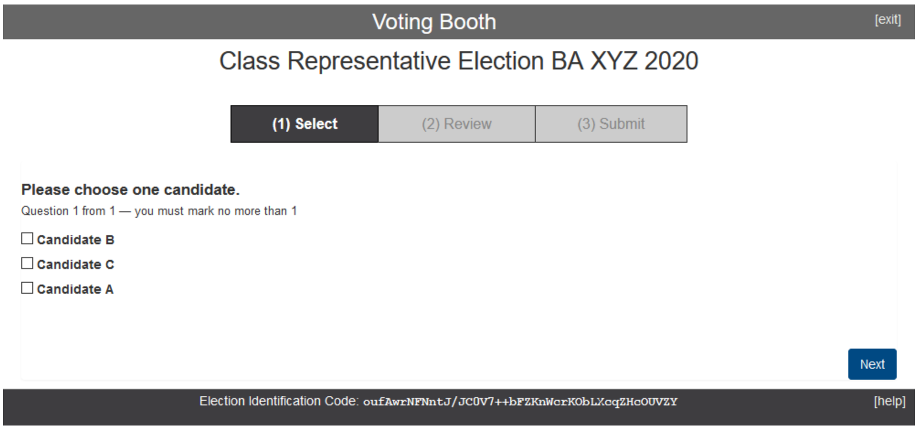


Figure 2.2: Screenshot of a Helios voting booth interface, showing candidate selection within a select-review-submit workflow. Source: Helios voting help page (MCI4ME, 2020).

Electronic voting systems provide another classic example of social choice mechanisms implemented as public-facing software. Helios (Adida, 2008) demonstrates how a web-based voting system can support election creation, secret ballot casting, tallying, and public auditability through cryptographic verification without requiring users to trust the server blindly. Figure 2.2 illustrates how the Helios voting booth presents candidate selection as a step-based interaction with select, review, and submit stages. The ACT's eVACS system (Elections ACT, 2024; Software Improvements, 2005) shows a different deployment context: a government election system combining electronic voting and electronic counting for Hare-Clark elections, with attention to accessibility, source-code scrutiny, and election integrity. These systems address higher-stakes requirements than board game allocation, especially secrecy, verifiability, and audit, but they reinforce the same broader point that applied social choice depends on interfaces, procedures, feedback, and trust mechanisms as well as formal aggregation rules.

Taken together, these systems show that the practical deployment of social choice systems depends not only on the underlying allocation rule, but also on interaction design. Users encounter these mechanisms through forms, valuation inputs, examples, schedules, explanations, verification steps, and result screens. This is particularly relevant to the present thesis, which adopts a solver-based allocation model while aiming to embed it within an interactive, user-facing decision support system.

2.2 Human-Computer Interaction (HCI)

2.2.1 HCI Foundations

Human-Computer Interaction (HCI) is a multidisciplinary area concerned with the relationship between people and computational systems. HCI does not treat interface design as a purely technical problem. It examines how people understand, use, and experience interactive technologies, and how those technologies can be designed to better support human goals. This perspective is particularly important here because the system is not only required to compute collective outcomes, but also to make the decision process understandable and usable for groups in a casual social setting.

The emergence of HCI as a distinct research area is closely tied to the spread of personal computing in the 1980s and to early work on situated action (Suchman, 1987) and human-machine interaction. As computing moved from specialist environments into workplaces and homes, researchers increasingly focused on how interface design shaped users' ability to complete tasks, interpret system behaviour, and recover from errors. Foundational HCI literature placed substantial emphasis on usability, interaction guidelines, and user-centred design methods for desktop and software systems (Carroll, 2003; Dix et al., 2004). Over time, this focus broadened from task performance alone toward user experience (UX), recognising that interactive systems are also evaluated through affective, aesthetic, and experiential qualities (Norman, 2004; Desmet and Hekkert, 2007; Hassenzahl et al., 2010). For systems that mediate group choices, these broader UX concerns are especially relevant. An allocation may be computationally valid, yet still fail if users experience the process as confusing, opaque, or socially awkward.

Key components of usability in HCI, crucial for effective interface design, include:

Effectiveness: Accuracy and completeness with which users achieve specified goals.

Efficiency: Resources expended in relation to the accuracy and completeness of goals achieved.

Satisfaction: User's comfort with and positive attitudes towards the use of the system.

These dimensions are widely used in HCI literature to describe whether an interactive system supports users in achieving their goals in a practical and acceptable way (Dix et al., 2004; Carroll, 2003). A user-centred design approach asks not only whether a system produces a correct output, but also whether users can reach that output with reasonable effort and confidence. This distinction is central to the present project. A board-game allocation system needs to satisfy algorithmic constraints while also helping participants understand what is happening, contribute preferences comfortably, and accept the resulting group decision as usable in practice.

Requirement Identification Methods

In HCI, requirements are often derived from accounts of users, tasks, and settings, not from technical functionality alone. For this thesis, requirement identification draws on three main types of evidence. First, analysis of the prior BGVS system identifies existing interaction assumptions and limitations. Second, observation of board game selection practice can reveal coordination work that users may not describe explicitly, such as who introduces candidate games, who waits, who explains rules, and when a decision stalls.

Third, questionnaires provide broader self-reported evidence about habits, frustrations, selection criteria, and desired support.

Personas, scenarios, and user stories are then useful representations for translating this evidence into design requirements. Personas support reasoning about archetypal users and their goals (Cooper, 1999; Pruitt and Adlin, 2006). Scenarios describe situated activities, constraints, and possible breakdowns in use (Carroll, 2000), while user stories express requirements from a user's perspective in a lightweight form used in agile development (Cohn, 2004). A peer-reviewed CHI example of these representations being used together is the TAC Toolkit (Nadal et al., 2022), where personas, multi-choice scenarios, and positive user stories helped designers reason about technology acceptance across time. In the present thesis, these representations are useful because board game selection involves participant capabilities, not a single stable role structure. The same person may join quickly, propose a candidate game, explain rules, help advance the room, or simply submit preferences during the same session.

Short open-ended questionnaire responses can provide useful requirement signals when analysed cautiously through qualitative content analysis (Hsieh and Shannon, 2005), but they should not be overstated as substitutes for richer interview or observation data (O'Cathain and Thomas, 2004). These methods support a defensible movement from evidence to requirements concerned with preference expression, participant agency, shared room control, constraint visibility, and social acceptability.

Design Process Framing

The Double Diamond (Design Council, n.d.) is better understood as a high-level design-process frame, not as a requirement identification method. The Design Council describes it as a model of divergent and convergent design activity, organised into four phases: Discover, Define, Develop, and Deliver. Here, the framework helps describe how different activities relate over time. Prior-system review, observation, and questionnaire work broaden understanding of the problem. Requirement synthesis and scenarios narrow that understanding into a design direction, prototype construction opens alternative solution paths, and implementation and evaluation narrow those paths into the PickNPlay and Boardot systems. The framework is used as a process lens, not as evidence by itself.

Prototyping and Design Iteration

Prototypes are design artefacts used to explore and refine requirements, not independent requirement sources. Low-fidelity prototypes, such as paper sketches, wireframes, and rough storyboards, are useful for examining interaction flow, terminology, and alternative designs before implementation. High-fidelity prototypes provide a closer approximation of final interaction behaviour and are more suitable when researchers need feedback on visual feedback, cognitive load, perceived control, and realistic task performance. Prototyping research treats prototypes as selective filters that foreground some aspects of a design while suppressing others, so fidelity should be chosen according to the design question being investigated (Rettig, 1994; Rudd et al., 1996; Lim et al., 2008). For this thesis, this distinction separates the evidence used to derive requirements from the prototypes used to test, challenge, and refine them.

Evaluation and Analysis Methods

Evaluation in this context should go beyond checking whether the allocation algorithm returns a feasible result. Work on recommender and explanation systems shows that algorithmic performance alone does not capture whether users find a decision-support system understandable, trustworthy, useful, or controllable (Konstan and Riedl, 2012; Tintarev and Masthoff, 2012). A user study provides a structured way to observe participants using the systems, collect questionnaire responses, and follow up with interviews about how they understood and experienced the decision process. In-lab testing is useful for prototype evaluation because it allows tasks, observations, breakdowns, and immediate feedback to be captured under comparable conditions (Nielsen, 1993). Closed questionnaire items and descriptive measures can capture aspects of usability, perceived fairness, trust, and control. For usability specifically, the System Usability Scale (SUS) (Brooke, 1996) provides a compact standardised instrument for assessing perceived system usability after use. When comparing the two interaction frameworks, a non-parametric test such as the Mann-Whitney U test (Mann and Whitney, 1947) is appropriate for comparing independent groups when the data are ordinal or when normality cannot be assumed.

Qualitative methods then complement these measures by showing why participants responded as they did. Observation can capture group behaviour during use, interviews can elicit interpretations of the process, and Thematic Analysis (TA) (Braun and Clarke, 2006, 2019; Naeem et al., 2023) offers a structured way to identify recurring patterns across interview transcripts, observation notes, and participant comments.

Longitudinal studies provide a further evaluation strategy when the research question concerns how interaction changes after repeated use, beyond a single prototype session. In HCI, this matters because initial impressions often differ from later experience. Early use may be shaped by novelty, learning effort, or confusion, while later use can reveal routines, trust calibration, fatigue, appropriation, and workarounds. Karapanos et al. (Karapanos et al., 2009) frame user experience as something that develops over time, showing that the qualities supporting early adoption may differ from those that sustain longer-term use. Methodologically, longitudinal HCI studies can combine repeated observations, interviews, log data, and diary-style self-reports. Diary studies (Rieman, 1993) are especially useful for capturing situated events between researcher visits and linking later interviews to concrete episodes of use. Retrospective methods such as the UX Curve (Kujala et al., 2011) ask participants to reconstruct how their experience changed over months of real use, making long-term satisfaction, frustration, and perceived value visible instead of treating usability as a one-off judgement. For this thesis, such methods are most relevant as background and future evaluation context. PickNPlay and Boardot can be studied in short in-lab sessions, but repeated board-game nights would better reveal whether participants continue to understand, trust, and accept the decision process after the novelty of the interface has passed.

These methods justify treating PickNPlay and Boardot not only as technical systems, but as interaction designs whose success depends on users' understanding, effort, trust, and social experience.

2.2.2 HCI in Social Choice and Group Decision Support

The preceding section described social choice systems as mechanisms for representing preferences, computing feasible outcomes, and embedding allocation rules in software. From an HCI perspective, however, such a system is not only the rule or solver that produces an allocation. It is also the interaction process through which people understand the available alternatives, express their preferences, interpret the result, and decide whether the procedure was acceptable. This makes interface design a central part of applied social choice: the same algorithmic mechanism can be experienced very differently depending on how preferences are elicited, how constraints are explained, and how the final outcome is presented to the group.

This issue is particularly visible in group decision-support systems. DeSanctis and Gallupe (DeSanctis and Gallupe, 1987) describe group decision support systems as interactive computer-based systems that combine communication, computation, and decision-support technologies to help groups work through unstructured decision problems. This framing is useful for the present thesis because board game selection is not a purely individual recommendation problem. Participants need to coordinate around shared constraints, such as player counts, teaching ability, play time, and willingness to join particular people or games. The interface needs to support both individual preference input and group-level awareness of what the decision process is doing.

Shared control is a direct consequence of this group framing. If the interface gives only one person authority to construct the candidate pool, move the session forward, or recover from mistakes, then the system can reproduce the same coordination burden it is meant to reduce. For a co-located board game group, relevant knowledge is often distributed: one participant may own a game, another may know how to teach it, and another may notice that the group has changed size. A group decision-support interface needs to make the state of the decision visible enough that participants can coordinate around it, while keeping actions understandable enough that responsibility can be shared without making the room chaotic.

The social choice systems reviewed above also show why HCI matters when formal allocation mechanisms become end-user tools. Users do not encounter a theorem, optimisation model, or solver directly, but a sequence of interface choices, explanations, forms, and outputs. In casual social settings, this interface layer becomes even more important because users may not want to study the mechanism in detail, but still need enough feedback to feel that their preferences were represented and that the result is reasonable.

For this reason, the present thesis treats social choice system design as an interaction problem as well as an algorithmic one. A useful background for this framing needs to account for how people provide input, how much effort the interaction demands, how initiative is distributed between users and the system, and how users retain confidence in mediated decisions. These issues are introduced below as general HCI concepts before they are applied to specific system designs in later chapters.

2.2.3 Interaction Modalities, Cognitive Load, and Perceived Control

In group decision support, an interaction modality is not only the device or channel through which users interact with a system. It also describes how the system organises preference input, communication, coordination, and decision feedback among group members. This broader view follows the GDSS framing (DeSanctis and Gallupe, 1987) introduced above: a group decision-support system combines communication, computation, and decision-support functions to help a group handle an unstructured decision problem. The modality of interaction shapes what users need to do, what information becomes visible to the group, and how the decision process is experienced.

A central distinction is whether the system relies on explicit or implicit input. Explicit input asks users to state preferences directly, for example by ranking options, assigning scores, selecting constraints, or confirming choices. This can make the user's intention easier to inspect, but it also adds work to the decision process. In contrast, implicit feedback approaches (Kelly and Teevan, 2003) infer preferences from behaviour that occurs during ordinary system use, reducing the need for separate feedback actions but also making the interpretation of user intention less direct. Cognitive load theory (Sweller, 1988) is useful here because it highlights the limited capacity of working memory during problem solving. In a group decision-support setting, an interface that asks participants to compare many alternatives, remember constraints, and anticipate group outcomes may become burdensome even when the underlying algorithm is sound.

Another relevant distinction concerns initiative. Some systems are mainly participant-initiated, requiring users to decide when and how to provide information. Others are facilitator-led, guiding the group through a sequence of prompts, reminders, or decision steps. Mixed-initiative interaction (Horvitz, 1999) describes a middle ground in which human users and computational agents can each take initiative when they are better positioned to contribute. This concept is useful for understanding interactive decision-support systems because it separates the question of who computes the outcome from the question of who steers the interaction at each moment.

As systems take on more of the work of eliciting preferences, guiding attention, or computing outcomes, transparency and perceived control become more important. Prior HCI work on algorithmic interfaces shows that transparency can affect user trust, particularly when automated outcomes differ from what users expect (Kizilcec, 2016). At the same time, explanation and intelligibility research cautions that making an automated system understandable is not simply a matter of exposing more internal detail. Explanations need to support useful mental models and allow users to see when intervention is possible (Abdul et al., 2018; Heer, 2019). For group decision support, the design challenge is to reduce unnecessary interaction burden while still giving participants enough feedback, contestability, and control to regard the process as acceptable.

2.3 Boardroom BGVS System

2.3.1 Backend — Algorithms, Implementations

Boardroom (Verrell, 2025), formally the Board Game Voting System (BGVS), is the immediate previous work for the present project and provides the main technical and interaction baseline for this thesis. It already frames board game selection as a constrained player-to-game assignment problem, not as a simple poll or recommendation task. Its backend should be understood as a room resolver: after a room has been prepared and participants submit rankings, the resolver receives the current room state and returns a list of player-game assignments. In the archived implementation, the frontend stores the room in Firebase Realtime Database, then posts the preference lists, per-game lower and upper quotas, and optional teaching information to a Python Firebase Function. The function runs the Z3 optimisation model and returns assignments as pairs of player and game indices, which are then written back into the room record for the results screen. This makes the backend largely stateless: persistent coordination is handled by the room database, while the algorithm itself solves one snapshot of the voting session at a time.

The solver formulation uses a Boolean decision matrix. For each player p and game g, a variable $a_{p,g}$ indicates whether player p is assigned to game g. The first hard constraint requires every player to be assigned to exactly one game:

$$\sum_{g=0}^{m-1} a_{p,g} = 1.$$

The second hard constraint enforces each game's playable player range while allowing games with insufficient demand to remain unassigned. If q_g^{min} and q_g^{max} are the minimum and maximum player counts for game g, then the number of assigned players must either be zero or fall inside that interval:

$$\sum_{p=0}^{n-1} a_{p,g} = 0 \quad \vee \quad q_g^{min} \leq \sum_{p=0}^{n-1} a_{p,g} \leq q_g^{max}.$$

This is the key difference between Boardroom and a conventional single-winner vote. The backend does not simply select the most popular game. It decides which subset of games can run and how players should be distributed across them.

Player preferences are represented as complete rankings over the selected games. The participant interface can move unwanted games into a visually separated soft exclusion area, but the implementation still appends those games to the end of the submitted ranking instead of treating them as hard vetoes. The exclusion interaction should be read as a soft low-priority signal. The objective maximises aggregate preference satisfaction over all assignments. In the source implementation, the score for assigning player p to game g is computed from the rank $r_{p,g}$ as

$$s_{p,g} = m - r_{p,g} - 2^{r_{p,g}},$$

and Z3 maximises $\sum_{p,g} s_{p,g} a_{p,g}$. The exponential term makes low-ranked games substantially less attractive to the solver, while the additional linear rank term gives finer separation between adjacent positions in the ranking. This scoring model is pragmatic, not axiomatically derived: it encodes a strong preference for satisfying high-ranked choices, but it does not expose a formal fairness guarantee such as envy-freeness or proportionality.

Boardroom also models teaching ability as a soft practical playability signal. Participants may mark games they can teach, producing a `teachMap` that maps each player to the games they are willing to explain. For each game with at least one participant able to teach it, the backend adds a soft constraint encouraging at least one such participant to be assigned to that game. The implementation deliberately adds this soft constraint after the main preference-maximisation objective, so teaching coverage acts as a tie-breaker among high-satisfaction assignments instead of dominating the primary preference objective. This design choice reflects a practical trade-off: a game with someone able to teach it is easier to run, but assigning someone to an undesirable game only because they can explain the rules may undermine the purpose of preference voting.

As a baseline for the present work, Boardroom's backend demonstrates that a solver-based formulation can handle the core structural constraints of board game allocation: one assignment per player, unassigned games, minimum and maximum player counts, and a secondary soft teaching signal. Its limitations are equally important. The algorithm returns only the final assignment and a success flag, so users receive little insight into why particular trade-offs were made. It also treats the submitted ranking as the main preference representation, leaving richer interaction needs such as vetoes, ties, negotiation, explanation, and iterative adjustment outside the backend model. These gaps motivate PickNPlay and Boardot, which retain the value of solver-backed allocation while placing more emphasis on interaction, transparency, and user control.

2.3.2 Frontend — Interaction and UX

The Boardroom frontend (Verrell, 2025) made the solver accessible through a lightweight room-based workflow. Its interaction model separated users into two fixed practical roles: a room administrator who prepared and controlled the session, and participants who joined the room to submit preferences. As shown in Figure 2.3, the room administrator first created a room, imported or added candidate games, selected the games to include, and then waited while participants joined through a room link, QR code, or room ID.

Once voting was complete, the administrator clicked resolve, the system ran the backend algorithm, and both administrator and participants could view the resulting assignments.

The administrator-facing interface concentrated most setup and control in one screen sequence. The room creation flow used a searchable table of games, populated either from BoardGameGeek data or manual entry, with selected rows forming the pool for the voting session. This kept voting relatively fast when one host had already prepared the candidate list: no game library or accounts were required. However, this workflow did not match all casual board game settings, where several participants may bring games, suggest options during discussion, or add candidates opportunistically at the table. Boardroom treated candidate-pool construction as administrator work instead of collaborative group work, so participants entered the process mainly after one person had already created the room and selected the options. In the voting-room view, the administrator could still adjust each game's minimum and maximum player counts before resolving the vote. This is an important UX detail because player-count quotas are not merely static metadata: in an actual board game session, a group may decide that a game is only worth running at a preferred count, or may relax a capacity constraint when the available group size changes.

For participants, Boardroom used a mobile-oriented ranking interface. Participants entered a name, then reordered game cards by dragging them from most to least preferred. The cards included game thumbnail images, playing-time information, a control for marking whether the participant could teach a game, and a mechanism for moving

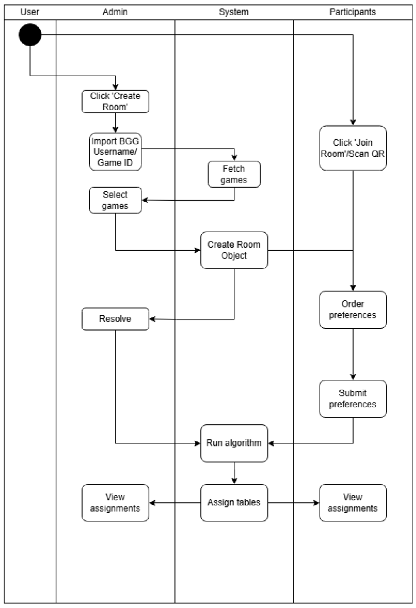


Figure 2.3: Boardroom activity flow, showing the separation between administrator, system, and participant actions. Source: Boardroom (Verrell, 2025).

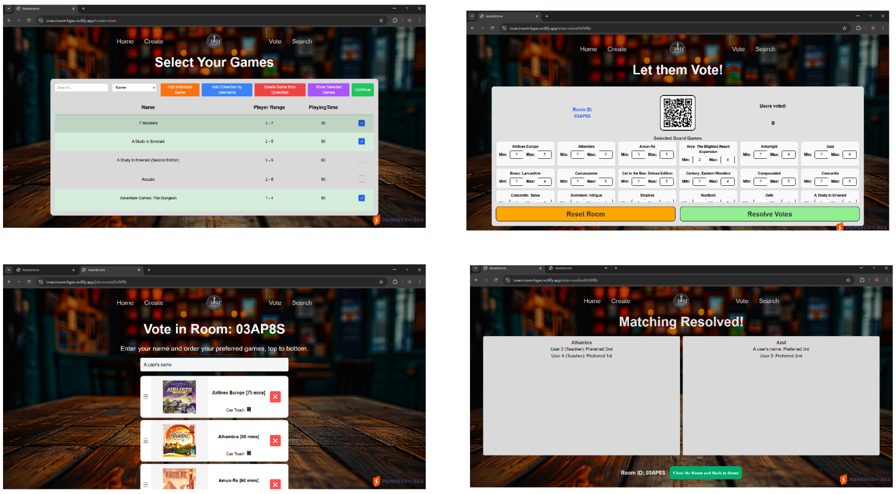


Figure 2.4: Key Boardroom interface states: game selection, administrator voting room with QR code and quotas, participant preference ranking, and resolved assignments. Source: Boardroom (Verrell, 2025).

games into a soft exclusion or low-priority area. This design translated the backend's ranked preference input into a direct manipulation task: users did not need to understand the solver, only to arrange the cards in the order they preferred. After submission, participants saw a waiting state and were redirected to the results once the administrator resolved the room.

The resolved-results screen presented assignments grouped by game, not by participant. Only games with assigned players were displayed, which reduced clutter from games that were not selected to run. Each assignment also showed the participant's preference position for the assigned game, and teaching status could be marked where relevant. This gave users some immediate evidence that the final allocation was connected to their submitted rankings, but the explanation remained outcome-level: the interface showed what was assigned, but not why the solver preferred one feasible allocation over another.

As a UX baseline, Boardroom solved the main access problem well: it lowered the coordination cost of running a preference-based vote in a casual setting, supported phones through QR-code entry, and exposed the minimum controls needed to run the backend algorithm. However, its interaction model was still linear and role-separated where one person prepared and controlled the room, and the others primarily submitted rankings and waited. This fixed setup role is a limitation of the previous baseline. It makes the frontend a useful reference point for the requirement analysis in Chapter 3, where fixed setup responsibility, ranking interaction, and outcome-level feedback are considered as design gaps.

2.4 Review of Related Systems

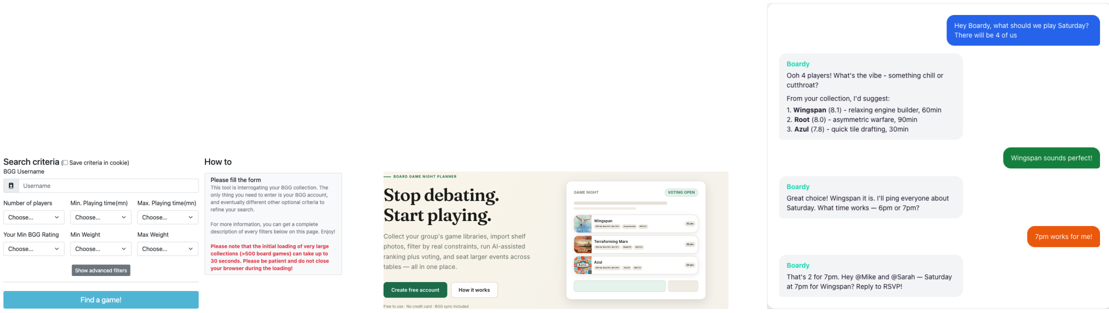
2.4.1 Board Game Selection Systems

Selecting a board game is a recurring coordination problem in hobby groups. This challenge stems from the need to balance the desired criteria such as the number of available players, expected duration, game ownership, familiarity with rules, player preferences, and the social fit of the occasion. Board game selection systems cover a broad range of interventions, from collection filters and randomisers, to recommendation algorithms, group voting tools, AI-assisted planning tools, and physical meta-games that make the act of choosing playful.

A first class of systems treats selection primarily as a filtering problem over an available collection. Cardboard Butler (Kristoffersen, 2018), Gameshelf (GameShelf.io, 2026), BGG Best Games for Player Count (Hisfantor, 2023), Board Game Selection Tool (Furlong, 2016), Board Game Pick (Board Game Pick, 2026), Board Game Chooser (Boneff-Peng, 2023), DiceDecider (DiceDecider, 2026), and CSV-based board game pickers (Zimmer, 2026) all use BoardGameGeek data or exported collections to narrow the candidate set by properties such as player count, play time, ownership, category, rating, complexity, or play history. BoardGameGeek (BoardGameGeek, 2026) is an especially common data source, although its "collection" concept is broader than physical ownership: it can include games a user owns, has played, has rated, has commented on, or otherwise wants to track. These tools are useful because they reduce a large library into a plausible shortlist, but the decision model is usually individual, filter-driven, or random. The social negotiation still happens outside the tool.

Group voting and planning tools move closer to the problem addressed here. What2Play (What2Play, 2026b,a), for example, supports event setup, BoardGameGeek synchronisation, ranked voting, host controls, feedback records, and even multi-table seating for larger events. These features acknowledge that game choice is not merely a catalogue query, but a multifaceted coordination process involving attendance, timing, availability, and group preference. However, the model still mainly guides groups from a shortlist to a final game, or from an event plan to table seating. While this model performs better than simple randomisers, it remains unclear how much control participants have over constraints and how well the system explains trade-offs between feasible options.

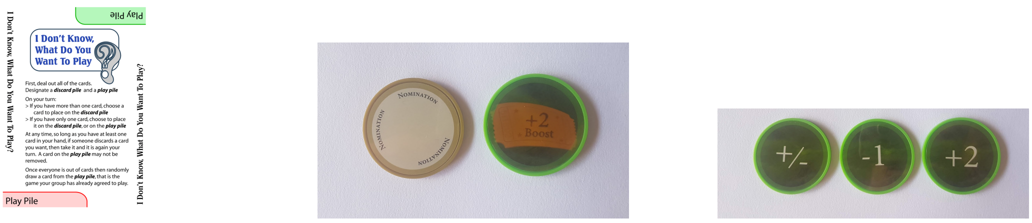
A newer set of tools frames planning labour itself as the site of intervention. Boardy (Rocket Power Software, LLC, 2025) advertises an SMS-native AI game-night assistant that can recommend games based on player count, time, a group's collection, and past preferences, while also handling reminders, RSVPs, availability polls, collection management, and play tracking. Game Night Picks (Game Night Picks, 2026) similarly presents itself as a planning platform for public and private game nights, including scheduling, suggestions, voting, personal preferences, expertise levels, play records, and collection management. These systems are evidence that the practical labour of running a game night matters as much as the final recommendation. At the same time, their public descriptions do not provide enough detail to evaluate the underlying decision rules, so their support for transparency, contestability, and constraint negotiation remains difficult to assess. Figure 2.5 illustrates how digital systems expose different parts of the game-selection process: filter-driven shortlist construction, event-level voting and recommendation, and conversational AI facilitation.



(a) Filtering. (b) Planning and voting. (c) AI facilitation.

Figure 2.5: Examples of digital board-game selection and planning systems. Source: Board Game Pick (Board Game Pick, 2026), What2Play (What2Play, 2026b), and Boardy (Rocket Power Software, LLC, 2025).

Physical meta-games approach the same problem from a different direction. Two separate examples are *I Don't Know, What Do You Want to Play?* (Ludopedia, 2026), a 2007 web-published card game in which a group creates cards from its own collection and plays them down to a final random choice, and *Game to Pick a Game: The Gateway Edition* (Board Game Oracle, 2026), a game from Chip Theory Games explicitly about selecting what to play. These examples do not try to optimise a recommendation. Instead, they turn indecision into a shared activity, using rules, piles, cards, tokens, and turn-taking to make the choice feel less like a stalled conversation. Figure 2.6 shows how such systems make preference expression tangible through pile-based elimination, nomination, boosting, and bidding.



Pile-based card play. Nomination and boost. Bidding tokens.

Figure 2.6: Physical meta-games can turn game choice into an explicit interaction procedure instead of a hidden recommendation. Source: *I Don't Know, What Do You Want to Play?* listing (Ludopedia, 2026); Game to Pick a Game (Expand Your Game, 2019) photographs.

Academic board game recommender systems address another part of the space. Ion (Ion, 2018) develops and evaluates collaborative, content-based, and hybrid recommendation approaches using BoardGameGeek data, with the goal of suggesting games that a user is likely to enjoy. Zalewski, Ganzha, and Paprzycki (Zalewski et al., 2019) likewise describe a recommender approach based on clustering board games and building user profiles over those clusters. This work is valuable for modelling item similarity and preference prediction, but it mainly treats board game choice as a recommendation task. It does not directly address the interaction problem of helping a present group inspect, revise, and accept a collective allocation.

Across these examples, the existing systems divide the problem into different pieces. Filters narrow the catalogue, recommenders predict relevance, voting tools aggregate preferences, planning tools reduce coordination work, and physical games make the choice itself social. The present thesis is concerned with the space between these approaches: how a group can express preferences and constraints, receive a computed allocation, and still understand and adjust the decision process.

2.4.2 Gap: Collective Allocation as an Interaction Problem

The main gap in the reviewed systems is that they usually optimise one part of the activity while leaving the collective allocation process implicit. A filter can say which games fit four players and ninety minutes, a recommender can predict which games resemble a user's favourites, a voting tool can identify a popular candidate, and a planning tool can reduce coordination work around attendance and reminders. None of these functions alone explains how a present group should resolve competing rankings, player-count quotas, teaching ability, and social acceptability when several feasible outcomes exist.

This distinction matters for PickNPlay and Boardot, which build on the Boardroom baseline as group decision-support systems, not only recommendation tools. The relevant input is not just a catalogue and a user profile, but a situated group with expressed rankings, available games, lower and upper quotas, and knowledge about who can teach particular games. The output is not simply "the best game", but a practical and defensible arrangement of people and games under constraints. The Boardroom baseline showed that solving the assignment alone does not make the process transparent or adjustable. This thesis explores how filtering, preferences, constraint solving, explanation, and adjustment can help groups reach decisions they understand and trust.

2.5 Related Background Considerations

2.5.1 Anonymity

Here, anonymity is defined as reducing the visibility of who submitted which preference, not as the formal social-choice axiom of anonymity (Brandt et al., 2016), where a rule should treat voters symmetrically regardless of identity. This distinction matters because a board game allocation system can be anonymous while still making individual rankings visible to other people through the interface.

Anonymity may become relevant in future versions because small social groups can create pressure around rankings, exclusions, or unpopular choices. A participant may be reluctant to rank a friend's favourite game low, or to reveal that they do not want to join a particular activity. However, full anonymity is not treated as a core requirement in the present work because the final outcome still needs to identify who is playing which game, and because teaching availability and group coordination require some information to remain visible. The design problem for this thesis is calibrated visibility, not full anonymity: the system should show enough preference and constraint information to make outcomes intelligible without exposing so much that participants feel socially pressured.

2.5.2 Privacy

Privacy is used here as a related but broader concern: it concerns how identity, preference, and game-related information flows through the system, not simply whether data is hidden. This follows the idea of contextual integrity (Nissenbaum, 2004), where privacy depends on whether information gathering and sharing are appropriate to the social context and its norms.

For board game selection, preference rankings, excluded games, teaching ability, and play history can reveal taste, confidence, expertise, or social comfort. These are not usually high-stakes data in the same way as financial or medical records, but they can still affect how comfortable people feel when contributing honestly. Full privacy-preserving allocation is outside the scope of this thesis. However, preference disclosure remains a design concern because explaining outcomes can create social pressure if individual rankings, exclusions, or teaching claims are exposed unnecessarily. The present work focuses on preference collection, allocation, explanation, and group usability while treating privacy-preserving mechanisms as future work.

Chapter 3

Requirement Analysis and System Overview

The previous chapter established the theoretical and practical background for this project: board-game selection can be understood as a constrained group-allocation problem informed by social choice, but existing systems often leave the human process of expressing, negotiating, and accepting preferences under-supported. This chapter turns that background into requirements for PickNPlay and Boardot. It does not present the later evaluation of the implemented systems. Instead, it uses formative evidence from questionnaire responses, baseline-system analysis, early observation, public community examples, low-fidelity prototyping, and design critique to derive, refine, and triangulate the requirements that informed PickNPlay and the exploratory Boardot extension. The chapter's central requirement-analysis move is to replace fixed administrator/participant roles with participant capabilities that can shift during a session: joining, contributing, teaching, interpreting, and helping the group resolve.

Section 3.1 first characterises board game selection as a situated group activity, not a simple recommendation task. Section 3.2 then describes the evidence sources and translation process used to derive requirements. Section 3.3 presents the preliminary questionnaire analysis. Section 3.4 adds contextual evidence from public board-game community discussions. Section 3.5 identifies limitations in the earlier BGVS system and in informal selection practices. Section 3.6 defines the relevant participant capabilities, scenarios, and user stories. Section 3.7 describes low-fidelity prototype work, wireframes, and formative design critique. Sections 3.8–3.9 then define the design goals and system requirements. Finally, Section 3.11 summarises how these requirements guide PickNPlay as the primary participant-directed framework and Boardot as an exploratory AI-mediated facilitator layer over the same room workflow. A central design commitment across both is that the room is shared: creating or advancing a room is a technical action, not automatically a persistent social role with special authority.

3.1 Problem Setting: Board Game Selection as Constrained Group Allocation

In a casual board game session, the group is not only choosing a popular game. Participants need to jointly decide which games are available, how many people can play each game, whether someone can teach the rules, how much time the group has, and how strongly different participants prefer different options. The practical outcome may also involve splitting a group across several tables instead of selecting a single winner. This makes the problem closer to constrained group allocation than to a conventional one-person recommendation task.

Here, the outcome is not simply a single winning game. The outcome is a feasible assignment of participants to one or more game tables. Each candidate game has feasibility constraints, especially minimum and maximum player counts, and may also depend in practice on whether at least one participant can teach or explain the game. Teaching ability is treated as a soft practical playability signal, not as a universal hard feasibility condition. Participants provide preference information over candidate games, and the system needs to transform those individual inputs into a collective assignment that balances feasibility, preference satisfaction, and social acceptability.

From a social choice perspective, each participant contributes preference information and the system computes a collective outcome. From an HCI perspective, however, the quality of the system depends on more than the mathematical validity of that outcome. Participants need to know what they are being asked to provide, the group needs shared control over session constraints, and everyone needs enough feedback to treat the result as fair and actionable. A correct allocation can still fail as an interaction if participants experience it as opaque, socially awkward, or disconnected from the discussion that normally shapes game choice.

The requirements distinguish between computational fairness and interactional fairness. Computationally, the system should produce feasible assignments under player-count constraints and prefer allocations with stronger aggregate preference satisfaction. Participants should be able to see how preferences and constraints contributed to the outcome, identify whether a result needs revision, and understand that no single participant controlled the process implicitly. The project prioritises the perceived procedural fairness above solving every formal fairness criterion for allocation.

The requirement analysis treats board game selection as a socio-technical workflow with four recurring tensions:

Constraint versus preference: A game can be highly preferred but infeasible because of player count or difficult to play in practice because of the timing, complexity, or teaching coverage.

Distributed control versus role concentration: A room may be initiated by one person, but the setup, progress, and decision control should not depend on that person acting as a permanent controller.

Automation versus agency: A solver or AI-mediated facilitator can reduce coordination burden, but too much automation can make participants feel that the decision is imposed on them.

Individual input versus group acceptance: Participants submit individual preferences, while the final decision needs to be accepted by the group as a shared social outcome.

These tensions motivate the requirements below. The system is required to collect rankings and output assignments, and support understandable, socially acceptable decision-making.

3.2 Requirement Analysis Method

Following the requirement identification methods reviewed in Section 2.2.1, the requirements were developed through an iterative user-centred process. Personas, scenarios, and user stories were used as practical representations for turning observed situations into system requirements. The overall process was framed by the Double Diamond model introduced in Chapter 2. The project first broadened understanding of the problem through prior system analysis, early observation, and questionnaire data. It then narrowed that understanding into baseline gaps, participant capabilities, and design goals, broadened again through prototype development for the direct PickNPlay workflow and the exploratory Boardot facilitation layer, and finally narrowed into the system requirements and implementation choices evaluated later. The resulting requirements are not treated as fixed specifications discovered once at the start of the project. Instead, they are working design commitments that were refined across the development and study stages.

3.2.1 Evidence Sources

Table 3.1 summarises the main evidence sources used in this chapter. These sources play different roles. Some identify problems in the previous system, some describe current group practices, and some help translate design intentions into implementation requirements.

The evidence sources are not treated as having equal methodological status. The questionnaire and Stage 1 materials provide project-specific empirical evidence. The public forum examples are used as contextual triangulation, not as systematically sampled qualitative data. Personas, scenarios, wireframes, and artefact boards are design synthesis artefacts, not independent empirical findings. Low-fidelity prototyping and formative design critique helped refine interaction decisions, but they are not treated as substitutes for participant evidence. Stage 1 is treated as baseline formative inquiry: it informs what the new system should address, but it is not the evaluation of the implemented PickN-

Table 3.1: Evidence sources used to derive and refine requirements.

Evidence source	**Role in requirement analysis**
Review of BGVS and its thesis documentation	Identifies baseline workflow, solver assumptions, and limitations in explanation, interaction, and participant control.
Stage 1 formative observation of existing board game selection practice	Provides situated evidence about how groups introduce games, discuss options, and use BGVS in a real session.
Early-stage preliminary questionnaire	Collects self-reported current practices, selection criteria, frustrations, and feature suggestions from board game players.
Public board-game forum discussions	Triangulates that large-group selection, shared candidate pools, player-count constraints, vetoes, and coordination burden also appear in public hobbyist discussions outside this study.
User stories and personas	Converts recurring needs into concrete interaction scenarios for room participants, game proposers, and participants able to teach or explain games.
Low-fidelity prototyping and formative design critique	Supports iteration on room setup, voting flow, mobile interaction, visual feedback, and Boardot facilitation before full implementation.

Play or Boardot systems. Requirements are strongest where multiple sources converge, especially where questionnaire responses, observation, and baseline-system analysis point to the same problem.

3.2.2 From Evidence to Requirements

The analysis followed a four-step translation process. First, the evidence sources were used to identify recurring problem signals: single-person setup labour, limited preference expressiveness, opaque allocation results, and weak support for discussion. Second, these signals were rewritten as participant capabilities and scenarios instead of isolated feature requests. Third, the scenarios were condensed into design goals that describe what the system should make easier, clearer, or more controllable. Fourth, the design goals were converted into functional, interaction, and explanation requirements that could guide the implementation of PickNPlay and Boardot.

This process mirrors the general HCI movement from personas and scenarios toward user stories and system requirements discussed in Chapter 2, but it is adapted to this project. The emphasis is not on producing a full interface specification at this stage. Instead, the aim is to make the design rationale explicit to ensure participant-facing requirements remain traceable to a problem in the prior system or a signal from early empirical work. The study-specific need to evaluate PickNPlay while also probing an AI-mediated station extension is documented separately as a research constraint.

The labels used below are project-specific traceability notation, not a separate formal method. Goal-oriented requirements engineering (van Lamsweerde, 2001) motivates keeping high-level goals explicit before refining them into actionable requirements, while requirements traceability (Gotel and Finkelstein, 1994) motivates preserving links between requirements, their sources, and later design artefacts. Here, DG labels denote design goals, FR labels denote required system functions, IR labels denote interaction qualities and workflow expectations, and ER labels denote explanation, trust, and preference-disclosure requirements. Research constraints are labelled separately as RC because they support the study design, not the participant experience. This separation is useful because PickNPlay and Boardot support shared room control, lightweight group interaction, understandable outcomes, and later empirical evaluation.

3.3 Preliminary Questionnaire Analysis

The early-stage questionnaire was primarily used as a requirement-scoping instrument. Twenty-seven visible consenting responses were collected between 17 October 2025 and 1 May 2026. The questionnaire contained structured items about board game habits, group size, selection methods, satisfaction, frustration, and the perceived risk that quieter preferences might be overlooked. It also included two open-ended questions about selection challenges and desired features for a game selection application. The questionnaire remained open during the project, so the current visible export is reported here as formative context, not as a strict chronological record of every design decision. Earlier responses informed requirement scoping directly, while later responses were used as additional contextual support for signals already being investigated through the baseline review and formative studies.

Closed responses were summarised descriptively, while short free-text responses were pilot-coded using the qualitative content analysis approach discussed in Section 2.2.1. The purpose of this coding was to identify candidate requirement signals, not to claim final themes or saturation. Table 3.2 summarises the main descriptive findings and how they informed the requirement analysis.

The results suggest a balanced design problem. Existing selection practices are not simply broken. Many respondents were satisfied with open discussion, and several reported little or no difficulty. At the same time, the structured and free-text responses show recurring situations where informal discussion becomes strained, especially when player count changes, available time is limited, game complexity differs across participants, or a few people carry the work of proposing and explaining games. For this chapter, the questionnaire contributes early requirement evidence for constraint visibility, lightweight preference input, and shared room contribution. Fuller evaluative theme development is reserved for Chapter 5.

Table 3.2: Preliminary questionnaire findings used for requirement scoping. Counts describe the current visible export only and are not treated as final prevalence claims.

Aspect	Preliminary result	Requirement implication
Participant context	14 of 27 respondents reported playing once a week; 24 of 27 reported regular group sizes of four to eight people.	The data mostly reflects active board game players working in small-to-medium group settings, not isolated individual recommendation use.
Current selection practice	22 of 27 respondents reported open discussion or negotiation until consensus. Satisfaction was generally positive, with 20 of 27 rating the current method as 4 or 5 on the satisfaction item.	The system should augment existing discussion instead of replacing it with a fully automated decision process.
Selection criteria	The most frequently selected criteria were playtime (20 selections), complexity or weight (16), and supported player count (14).	Game context and feasibility constraints need to be visible before voting and meaningful to the allocation process.
Friction and fairness risk	Responses to frustration and overlooked-preference items were mixed: 9 of 27 rated selection as highly frustrating or time-consuming, and 8 of 27 gave high ratings for overlooked-preference risk.	The system should target specific breakdowns without assuming that all groups experience game selection as a severe problem.
Open-ended responses	24 respondents described a challenge and 21 suggested a feature. Pilot codes included player-count fit, time constraints, complexity fit, preference conflict, quieter voices, proposer burden, filtering, veto/fairness support, remembered preferences, and random discovery.	These signals support lightweight filtering, flexible preference expression, shared setup, and transparent result feedback as candidate requirements.

3.4 Public Evidence from Board-Game Communities

Public board-game community discussions were used as contextual triangulation, not as a formal empirical dataset. They are useful because they show the same coordination problem being articulated outside this study. Large groups need to identify which brought games are available, decide whether to split into several tables, negotiate who takes responsibility for moving the decision forward, and determine how individual preferences or vetoes should be represented. These sources are included to show ecological plausibility, not prevalence. They do not establish how common these problems are across board-game communities.

A public BoardGameGeek recommendations thread from February 2026 (BoardGameGeek forum community, 2026) raises a scenario closely aligned with this thesis which is a large group with multiple brought games, and a request for tooling to help choose what the group should play. Responses in the thread discuss group splitting, reluctance to take initiative, differences in player taste, player-count settings, and the need to satisfy most people. Later replies add rota-based selection, hard constraints such as timing, table grouping and teaching ability, ranked preference input, optimisation around least-disappointed players, shared-pool assumptions, and inclusive group formation. Figure 3.1 shows anonymised browser captures prepared from the thread context. The figure is included as contextual evidence only, not participant data or part of the later thematic-analysis corpus.

Other public community sources point to the same requirement signals. BoardGameGeek's game-day hosting guidance (BoardGameGeek wiki contributors, n.d.b) emphasises expected attendance, flexibility around player numbers, multiple tables, and the need for a capable explainer at each table. A BoardGameGeek wiki discussion of frequent recommendation requests (BoardGameGeek wiki contributors, n.d.a) notes that large-group strategic play is often constrained by player count, play time, and depth, and that splitting into smaller groups is often the practical answer. A Reddit discussion asking for a tool to help a group decide what to play (Reddit r/boardgames community, 2023) similarly describes shared pools of owned games, voting, vetoes, suggestions, one-person maintenance burden, and multi-group assignment. These sources support the same direction as the questionnaire and Stage 1 evidence: PickNPlay and Boardot should support shared candidate construction, lightweight preference expression, player-count feasibility, group splitting, visible progress, and explainable outcomes without relying on one participant to coordinate everything.

3.5 Baseline Gaps in BGVS

As reviewed in Section 2.3, the earlier BGVS system provides the main baseline for this project. It demonstrated that a solver-backed board game allocation system is feasible: a single designated user can create a room, select candidate games, collect participant rankings, and resolve the session into assignments under player-count constraints. This

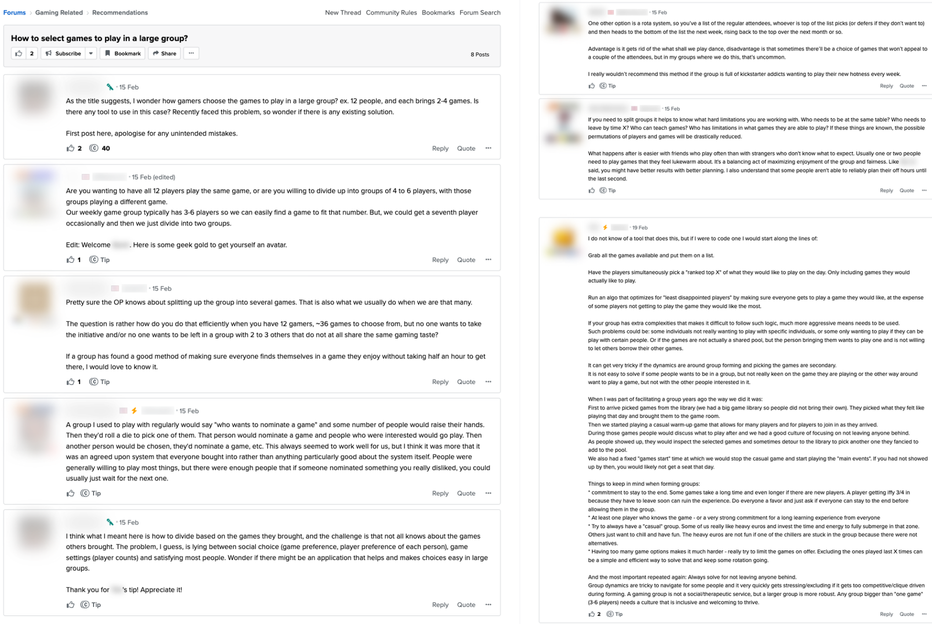


Figure 3.1: Anonymised combined browser capture from a public BoardGameGeek thread about selecting games for a large group. Left: the initial problem and early replies; upper right: rota selection, hard constraints, timing, teaching, and fairness; lower right: ranked input, least-disappointed-player optimisation, shared pools, fixed start times, and inclusive group formation. Account names, avatars, and visible personal-name references were blurred before capture.

makes BGVS a strong starting point because it already frames game choice as a collective allocation problem, not only as catalogue filtering.

Stage 1 was a formative baseline inquiry into BGVS use and existing board-game selection practice. It added situated evidence about how this baseline workflow was experienced in use. The Stage 1 evidence used in this section comprised two focus-group transcripts and fourteen completed System Usability Scale (SUS) responses: seven Group 1 responses and seven Group 2 responses. Audio recordings, participant photos, consent documents, and NVivo project files were used only as audit artefacts, not as primary evidence for claims in this section. The SUS material was not treated as a standalone usability evaluation because it came from two small formative sessions, not from a controlled summative study. Instead, the response patterns and written annotations were read alongside the focus-group transcripts to identify where participants experienced friction in the baseline workflow. Across these sources, the strongest signals concerned setup responsibility, the effort of expressing preferences, and difficulty interpreting whether the final allocation reflected preferences, constraints, or setup errors.

The two Stage 1 groups had different evidentiary shapes. Group 1 was a larger co-located board-game session with seven SUS responses and a focus-group discussion involving experienced board-game players, some of whom had seen earlier versions of the system or were familiar with the recurring game-night context. Group 2 contributed seven SUS responses, with two participants also taking part in the focus-group discussion used for qualitative excerpts. Quoted speakers are cited using group-local participant pseudonyms, such as G1-P2 and G2-P1, so the excerpt evidence remains traceable without identifying participants across the whole project.

Figure 3.2 summarises the quantitative SUS evidence used in this section. Group 1 scores clustered around a moderate descriptive mean, while Group 2 scores were lower overall and more widely spread. Because the sample is small and formative, the figure is used only to show descriptive baseline friction, not to compare groups or evaluate final usability.

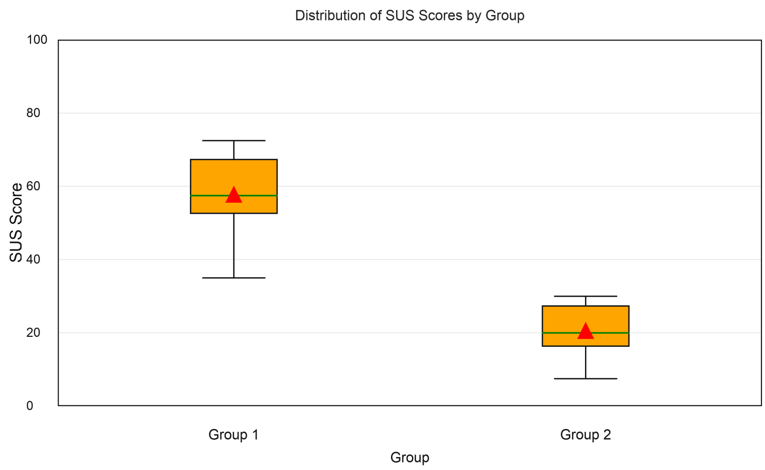


Figure 3.2: Descriptive Stage 1 SUS score summary for BGVS. The scores are treated as formative requirement evidence, not as a final usability evaluation.

Figure 3.3 summarises the corresponding qualitative signals from the two focus-group transcripts. The figure shows how both groups surfaced similar requirement problems, although with different emphases. These signals were used to organise the four baseline gaps below, not presented as final thematic-analysis themes.

Together, the quantitative and qualitative evidence suggested four gaps that limit BGVS as an end-user interaction system.

3.5.1 Single-User Room Control

The BGVS workflow placed most control with the person who set up the session. That person created the room, selected games, adjusted game constraints, waited for votes, and triggered the final resolution. This is practical when one person has already prepared a candidate list and can manage a central device, but it concentrates setup labour and de-

Baseline gap	Group 1 focus group	Group 2 focus group
Single-user room control	G1-P2 joked that the room creator became “the dictator”; later discussion described slow game entry and manual recovery after a setup error.	G2-P1 said the room creator “cannot vote again”; G2-P2 noted that “only one person” could add games.
Preference/context limits	G1-P9 wanted to say “I don’t know what this is”; G1-P1 questioned strict ordering when several games were similarly acceptable.	G2-P1 said a player “cannot know each game” without descriptions; G2-P2 said descriptions were “not enough”.
Outcome explanation/trust	G1-P7 questioned assignments where nobody could teach the game; G1-P2 said users “can’t inspect the votes”.	G2-P2 described “questions about the process” when lower-preference outcomes appeared.
Weak social negotiation support	G1-P2 valued reducing the influence of whoever was “loudest and most decisive”, but the group still negotiated manually.	G2-P2 said the group discussed entering the platform instead of “selecting the game”.

Figure 3.3: Qualitative Stage 1 focus-group signals used to organise the BGVS baseline gaps. The cells summarise pilot-coded requirement signals, not final thematic-analysis themes.

cision pacing in one place. The frontend effectively treated candidate-pool construction as single-person setup work instead of collaborative group work, so participants entered the process mainly after one person had created the room and selected the options.

Stage 1 made this concentration visible in participant accounts. In Group 1, G1-P2 joked that the setup actor became “the dictator” because that person effectively described what the group was playing. Later in the same session, G1-P2 called initial game entry “the most painful thing about the app”, explaining that small setup mistakes were hard to recover once participants had already voted. Group 2 made the same issue explicit from a smaller-session perspective: G2-P1 said that “who created room, that person cannot vote again”, while G2-P2 said that “only one person” could add games, leaving others waiting or trying to work out how to join. These excerpts support the gap as a participant-facing issue, not only as an implementation inconvenience.

This role model does not fit all casual board game settings. Several participants may bring games, know different candidate games, suggest options during discussion, or add candidates opportunistically at the table. In those cases, requiring one person to create the room and manage the candidate pool, often through an interface better suited to a laptop than to quick phone-based setup, becomes part of the coordination problem instead of a solution to it. The questionnaire and observation data both suggest that this concentration matters: some groups already rely on one person to bring, shortlist, or explain games, and a system can either reinforce or redistribute that labour. A new system should treat room control as distributed group work. Anyone may initiate a room, but the room itself should remain a shared object that participants can contribute to, monitor, revise, and advance together.

This does not mean that shared control should be uncontrolled. Candidate-game entry is a form of agenda control because it shapes which outcomes are possible before preference input begins. The requirement is to avoid assuming a permanent privileged controller role while making critical actions, such as adding games, starting voting, resolving the room, or revising mistaken input, visible, recoverable, and where necessary confirmable by the group. Shared control also introduces a governance problem: actions such as starting voting, removing games, returning to setup, or resolving the room affect all participants. The design should balance distributed agency with lightweight safeguards so that room-level changes are visible to the group, reversible where possible, and confirmable where they change the state of the whole room.

3.5.2 Limited Preference Expressiveness

The participant voting interface asked users to produce an ordered ranking. Ranking is a useful preference representation for a solver, but it does not capture all of the nuance in casual game choice. Participants may be indifferent between several games, strongly opposed to one game, happy to play anything, or mainly care about time and complexity instead of a named game. The earlier exclusion area also behaved as a low-preference ordering, not a hard veto. This gap motivates requirements for more expressive, lightweight preference input.

The Stage 1 transcripts support this gap directly. G1-P1 suggested that the system should not require a "strictly ordered rating system" when several games were similarly acceptable, and G1-P2 described a common preference condition as wanting to avoid games "more than an hour" instead of wanting one named game. Several comments pointed toward scoring, filters, or an explicit "not applicable" style option instead of a single total order. Participants also struggled to rank unfamiliar games: G1-P9 wanted a way to say "I don't know what this is", G2-P1 said that "as a player, I cannot know each game" without a small description, and G2-P2 said that the game descriptions were "not enough" unless everyone already knew the games. The issue was not only the mathematical preference format, but the interaction cost of asking people to produce precise rankings without enough context.

These cases are conceptually distinct. Indifference means several games are equally acceptable, unfamiliarity means the participant lacks enough information to rank the game confidently, low preference means the participant would prefer not to play a game but could accept it, hard unacceptability or a veto means the participant treats the option as unsuitable, and constraint preference means the participant mainly cares about features such as playtime, complexity, or teachability, not the game itself. The requirements should avoid collapsing all of these cases into a single low-rank position.

3.5.3 Outcome Without Sufficient Explanation

BGVS displayed the final allocation grouped by game and showed the preference position associated with each assignment. This gave participants some evidence that rankings affected the result, but it did not explain why the solver selected one feasible allocation over another. As discussed in Section 2.2.3, transparency and perceived control matter most when they help users form useful expectations and see when intervention is possible. For this project, the relevant requirement is not to expose every solver detail, but to make the main trade-offs legible: which constraints mattered, whose high preferences were satisfied, and why some games were not selected.

Stage 1 showed that preference positions alone were not enough to make surprising results acceptable. Participants questioned why they were assigned away from a higher-ranked game, why a game could be assigned to one person when the social situation clearly called for a group, and whether the "can teach" input actually affected the allocation. G1-P7 described an assignment problem where "nobody of us knows the rules", and later argued that the system should warn the group when nobody in an assigned group can teach the game. G1-P2 also noted that users "can't inspect the votes", making it hard to tell whether a surprising result came from the algorithm, a mistaken preference, or a setup error. In Group 2, G2-P2 similarly described having "questions about the process" after seeing lower-preference outcomes. At the same time, participants recognised value in fast resolution and in reducing the influence of the loudest or most decisive person. This suggests a targeted explanation requirement: the system should preserve the speed and impartiality of computation while showing enough of the relevant constraints and trade-offs for participants to judge whether the result is socially workable.

3.5.4 Weak Support for Social Negotiation

In actual board game sessions, the group often discusses games before, during, and after formal preference input. The earlier system treated voting as a mostly linear process: submit rankings, wait, resolve. This leaves limited room for negotiation, shared understanding, or facilitated discussion. Since board game choice is a group activity, the system should support the social process around the computation, not only the computation itself.

The Stage 1 groups described their normal practice as open discussion, consensus seeking, show-of-hands checks, and people physically proposing boxes at the table. G1-P2 described the usual process as piling up boxes until someone strongly proposed a game, while G1-P7 noted that people who do not voice opinions strongly may not get what they want. This informal process has its own weaknesses, especially when quieter participants do not voice preferences or when a larger group may need to split across multiple games. However, the baseline system did not consistently improve that social process. In Group 2, G2-P2 said the group spent effort deciding how to enter the platform instead of "selecting the game". In Group 1, G1-P2 said the system became more useful when there were "12 people" and "20 games", because otherwise selection often came down to whoever was “loudest and most decisive”. The requirement implication is not to replace discussion with voting, but to support discussion with shared awareness, lightweight facilitation, and explanations that the group can use while negotiating the final decision.

3.6 Participant Capabilities, Scenarios, and User Stories

The requirements were organised around participant capabilities, not fixed social roles. In the proposed systems, everyone in the room is a participant, and any participant may also initiate setup, propose games, explain rules, or help advance the session.

Joining participant: A participant who wants to join quickly, understand enough about each option to express a preference, and avoid slowing the session down. This user may not know every game and may not want to learn a complex interface.

Contributing participant: A participant who adds candidate games, corrects game information, scans physical boxes, or helps move the room from setup to voting and resolution. This is an activity that any participant can perform, not a persistent authority role.

Experienced player able to teach: A participant with deeper game knowledge who can explain rules, identify whether a game fits the group, and recognise when an allocation is technically feasible but socially poor. This user needs the system to account for teachability and game suitability without dominating the group’s preferences.

These user groups lead to four core use scenarios.

1. Participants arrive at different times, so the group needs to identify which games remain playable as player numbers change.
2. Several participants know different subsets of the available games and need enough information to express preferences.
3. A larger group may need to split across two or more games while preserving fairness and avoiding undesirable assignments.
4. The group wants help reaching a decision without giving up the feeling that participants still shaped the outcome.

These scenarios are used as a bridge between the abstract research questions and the concrete system requirements. They were also converted into lightweight user stories so that the requirements could be written from the perspective of what participants need to do during a real session:

Join and understand: As a joining participant, I want to enter the room quickly and see enough game context so that I can submit preferences without slowing the group down.

Contribute candidates: As a contributing participant, I want to add or correct candidate games so that the room reflects what the group can actually play.

Represent flexible preferences: As a participant with weak or uncertain preferences, I want to express indifference or low preference without being forced into a misleading strict ranking.

Teach and interpret: As an experienced player, I want teaching ability and practical game fit to be visible so that the final allocation is playable, not only mathematically feasible.

Resolve together: As a room participant, I want progress, readiness, and result feedback to be visible so that the group can decide when to move forward and understand the outcome.

These stories deliberately describe activities, not permanent user roles. The same person may join, contribute a game, explain rules, and help resolve the room in the same session.

The early shared-table sketch in Figure 3.4 was used to keep this scenario framing concrete. It visualises the group as co-present around the physical game collection, with each participant contributing preference information on their own device. This artefact supported the decision to treat preference input, candidate-game awareness, and teaching knowledge as distributed participant capabilities, not as tasks owned by a separate controller.

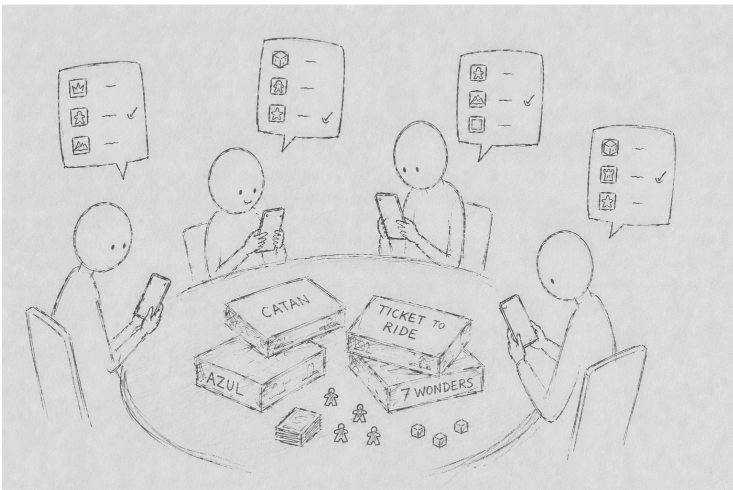
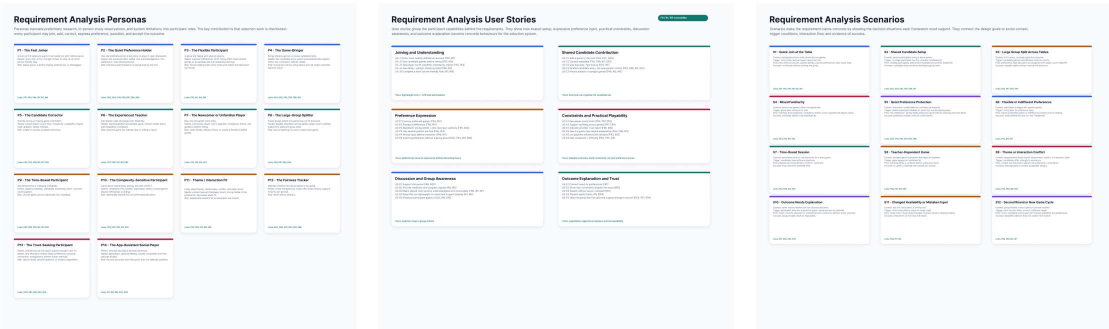


Figure 3.4: Early shared-table sketch showing participants expressing preferences around a physical game collection. The sketch was used as a scenario artefact for reasoning about co-located participation, distributed game knowledge, and shared candidate awareness.

The personas, user stories, and scenarios were maintained as design artefacts during the requirement analysis. Figure 3.5 shows the current artefact boards used to check that the requirements remained grounded in participant capabilities instead of a fixed control role. These boards are not presented as final qualitative themes. Instead, they document the intermediate synthesis used to move from evidence sources to design goals and requirements.



(a) Personas (b) User stories (c) Scenarios

Figure 3.5: Requirement analysis artefact boards used to organise personas, user stories, and scenarios. The boards summarise the intermediate design evidence that supports the participant-capability framing in this chapter.

3.7 Low-Fidelity Prototyping and Formative Design Critique

The questionnaire, baseline review, and participant artefacts identified broad requirement signals, but they did not by themselves show how those signals should appear in an interface. Low-fidelity prototyping and formative design critique were used as a design step between problem discovery and system implementation. This work made the requirements concrete enough to examine the sequence of actions, the language used in the interface, and the division of responsibility between direct participant interaction in PickNPlay and AI-mediated facilitation in Boardot.

3.7.1 Low-Fidelity Prototype and Wireframes

The first design artefacts were low-fidelity prototypes and wireframes, not fully implemented screens. These artefacts focused on the main decision flow: joining a room, building the candidate game pool, viewing game context, submitting preferences, monitoring group progress, and inspecting an outcome. For Boardot, the sketches also explored how an AI-mediated facilitator could introduce the process, summarise candidate games and constraints, and prompt the group without appearing to take ownership of the decision.

Figure 3.6 shows paired storyboards for PickNPlay and Boardot. The PickNPlay storyboard explored a direct participant-interaction model: a group creates a room, contributes candidate games, ranks or rates games on phones, sees live comparison, and receives group assignments. The Boardot storyboard explored a mediated model: participants gather around a central station, scan games and join codes, vote from phones, and ask Boardot to guide or resolve the session. The contrast helped clarify that both systems should preserve the same allocation task while distributing interaction burden differently.

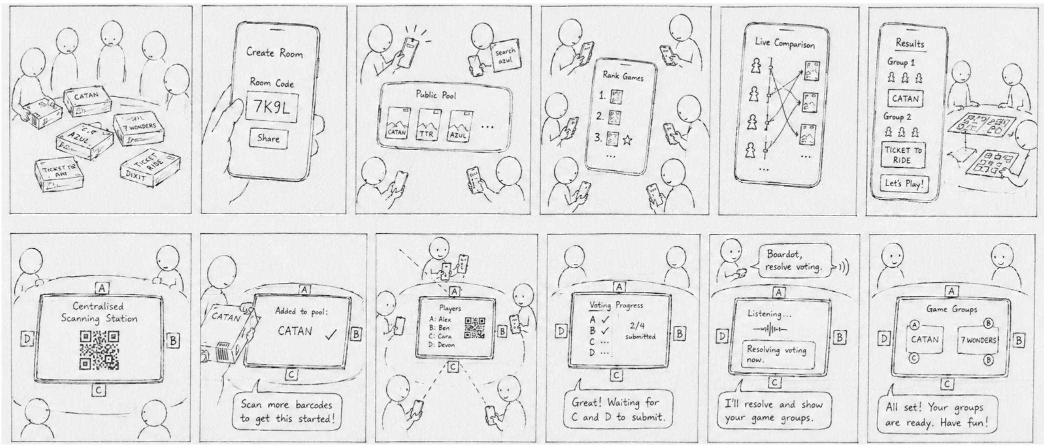


Figure 3.6: Low-fidelity storyboards comparing PickNPlay and Boardot interaction flows. The sketches were used to reason about direct participant interaction, AI-mediated facilitation, shared progress awareness, and the moment when a computed allocation becomes a socially actionable plan.

The mobile wireframe in Figure 3.7 was used to check interaction structure, not visual polish. Several design questions were examined at this level. These included whether joining could remain free from user-facing account registration and phone-friendly, whether candidate games and constraints were visible before voting, whether any participant could add or correct information, whether room progress was understandable without a fixed controller, and whether the result view could connect assignments back to preferences and constraints. Treating these questions as wireframe-level issues helped avoid committing too early to a specific layout while the requirement logic was still changing.

The main outcome of this stage was a clearer participant-capability framing. Early layouts that implicitly centred one setup person were revised toward shared room state and group-facing actions. Game information was moved earlier in the flow so that unfamiliar participants could vote with enough context. Preference input was kept lightweight because the low-fidelity screens made the cost of ranking many games on a phone more visible. The artefacts also made candidate capture a time-sensitive interaction problem. If adding physical games is slow, the system can delay play before preference input even begins. They provided formative support for the later requirements around shared candidate entry, rapid game capture, game context before voting, low-friction preference

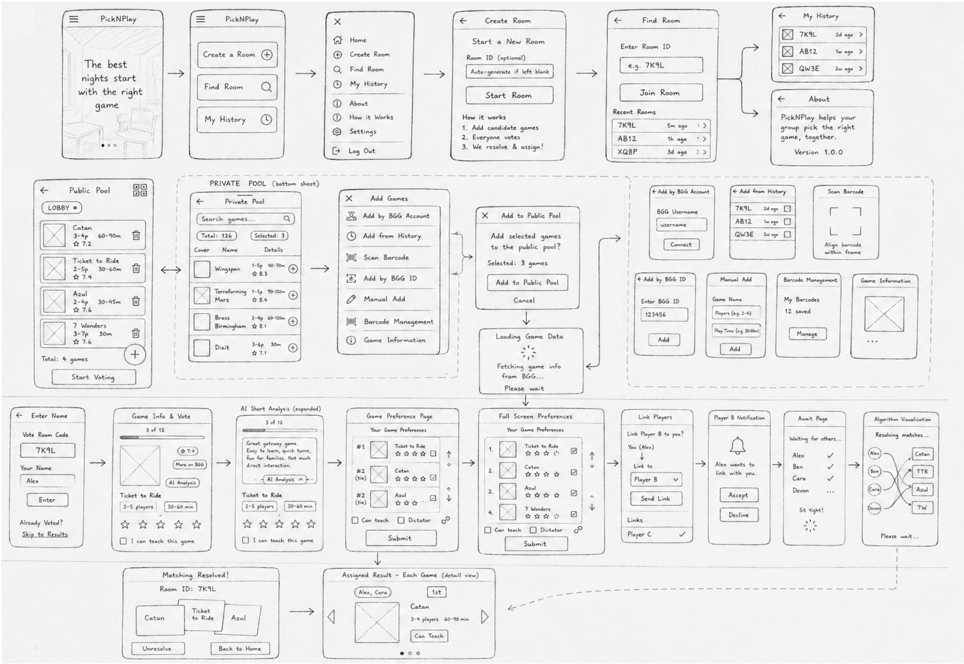


Figure 3.7: Low-fidelity PickNPlay mobile wireframe flow. The flow explored room creation and joining, shared candidate-pool construction, multiple game-entry paths, game information before voting, preference submission, progress awareness, player linking, algorithm visualisation, and resolved assignments.

input, progress awareness, and explanation of allocation results.

3.8 Design Goals

Six design goals organise the participant-facing requirements for PickNPlay and Boardot.

DG1: Distribute coordination burden. The system should reduce the amount of labour concentrated on one organiser, especially game entry, preference collection, and progress management.

DG2: Preserve participant agency. Participants should feel that their input matters and that the system supports group decision-making instead of replacing it.

DG3: Support expressive but lightweight preference input. Preference input should be fast on mobile devices while still allowing the system to represent low preference and indifference, and to surface unfamiliarity, hard unacceptability or veto, and practical constraints where these distinctions affect allocation or interpretation.

DG4: Make allocation trade-offs legible. The system should communicate enough about the result for participants to understand why the outcome is feasible, socially plausible, and open to revision when necessary.

DG5: Support shared room control. The system should make room-level state changes visible, recoverable, and where necessary confirmable, so that authority over the session is not silently concentrated in one participant.

DG6: Minimise setup and interaction time. The system should support quick joining, rapid candidate-game capture, and short mobile interactions so that the tool speeds up selection instead of becoming another source of delay.

These goals intentionally balance computational and experiential concerns. A system that optimises assignments but leaves participants confused would fail DG2 and DG4. A system that is socially pleasant but ignores feasibility constraints or forces one participant to manage the session would fail DG1 and DG3. A system that is theoretically expressive but slow to set up in a live session would fail DG6. The requirements below translate this balance into design commitments. Single-person setup and slow manual candidate entry motivate shared room state and lightweight setup support. Strict ranking and weak exclusion semantics motivate tiered preference input and visible preference-boundary signals, outcome displays without sufficient rationale motivate explanation requirements, and informal discussion around the tool motivates shared progress cues and bounded facilitation rather than full automation.

3.9 System Requirements

The participant-facing requirements are grouped into functional requirements, interaction requirements, and explanation, trust, and privacy requirements. Each group is traceable to the evidence sources and design goals above.

3.9.1 Functional Requirements

The functional requirements define the system capabilities needed for the participant workflow. They are primarily derived from the BGVS baseline, the questionnaire, and the participant scenarios.

FR1: Room-based participation. The system shall allow participants to create or join a shared game-selection session without user-facing account registration or a permanent privileged controller role.

FR2: Shared candidate game entry. The system shall support rapid addition of candidate games from a collection, search, manual entry, or barcode-based lookup where available, so that the game pool can be built by the group, not by one assigned person.

FR3: Game constraints and context. The system shall represent each game with editable player-count constraints and relevant contextual information such as playtime and complexity.

FR4: Preference collection. The system shall collect participant preferences in a structured form, such as ratings, tiers, or rankings, suitable for solver-backed allocation.

FR5a: Tiered preference states. The system shall support low preference and indifference through a lightweight preference-input model suitable for solver-backed allocation.

FR5b: Preference-boundary signals. The system shall make unfamiliarity and hard unacceptability or veto visible as interaction states or explanation signals, even where the current solver does not enforce them as separate hard constraints.

FR6: Participant teaching ability. The system shall support information about which participants are able to teach or explain a game, so that teaching coverage can be preferred, warned about, or explained where it affects practical playability.

FR7: Feasible group allocation. The system shall compute player-count-feasible assignments instead of only selecting a single most popular game.

FR8: Shared session state and revision. The system shall expose enough session state for participants to review candidate games, revise preferences or constraints, manage room phase changes, re-run resolution, or recover from mistaken input where feasible.

3.9.2 Interaction Requirements

The interaction requirements define how the system should be experienced during a board game session. These requirements are especially important because the setting is casual and time-sensitive: the tool should not become more disruptive than the decision problem it is meant to support.

IR1: Lightweight participant flow. The participant workflow, including room entry, candidate contribution, and preference submission, shall be short enough to complete comfortably on a phone during a social game session.

IR2: Group progress awareness. The system shall show clear session progress, including the visible candidate pool, who has joined, submission status without exposing preference content, unresolved readiness states, and when the group is ready to resolve.

IR3: Game context before voting. The interface shall provide game context before preference submission, especially for participants unfamiliar with the candidate games.

IR4: Low-friction preference input. The voting interaction shall reduce the burden of ordering many games, especially on small screens.

IR5: Shared room control. The system shall make critical room actions visible, understandable, recoverable, and where necessary confirmable by the group while avoiding a fixed permanent control role.

IR6: Support for discussion. The system shall provide shared artefacts, progress indicators, and result explanations that participants can use during discussion instead of treating preference submission as an isolated individual task.

IR7: Recoverable errors. The interface shall make errors recoverable, including wrong votes, accidental game selections, or changed player availability.

3.9.3 Explanation, Trust, and Privacy Requirements

The explanation, trust, and privacy requirements define what the system should communicate about the decision process. They are grounded in the need to make algorithmic outcomes legible without overwhelming users with implementation detail or unnecessarily exposing socially sensitive preference information.

ER1: Outcome-preference connection. The result view shall show selected games, assigned participants, relevant constraints, and enough preference connection for participants to understand why the allocation is plausible.

ER2: Constraint visibility. The system shall indicate major constraints that shaped the result, such as player-count limits or the absence of a participant able to teach a game.

ER3: Useful transparency. The system shall avoid exposing solver detail that does not help participants interpret or act on the result.

ER4: Trade-off awareness. The system shall support lightweight awareness of major trade-offs, such as selected games, unselected games, binding constraints, or preference compromises, where this helps the group understand the outcome.

ER5: AI role clarity. The AI-mediated framework shall make clear whether Boardot is facilitating discussion, summarising available information, explaining a computed allocation, proposing a next action, or offering a recommendation. Boardot shall not silently alter preferences, constraints, candidate games, or allocation logic, and shall not obscure the distinction between AI facilitation and the underlying allocation mechanism.

ER6: Preference-disclosure management. The system shall explain outcomes using enough preference and constraint information for the group to understand the result, while avoiding unnecessary exposure of individual preference content where this could create social pressure.

3.10 Research and Evaluation Constraints

The requirements above describe what participants need from the systems. A separate set of constraints came from the research design. These constraints shaped the implementation and study setup, but they are not treated as participant-facing requirements because they are motivated by the need to evaluate PickNPlay and to examine Boardot as an exploratory extension under comparable task conditions.

RC1: Comparable task conditions. PickNPlay and Boardot should preserve the same core decision task, candidate game pool, preference input, constraint representation, and allocation logic so that Boardot can be examined as a facilitation layer, not as a different underlying problem.

RC2: Study-compatible workflow. The systems should support controlled in-person and online study sessions with a consistent core workflow.

RC3: Evaluation evidence capture. The systems should allow interaction data, questionnaire responses, and participant comments about understanding, trust, fairness, and control to be collected for later usability and qualitative analysis.

3.11 From Requirements to System Overview

The participant-facing requirements and research constraints led to PickNPlay as the primary framework and Boardot as an exploratory facilitator layer over that framework. Both address the same underlying social choice problem, but they distribute agency and interaction burden differently.

3.11.1 PickNPlay: Direct Participant Interaction

PickNPlay is designed as a participant-directed web system. Participants join a room, inspect candidate games, express preferences, and view a computed result. The implementation treats the room as shared state, not as an object owned by any one participant. Room creation opens the session, but subsequent setup, voting, progress monitoring, and recovery actions are framed as group-facing controls. This framework addresses the requirements by giving users direct control over preference input and by making the allocation process visible through interface feedback and result visualisation. It is especially relevant to RQ1 and RQ2 because it tests whether interface design can make the social choice process understandable and acceptable while reducing organiser-centred control through shared room state and visible participant contribution.

PickNPlay emphasises explicit participant input, rapid candidate capture, shared session state, lightweight room control, feasible group allocation, and a visible connection between preferences, constraints, and outcomes.

3.11.2 Boardot: AI-Mediated Facilitation

The formative evidence directly motivated shared room control, lightweight preference input, and clearer result explanation. Boardot is not treated as an independently required solution, but as an exploratory design probe. It tests whether AI-mediated facilitation and a shared station can address the remaining coordination burden around the same PickNPlay workflow.

Boardot builds on top of PickNPlay instead of replacing it with a separate decision process. It keeps the same underlying room, candidate game pool, preference input, constraint representation, and allocation logic, but adds an AI-mediated facilitation layer around that workflow. This extension is intended to address problems that remain difficult for a direct interface alone, including prompting the group to move through setup and voting, summarising candidate games and constraints, supporting discussion when participants have uneven familiarity with the options, and reducing the coordination burden that might otherwise fall to whichever participant is informally leading at that moment.

In this sense, Boardot tries to solve a different interaction problem from PickNPlay. PickNPlay makes the social choice process visible and directly controllable. Boardot investigates whether an AI-mediated facilitator can make that same process easier to run in a live group setting. This does not remove the need for participant agency. Instead, Boardot should help the group coordinate, compare options, and interpret the result while leaving the final decision recognisably under group control.

Because an AI-mediated facilitator can influence the group by framing options or suggesting when to move forward, Boardot should remain bounded as a mediator, not a hidden decision-maker. It should not silently alter the candidate pool, participant preferences, constraints, or allocation logic. Its suggestions should remain contestable through ordinary group actions such as revising input, asking for explanation, or re-running the allocation.

Boardot is especially relevant to RQ3 because it probes whether bounded AI-mediated facilitation changes users' perceived control, trust, fairness, and social cohesion when compared with the same PickNPlay allocation workflow. Its requirements emphasise facilitation boundaries, explanation of AI roles, and support for social discussion. In particular, ER5 is critical because participants should be able to distinguish between Boardot as an AI-mediated facilitator and the allocation mechanism as the procedure that turns preferences and constraints into outcomes.

Chapter 4

System Design and Architecture

This chapter describes the design of PickNPlay, the proposed framework for participant-directed group board-game allocation, and Boardot, an exploratory station-mediated facilitator layer used to study shared display and bounded AI facilitation. The central design problem is to transform a shared candidate game pool, participant preferences, and practical constraints into a feasible assignment of players to games.

Section 4.1 first describes the shared architecture and room-state model that connect the implemented system to the requirements from Chapter 3. Section 4.2 then defines the allocation model and Z3-backed solver interface. Section 4.3 presents PickNPlay as the main participant-facing framework, focusing on public/private pools, rating-to-tier preference input, revision, and result explanation. Section 4.4 describes Boardot as an exploratory station-mediated facilitator. Finally, Section **??** summarises how these design decisions support the primary evaluation of PickNPlay and the exploratory role of Boardot.

4.1 Shared Architecture and Room State

The central architectural decision is that PickNPlay is the core room-based allocation framework, while Boardot is not a separate allocation system. Both use the same room state, preference-input pipeline, and Z3-backed resolver. PickNPlay makes the pipeline available through participant phones or browsers. Boardot adds a shared station with table-position markers, voice interaction, and generated facilitation text. This makes Boardot an exploratory change in facilitation and presentation, not a change to the underlying social-choice mechanism.

The room is the central state object. It stores the candidate games, participants, submitted preferences, practical constraints, voting progress, and final assignments. Participant clients and the Boardot station both read from and write to this room state, so changes to the candidate pool, voting phase, or resolved allocation become visible across the group. This builds on the Boardroom BGVS architecture (Verrell, 2025) from Chapter 2, which separated live room state from a stateless resolver while representing games, players, rankings, player-count quotas, teaching information, and assignments.

The implementation separates live coordination, persistent support data, and computation. Realtime Database stores the active room state, Firestore stores reusable participant and game-support data, Cloud Functions provide resolver and language-generation endpoints, and BoardGameGeek and Gemini supply contextual metadata and optional text. The allocation outcome is produced by applying the resolver to the room snapshot. The frontend is implemented with React (Meta Open Source, 2026) and Vite (Vite Contributors, 2026) and deployed through Netlify (Netlify, 2026). The data and function layer uses Firebase services (Google Firebase, 2026), and optional language support uses Gemini (Google AI for Developers, 2025).

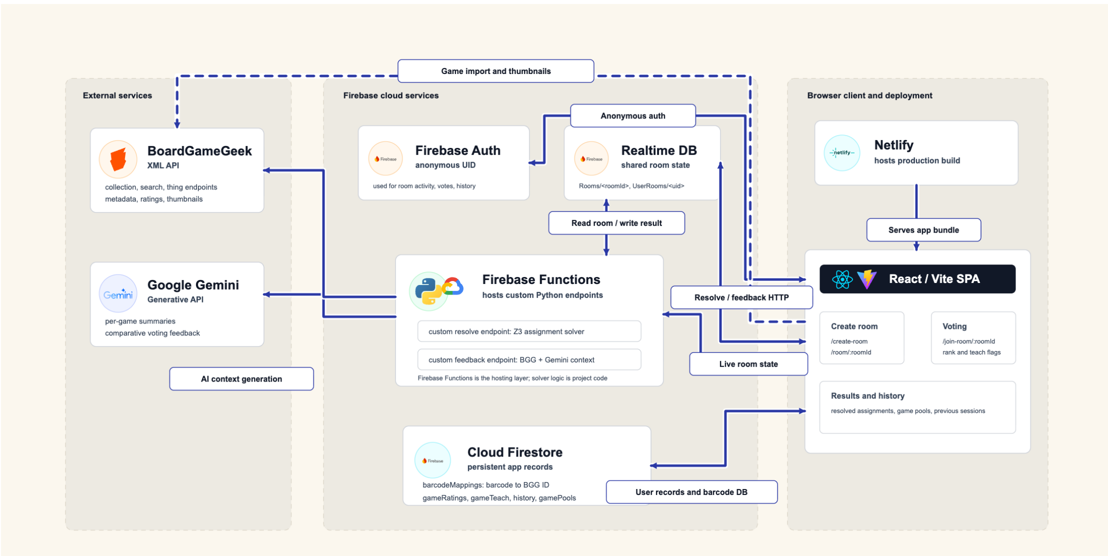


Figure 4.1: PickNPlay backend architecture and data flow. The diagram separates live room state, persistent support records, and computational services.

The privacy and disclosure boundary follows the same separation. The live room record contains the information needed for the current session, including public candidate games, participant presence, submitted room votes, practical constraints, progress state, and assignments. Reusable ratings, saved pools, and teach flags are stored as per-user support data, not as a public preference profile. When these records prefill a later vote, the participant still submits an active room-specific vote before that data becomes part of the shared allocation input. Generated game-context feedback receives candidate game identifiers and names, player count, and don't-care count. Boardot dialogue receives bounded visible room context such as phase, event, counts, game names, selected game names, and an optional spoken question. These generated-text paths are not given full individual rankings as a public profile and cannot modify votes or choose assignments.

The room moves through three visible phases: lobby, voting, and resolved. In the lobby, participants join the room and construct the public candidate pool. In the voting phase, participants submit preferences and practical constraints such as teaching ability, linked-player requests, and top-tier must-include choices. In the resolved phase, the system displays selected games and player assignments. These phases are deliberately simple so that the state of the group decision remains legible during a social play session.

Because several clients can act on the same room, room-level changes are phase-gated and guarded by lightweight action locks. Candidate-pool edits occur in the lobby, preference submissions are collected in the voting phase, and assignments are written after the frontend fetches a fresh room snapshot and the resolver processes that snapshot. Actions such as starting voting and resolving set an action-in-progress flag so other clients can avoid issuing the same transition simultaneously. This is a pragmatic consistency model, not a full collaborative editing protocol. The current implementation does not attach version identifiers to resolver payloads or prove that every stale-client race is impossible. Table 4.1 summarises how the main design decisions trace back to the requirements from Chapter 3.

Table 4.1: Traceability from design decisions to requirements.

Design decision	Requirements	Thesis rationale
Shared room state	FR1, FR8, IR2, IR5, DG1, DG5	Makes the decision process visible and editable by the group instead of private to one organiser.
Public/private game pools	FR2, FR3, DG6, IR3	Supports exploration before changing the shared candidate set.
Rating-to-tier input	FR4, FR5a, IR4, DG3	Reduces ranking burden while preserving weak preference orders and ties.
Solver-backed allocation	FR3, FR7, ER2	Enforces player-count feasibility and assignment constraints for a multi-game outcome.
Result explanation and recovery	FR8, IR7, ER1, ER2, ER3, ER4, ER6	Connects outcomes to preferences and constraints while keeping revision and reset actions visible.
Boardot station facilitation	RC1, IR6, ER5, DG2, RQ3	Explores an AI-mediated facilitator layer without changing the allocation mechanism.

This shared-state architecture has an HCI purpose. It turns candidate games, voting progress, and allocation outcomes into visible group artefacts instead of private state on one device. The system still uses computation to resolve a constrained allocation, but the surrounding architecture keeps the process socially inspectable. Candidate games can be discussed before voting, progress can be checked during voting, and results can be interpreted after resolution.

4.2 Allocation Model and Solver Interface

4.2.1 Resolver Contract

The allocation layer is implemented as a Python Cloud Function backed by Z3. The frontend does not run the optimisation model locally. Instead, it sends a compact room snapshot containing candidate games, player-count quotas, submitted preference structures, practical constraints, and any don't-care players. The resolver returns assignment pairs or a failure response, which the frontend writes back into the shared room. Table 4.2 defines which user-facing inputs enter the solver as hard constraints, soft signals, preference representations, or current implementation boundaries.

Table 4.2: Solver treatment of user-facing inputs.

User-facing input	Solver treatment	Status	Explanation implication
Minimum and maximum player count	Enforced through game quota feasibility.	Hard constraint	A game can run only with a valid assigned player count.
One assignment per player	Each player has exactly one selected assignment variable.	Hard constraint	Every participant considered by the resolver receives one game assignment.
Preference tiers	Converted into ranks and scores.	Optimisation objective	Higher tiers are preferred but do not guarantee assignment.
Equal ratings and chain links	Represented as same-rank tied tiers.	Preference representation	Participants can express indifference without inventing a false order.
Teaching ability	Prefers selected games to include a declared participant able to teach where available.	Soft practical playability signal	Teaching improves practical playability but is not required for feasibility.
Mutual linked-player request	Enforced as co-assignment only when both users request the link.	Hard constraint	Mutual social constraints are respected.
One-way linked-player request	Ignored by the resolver.	Not enforced	Prevents unilateral grouping pressure.
Top-tier must-include	If the game is selected, the requester must be assigned to it.	Hard conditional constraint	A strongly advocated game is not selected without the requester.
Don't-care player	Added with equal score for every game.	Feasibility participant	Helps fill feasible table sizes without preference pressure.
Hard veto or unfamiliarity	Not represented as a separate submitted Z3 state.	Current boundary	Handled through low ranking, revision, or future work, not as an enforced veto.

This contract follows the architectural pattern described in Chapter 2: the database coordinates the room, while the resolver solves one consistent snapshot (Verrell, 2025). The contract also makes the algorithmic boundary explicit. The solver sees game quotas, submitted preference orderings, tier information, teaching data, linked-player data, top-tier must-include flags, and don't-care player count. It does not see the visual layout of the interface, the station marker positions, the Boardot dialogue, or the generated game-context feedback. Those elements can affect how participants understand and submit input, but they do not enter the optimisation model.

4.2.2 Decision Variables and Hard Constraints

Let P be the set of players considered by the resolver, including participants who submit preferences and any don't-care players, and let G be the candidate games in the room. The solver creates a Boolean variable $a_{p,g}$ for each player $p \in P$ and game $g \in G$. The variable is true exactly when player p is assigned to game g. In the arithmetic constraints below, Boolean assignment variables are treated as 0/1 indicators.

The first hard constraint is that every player must be assigned to exactly one game:

$$\sum_{g \in G} a_{p,g} = 1 \quad \forall p \in P.$$

The second hard constraint is quota feasibility. For a game g, let $N_g = \sum_{p \in P} a_{p,g}$, q_g^{min} be the minimum player count, and q_g^{max} be the maximum player count. The solver enforces:

$$N_g = 0 \vee q_g^{min} \leq N_g \leq q_g^{max} \quad \forall g \in G.$$

This means a game is either not selected at all, or it receives enough players to be playable without exceeding its maximum player count. This is the core difference between PickNPlay and a single-winner vote. The system is not asking which game is most popular overall. It is deciding which subset of games can run and which participants should play each selected game.

The quota constraint also explains why the game metadata matters architecturally. Minimum and maximum player counts are not only displayed as context. They become solver inputs. A candidate game with a maximum of four players can be selected for at most four assigned participants. A game with a minimum of three players cannot be selected for a pair unless the constraints are changed before resolution.

4.2.3 Preference Objective

The resolver receives preference information as an ordered structure that can include tied tiers. Equal-tier games share the same rank, and the backend normalises this structure before scoring. Ties use competition ranking: if two games are tied in the first tier, both receive the best rank and the next tier receives the third rank instead of the second. For example, if a participant gives three games the same top rating, all three receive rank 0, and the next tier receives rank 3. This preserves the participant's statement that tied games are equally preferred while still separating lower tiers from a broad top tier.

For a participant who submits preferences, let $r_{p,g} = 0$ denote player p's most preferred tier for game g, and let m be the number of games. The backend assigns a score to each game using:

$$s_{p,g} = m - r_{p,g} - 2^{r_{p,g}}.$$

The optimiser maximises:

$$\max \sum_{p \in P} \sum_{g \in G} s_{p,g} a_{p,g}.$$

This scoring model strongly favours satisfying high-ranked choices and penalises lower-ranked assignments increasingly sharply. Because the exponential term can make lower-ranked games receive negative scores, the objective both rewards high-ranked assignments and penalises low-ranked assignments. A negative score does not make an assignment infeasible. It only makes that assignment less attractive than other feasible alternatives. The model is a pragmatic optimisation objective, not a complete social-choice fairness theory. It should not be read as a proof of envy-freeness, proportionality, strategy-proofness, or any other formal fairness guarantee. This caveat is important because SMT (Barrett et al., 2021), Z3 (de Moura and Bjørner, 2008), νZ (Bjørner et al., 2015), and related constraint-programming tools (Rossi et al., 2006) are flexible modelling frameworks, not automatic guarantees of social legitimacy.

Don't-care players are treated differently. They are added to the player set so that the resolver can produce groups of feasible size, but they receive the same score for every game. They influence feasibility and table size without expressing preference pressure. A don't-care player represents a participant included in the current headcount who has explicitly agreed to be assigned wherever needed, or who does not wish to submit preferences. It should not be used to silently include absent or unconsulted participants. For this reason, don't-care entries are shown separately from submitted voters in the waiting and result views.

4.2.4 Teaching, Linking, and Must-Include Constraints

Teaching ability is implemented as a soft practical playability signal. If at least one participant says they can teach a game, the optimiser prefers assignments where that game is either not selected or has at least one participant able to teach assigned to it. This is not a hard feasibility constraint. The choice is deliberate: in small groups, requiring a participant able to teach for every selected game can make otherwise reasonable allocations impossible. In the implementation, teaching coverage is represented as a Z3 soft constraint with weight 1, added after the main preference objective. It can guide optimisation among feasible solutions, but violating it never makes an assignment infeasible.

The linked-player mechanism is implemented conservatively. The frontend lets a participant request to link with another participant, but the backend only enforces links that are bidirectional. If player A links player B but player B does not link player A, the one-way request is ignored by the solver. If both participants link each other, the solver enforces co-assignment by requiring linked participants to be assigned to the same game. This protects participants from being unilaterally forced into another person's group while still supporting mutual social constraints.

The top-tier must-include control is best understood as a practical constraint, not a theoretical voting-rule claim. If a participant enables this control for a top-tier game, then that game may only be selected if the participant is assigned to it. Otherwise, the game can be left unselected. This addresses a common board-game situation where a participant may bring or strongly advocate for a game, but only wants it chosen if they are included in the group that plays it.

These additional constraints show why the resolver is separated from the interface. The interface can present teaching, linking, and must-include choices in socially meaningful terms, while the backend converts them into constraint and optimisation terms. The boundary is still explicit: the solver can enforce or prefer only the fields included in the resolver payload.

4.2.5 Known Algorithmic Boundaries

The current implementation partially satisfies the flexible-preference requirements from Chapter 3. It supports low-priority placement through lower-rated or lower-ranked tiers and indifference through tied tiers. However, hard veto and unfamiliarity are not currently represented as separate submitted fields in the Z3 model. Where participants need to avoid an outcome, the current workflow relies on lower rankings, discussion and revision, or removing unsuitable candidate games before resolution instead of a separate hard-veto state. These remain known implementation boundaries.

4.2.6 Failure Handling and Revision

When the resolver cannot find a feasible allocation, it returns a failure response instead of a partial assignment. The interface treats this as a revision point, not as a final outcome. Participants can return to the lobby, adjust candidate games or player-count constraints, revise linked-player or top-tier must-include choices, add don't-care players where appropriate, or re-run resolution after changing the room state. The current system does not compute a formal minimal unsatisfiable core, so failure explanations are practical revision prompts, not proofs of infeasibility.

4.3 PickNPlay Interaction Design

PickNPlay is the direct participant-interaction framework. It exposes candidate-pool construction, preference input, revision, resolution, and result interpretation through participant devices, treating allocation as a shared decision supported by a solver.

4.3.1 Shared Room Control

The PickNPlay flow begins with room creation or room joining. A room can be created with a custom or generated code, and participants can join using the room ID or QR code. The creator is technically the first person to create the room, but the interface avoids making that person a permanent controller. Once the room exists, connected participants see candidate games, voting progress, and results through views derived from the same room state.

Shared control is implemented through visible phase changes and recoverable actions. Starting voting asks participants to confirm that the public pool is ready. Resolving the room is shown as a group action, not a background event, and resetting votes, returning to the lobby, removing mistaken submissions, and unresolving results remain available as explicit controls. The design problem here is agency. Participants should be able to see when the group decision is changing state and should have a path to correct mistakes before treating the allocation as final.

4.3.2 Candidate-Pool Construction

The lobby separates private exploration from public commitment. The private pool acts as a participant's staging area where users can import games, add games individually, scan a barcode, reuse saved pools, or manually create a game record. The public pool is the room's shared candidate list. Moving games from private to public is a deliberate action, usually followed by confirmation of player-count details. This design matches the social practice of shortlisting. Participants often need to inspect possible games before agreeing that they belong in the shared voting set.

The public pool is also where feasibility begins. Each public game carries minimum and maximum player counts, and participants can correct or remove unsuitable games before voting. This matters because the same player-count information later becomes the resolver's lower and upper quota input. BoardGameGeek metadata and optional generated game-context feedback are used as context for unfamiliar games, not as recommendations that determine the outcome. This generated text supports game understanding. It does not recommend how a participant should vote and does not enter the resolver. The participant still forms and submits their own preference order, and the solver uses that explicit input.

4.3.3 Rating-to-Tier Preference Input

The voting interaction uses a two-step preference process. First, participants rate each game from one to five stars. This is faster on a phone than manually ranking many game cards from the beginning. Second, the system converts ratings into a preference order: higher ratings come earlier, and games with the same rating become tied preference tiers. The participant can then fine-tune the generated order through drag-and-drop, arrow controls, a full-screen list, and chain buttons that explicitly tie adjacent games.

This design addresses the tension between expressive input and interaction burden. The resolver needs structured preference information, but the user should not have to perform a long and fragile ranking task on a small screen. Rating first and refining later gives participants a quick path while still producing the tiered preference structure used by the backend. It also gives indifference a concrete interaction. If two games are equally acceptable, the participant can leave them tied instead of inventing a false preference.

The same voting flow collects the practical constraints described above, including teaching ability, mutual linked-player requests, and top-tier must-include choices. These inputs are deliberately framed as social constraints, not as raw optimisation parameters. In a station-linked room, participants also select a table marker before submission so the station can connect voting progress to the physical seating arrangement.

Preference data is also reused carefully. The app can prefill ratings and teach flags from Firestore per-user caches keyed by game ID, reducing repeated input across rooms. However, the actual room vote is still submitted into the live room when the participant votes. This distinction avoids confusing historical convenience data with active consent in the current session.

The same support-data layer also stores per-user session history and previous-played pools after resolution, allowing participants to review past sessions or reuse earlier candidate sets without making them part of a new room until they explicitly import them.

4.3.4 Waiting, Revision, and Result Explanation

After submitting preferences, participants reach a waiting view. This view shows the submitted roster, pending participants, candidate games, don't-care count, and resolve controls. The group can edit votes, remove mistaken submissions, add don't-care players, reset votes, return to the lobby, or resolve the room. The interface can warn when the current participant count, don't-care count, or candidate-game quotas make resolution unlikely to succeed, but it keeps the action visible to the group instead of hiding the decision in a backend process.

During resolution, PickNPlay can show an algorithm visualisation before presenting a static result. The visualisation provides a simplified representation of games, participants, preference links, linked-player relationships, must-include state, don't-care player nodes, and assignment movement. It is not a low-level Z3 trace. Its role is explanatory: participants see that the system is constructing feasible player-game groups from preferences and constraints without exposing unnecessary solver internals (Kizilcec, 2016; Abdul et al., 2018). The visualisation is explanatory, not diagnostic: it does not prove optimality to the user, show the complete solver search trace, or compare all feasible alternatives.

The final results view groups assigned players under selected games. For each participant who submitted preferences, it can show their preference rank for the assigned game, whether the assignment came from a tied tier, whether they can teach the game, and whether the row corresponds to the current viewer. Don't-care players are labelled as don't-care entries. In station-linked rooms, the phone result can also help the participant identify themselves through their table marker identity. The result design connects outcomes to preferences and constraints without displaying every participant's full ranking as a public table.

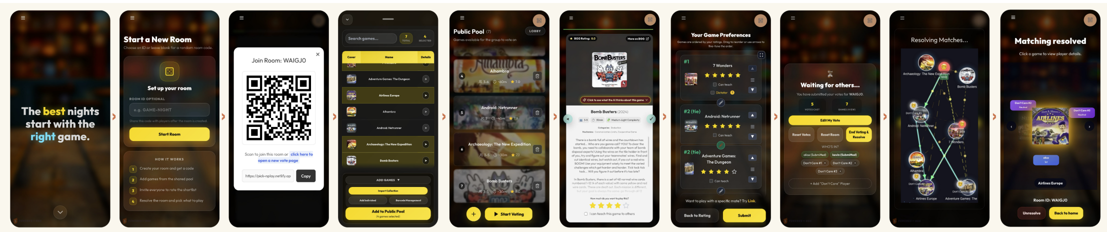


Figure 4.2: Current PickNPlay interface flow. The composite screenshot shows the direct participant journey from room creation and QR joining, through shared candidate-pool construction, game-context inspection, tiered voting, waiting and resolution, to the final matched assignments.

The interaction flow is also available as an interactive wireframe for readers who want to inspect the screen transitions directly: https://kevinzzzhu.github.io/projects/picknplay-wireframe/. This supplementary wireframe is design documentation, not evaluation evidence.

4.4 Boardot Station and AI Facilitation

Boardot presents the PickNPlay workflow through a shared station for co-located groups. It is an AI-mediated facilitator, not an authority-bearing decision maker or recommender. It is implemented as a station route that first creates a station identifier and displays a QR code for room creation. Once a phone links a room to the station, the station subscribes to the live room state and changes its display according to the room phase. Preference entry remains on personal devices, while the station provides a public surface for room entry, candidate games, voting progress, table-position markers, prompts, and final assignment clusters.

Figure 4.3 shows the physical prototype used to instantiate the station concept. Its purpose is not to introduce a new allocation mechanism, but to make the shared decision process publicly visible and addressable within the play space.

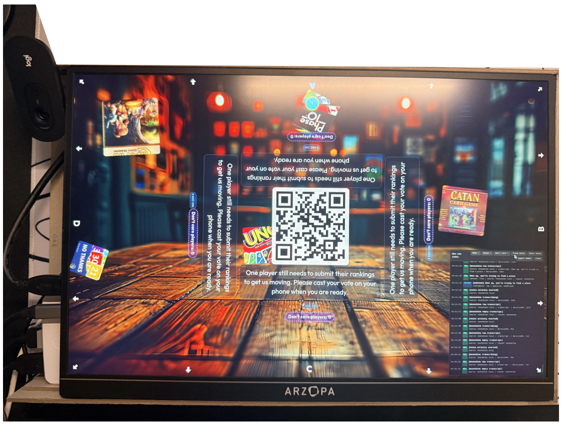


Figure 4.3: Boardot station prototype built with cardboard. The screen, camera, speaker, and microphone are combined in one table-facing station setup for shared display and voice/audio facilitation.

4.4.1 Shared Station and Co-located Awareness

The station follows the same lobby, voting, and resolved phases described earlier, but filters the shared room state for a public table display. In the lobby, the central QR code is the main action target, floating game thumbnails show the emerging public pool, and scanner/microphone status pills show whether physical game capture and voice input are available. During voting, the QR code remains available for late joiners, while the station shows voting progress and the current number of don't-care players. After resolution, the station switches to a public result display instead of asking participants to inspect only their phones.

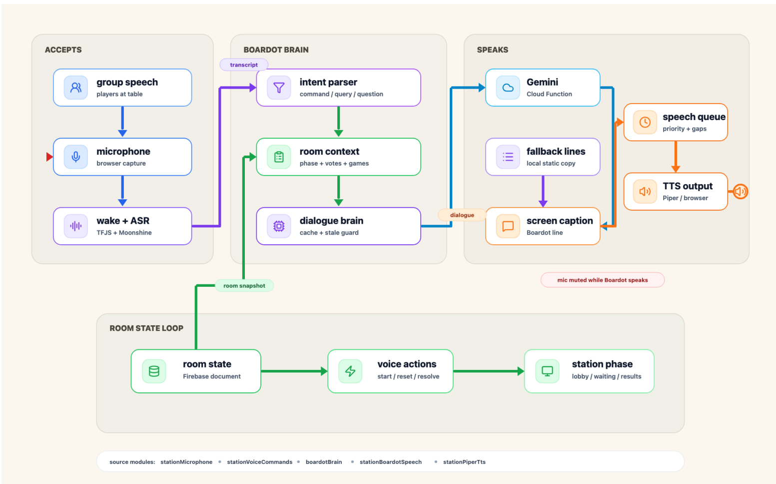
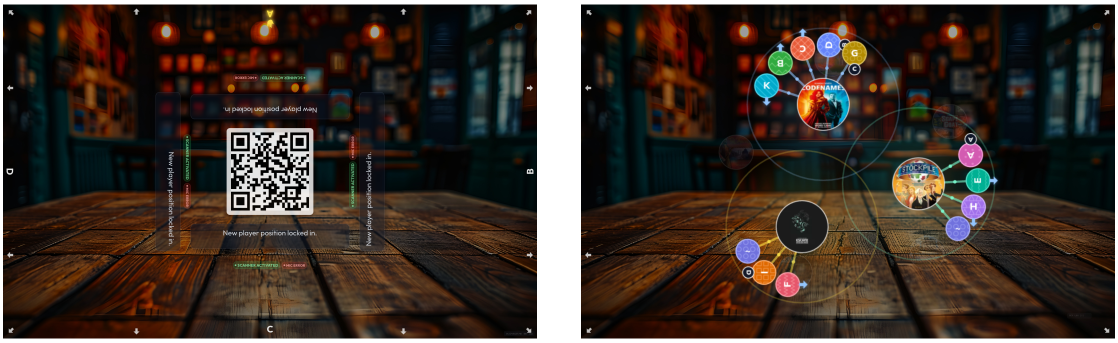


Figure 4.4: Boardot station architecture and data flow.

Figure 4.5 shows two current station states captured at 2560 × 1600 resolution. The lobby state illustrates the around-the-table QR, marker, scanner, and concise-prompt layout. The result state illustrates the station-specific “Battle Power” visualisation, in which selected games and assigned participants are arranged as public clusters.



(a) Lobby and joining state.

(b) Result visualisation state.

Figure 4.5: Boardot station interface states. The screenshots are implementation evidence for the station layout and public result presentation, not evaluation results.

The station is designed for an around-the-table viewing situation, not a single front-facing viewer. The central area is kept for the current phase, concise prompts, and shared actions, while persistent information is expressed through glanceable visual cues such as QR placement, game thumbnails, edge markers, vote-status indicators, and result clusters. This decision follows the cognitive-load framing (Sweller, 1988) in Chapter 2: because working memory is limited during problem solving, the shared display should avoid unnecessary reading burden while retaining feedback, contestability, and control (Abdul et al., 2018).

Boardot uses table-edge markers to connect the shared display with participants’ physical positions. Participants select a marker from their phone, and the station uses markers to show joining and voting progress spatially. The identity-token design combines colour with shape and pattern so participants can recognise themselves from different viewing angles without relying only on small or rotated text. Using shape and pattern as well as colour also reduces reliance on colour alone, which is important for accessibility and for viewing the station from different angles. This gives participants visible feedback about who has joined, who is pending, and where final assignments appear around the station (Kizilcec, 2016; Abdul et al., 2018). The markers do not affect the resolver. They help participants relate the public display to their own submitted input and final assignment.

The result display uses the station variant of the algorithm visualisation instead of the phone result table. Its “Battle Power” framing is intentionally presentational: it stages selected games one at a time and then locks in the assigned table clusters. The label is a playful visual metaphor and should not be interpreted as a participant score, game quality score, or fairness metric. The visualisation helps a group gathered around the station see that multiple games may be selected at once, but it is not a second allocation model and does not change the assignments already returned by the resolver.

4.4.2 Barcode as Setup Support

The broader PickNPlay workflow already supports barcode-assisted candidate entry. In Boardot, the station centralises this shortcut for co-located setup: the camera runs a background scanner, a recognised barcode is mapped to a BoardGameGeek identifier, and the resulting game is added to the public candidate pool. Duplicate or unmapped scans are treated as feedback events, not allocation events. When a barcode cannot be mapped, participants can still use the ordinary PickNPlay search or manual-entry paths. Participants can also remove unsuitable entries before voting, so scanning speeds up setup without bypassing group control.

Unmapped barcodes are not added automatically; they must first be registered through a barcode-mapping workflow that links the scanned product code to a BoardGameGeek game identifier.

4.4.3 Bounded AI Facilitation

Boardot's voice and dialogue layer is designed to facilitate group progress without taking allocation authority away from the shared room workflow. The station parses short spoken commands into phase-bounded intents, such as starting voting in the lobby, resolving during voting, resetting votes, or adding a don't-care player. It can also answer simple questions about what to do next, how voting works, how markers work, and how many people have voted. Phase-changing or destructive actions remain explicit and confirmable, and unavailable actions are rejected in the current phase instead of silently changing the room.

Generated dialogue does not alter candidate games, preferences, constraints, room status, or assignments. Figure 4.4 shows this separation between dialogue generation and room-state authority. The station may display or speak Gemini-generated text, but allocation remains the responsibility of the Z3-backed resolver described in Section 4.2. Boardot dialogue requests use bounded visible room context, such as phase, event, counts, game names, selected game names, and an optional spoken question. They do not send full individual rankings for Gemini to infer hidden preferences or choose assignments. The active implementation also keeps speech recognition for commands on the station side, using generated dialogue for short contextual prompts and answers instead of hidden preference inference. Participants retain visible controls for revision, reset, resolution, and result inspection. This separation responds to the automation concern discussed in Chapter 2. Systems can reduce effort while also reducing agency if users cannot tell what automation has decided or whether they can contest it (Horvitz, 1999; Heer, 2019). Boardot is a facilitator for the shared decision process, not an autonomous recommender or hidden allocation authority.

Chapter 5

User Study and Evaluation

This chapter presents the empirical evaluation of PickNPlay as the proposed participant-directed framework for group board-game allocation, and Boardot as an exploratory station-mediated extension of the same workflow. The evaluation connects the implementation described in Chapter 4 with participant evidence about usability, understanding of the allocation process, trust, perceived fairness, control, and the social experience of reaching a group decision. The study was designed as an HCI evaluation of interaction and experience, not as a test of a new allocation algorithm.

The chapter is organised as follows. Section 5.1 describes the study setup, ethics, recruitment, procedure, observation, questionnaires, and interviews. Section 5.2 reports the quantitative questionnaire analysis, beginning with descriptive SUS statistics before any inferential comparison. Section 5.3 presents the thematic-analysis approach and qualitative findings. Earlier formative material from Chapter 3, including Stage 1 baseline evidence and the early-stage questionnaire, is used as design and motivation context, not re-analysed as implemented-system evaluation evidence. Chapter 6 then integrates the quantitative and qualitative results against the research questions.

5.1 User Study

The user study examined the implemented systems after participants had used them in board-game selection sessions. As described in Chapter 4, Boardot shared PickNPlay's room state, preference-input pipeline, and Z3-backed resolver, while changing how the group interacted with that workflow. The study evaluated PickNPlay as the main participant-directed framework and examined whether station-mediated AI facilitation changed participants' experience of the same underlying allocation task.

The study used the HCI evaluation methods introduced in Section 2.2.1: observation during system use, post-use SUS responses, and semi-structured interviews. Because the study used small, situated groups instead of a large representative sample, the results were treated as exploratory HCI evidence. Across the implemented-system evaluation sessions, the immediate post-use SUS dataset contained 42 session-level participant responses: 37 from in-person sessions and 5 from the online PickNPlay session. Some participants had prior exposure to earlier project stages, so these counts are treated as session-level evaluation responses, not as a claim about unique independent participants. This scale was consistent with small-sample HCI evaluation practice, where sample size is shaped by method, study setting, and research purpose instead of by one universal threshold (Caine, 2016).

5.1.1 Configuration and Setup

The evaluation sessions were configured around a realistic group board-game selection task. Participants used their own phones or browsers to join a shared room, inspect the candidate game pool, submit preferences, and view the computed allocation. This setup preserved the participant-facing structure described in Chapter 4: the room acted as the shared state object, preference entry remained individual, and the final allocation was produced by the same resolver across the systems.

The in-person configuration used the systems in a co-located group setting. For PickNPlay, participants interacted directly with the web interface on their own devices while the group moved through room setup, voting, and result inspection. For Boardot, the same room workflow was linked to a shared station. The station displayed room entry information, candidate games, voting progress, participant marker positions, and Boardot facilitation prompts, while participants still entered preferences through their own devices. This configuration allowed the study to examine Boardot as a station-mediated extension of PickNPlay without changing the underlying allocation task.

The online configuration used the web-based PickNPlay workflow remotely with participants joining the shared room through their own devices. This preserved the core task of constructing a candidate pool, submitting preferences, and inspecting an allocation, but removed the physical station and co-located table context. Online sessions were treated as evidence about the web interface and remote group use, not as direct evidence about the embodied Boardot station setup.

5.1.2 Ethics Consideration

The study was conducted under ANU Human Research Ethics Committee approval, Protocol H/2025/0501. Participants received the Participant Information Sheet and consent form before taking part, and consent was confirmed before the study activity began. Participation was voluntary: participants could decline to take part, refuse to answer interview or survey questions, stop an interview, and withdraw within the approved withdrawal period. If a participant withdrew, their data was to be destroyed by default and excluded from analysis.

The main ethical concerns were privacy, consent for recording, and the risk of re-identification in small groups. The study collected questionnaire responses, observation notes, app interaction data, and interview or session recordings where consent had been given. These materials were handled as research data, not as public artefacts. Participant references are de-identified using labels such as P1 or group-level labels, and quotations are only used where they can be presented without unnecessary identifying detail.

Data management followed the approved protocol. Research data was stored on local storage with access restricted to the immediate research team, and analysis copies were kept on password-protected computers. Audio material was transcribed locally on the researcher's laptop using Apple's built in transcription functionality, which was selected because the transcription process did not transmit audio to external servers. These constraints were especially important because the study involved small gaming groups, where specific comments, interaction patterns, or contextual details could otherwise make participants easier to recognise.

5.1.3 Participant Recruitment

Participants were recruited through a purposive and convenience-based strategy suited to an exploratory HCI study. The approved recruitment criteria required participants to be at least 18 years old, have experience playing board games, and be comfortable using mobile or web applications. Recruitment was attempted through campus posters, online board-game forums and posts, direct invitations to existing contacts, and the ANU board game club. Interested participants received the Participant Information Sheet and consent form before participation, and eligibility and consent were checked before the study activity began.

The resulting groups were not intended to be demographically representative of all board-game players. Instead, they were selected because they could plausibly evaluate a group board-game selection system in use. Some groups participated across multiple stages of the project. One technically experienced ANU-affiliated group had prior familiarity with BGVS and continued through later PickNPlay and Boardot sessions. Another group was recruited through a direct personal approach and also participated across stages. Additional later-stage groups were recruited through online contact and through the ANU board game club. This continuity was useful for comparing how participants responded to successive system versions, but it also meant that some later-stage comments reflected prior exposure instead of first-use impressions only.

Data coverage differed by instrument. All participants in the relevant sessions completed the SUS questionnaire, but not every participant in every group took part in interviews. The quantitative usability summaries describe the full questionnaire respondent set for each session, while interview-based claims are treated as evidence from the interviewed subset.

(a) Poster placement near the board-game shelves.

(b) Poster placement on a campus noticeboard.

(c) Closer noticeboard view of the study poster.

Figure 5.1: Examples of campus recruitment poster placement.

5.1.4 Study Process

Each implemented-system evaluation session used the same four-stage structure. The main procedural difference was whether the selection task was conducted through the participant-directed PickNPlay web workflow or through the station-mediated Boardot workflow.

1. **Introduction and consent**: Participants were briefed on the study activity, confirmed eligibility and consent, and were reminded that they could decline questions or withdraw under the approved ethics procedure.

2. **Tutorial and familiarisation (about 5 min)**: Participants were introduced to the workflow used in that session. PickNPlay sessions covered room joining, candidate-game inspection, preference and constraint entry, and result interpretation. Boardot sessions additionally introduced the shared station, table markers, voting-progress display, and facilitation prompts.

3. **Group board-game selection task (10–15 min)**: Participants used their own phones or browsers to complete the selection task with their group. In PickNPlay sessions, this centred on the web room workflow: constructing or reviewing the candidate pool, submitting preferences, and inspecting the computed allocation. In Boardot sessions, preference entry still occurred on personal devices, while the shared station displayed room entry information, candidate games, voting progress, marker positions, facilitation prompts, and final assignment clusters.

4. **Post-session evaluation (15–20 min)**: Participants completed the SUS questionnaire and, where they agreed to do so, participated in a semi-structured inter-

view about usability, understanding, trust, fairness, and the effect of the system on group decision-making. Boardot interviews additionally asked about the central AI host, including whether participants followed, ignored, or debated its guidance.

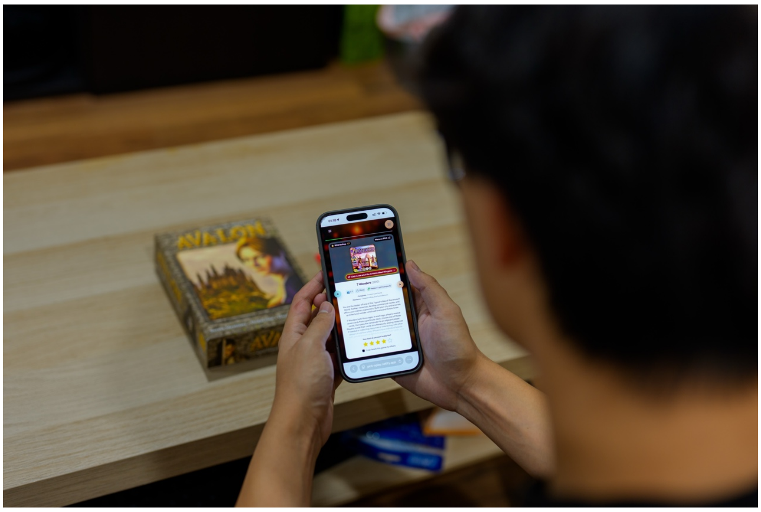


Figure 5.2: PickNPlay evaluation session with the mobile interface.

Observation Study

Observation was used to capture how the systems were used during the live selection task. The focus was on the interaction process, not on measuring a controlled behavioural outcome. During each session, the researcher attended to how participants joined the room, understood the candidate game pool, entered preferences and constraints, coordinated with one another, interpreted progress indicators, and responded to the final allocation. Usability obstacles, confusion points, recovery actions, and moments where the group discussed or challenged the system output were treated as relevant observation signals.

For Boardot sessions, observation also covered the station-mediated parts of the workflow: how participants used the shared display, whether table markers helped connect physical positions with digital state, how facilitation prompts entered the discussion, and whether participants treated Boardot as guidance, not as the allocation authority. For online sessions, the observation scope was narrower and centred on the web-based room workflow and remote group interaction. Observation notes and session records were used as contextual evidence for interpreting questionnaire and interview responses, not as a standalone quantitative log of participant behaviour.

5.1.5 Post-Evaluation Study

Immediately after the group selection task, participants completed post-evaluation instruments that combined structured usability measurement with reflective feedback. All participants in the relevant sessions completed the immediate SUS questionnaire, while participants were invited to take part in semi-structured interviews where they agreed to do so. This combination allowed the study to connect descriptive usability scores with participants' explanations of how they understood the interface, the allocation result, and the system's effect on group decision-making.

The post-evaluation materials were interpreted alongside the observation evidence, not as isolated measures. Questionnaire responses provided a comparable indication of perceived usability after system use, while interviews helped explain why participants trusted, questioned, or misunderstood parts of the workflow. Where delayed follow-up material was collected from Stage 2 Group 1, it was treated as a small longitudinal study component that captured remembered reflection on the earlier PickNPlay session, not as a replacement for the immediate post-task response.

Interview

Semi-structured interviews were used to elicit qualitative feedback about participants' experience with the system used in that session. A shared interview protocol guided the discussion through open-ended prompts about first impressions, comparison with usual board-game selection practice, ease of learning, confusing or unnecessary features, perceived fairness, trust in the result, and possible improvements. This structure kept the interviews comparable across groups while still allowing participants to explain issues that emerged from their own session.

The interview protocol was adapted to the system condition. PickNPlay interviews focused on the web workflow, preference entry, result interpretation, and whether participants could imagine using the system with their own group. Boardot interviews included additional prompts about the central AI host, asking how its guidance shaped discussion, whether participants felt more or less in control, and whether the group followed, ignored, or debated its prompts. Where participants had experienced more than one system across the study, interviews also allowed them to compare interaction styles across stages.

Interview material was used to explain and contextualise the questionnaire and observation data, not to claim that every participant shared the same view. Not every participant in every group took part in interviews, so interview-based claims in the later qualitative analysis are treated as evidence from the interviewed subset. Where consent had been given, interviews or sessions were audio-recorded and transcribed for later analysis. Transcripts and quotations were handled under the de-identification and data-management constraints described in the ethics section.

Questionnaire

The main post-use questionnaire instrument was the System Usability Scale (SUS), a ten-item standard usability questionnaire using five-point agreement responses (Brooke, 1996). Participants completed the SUS after using the relevant system in the selection task. The questionnaire was used to produce a comparable perceived-usability measure for the evaluation sessions, while interpretation of those scores was kept exploratory because of the small, situated participant groups described above.

The aggregate SUS score was the primary quantitative measure. Individual items were inspected diagnostically to identify whether usability concerns related mainly to acceptance, ease of use, complexity, learnability, confidence, or consistency, but item-level patterns were used only to interpret the aggregate scores rather than as separate scales.

Responses were later transformed into the standard 0–100 SUS score using the scoring procedure described in Section 5.2. The resulting scores were used descriptively, not as evidence that the small participant groups represented all board-game players.

Where follow-up questionnaire material was collected after Stage 2, the approved follow-up instrument included delayed SUS items and additional reflection items. The manually supplied follow-up dataset used in the current analysis contained the five reflection items only, so it was not scored as SUS. These items asked about remembered understanding, trust and explanation needs, usefulness for larger groups, possible independent use, and remaining unclear parts of the system. They were analysed separately from the immediate SUS responses as the Stage 2 Group 1 longitudinal follow-up component.

5.2 Quantitative Analysis

The quantitative analysis started with descriptive statistics. This matched the study design: the participant groups were small, recruitment was mixed, and some participants had prior exposure to earlier stages of the project. SUS scores are reported as descriptive evidence of perceived usability after system use. Cross-stage Mann–Whitney U tests are used only as exploratory checks: Stage 1 remains introduced in Chapter 3, while this chapter compares its baseline SUS scores with Stage 2 PickNPlay and compares Stage 2 PickNPlay with Stage 3 Boardot.

5.2.1 SUS Scoring and Descriptive Statistics

SUS responses were scored on the standard 0–100 scale, using the odd-item and even-item transformations described with the SUS instrument. The analysis used de-identified markdown tables prepared for each stage and group. Rows were included only when all ten item responses were available or a valid SUS score was provided. The Stage 2 PickNPlay analysis contained 29 complete immediate post-use SUS responses across four groups, or 290 completed SUS item responses in total. Stage 2 Group 3 was conducted online, while the other Stage 2 groups were conducted in person. The Stage 3 Boardot analysis added 13 complete immediate post-use SUS rows across two groups, both treated as observed Boardot evaluation data.

Table 5.1: Stage 2 PickNPlay SUS descriptive statistics.

Stage 2 group	N	Mean	Median	SD	Min–max
All Stage 2 PickNPlay	29	79.14	82.50	10.44	52.50–97.50
Group 1	8	76.56	76.25	9.35	57.50–87.50
Group 2	7	86.07	87.50	7.48	72.50–97.50
Group 3 (online)	5	79.00	82.50	8.59	67.50–87.50
Group 4	9	76.11	77.50	12.94	52.50–92.50

Stage 2 PickNPlay SUS Results

The participant-level Stage 2 table is retained in Appendix B.2. The main chapter focuses on the group-level pattern. Table 5.1 shows that Stage 2 PickNPlay scores were generally positive, with a descriptive mean of 79.14 and a median of 82.50, and most scores above 70.00. Group means ranged from 76.11 to 86.07, with Group 2 producing the highest descriptive mean and Group 4 showing the widest score range. Because the groups differed in recruitment route, prior exposure, and session context, the table describes the observed responses without claiming that one group condition caused higher usability.

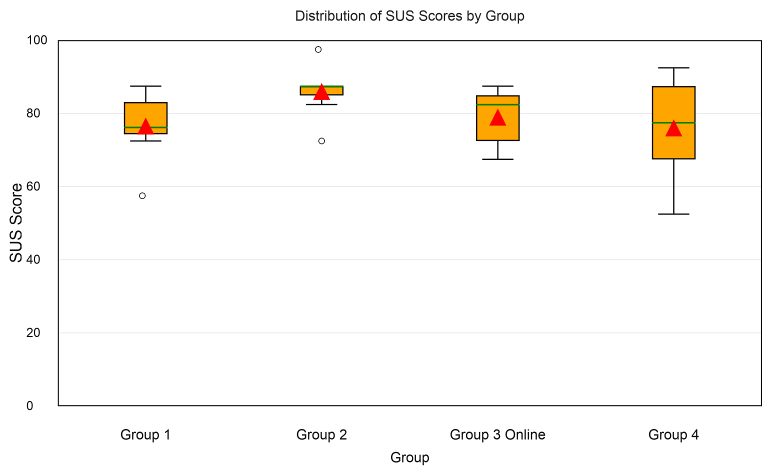


Figure 5.3: Stage 2 PickNPlay SUS distribution by group.

Figure 5.3 visualises the group-level spread. Group 2 clustered tightly around high scores, with an interquartile range of 2.50, while Group 4 had the widest spread, with an interquartile range of 20.00 and the lowest observed score. Group 3 online had a similar median to the overall Stage 2 median, but its small sample size and remote setting mean it should be read as descriptive remote-use evidence, not as a direct comparison with the in-person groups.

The item-level responses help explain the aggregate score without replacing it. Positively phrased items generally clustered around agreement: for example, Q9 had a raw mean of 4.21 and no responses below neutral, indicating relatively high confidence after use. For the negatively phrased items, raw responses mostly clustered around disagreement. Q4 was especially consistent, with 28 of 29 participants selecting either 1 or 2 for needing technical support. Q2 and Q8 showed more residual uncertainty: both had two participants selecting 4, meaning that a small subset still perceived unnecessary complexity or cumbersomeness. This pattern is consistent with the overall SUS result: PickNPlay was generally rated as usable, but complexity and interaction effort remained useful targets for design interpretation.

Stage 2 Group 1 Longitudinal Follow-Up Questionnaire

Stage 2 Group 1 also completed a longitudinal follow-up questionnaire after the original PickNPlay session. The available follow-up markdown data contained six participants and five reflection items scored on a 1–5 agreement scale, so it was analysed separately from SUS and retained as a small delayed-reflection component. The item-level table is reported in Appendix B.3. In the main chapter, the key point is that participants most strongly agreed that PickNPlay would be useful for larger groups or many possible games and that they could imagine using it without the researcher present. The negatively framed item about remaining unclear parts had a lower median of 2.00 but still ranged from 1 to 4, so some delayed uncertainty remained in the follow-up group.

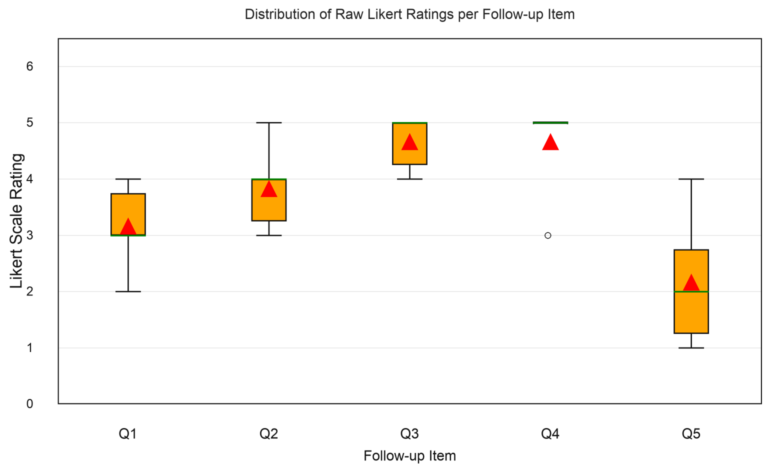


Figure 5.4: Stage 2 Group 1 follow-up response distributions.

Figure 5.4 shows the same delayed-reflection pattern visually. The trust-and-explanation item also had a positive central tendency, suggesting that explanation remained relevant to retrospective trust.

Table 5.2: Stage 3 Boardot SUS descriptive statistics.

Stage 3 group	N	Mean	Median	SD	Min–max
All Stage 3 Boardot	13	79.62	80.00	8.35	67.50–92.50
Group 1	4	85.63	85.00	6.58	80.00–92.50
Group 2	9	76.94	75.00	7.88	67.50–92.50

Stage 3 Boardot SUS Results

The Stage 3 Boardot SUS analysis used the two markdown data references for Stage 3 Group 1 and Group 2. Together, these files contained 13 complete SUS rows. The participant-level table is reported in Appendix B.4. Table 5.2 keeps the main chapter focused on the descriptive pattern. Across all Stage 3 rows, Boardot had a mean SUS score of 79.62 and a median of 80.00. Group 1 had a higher descriptive mean than Group 2, but the group sizes were small, so this difference should not be interpreted as a reliable group effect.

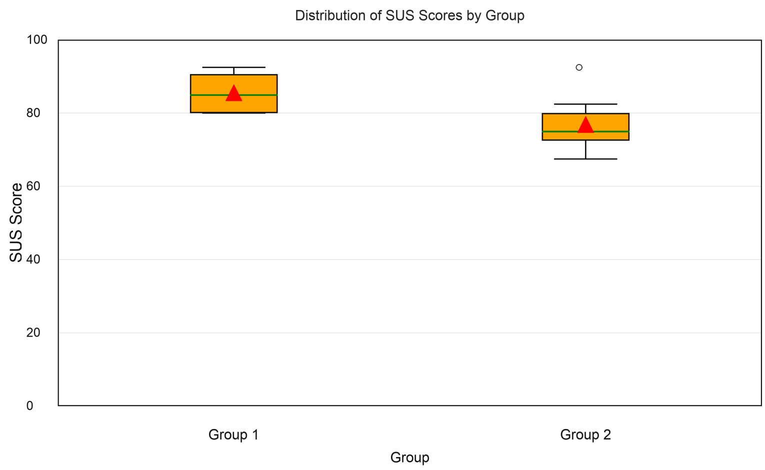


Figure 5.5: Stage 3 Boardot SUS distribution by group.

Figure 5.5 shows that the current Boardot scores were concentrated in the upper part of the SUS scale, with all Group 1 responses at 80.00 or above and Group 2 ranging from 67.50 to 92.50. The all-stage median of 80.00 is close to the Stage 2 PickNPlay median of 82.50, suggesting that the station-mediated version did not obviously reduce perceived usability in the current descriptive data.

The raw Stage 3 item responses gave the same diagnostic pattern without needing a separate main-text figure. Responses to the positively phrased items generally clustered toward agreement, especially Q9 on confidence using the system. For the negatively phrased items, Q8 on cumbersomeness showed the lowest adjusted contribution in the

Table 5.3: Exploratory cross-stage Mann–Whitney U results.

Comparison	**N**	**Mean**	**Median**	**U**	**p**	r_{rb}
Stage 1 BGVS vs Stage 2 PickNPlay	14 / 29	39.29 / 79.14	32.50 / 82.50	17.00	$< .001$	-0.916
Stage 2 PickNPlay vs Stage 3 Boardot	29 / 13	79.14 / 79.62	82.50 / 80.00	193.00	.913	0.024

item summary, so interaction effort remained a relevant design concern even though the overall Stage 3 SUS distribution was positive.

5.2.2 Mann–Whitney U Test

Two Mann–Whitney U tests were run as exploratory independent-samples comparisons. The first compared the Stage 1 BGVS baseline SUS scores introduced in Chapter 3 with the Stage 2 PickNPlay SUS scores reported above. A separate Stage 1 result subsection is not repeated here because Stage 1 functions as formative baseline evidence in the thesis structure. The second comparison tested the current Stage 2 PickNPlay and Stage 3 Boardot SUS distributions.

Figure 5.6 shows the three stage-level SUS distributions used as visual context for the pairwise Mann–Whitney U tests. It is not a separate three-group significance test.

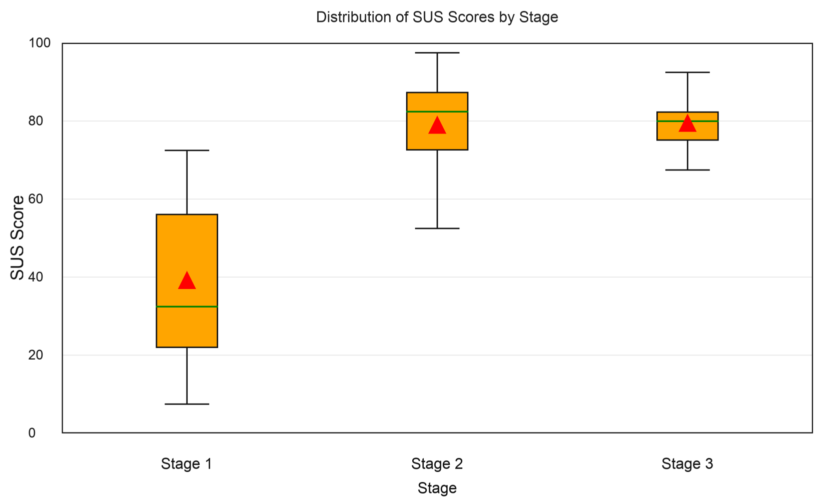


Figure 5.6: Stage 1 BGVS, Stage 2 PickNPlay, and Stage 3 Boardot SUS distributions.

The Stage 1–Stage 2 comparison supports the design narrative already developed across Chapters 3 and 4: PickNPlay was rated much more usable than the baseline formative workflow. However, this should not be read as a controlled experiment because

Stage 1 had a different methodological role, and the samples were small. The Stage 2–Stage 3 comparison did not indicate a detectable difference in SUS scores in the current data. This suggests that Boardot's station-mediated interaction did not obviously harm usability, but any stronger claim would require a larger independent sample, clearer participant-independence checks, and a study design intended for controlled comparison.

5.3 Thematic Analysis

The qualitative analysis explains how participants made sense of the implemented Stage 2 PickNPlay and Stage 3 Boardot systems. It is reported separately from the SUS analysis because the interview, session, and follow-up material was used to interpret participants' experiences of setup, preference entry, shared decision-making, result acceptance, and station-mediated facilitation, not to measure usability as a single score. Stage 1 material remains formative evidence for the requirements in Chapter 3 and is not re-analysed here as implemented-system evaluation evidence.

5.3.1 Thematic Analysis Process

The analysis followed a staged thematic-analysis workflow adapted from Naeem et al.'s process of familiarisation, keyword selection, coding, theme development, conceptual interpretation, and synthesis (Naeem et al., 2023). The workflow also follows the Braun and Clarke framing introduced in Chapter 2, where themes are treated as analytic patterns developed through interpretation, not as counts alone (Braun and Clarke, 2006, 2019). In this study, an explicit sub-theme stage was added between refined codes and themes so that lower-level interaction evidence remained traceable before thesis-level claims were written.

Corpus, Transcription, and Familiarisation

The included qualitative corpus contained 11 implemented-system transcript sources, with nine Stage 2 PickNPlay sources and two Stage 3 Boardot sources. Stage 2 included in-person PickNPlay interviews, one online translated PickNPlay interview, the Stage 2 comparison transcript, and the Stage 2 follow-up material confirmed as belonging to Group 1. Stage 3 included Boardot Group 1 and Group 2 interview/session material. Raw files in the study folder were not overwritten. Instead, cleaned derived analysis files were prepared with de-identified speaker labels, source paths, stage labels, system labels, group labels, and translation status.

The familiarisation stage involved reading the cleaned transcripts and recording corpus boundaries before coding. Stage 2 and Stage 3 were processed through the same workflow, but their evidence was kept analytically separate. The Stage 2 Group 3 online transcript was analysed only in translated form, and Stage 3 Group 2 material was treated as observed Boardot evaluation evidence.

Quote Extraction and Keyword Selection

Instead of extracting only illustrative examples, the analysis first built an exhaustive quote bank from the usable transcript corpus. This produced 397 included quote units, with 334 from Stage 2 PickNPlay and 63 from Stage 3 Boardot. A further 311 passage units were excluded with reasons such as administrative talk, off-task talk, inaudible or unclear material, or content not relevant to the research questions. Each included quote unit retained its stage, system, group, participant or speaker label, source artifact, and derived transcript reference. The quote identifiers used in the findings below, such as Q0035, refer to the internal de-identified quote ledger. The quote-extraction audit board is retained in Appendix B.2 rather than repeated in the main findings chapter.

Keywords were then assigned as short content descriptors before coding. This step helped separate descriptive signals, such as setup friction, rating semantics, fairness explanation, or shared display attention, from later analytic codes. Provenance labels such as translation status or consent handling were retained as metadata, not treated as keywords. The keyword-selection board is retained in Appendix B.2.

Initial and Refined Coding

Initial coding was inductive. Each quote unit could receive more than one initial code when it spoke to multiple interaction issues, such as game information, preference mechanics, and result trust. The initial codebook contained 46 initial code labels and 1081 quote-to-initial-code assignments. The initial-code audit board is retained in Appendix B.2. The main chapter keeps the later refined-code and theme layers because they are closer to the findings structure.

The code refinement stage consolidated the initial codes into 22 refined codes. Similar codes were merged, overly broad codes were split, and ambiguous labels were renamed so that each refined code described a coherent interaction pattern. For example, evidence about camera access, room joining, and setup status was kept distinct from evidence about fairness explanation or social negotiation. Refinement decisions were documented in the working codebook so that later theme claims could be traced back to quote-level evidence. Figure 5.7 shows the refined-code layer used before sub-theme grouping.

(a) Stage 2 PickNPlay refined-code layer.

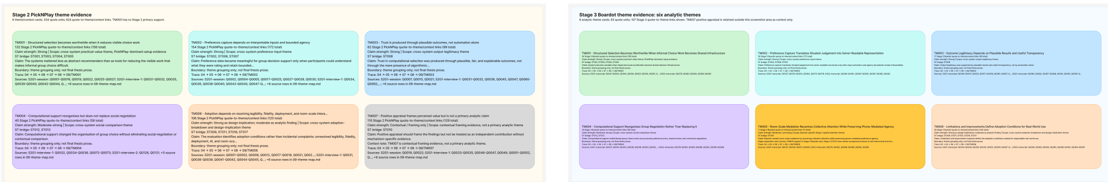
(b) Stage 3 Boardot refined-code layer.

Figure 5.7: Refined-code audit boards.

Sub-theme, Theme, and Synthesis Development

The 22 refined codes were clustered into 17 sub-themes and then into six primary analytic themes plus one contextual framing theme. FigJam was used as an affinity-mapping and audit-support board for the hierarchy from quote units through keywords, initial codes, refined codes, sub-themes, themes, and synthesis. The board was not treated as evidence by itself. It was used to make the analysis process inspectable and to preserve the separation between Stage 2 PickNPlay and Stage 3 Boardot evidence. Figure 5.8 shows the final theme layer with the stage-specific evidence kept in separate areas.

The final themes were reviewed for claim strength and boundary conditions before being translated into thesis prose. Cross-system themes were used only where both stages contained relevant evidence and where the source labels remained visible in the ledger. Boardot-specific evidence was written as Stage 3 evidence only, especially for the theme about a shared station recentring collective attention. The conceptual-synthesis artefact listed in Appendix B.2 records how the theme map was connected to the findings structure below.



(a) Stage 2 PickNPlay theme layer.   (b) Stage 3 Boardot theme layer.

Figure 5.8: Theme-development audit boards.

5.3.2 Overview of Qualitative Findings

The qualitative findings are organised as six analytic themes. A broader pattern of positive appraisal was retained as contextual framing, not elevated into a theme by itself, because positive comments did not explain what made the systems useful, fragile, or different from existing group-selection practices. The themes instead describe the interactional conditions under which PickNPlay and Boardot worked as computational support for group choice. They reduced visible coordination work, translated situated preferences into computable representations, produced plausible outcomes, reorganised negotiation, recentred attention in the room, and exposed the adoption work still required for real use.

- **Shared decision infrastructure,** focuses on how PickNPlay and Boardot became useful when they turned informal coordination work into shared room state, candidate entry, progress visibility, preference capture, and allocation.
- **Preference representation,** describes how participants translated situated judgement into solver-readable ratings, ties, and submissions, while Boardot preserved

phone-mediated input rather than changing the underlying preference algorithm.

- **Outcome legitimacy,** concerns how participants accepted results when they appeared plausible and explainable enough to act on, without treating perceived legitimacy as a formal fairness proof.
- **Socio-technical negotiation,** highlights how computation redistributed negotiation across private input, shared progress, result inspection, and revision, with Stage 2 comparison material used only as comparison evidence.
- **Room-scale mediation,** captures Boardot's Stage 3-specific contribution of making allocation a shared spatial event while still preserving phone-mediated individual agency.
- **Adoption conditions,** consolidates limitations and improvement feedback around legible controls, recoverable setup, clear AI roles, robust voice input, physical persistence, and prototype fidelity.

5.3.3 Structured Selection Becomes Worthwhile When Informal Choice Work Becomes Shared Infrastructure

The first theme concerns the point at which computational structure becomes interactionally worthwhile. Participants did not value PickNPlay and Boardot simply because they were automated or novel. They valued them when the systems converted dispersed social coordination into shared decision infrastructure by creating a room, building a candidate game pool, collecting individual preferences, tracking progress, and producing a feasible allocation quickly enough for play to begin.

This was strongest in Stage 2 PickNPlay. Participants repeatedly connected usefulness to the redistribution of setup work. Concurrent candidate-game entry was described as faster than previous approaches (Q0035, Stage 2 PickNPlay, S2G1), and multiple people entering games into the same room was identified as a major benefit (Q0133, Stage 2 PickNPlay, S2G1). The translated online evidence points to the same mechanism from a remote context. The system was "more organised" than usual discussion because everyone could enter a room, list preferences, and then receive a result (Q0294, Stage 2 PickNPlay, S2G3 translated).

This evidence evaluates several requirements from Chapter 3. DG1, DG5, and DG6 framed the system as a way to distribute coordination burden, treat the room as shared state, and keep setup short. FR1, FR2, and IR2 made this concrete through room-based participation, shared candidate entry, and visible group progress. The Stage 2 evidence suggests that these goals were meaningful to participants, but only when the system genuinely moved work away from a single organiser and into shared, visible interaction. In Chapter 4 terms, the value came less from allocation as an isolated solver result and more from the surrounding room-state architecture, including shared candidate games, visible progress, active preference submission, and a solver result that could be inspected as a group outcome.

Stage 3 Boardot extends this finding by showing the cost of making that infrastructure physical. The request for Boardot to be "set up forever in this room" and integrated into the table (Q0395, Stage 3 Boardot, S3G2) is not merely a deployment complaint. If a station is repeatedly assembled, joined, and explained, it risks becoming new coordination work instead of reducing existing work. A shared station can reduce coordination work only if it absorbs its own setup cost across repeated sessions.

The boundary of this theme is that structured selection is not always better than informal choice. The systems appear most useful when coordination work is high, such as when multiple people, candidate games, feasibility constraints, or repeated sessions have to be managed together. For small groups or familiar games, ordinary discussion may remain cheaper than system setup. The design implication is to make structured selection lightweight, recoverable, and reusable, so that the system's coordination infrastructure appears only when it repays the work it asks from participants.

5.3.4 Preference Capture Translates Situated Judgement into Solver-Readable Representation

The second theme concerns how individual preferences became computational input. PickNPlay and Boardot did not simply collect preferences that already existed in a stable form. They asked participants to translate situated judgement into structured, solver-readable representations, including ratings, ties, linking requests, and room-specific submissions. The quality of that translation depended on whether participants understood the games, the meaning of the controls, and the boundaries of what the system would do with their input.

Stage 2 PickNPlay provided the strongest evidence that game knowledge shaped the meaning of preference data. One participant described uncertainty when rating games they had not seen before, even after reading enough to get some sense of them (Q0050, Stage 2 PickNPlay, S2G1). Another participant wanted "one or two sentences" as a short introduction instead of a full description (Q0290, Stage 2 PickNPlay, S2G2). These comments suggest that game information is not a secondary content feature. It is part of the preference-capture apparatus because ratings are weak evidence when participants do not know what they are rating.

The rating mechanics also needed to be legible as preference representation. A Stage 2 participant said it was not clear what the stars meant once a game was inside a segment (Q0089, Stage 2 PickNPlay, S2G1), and another participant found the tie/linking mechanism confusing (Q0315, Stage 2 PickNPlay, S2G4). These are not only interface clarity issues. They affect the integrity of solver input, because equal ratings and chain links become tied tiers in the allocation model described in Chapter 4. If participants misunderstand these controls, the system may compute over a representation that looks precise but encodes the wrong kind of indifference or priority.

This finding links directly to the requirements in Chapter 3. DG3 framed the need for expressive but lightweight preference input, while FR4, FR5a, FR5b, IR3, IR4, and

ER6 required preference collection, tiered preference states, preference-boundary signals, game context before voting, low-friction input, and disclosure management. Chapter 4 implemented this through a rating-to-tier pipeline that preserves weak preference orders and ties while avoiding the burden of full ranking on a phone. The evaluation suggests that this direction was appropriate, but it also confirms the boundary around unfamiliarity and low-confidence ratings. Future versions should treat uncertainty as an explicit interaction state or explanation signal instead of collapsing it into ordinary low preference. Hard veto remains a design requirement and known implementation boundary, not a solved part of the evaluated pipeline.

Stage 3 Boardot adds an agency boundary, not a new preference mechanism. Participants still entered and adjusted preferences on personal phones. One participant said decision control did not change because they were still making decisions on their own phones instead of relying on the station (Q0392, Stage 3 Boardot, S3G2). Another described the system as preserving private individual decision-making compared with open discussion (Q0381, Stage 3 Boardot, S3G2). Stage 3 also shows that preserving phone-mediated agency is not enough by itself. One participant wanted star adjustments during setup, not only after voting began (Q0376, Stage 3 Boardot, S3G1). This is important for the Boardot interpretation. The station changed the surrounding presentation of the process, but it did not take over preference expression.

The boundary of this theme is that it does not claim the implemented preference model was wrong or unusable. Participants often completed the workflow successfully, and the model gave the resolver usable preference structure. The finding is more specific. Preference data are only as meaningful as the participant's understanding of the game, the input semantics, and the disclosure and revision boundaries around that input. The design implication is to treat preference entry as representational work, not merely as form completion.

5.3.5 Outcome Legitimacy Depends on Plausible Results and Useful Transparency

The third theme concerns how participants came to treat system outputs as legitimate. PickNPlay and Boardot did not earn trust simply by producing an automated result. Participants read the outcome through visible or inferable cues that the allocation was plausible. Preferences appeared to have been submitted, player-count constraints appeared to matter, and the result produced playable groups instead of an arbitrary allocation.

Stage 2 PickNPlay shows this most clearly through the tension between opacity and acceptance. One participant described the algorithm as "completely opaque" while still saying that the result seemed fair (Q0115, Stage 2 PickNPlay, S2G1). Another participant said they were not sure how the system reached the result because it did not really explain the process (Q0282, Stage 2 PickNPlay, S2G2). These comments show that perceived fairness was not the same as algorithmic understanding. Participants could accept an outcome as plausible while still wanting a clearer account of the trade-offs behind it.

The translated online evidence gives a more explicit account of why the result could feel legitimate. The participant was not especially worried about fairness because the system used everyone's preferences and the number of players to calculate the result, but they still said transparency could be stronger (Q0302, Stage 2 PickNPlay, S2G3 translated). A related translated quote positioned system trust against ordinary group discussion. The system was trusted more because it followed rules instead of reflecting whoever was most powerful in the discussion (Q0303, Stage 2 PickNPlay, S2G3 translated). This supports a procedural interpretation of trust. The system was not trusted because it was mysterious automation. It was trusted when it appeared to follow a consistent, preference-sensitive procedure.

This finding maps directly onto DG4 and the explanation requirements in Chapter 3. ER1 to ER4 asked the result view to connect assignments to preferences and constraints without overwhelming participants with solver detail. Chapter 4 implemented this through result explanation and a simplified visualisation, not a low-level Z3 trace. The evaluation suggests that this level of explanation is directionally appropriate, but incomplete. Participants wanted enough rationale to see why the allocation followed from their inputs, not just the final grouped output.

Stage 3 Boardot adds evidence about public and spatial presentation of the resolution episode, not deeper algorithmic transparency or a different allocation logic. A participant described the table marker as linking location to the resolution process (Q0387, Stage 3 Boardot, S3G2), and later described the shared screen and process as bringing people together around the result (Q0388, Stage 3 Boardot, S3G2). These comments suggest that Boardot can make the allocation episode more publicly observable and spatially situated. However, this should not be overclaimed as deeper algorithmic transparency. The station made the resolution episode more visible, while the same underlying allocation still required explanation of preferences, constraints, and trade-offs.

The boundary of this theme is central. The evidence supports perceived procedural legitimacy, not formal fairness proof or demonstrated algorithmic comprehension. Future design work should focus on useful transparency that shows selected and unselected games, binding constraints, preference compromises, teaching or don't-care factors, and revision paths. The design goal is not to expose every solver step, but to give groups enough explanation to decide whether the outcome is acceptable enough to discuss, accept, or revise.

5.3.6 Computational Support Reorganises Group Negotiation Without Replacing It

The fourth theme concerns the relationship between computation and social negotiation. Participants did not treat PickNPlay or Boardot as systems that simply replaced group discussion. Instead, the systems reorganised the decision process. Preferences were first entered individually, the room state made progress and candidate games visible, and the final allocation became an object that the group could interpret together.

Stage 2 PickNPlay evidence shows why this matters. One participant acknowledged that the system could not "please everyone" but still valued the fast process because the group could start playing instead of continuing to decide (Q0272, Stage 2 PickNPlay, S2G2). The translated online evidence makes the social mechanism sharper. The system felt fairer because selection was not about who could talk best (Q0298, Stage 2 PickNPlay, S2G3 translated). Similarly, a Stage 2 participant contrasted the system with open discussion, where some players' preferences could be ignored (Q0312, Stage 2 PickNPlay, S2G4). These comments show that computation did not remove group judgement. It changed the conditions under which individual preferences entered the collective decision.

This is the technical and HCI significance of the shared preference pipeline, and it links to DG1, DG2, DG5, IR5, IR6, and IR7 from Chapter 3. Chapter 4's resolver depends on submitted preference structures and feasibility constraints, but those inputs are produced through social interaction. Participants decide what they know, what they want, whether to express indifference, and whether to accept the result. The allocation model can compute across submitted data, but it cannot fully represent every contextual reason a group might prefer one evening's compromise over another. That is why the result view, explanation, and revision pathway remain part of the interaction design, not optional polish.

Stage 3 Boardot makes the reorganisation more visible without changing the underlying allocation logic. A participant valued that everyone watched the ranking or review together on the same screen (Q0337, Stage 3 Boardot, S3G1). Another said Boardot did not change their individual expression but added a social aspect by bringing people into discussion (Q0385, Stage 3 Boardot, S3G2). This is a different claim from saying that Boardot made the algorithm more powerful. Boardot changed the social staging of the decision. The station gave the group a shared review surface while phones preserved individual preference entry.

This theme also marks how comparison evidence is used. The Stage 2 Group 2 comparison transcript is valid evaluation material, but here it is used as comparison evidence about how participants contrasted approaches to selection. It should not be used to claim that computational support removed negotiation. Rather, it helps show that PickNPlay and Boardot shifted where negotiation happened and what information the group negotiated around.

The boundary is that TM004 should not overstate computational authority. A Stage 3 participant explicitly said decision control did not change because people were still making decisions on their own phones instead of relying on the station (Q0392, Stage 3 Boardot, S3G2). The systems support private preference expression, shared progress awareness, outcome explanation, and revision. They do not assume that an allocation result is the end of the decision.

5.3.7 Room-Scale Mediation Recentres Collective Attention While Preserving Phone-Mediated Agency

The fifth theme is the main Boardot-specific finding. Boardot's distinctive contribution was to change how the group attended to the decision process while leaving the preference computation unchanged. By placing a shared station in the room, Boardot made the allocation episode visible as a collective event while leaving individual preference input on personal phones.

Several Stage 3 comments described the shared station as the difference. One participant said the centralised station was the most different element compared with previous use (Q0379, Stage 3 Boardot, S3G2). Another described the ranking or review stage as something everyone watched together on the same screen (Q0337, Stage 3 Boardot, S3G1). This matters because the same underlying allocation model can be experienced differently depending on where the process is displayed. A phone-only system can collect individual inputs efficiently, but a room-scale station can turn the resolution phase into a shared object of attention.

The location and marker evidence strengthens this spatial interpretation. Participants described the marker and phone background as helping bring the group together (Q0386, Stage 3 Boardot, S3G2), and the shared screen plus location cues as giving a sense of the resolution process and bringing people together (Q0388, Stage 3 Boardot, S3G2). These comments are not merely about visual novelty. They suggest that spatial cues helped participants connect abstract allocation results to people and places in the room.

At the same time, Boardot did not replace phone-mediated agency. A participant explicitly said that decision control did not change because preferences were still made on personal phones instead of by relying on the station (Q0392, Stage 3 Boardot, S3G2). This is the key design tension. Boardot recentres collective attention, but it does not centralise preference entry. The mixed-device arrangement is valuable because it separates private preference articulation from public resolution.

This finding links directly to the Boardot design in Chapter 4. The station, markers, station prompts, and bounded dialogue are best interpreted as presentation and orientation infrastructure around the shared room state, not as a second allocator. Boardot uses the same room state, preference pipeline, and resolver as PickNPlay. The station displays candidate games, progress, prompts, markers, and final clusters, but it does not change submitted preferences or solver constraints. The design implication is to preserve this division of labour. Phones should remain the space for private preference input and revision, while the station should support group awareness, progress, explanation, and spatial orientation.

The boundary is that TM005 is a moderate, promising finding, not the strongest theme in the thesis. It is supported by 11 primary Stage 3 Boardot quote units and should not absorb Stage 2 comparison evidence as if it were Boardot evaluation evidence. It is still important because it gives Boardot a distinct analytic role by spatially mediating collective attention around constrained group allocation.

5.3.8 Limitations and Improvements Define Adoption Conditions for Real-World Use

The sixth theme consolidates limitations and improvement feedback as adoption conditions. Across both PickNPlay and Boardot, participants identified breakdowns that were smaller than the main design claims but still central to whether the systems could be used in ordinary sessions. These included unclear controls, workflow interruptions, AI visibility, prototype fidelity, physical setup burden, and Boardot's room-scale voice fragility. These adoption layers were not equally evidenced. Interaction legibility and prototype reliability were broad cross-system concerns, while persistent deployment and Boardot voice robustness were narrower but important Stage 3 conditions.

Stage 2 PickNPlay shows the importance of interaction legibility and recovery. A participant trying to add a game through the camera/browser flow was taken out of the room and had to enter again (Q0042, Stage 2 PickNPlay, S2G1). Another participant said the linking and rating interaction needed more instruction (Q0280, Stage 2 PickNPlay, S2G2). These are not isolated usability notes. They expose adoption risks in a system whose value depends on fast, shared coordination. If joining, scanning, linking, and rating are brittle, the system can recreate the coordination burden it was meant to reduce.

AI-related evidence should be kept conditional. In Stage 2, a participant saw the AI summary but did not click it, partly because it was interpreted like an ordinary description (Q0291–Q0292, Stage 2 PickNPlay, S2G2). This suggests that generated support needs discoverability, timing, and role clarity before it can contribute to game understanding or confidence. It should not be treated as automatically valuable because it is labelled AI.

Stage 3 Boardot adds a different adoption condition around voice and room-scale robustness. One participant said they had not tried many prompts because noise and overlapping talk made the interaction difficult (Q0390, Stage 3 Boardot, S3G2). The same participant suggested that listen/speak functionality might need to move to individual phones because many people talking made it difficult for the system to listen to the correct person (Q0391, Stage 3 Boardot, S3G2). These comments suggest that future voice facilitation should be treated as a room-scale interaction problem involving noise, turn-taking, and source ambiguity, not merely as a speech-recognition feature request.

Deployment and prototype fidelity also shaped adoption judgements. Boardot was seen as more useful if it could remain integrated into the room or table instead of being set up each time (Q0395, Stage 3 Boardot, S3G2). The same corpus includes caution that the prototype was not yet fully presented at a level that would support confident judgement (Q0397, Stage 3 Boardot, S3G2). Prototype fidelity did not invalidate the evaluation, but it limited how confidently participants could judge repeated real-world use. For a room-scale system, physical persistence and fidelity are part of the interaction design.

This theme links back to DG2, DG6, IR3, IR4, IR7, ER3, and ER5 in Chapter 3, and to the implementation boundaries described in Chapter 4. These boundaries include camera and joining flows, preference-control semantics, AI-generated descriptions, station prompts, voice interaction, and physical setup. Its purpose is to define the adoption work required for the earlier benefits to hold, not to compete with the main conceptual themes. The design implication is to make the infrastructure dependable at the edges through legible controls, recoverable workflow, clear AI roles, robust room-scale interaction, persistent deployment, and sufficient prototype fidelity for credible real-use judgement.

5.3.9 Qualitative Synthesis

Taken together, the six themes show that PickNPlay and Boardot should be understood as systems for organising group choice work, not as allocation engines alone. The allocation model matters, but participants encountered it through a broader interactional arrangement. This arrangement included shared room state, candidate-pool construction, private preference representation, public progress, result explanation, group discussion, and recovery. PickNPlay provides the stronger evidence for setup, preference capture, and outcome-legitimacy conditions. Boardot adds a more bounded but important finding. Shared spatial infrastructure can recentre group attention around the allocation process while leaving individual preference input on personal devices.

These qualitative findings also clarify the SUS results reported earlier in this chapter. The positive SUS distributions indicate that participants generally found the implemented systems usable, but they do not explain why the systems were accepted or where adoption remained fragile. The thematic analysis shows that usability depended on visible reductions in choice work, meaningful preference representation, plausible and explainable outcomes, preservation of social negotiation, and practical adoption conditions. Chapter 6 discusses this synthesis together with the quantitative results in relation to the three research questions.

Chapter 6

Conclusion

This thesis began from a practical HCI problem. Groups often need help making constrained collective decisions, but computational support can easily become another source of opacity, control imbalance, or social friction. Board-game selection made this problem concrete. Participants had to coordinate preferences, player-count constraints, teaching knowledge, game familiarity, and the possibility of splitting across multiple tables. In this setting, the design challenge was not only to compute a feasible allocation, but to make that allocation understandable and socially usable.

The main conclusion is that solver-backed group allocation can be made more acceptable in this setting when it is designed as shared interaction infrastructure, not as a hidden decision engine. PickNPlay provides the primary evidence for this claim. Shared room state, collaborative candidate-pool construction, lightweight preference input, visible progress, and result explanation helped participants treat the computed allocation as something they could inspect, question, and use. Boardot extends the argument more cautiously by showing that the same allocation workflow can be spatially mediated through a shared station, while individual preference expression remains on personal devices. The station changed how the group attended to the decision process, not what the resolver computed.

This chapter is organised as follows. Section 6.1 first answers the three research questions. Section 6.2 then states the thesis contributions and significance. Section 6.3 discusses the limits of the evidence and identifies future work. The emphasis is on synthesis. The quantitative findings describe perceived usability in prototype-oriented, situated study sessions, while the qualitative findings explain the interactional conditions under which that usability became meaningful. The claims should be read as evidence about perceived usability, procedural legitimacy, and interaction design conditions, not as population-level proof of effectiveness.

6.1 Findings per Research Question

6.1.1 RQ1: Transparency, Fairness, and Acceptable Allocation

RQ1 asked how the social choice process could be communicated transparently and accessibly so that participants understood the procedure and regarded the allocation outcome as fair and acceptable. The quantitative results provide the first part of the answer. PickNPlay was rated much more usable than the formative BGVS baseline. The exploratory Mann–Whitney comparison between Stage 1 BGVS and Stage 2 PickNPlay showed a large descriptive shift in SUS scores, from a Stage 1 mean of 39.29 and median of 32.50 to a Stage 2 mean of 79.14 and median of 82.50 (Table 5.3). This does not prove a controlled causal effect, because Stage 1 had a formative role and the samples were not designed as independent controlled groups, but it supports the design narrative that the implemented PickNPlay workflow was rated as easier to use than the earlier baseline.

The more important result is what transparency meant in practice. The study does not support the idea that participants needed full visibility into the optimiser, and the system did not provide a complete Z3 trace. Instead, TM002 and TM003 show that transparency worked across the interaction. Participants needed enough game context to make preferences meaningful, enough clarity about rating and tier semantics to trust their own input, and enough result feedback to see that player-count constraints and submitted preferences had shaped the allocation. Transparency was not a single screen or explanation feature. It was a relationship between preference capture, constraint visibility, outcome presentation, and revision.

This also changes how fairness should be interpreted. The thesis does not claim that PickNPlay proves formal fairness properties such as envy-freeness, proportionality, or strategy-proofness. The evidence concerns perceived procedural legitimacy. Some participants accepted results as fair enough even while describing the algorithm as opaque, while others wanted clearer explanation of why a particular allocation had been chosen. This indicates that participants judged fairness through whether the procedure appeared consistent, preference-sensitive, and constrained by practical playability, not through direct verification of the optimisation model.

The six-person Stage 2 Group 1 follow-up responses reinforce this boundary without broadening it into a general adoption claim. Participants most strongly agreed that PickNPlay would be useful for larger groups or many possible games and that they could imagine using it independently, both with means of 4.67 and medians of 5.00. However, the item about understanding how preferences were used had a lower mean of 3.17, while the item about trusting PickNPlay more if it showed why one allocation was chosen had a mean of 3.83 (Appendix Table B.3). The system was perceived as usable and practically acceptable in these sessions, but explanation remained part of the core design problem, not an optional enhancement.

The answer to RQ1 is therefore that transparent social choice support should be designed as useful transparency, not complete algorithmic exposure. For this domain, useful transparency means showing how submitted preferences, tied or low-ranked games, playable player-count constraints, teaching or don't-care factors, selected games, unselected games, and revision paths relate to the outcome. This aligns with DG4 and the explanation requirements in Chapter 3, and with the simplified visualisation and result-view design in Chapter 4. The technical allocator is necessary, but participants need an interaction layer that makes the allocation plausible, inspectable, and revisable enough for group use.

6.1.2 RQ2: Participation, Control, and Negotiation

RQ2 asked how interface design could reduce organiser-centred control and support more balanced participation while preserving negotiation. The quantitative results do not measure balanced participation directly, but they suggest that the PickNPlay workflow was usable enough not to reintroduce organiser dependence through technical difficulty. Stage 2 PickNPlay had a mean SUS score of 79.14 and median of 82.50 across 29 responses (Table 5.1). The item-level responses also indicated low perceived need for technical support. In total, 28 of 29 participants disagreed or strongly disagreed that they would need technical assistance. This matters because a system intended to reduce organiser burden cannot depend on a skilled organiser to run it.

The qualitative findings explain what kind of participation the interface supported. TM001 shows that PickNPlay was useful when it converted tasks previously handled informally, or by one organiser, into shared decision infrastructure. These tasks included joining a room, building the candidate pool, checking progress, submitting preferences, and reaching a feasible allocation. This directly responds to the BGVS baseline gap identified in Chapter 3, where room setup and resolution were concentrated in a fixed administrator role. PickNPlay did not make every participant equally influential in every moment, but it made more of the decision state and more of the available actions visible to the group.

This is the important distinction for RQ2. Reducing organiser-centred control is not the same as removing human judgement or social influence. Candidate-pool construction still shapes which outcomes are possible. Game familiarity still affects whether preferences are meaningful, and participants still need to discuss whether a computed result suits the occasion. The system's contribution was to redistribute parts of the work that can reasonably be shared, including adding or reviewing candidate games, expressing preferences privately, seeing who has submitted, resolving when the room is ready, and revising when the outcome or setup is not right.

TM004 then shows how negotiation was preserved. PickNPlay did not replace discussion with a final automated answer. It changed where negotiation happened and what the group negotiated around. Before voting, the group could negotiate the candidate pool and constraints. During voting, participants could express preferences individually without relying on the loudest or most decisive person to represent them. After resolution, the computed allocation became a shared artefact that participants could interpret, accept, question, or revise. This is why the six-person Stage 2 Group 1 follow-up evidence matters cautiously. Participants strongly agreed that PickNPlay would be useful for larger groups or many possible games, which is precisely where informal discussion becomes more costly and uneven.

The answer to RQ2 is therefore that balanced participation is supported by distributing the interactional work of selection, not simply by automating the final choice. Interface design should turn organiser tasks into visible shared capabilities, including contributing options, checking readiness, submitting preferences, understanding progress, resolving, and recovering from mistakes. At the same time, the interface should preserve negotiation after the allocation, because the computed result is an artefact for group decision-making, not the end of the decision.

6.1.3 RQ3: Direct Interaction and AI-Mediated Facilitation

RQ3 asked how direct participant interaction and AI-mediated facilitation affected perceived control, trust, fairness, and social cohesion during group decision-making. The Stage 3 Boardot SUS results show that adding the station-mediated layer did not obviously reduce perceived usability in the current data. Boardot had a mean SUS score of 79.62 and median of 80.00 across 13 responses (Table 5.2), close to the Stage 2 PickNPlay median of 82.50. The exploratory Stage 2–Stage 3 Mann–Whitney comparison found no detectable difference in the current exploratory sample, with $p = .913$ and $r_{rb} = 0.024$ (Table 5.3). This should not be read as evidence of equivalence, but it suggests that the shared station did not visibly undermine usability.

The study only partially answers the AI-mediated part of RQ3 because Boardot bundled several changes, including a shared display, spatial markers, station-side voice interaction, generated prompts, and prototype setup. The evidence speaks more directly about station-mediated shared attention than about the independent effect of AI facilitation.

The qualitative answer is more specific than a general comparison between “direct” and “AI-mediated” interaction. Direct participant interaction supported perceived control because preference expression remained on personal devices. Participants could still inspect games, enter ratings, adjust preference structure, and submit their own room-specific vote. This phone-mediated layer matters because it kept Boardot from becoming the decision-maker. Even in the station condition, the system did not ask participants to hand preference authority to an AI host or to the public display.

Boardot’s stronger contribution was to social attention, not to algorithmic trust or fairness. TM005 shows that the station changed the ecology of the same allocation process. The shared display, table markers, progress cues, and public result presentation made the resolution episode more observable and spatially situated. This supported shared orientation and co-present awareness in a limited but meaningful sense. The group could attend to the allocation together, while individuals still produced their preferences privately. Boardot suggests a useful division of labour for co-located group decision support.

Phones are effective for private preference articulation, while a shared station can help the group orient around progress, resolution, and the final allocation as a collective event.

The AI-mediated part of RQ3 needs a more cautious answer. TM006 shows that Boardot's voice and generated-text features were adoption conditions, not settled contributions. Voice prompts were fragile in a noisy group room, and participants suggested that listening or speaking might need to move closer to individual phones. AI-generated game support also depended on discoverability and role clarity. It was not automatically useful because it was labelled AI. The evidence does not show that AI facilitation independently improved perceived fairness, trust, or control. Instead, it suggests that AI facilitation should remain bounded, legible, and clearly separate from the allocation mechanism.

The answer to RQ3 is therefore conditional. Direct phone-mediated interaction and public station-mediated resolution can coexist productively when their responsibilities are separated. The station can make the allocation process feel more collective without taking away individual agency, but stronger claims about AI facilitation, voice interaction, trust improvement, fairness improvement, or social cohesion would require a more mature prototype and a study design focused specifically on those mechanisms.

6.2 Contributions and Significance

This thesis makes four connected contributions. The first is a framing contribution. It treats board-game selection as a constrained group-allocation interaction problem, not as a simple recommendation, filter, or vote. This framing matters because the group is not only choosing "the best game". It is constructing a playable arrangement of people and games under player-count constraints, uneven game knowledge, teaching needs, social preferences, and time pressure. The background chapter established the computational side of this problem through social choice, allocation, and constraint-solving literature, while the requirement analysis showed why the same problem also depends on shared control, explanation, and social acceptability. The conclusion from the thesis is that applied social choice systems in casual group settings need to be designed around the whole decision process, not only around the final aggregation rule.

The second contribution is PickNPlay as a participant-directed interaction framework for constrained board-game allocation. PickNPlay realises the requirements developed in Chapter 3 through shared room state, collaborative candidate-pool construction, rating-to-tier input, and a result view that connects assignments back to preferences and constraints. These elements reduce dependence on a fixed organiser, let the group shape the option set, reduce the burden of strict ranking while preserving useful preference structure, and make the result inspectable. PickNPlay also contributes a structured information flow between users and algorithm. The group builds a shared room state, participants convert situated preferences into submitted room votes, the resolver processes a bounded snapshot, and the result returns as an explainable and revisable group artefact. The Z3-backed resolver is important because it makes feasible multi-game allocation possible, but the HCI contribution is the way the resolver is embedded in a visible and revisable group workflow. In this sense, PickNPlay is not just an interface wrapped around an optimiser. It is an interaction structure for turning situated group preferences into a computable allocation and then returning that allocation to the group as something they can discuss.

The third contribution is the exploratory Boardot extension, which tests how the same allocation workflow changes when it is mediated through a shared room-scale station. Boardot should be understood carefully. It does not contribute a different allocation rule, and the evidence does not show that AI facilitation independently improved trust or fairness. Its contribution is narrower and more useful because it suggests that public display, spatial markers, and bounded facilitation can help recentre collective attention around the allocation process while preserving phone-mediated individual preference input. This distinction keeps PickNPlay and Boardot from being treated as interchangeable systems. PickNPlay contributes the core participant-directed framework. Boardot contributes evidence about how that framework can be staged in a co-located room.

The fourth contribution is the empirical synthesis of usability and qualitative evidence around these systems. The SUS results show that PickNPlay and Boardot were generally perceived as usable in the study sessions, while the thematic analysis explains what made that usability meaningful. Participants valued the systems when they reduced visible coordination work, supported meaningful preference representation, produced plausible outcomes, preserved space for negotiation, and made the decision process easier to attend to as a group. The same evidence also defines the boundary of the contribution. Explanation remained incomplete, unfamiliarity and veto were not fully represented in the solver model, Boardot's AI and voice features remained fragile, and the evaluation was exploratory, not a controlled demonstration of comparative effectiveness.

The thesis is significant because it gives an HCI account of how solver-backed allocation can become socially usable in a casual group setting. The constraint solver matters, but the contribution lies in how the allocation process is made understandable, revisable, and usable by a co-located group. For group decision support, the thesis suggests that distributing decision work means more than distributing button access. Participants need visible state, meaningful input semantics, recovery paths, and a result that can become a shared object of discussion. For social choice interfaces, the evidence indicates that perceived fairness and acceptability in this domain depend on useful transparency, not complete exposure of algorithmic internals. For room-scale AI facilitation, the thesis suggests that AI should be bounded as orientation and support around the decision process, not treated as a hidden decision authority.

Together, these contributions answer the larger problem introduced at the start of the thesis. Computational support can help groups make constrained collective decisions, but only when the interaction design preserves enough agency, explanation, and social negotiation for the computed outcome to become part of the group's own decision-making process.

6.3 Limitations and Future Work

The limitations of this thesis are best understood as boundaries around the kind of HCI contribution it makes. The project developed and evaluated PickNPlay as the main participant-directed framework for constrained board-game allocation, while Boardot was used as an exploratory extension for studying room-scale facilitation. The evaluation therefore supports claims about interaction design, perceived usability, procedural legitimacy, and adoption conditions in situated group use. It does not aim to provide population-level evidence about all board-game groups or a controlled comparison of every system component.

The future-work directions follow directly from these boundaries. Theme 6 in Chapter 5, adoption conditions, is especially important here because it shows that the remaining work is not simply interface polish. Participants' comments about legible controls, recoverable setup, AI role clarity, voice robustness, physical persistence, and prototype fidelity identify the conditions under which the framework could move from a study prototype into repeated use.

6.3.1 PickNPlay Evaluation Scope

The main evaluation limitation concerns the scope of inference. The Stage 2 PickNPlay evaluation involved four groups and 29 complete immediate SUS responses, with in-person and online use represented across the sessions. This gives useful cross-group evidence for the core system and is appropriate for a prototype-oriented HCI evaluation. However, the recruitment routes, group familiarity, and prior exposure to earlier project stages varied across groups. The results should therefore be read as situated evidence about how participants experienced PickNPlay in the study sessions, rather than as population-level claims about board-game players in general.

Stage 1 should be interpreted as formative context for the requirements and design problem, not as a controlled baseline condition. It helped identify why the earlier BGVS workflow needed stronger support for shared control, preference representation, explanation, and revision. Similarly, Boardot should be interpreted as an exploratory station-mediated extension over the same PickNPlay room workflow. Its evidence is useful for discussing shared attention and bounded facilitation, but the thesis does not depend on Boardot proving comparative effectiveness over PickNPlay.

This scope still supports the central thesis claim. PickNPlay was evaluated as an interaction framework for turning group preferences and constraints into an inspectable allocation process. The limitation is that the evidence supports bounded HCI design claims about use experience and procedural acceptability, rather than broad statistical generalisation or causal claims about individual features.

6.3.2 Measurement Limits

A second limitation concerns what the evaluation measures directly. SUS is useful because it gives a compact, comparable measure of perceived usability after system use. In this thesis, the Stage 2 and Stage 3 SUS results support the claim that PickNPlay and Boardot were generally experienced as usable in the study sessions. However, SUS does not directly measure whether participants understood the allocation mechanism, judged the process as fair, trusted the result, felt in control, or experienced stronger group cohesion.

Those constructs were examined mainly through interviews, observation, follow-up reflection, and thematic analysis. This is appropriate for the thesis because the research questions are about interaction experience and procedural acceptability, not only interface efficiency. At the same time, it means that claims about fairness, trust, control, and social cohesion should be read as qualitative HCI interpretations. The evidence shows how participants talked about and made sense of the systems, but it does not provide separate validated scales for each construct.

The study also cannot isolate the effect of each individual feature. Shared room state, candidate-pool construction, rating-to-tier input, result explanation, revision, and group discussion worked together as an interaction sequence. The evaluation therefore supports claims about the overall PickNPlay workflow and its design implications, rather than causal claims that one specific interface element independently produced trust, fairness, or social cohesion.

6.3.3 Implementation and Model Boundaries

PickNPlay currently represents board-game selection as a constrained allocation problem using submitted preferences, player-count constraints, teaching information, linked-player requests, and room-state revision. This was sufficient for evaluating the participant-directed framework, but the study also exposed places where the solver-facing model could be better translated into user-facing controls.

One boundary concerns how group-size constraints are expressed. The current interface exposes minimum and maximum player counts for candidate games, because these values map directly onto the resolver's feasibility constraints. However, observation suggested that some participants did not naturally want to control the outcome through minimum player counts. In practice, they often reasoned about the desired number of final game groups. For example, with nine participants, the group might want to choose between a 4/5 split and a 3/3/3 split. This is easier to understand as "how many games should we assign?" than as a set of minimum-player constraints on individual games. Future work should therefore explore higher-level allocation controls, such as specifying the desired number of selected games or preferred table-size patterns, while keeping the underlying solver constraints hidden unless users need to inspect them.

A second boundary is explanation depth. The current result view helps participants inspect the selected games and player assignments, and the workflow already allows participants to revise the room state and re-run resolution. However, the system does not yet fully explain alternative feasible allocations or the trade-offs that led to the selected result. For example, it does not show what would need to change for a different game-count pattern to be selected, which constraints were most binding, or how much preference compromise was involved for particular players. Future work should therefore extend explanation beyond final-result presentation toward trade-off explanation that helps groups decide whether to accept the result or revise it.

6.3.4 Privacy, Disclosure, and Anonymity

PickNPlay already separates private support data from room-specific allocation input. Reusable ratings, saved pools, and teaching information are stored as participant support data, while the allocation uses the vote and constraints submitted for the current room. Boardot also keeps generated-text features bounded: the AI layer receives visible room context and optional spoken questions, but it does not receive authority to modify preferences or choose assignments. These design choices reduce unnecessary disclosure and keep allocation authority separate from AI facilitation.

The limitation is that the study did not directly evaluate how much preference disclosure participants find acceptable during explanation and negotiation. This matters because stronger explanation can make an allocation more understandable while also making social pressure more visible. For example, a public explanation might reveal that one participant strongly disliked a game, that a teaching claim affected the result, or that a particular preference pattern blocked another allocation. In a co-located group, even de-identified explanations may be easy to infer from context.

Future work should therefore treat explanation and privacy as a joint design problem. One direction is to separate public group-level explanations from private participant-level explanations. The shared display or result screen could show aggregate reasons, selected and unselected games, binding player-count constraints, and possible revision paths without naming individual preference conflicts. Personal devices could show more detailed private explanations about how a participant's own preferences affected their assignment. More sensitive states, such as vetoes, teaching confidence, or unfamiliarity, should only be disclosed in ways that participants understand and consent to.

This also connects to anonymity in repeated or hybrid use. If PickNPlay is used across multiple sessions, or extended into Boardot, online board-game platforms, or AR/VR spaces, the system may accumulate more interaction traces, preference histories, camera inputs, voice events, and group context. Future versions should make it clear what is stored, what is temporary, what is visible to the group, and what is used only for computation. The goal is not full anonymity, which may be impossible in a co-located group, but calibrated disclosure: enough information for the group to trust and revise the allocation without unnecessarily exposing individual preferences or social positions.

6.3.5 Physical-Game Recognition and Metadata Capture

A further future-work direction is physical-game recognition and metadata capture. PickNPlay already supports several ways to construct the candidate pool, including search, manual entry, saved pools, and barcode-assisted lookup. This matches the requirement that candidate-game entry should be lightweight and shared rather than handled by one organiser. However, candidate-pool construction is still one of the places where setup effort can return, especially when a group has several physical games on the table and wants to start quickly.

Exploratory OCR work was implemented and tested during development, but it was not reliable enough to become part of the evaluated system. The main issue was that board-game boxes are visually inconsistent. Game names use stylised fonts, curved or decorative lettering, different languages, editions, expansions, glare, partial occlusion, and uneven camera angles. A local OCR model could sometimes detect useful text, but it also produced noisy or incomplete game names that required manual correction. For that reason, OCR should be treated as future work rather than as a Chapter 4 system contribution.

The ideal version of this feature would go beyond scanning one barcode at a time. With stronger computer-vision or multimodal AI models, a user could take a single image of several stacked or tabled game boxes, and the system could propose likely game matches. Those matches could then be checked against BoardGameGeek metadata, including game name, edition, player counts, and identifiers. The important interaction point is that recognition should suggest candidates, not silently add games to the public pool. Participants should confirm each recognised game before it becomes part of the shared room state.

Future work should therefore design this as a confidence-based and recoverable workflow. The system could show likely matches, confidence levels, ambiguous alternatives, and fallback search. It should cache confirmed matches where appropriate, preserve metadata provenance, and make it easy to correct wrong game names or editions before voting begins. Because this feature uses camera input and may rely on external recognition APIs, it also connects to privacy and disclosure. A practical design should make clear whether recognition is local or remote, what image data is stored or discarded, and when human confirmation is required.

6.3.6 Boardot as Persistent Room Infrastructure

Boardot should remain framed as an exploratory extension of PickNPlay, not as a second fully evaluated system. Its value in this thesis is that it shows how the same room state, preference-input pipeline, and resolver can be staged through a shared room-scale station. The strongest finding was about collective attention: the station made the voting and result process more visible as a group event while preserving private preference input on phones.

The next step for Boardot is to study it as persistent room infrastructure rather than as a session prototype. In the current study, the station had to be assembled, explained, and operated in a research setting. This makes sense for an exploratory prototype, but it does not show whether the station would become useful as part of an ordinary game-night environment. A persistent setup could change the experience substantially. If the display, camera, markers, microphone, speaker, and joining flow are already part of the table, the station may reduce setup burden and support shared attention more naturally.

Future work should therefore examine the physical and social durability of the station. This includes marker durability, table placement, screen visibility, audio placement, setup time across repeated sessions, and whether participants treat the station as shared infrastructure or as another device that someone has to manage. It should also test whether the station remains helpful once the novelty of the prototype has worn off. The relevant question is whether Boardot can become part of the group's ordinary coordination environment, not whether it produces a different allocation from PickNPlay.

Future work should also refine the existing station-side voice and facilitation layer under realistic room conditions. Boardot already supports listening and phase-bounded flow-control commands, but noisy co-located settings introduce turn-taking, source ambiguity, and robustness problems. Persistent deployment should therefore test whether visible listening state, clearer confirmation, or partly phone-mediated speaking and listening controls make the existing bounded facilitation easier to use.

6.3.7 Beyond Physical Board-Game Sessions

Although the thesis is grounded in co-located board-game selection, the interaction model is not limited to physical game nights. PickNPlay's core structure is a shared room workflow: participants construct a candidate pool, submit private preferences, see group progress, receive a feasible allocation, and revise when necessary. This structure could be adapted to online board-game platforms where the candidate pool is digital, participants are remote, and the allocation result might assign people to different online tables, rooms, or game instances.

This direction should be treated carefully because online play changes the social setting. In a physical session, the group can rely on table talk, visible bodies, and shared objects to negotiate meaning around the allocation. In an online session, the system may need to provide more explicit awareness cues, explanation, and coordination support because participants cannot rely on the same physical context. Future work could therefore examine how PickNPlay's shared room state works when embedded into online board-game communities or platforms, especially for larger groups that need to split across multiple games.

The same concept could also extend to AR or VR play spaces. In those settings, the shared allocation state could become spatial rather than screen-bound: candidate games, participant markers, table groups, and allocation explanations could be placed in a virtual or augmented room. This would extend the Boardot idea beyond a physical station. Instead of one shared display, the group could interact with a shared spatial representation of the decision process. The key design principle should remain the same: private preference expression should stay under participant control, while the shared environment supports group awareness, explanation, and revision.

These extensions are future work because they would introduce new interaction problems that the current thesis did not evaluate, including platform integration, identity and anonymity across remote groups, spatial interface design, and new forms of social presence. They are still relevant because they show that the thesis contribution is not only a board-game prototype. It is a model for combining private preference input, shared group state, solver-backed allocation, and explainable revision in settings where groups need to divide themselves across constrained activities.

Appendix A

Study Documentation and Instruments

This appendix records participant-facing study documents and blank study instruments for the PickNPlay and Boardot evaluation. It includes copies of the participant-facing templates, and it does not include raw participant responses, signed consent records, audio recordings, transcripts, or figures already presented in the thesis chapters.

A.1 Ethics and Project Records

The study was conducted under ANU human research ethics protocol H/2025/0501. The project records listed in Table A.1 define the approved study scope and participant-facing consent process.

Table A.1: Ethics and study governance records.

Document	Appendix use
Human Research Ethics Application	Ethics protocol record for H/2025/0501.
Project Description	Approved project description.
Participant Information Sheet	Participant-facing study information.
Consent Form	Participant consent template.

A.1.1 Participant Information Sheet

1 Researcher

Hello! I am Kevin Zhu, a Master of Computing (Advanced) student currently in my final year undertaking a research project under the supervision of Dr Michael Norrish and Dr Peter Höfner at the Australian National University (ANU), School of Computing. I invite you to take part in a study exploring an innovative web interface for a Social Choice Board Game App.

2 Project Title:

UX Design, Implementation and Evaluation for a Social Choice Board Game App

3 Outline of the Project:

The idea of this project is to explore how an interface can make social choice interactions easier for groups selecting board games. We will do this by providing you with a mobile app that implements social choice algorithms to help groups collectively decide which board game to play. You will input your preferences for different board games, and the system will use these preferences along with other group members' choices to suggest an optimal game selection. By observing your experiences with the app, we hope to discover which parts of the interface are intuitive and where there may be confusion, so that we can develop a more accessible, user-friendly tool for group decision-making.

Description and Methodology: This research follows an iterative design process with multiple study sessions throughout the development of the PickNPlay and Boardot app prototype. You may be invited to participate in different stages of this research:

Stage 1 (Early Research): Initial interviews to understand current weaknesses and requirements for board game selection tools.

Stage 2 (First Evaluation Cycle): Testing with the initial prototype where you will interact with our Social Choice Board Game App PickNPlay to help your group select board games for a game night. You will input your preferences for various board games, and the app will use social choice algorithms to suggest optimal games based on the group's collective preferences.

Stage 3 (Second Evaluation Cycle): Testing with another system Boardot, focusing on comparative evaluation of usability improvements.

In each stage, we will observe your interactions with the interfaces, note any questions or issues that come up, and gather your feedback on the app's design and usability. After using the app, we will conduct a short interview with you and ask you to fill out a brief survey. This combination of direct interaction, interview, and survey will help me understand both how you use the app in real time and how you feel about it afterwards.

Participants: We plan to recruit participants from board game clubs and online board game forums who are 18 years or older. Participants should have experience playing board games and be comfortable using mobile applications. We will recruit both for in-person study sessions and for online participation where players can use the app with their own groups and provide feedback remotely. During the study, participants will act as they normally would when selecting board games for their group. No specialised programming, AI, or social choice theory knowledge is required.

Use of Data and Feedback: During the session, we will collect your observations and experiences, which will be analysed as part of my academic research. This will include recordings of your interactions with the app, notes from our interview, and your survey responses. We will keep your personal details private, and none of your identifying information will appear in any publications resulting from this research. The findings may be included in my thesis or published in academic articles. At the end of the project, we are happy to provide my thesis to you.

4 Participant Involvement:

Voluntary Participation & Withdrawal: Your participation in this project is voluntary, and you may decline to take part or withdraw from this research without providing a reason within two weeks after participation. You can refuse to answer any interview or survey questions, and you are in charge of your own interactions with the board game app. If you withdraw from the project, by default, your data will be destroyed and will not be used.

What does participation in the research entail? Depending on which stage(s) you participate in, here is what you will do:

Stage 1 (Early Research):

1. Participate in an interview about your current board game selection process and any challenges you face (10–20 minutes).

Stage 2 & 3 (Evaluation Cycles):

1. Take part in a brief tutorial (about 5 minutes) to learn how to use the Social Choice Board Game App interface.
2. Participate in a group session where you will input your board game preferences and interact with the app to help select games for your group (10–15 minutes). We will take notes of your interactions with the app to understand how you use the interface.
3. Complete a short survey and participate in a semi-structured interview (15–20 minutes). We will ask you about your experience with the app, what you liked or disliked, and any suggestions you have for improvement.

You may be invited to participate in multiple stages if you are interested, but participation in any single stage is also valuable. Overall, each evaluation session should last approximately 20 minutes interacting with the app and 20 minutes for the interview. We will use voice memo recordings of the interviews to help us accurately capture your feedback, and these recordings will be securely stored and used only for research purposes, but not published. We will ask for your consent to include quotes from your interview in any publications resulting from this research, and we will only use your responses if you agree to it.

Location and Duration: The study will take place where you normally play board games with your group. You will use your own mobile devices with your existing gaming group. You only need to participate once, and the session will last approximately 40 minutes total (20 minutes app interaction, 20 minutes interview).

Risks: There are generally few risks or hazards involved in this project. The main risks are minimal and relate to technology use and group interaction. You may experience minor frustration if the app interface is confusing or if there are technical issues during the session. As this project will have participants there is a risk that you might be identified as a participant from the resulting publications. We will attempt to mitigate this risk by de-identifying any reproduction of statements made in interviews that you give us permission to include in research publications. This research is not intended to cover sensitive or uncomfortable topics, so you may refuse to answer any interview questions that you wish and may stop participation at any time if you are uncomfortable.

Benefits: While there is no direct personal benefit to participating, you may enjoy exploring a new tool for group decision-making with your gaming group. Your feedback will help us and other researchers design more user-friendly, accessible applications for group decision-making in gaming contexts. This could benefit future board game enthusiasts and groups who use similar tools for selecting games and making collective decisions.

5 Confidentiality:

During the project, only my supervisors (Dr Michael Norrish and Dr Peter Höfner) and I will have access to the data you provide, including questionnaires, interview recordings, app interaction recordings, and transcripts. We will store these materials securely both during and after the project. Where quotes or images are included in any publications resulting from this research, we will de-identify that information and attribute it to a pseudonym (e.g., "P1"), unless you give us explicit permission to use your name. We will take every reasonable precaution to protect your confidentiality as allowed by law.

Because there will be few participants in this study, it may be possible to re-identify you through specific details in your interview responses or app usage patterns. To help safeguard your privacy, we ask that you refrain from sharing sensitive personal information during your interviews.

Participant Responsibility for Confidentiality: Please note that by agreeing to participate in this study, you also agree to maintain confidentiality regarding your participation and that of others. In particular, you should not disclose or discuss any individually identifiable or sensitive information about other participants or the study sessions outside the research context.

6 Privacy Notice:

In collecting your personal information within this research, the ANU must comply with the Privacy Act 1988. The ANU Privacy Policy is available at `https://policies.anu.edu.au/ppl/document/ANUP_010007` and it contains information about how a person can:

- Access or seek correction to their personal information;
- Complain about a breach of an Australian Privacy Principle by ANU, and how ANU will handle the complaint.

7 Data Storage:

Where: Data will be stored securely on ANU Microsoft OneDrive accessible to Kevin, Michael and Peter and downloaded onto encrypted and password-protected computers for analysis.

How long: Data will be stored for a period of at least five years from the date of any publication arising from the research.

Handling of data following the required storage period: Data will be destroyed following the storage period.

8 Queries and Concerns:

Contact Details for More Information: Kevin Zhu by email: `kevin.zhu@anu.edu.au` or Dr Michael Norrish by email: `Michael.Norrish@anu.edu.au` or Dr. Peter Höfner by email: `Peter.Hoefner@anu.edu.au` or phone: (61) 478 007 381

9 Ethics Committee Clearance:

The ethical aspects of this research have been approved by the ANU Human Research Ethics Committee (Protocol H/2025/0501). If you have any concerns or complaints about how this research has been conducted, please contact:

Ethics Manager
The ANU Human Research Ethics Committee
The Australian National University
Telephone: +61 2 6125 3427
Email: Human.Ethics.Officer@anu.edu.au

A.1.2 Consent Form Template

1 Consent for Participation

I have read and understood the Information Sheet you have given me about the research project, and I have had any questions and concerns about the project listed here:

____

____

addressed to my satisfaction.

I agree to participate in the project.	YES ☐	NO ☐
I agree to interviews and interactions with the application being audio recorded.	YES ☐	NO ☐
I agree to being identified by pseudonym (e.g., P1) in research outputs.	YES ☐	NO ☐
I agree to maintain confidentiality and not disclose any individually identifiable or sensitive information about other participants or the study sessions.	YES ☐	NO ☐

Signature: .... Date: ....

A.2 Recruitment and Study Instruments

Table A.2 lists the recruitment and study instruments retained for auditability.

Table A.2: Recruitment and study instruments.

Instrument	Intended use
Recruitment poster	Poster used to advertise the study.
Early-stage questionnaire	Preliminary requirement-scoping questionnaire.
Study session script	Session procedure and researcher prompts.
Interview guide	Semi-structured interview prompts.
System Usability Scale questionnaire	Post-session usability questionnaire template.
Follow-up questionnaire	Delayed reflection questionnaire template.

A.2.1 Interview Guide

1 Stage 1: Initial Interview Questions (Early Research)

These interviews will be conducted after participants use the initial PickNPlay app prototype (Boardroom) to gather early feedback and identify areas for improvement:

- How does your group typically select which board game to play during game nights?
- What were your first impressions of the PickNPlay app?
- How did using the app compare to your usual board game selection process?
- What challenges or frustrations did you encounter while using the app?
- What did you like most about the app?
- What did you like least about the app?
- Was there any time that you were concerned about the fairness and validity of the voting result of the system?
- Were there any features you expected to see that were missing from the app?
- Were there any features that seemed unnecessary or confusing?
- How easy or difficult was it to learn to use the app?
- Would you consider using this app with your gaming group in the future?
- What improvements would you suggest for the app?
- What factors are most important to you when selecting a board game (e.g., playtime, complexity, number of players, theme)?
- How well did the app address these important factors?

2 Stage 2 Interview Questions

These semi-structured questions will be asked after participants use the system in a board game selection task. The wording should refer to the specific system used in that session.

- How does your group typically select which board game to play during game nights? (Ignore if not the first time in the group)
- What were your first impressions of the system?
- How did using the system compare to your usual/previous board game selection process?
- How easy or difficult was it to learn to use the system?
- What challenges or frustrations did you encounter while using the system?
- What did you like most about the system?
- What did you like least about the system?
- Were there any features you expected to see that were missing?
- Were there any features that seemed unnecessary or confusing?

- Was there any time you were concerned about the fairness, validity, or transparency of the result?
- Did you trust the recommendation or outcome provided by the system? Why or why not?
- How did your group react to using the system?
- Would you consider using a system like this with your gaming group in the future? Why or why not?
- What factors are most important to you when selecting a board game (e.g., playtime, complexity, number of players, theme)?
- How well did the system support those factors?
- Is there anything else you would like to share about your experience with the system?

2.1 Stage 2 Longitudinal Follow-Up Questions

These questions will be asked several weeks after Stage 2 with participants who already used PickNPlay. The aim is not to repeat the immediate usability interview, but to understand delayed reflection, changed understanding, trust, remembered usefulness, and possible future use.

Changed understanding and memory

- What parts of PickNPlay do you remember most clearly now?
- If you had to explain PickNPlay to someone who had not used it, how would you describe what it does?
- Has your understanding of how PickNPlay produced the final result changed since the Stage 2 session? If so, how?

Remembered usefulness and limitations

- Looking back, do PickNPlay or any parts of it seem genuinely useful for a real board game night?
- Which problems or confusing parts from Stage 2 still seem important now?
- Were there any problems from Stage 2 that now seem less important than they felt at the time?

Trust, fairness, and transparency

- How much would you trust PickNPlay's allocation result now? What would affect that trust?
- What information would help you better understand why PickNPlay selected one allocation rather than another?
- In what situation would you override or ignore the result suggested by PickNPlay?

Adoption and group practice

- If your group used PickNPlay repeatedly, do you think the system supports enough to speed up the learning and decision process?
- Do you think after using PickNPlay, the group and individuals can be more confident to express their preferences? (Compare to open discussion)

Design implications

- What one change would make PickNPlay more suitable for long-term or repeated use?
- Is there anything else you have thought about since the Stage 2 session that you did not mention at the time?

Stage 3 Boardot Short Interview

For returning participants who have already experienced BGVS and PickNPlay, the facilitator may use the following shortened set of prompts instead of asking all Section 2 and Section 3 questions. These prompts preserve coverage of usability, preference expression, group discussion, trust, fairness, transparency, control, and Boardot-specific AI-host mediation.

- Compared with BGVS and PickNPlay, what felt most different about using Boardot today?
- What parts of Boardot made the task easier or harder to learn and use?
- How well did Boardot help you express your preferences for the games?
- How did the shared station affect your group discussion or coordination?
- Did the table markers and voting-progress display help you understand what was happening? Why or why not?
- Did you trust the final assignment result? What made it feel fair, unfair, clear, or unclear?
- How did Boardot's host prompts or guidance affect the group? Did you follow, ignore, or debate them?
- Compared with PickNPlay, did Boardot make you feel more or less in control of the decision?
- If your group could choose BGVS, PickNPlay, Boardot, or no system, which would you use and in what situation?
- What one change would most improve Boardot?

Optional probes

- Was anything missing, unnecessary, or confusing?
- Did any part of the result need more explanation?
- Has your group's normal board-game selection process changed since the previous sessions?

3 Cross-System Comparison Questions (If Applicable)

These questions will be asked only if the participant has used another system earlier in the study (e.g., they have experience with Boardroom and are now using PickNPlay or Boardot).

- Compared to the other system(s) you used, what felt most different about this system?
- Which system better supported your group discussion? Why?
- Which system made it easier to express your preferences accurately? Why?
- Which system felt more fair or trustworthy? What made you feel that way?
- Which system gave you more control over the final outcome? Why?
- If your group could choose only one system to use in the future, which would you choose and why?

4 Survey Questions

We will administer the System Usability Scale (SUS), which is a standard written questionnaire of usability. There are ten questions with 5-point Likert responses from "Strongly Disagree" to "Strongly Agree". For the longitudinal follow-up, SUS will be reused with the same wording and treated as a delayed perceived-usability measure, not as a separate custom survey.

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

Additional follow-up reflection items, separate from SUS These items are not part of the System Usability Scale (SUS) and will not be included in the SUS score. They are used only to capture delayed reflection after Stage 2.

Participants will answer each item on the same 5-point scale: 1 = Strongly Disagree, 5 = Strongly Agree.

1. I now understand how PickNPlay uses group preferences to produce an allocation.
2. I would trust PickNPlay more if it showed why one allocation was chosen over alternatives.
3. PickNPlay would be useful when the group is large or has many possible games.
4. I could imagine using PickNPlay without the researcher present.
5. Some parts of PickNPlay still feel unclear when I think back to the study.

A.2.2 System Usability Scale Questionnaire

System Usability Scale (SUS)
PickNPlay Social Choice Board Game App

Participant ID: ____________________

Date: ____________________

Study Stage: ____________________

	Strongly Disagree				Strongly Agree
1. I think that I would like to use the PickNPlay app frequently	1	2	3	4	5
2. I found the PickNPlay app unnecessarily complex	1	2	3	4	5
3. I thought the PickNPlay app was easy to use	1	2	3	4	5
4. I think that I would need the support of a technical person to be able to use the PickNPlay app	1	2	3	4	5
5. I found the various functions in the PickNPlay app were well integrated	1	2	3	4	5
6. I thought there was too much inconsistency in the PickNPlay app	1	2	3	4	5
7. I would imagine that most people would learn to use the PickNPlay app very quickly	1	2	3	4	5
8. I found the PickNPlay app very cumbersome to use	1	2	3	4	5
9. I felt very confident using the PickNPlay app	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with the PickNPlay app	1	2	3	4	5

A.2.3 Follow-Up Questionnaire

System Usability Scale (SUS)

PickNPlay Social Choice Board Game App

Longitudinal Follow-Up / Delayed Reflection

Participant ID: ______________________

Date: ______________________

Study Stage: ______________________

Please answer the following questions based on your remembered experience using the PickNPlay app in the Stage 2 session, and any reflection since then. The questions below are the standard SUS items and should be answered using the 1–5 scale shown.

	Strongly Disagree				**Strongly Agree**
1. I think that I would like to use the PickNPlay app frequently	1	2	3	4	5
2. I found the PickNPlay app unnecessarily complex	1	2	3	4	5
3. I thought the PickNPlay app was easy to use	1	2	3	4	5
4. I think that I would need the support of a technical person to be able to use the PickNPlay app	1	2	3	4	5
5. I found the various functions in the PickNPlay app were well integrated	1	2	3	4	5
6. I thought there was too much inconsistency in the PickNPlay app	1	2	3	4	5
7. I would imagine that most people would learn to use the PickNPlay app very quickly	1	2	3	4	5
8. I found the PickNPlay app very cumbersome to use	1	2	3	4	5
9. I felt very confident using the PickNPlay app	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with the PickNPlay app	1	2	3	4	5

Additional follow-up reflection items, separate from SUS These items are not part of the System Usability Scale (SUS) and will not be included in the SUS score. They are used only to capture delayed reflection after Stage 2.

	Strongly Disagree				Strongly Agree
1. I now understand how PickNPlay uses group preferences to produce an allocation.	1	2	3	4	5
2. I would trust PickNPlay more if it showed why one allocation was chosen over alternatives.	1	2	3	4	5
3. PickNPlay would be useful when the group is large or has many possible games.	1	2	3	4	5
4. I could imagine using PickNPlay without the researcher present	1	2	3	4	5
5. Some parts of PickNPlay still feel unclear when I think back to the study.	1	2	3	4	5

A.3 Study Corpus Boundary

The analysis used formative Stage 1 materials, Stage 2 PickNPlay evaluation materials, Stage 3 Boardot evaluation materials, and cleaned thematic-analysis derivatives. Table A.3 records the corpus boundary without exposing participant-level files.

Table A.3: Corpus-level study materials used for analysis.

Corpus area	Appendix treatment
Preliminary questionnaire	Formative requirement-scoping evidence.
Stage 1 baseline evidence	Formative/background context for the earlier BGVS baseline.
Stage 2 PickNPlay evaluation	SUS and thematic-analysis source corpus; participant-level files excluded.
Stage 3 Boardot evaluation	SUS and thematic-analysis source corpus; participant-level files excluded.
Cleaned analysis derivatives	Derived transcript material with de-identified speaker labels.

Appendix B

Analysis and Implementation References

This appendix records analysis outputs, prototype links, and implementation references that support the thesis. It deliberately avoids repeating figures already included in the main chapters.

B.1 Quantitative Analysis Artefacts

The SUS and follow-up questionnaire analyses were generated from de-identified markdown inputs. Table B.1 lists the audit artefacts used to support Chapter 5.

Table B.1: SUS and follow-up analysis artefacts.

Artefact	Appendix use
SUS analysis workflow note	Analysis workflow, expected input schema, output descriptions, and small-sample caveats.
Stage 2 PickNPlay SUS outputs	Stage 2 PickNPlay descriptive statistics and plots.
Stage 3 Boardot SUS outputs	Stage 3 Boardot descriptive statistics and plots.
Stage 2 follow-up outputs	Stage 2 follow-up Likert questionnaire summaries.
Stage 2 and Stage 3 comparison outputs	Exploratory Mann-Whitney U comparison output.
SUS analysis script	Script used to normalise rows, compute statistics, create plots, and run optional Mann-Whitney U tests.

B.1.1 Participant-Level and Follow-Up Tables

The following tables retain participant-level and item-level detail moved out of Chapter 5. The main chapter reports the descriptive pattern; these tables support auditability.

Table B.2: Stage 2 PickNPlay participant SUS scores.

Group 1	SUS	Group 2	SUS	Group 3	SUS	Group 4	SUS
G1-P1	77.50	G2-P1	87.50	G3-P1	72.50	G4-P1	87.50
G1-P2	85.00	G2-P2	82.50	G3-P2	87.50	G4-P2	65.00
G1-P3	57.50	G2-P3	97.50	G3-P3	85.00	G4-P3	87.50
G1-P4	75.00	G2-P4	87.50	G3-P4	82.50	G4-P4	72.50
G1-P5	82.50	G2-P5	87.50	G3-P5	67.50	G4-P5	77.50
G1-P6	75.00	G2-P6	72.50	–	–	G4-P6	82.50
G1-P7	87.50	G2-P7	87.50	–	–	G4-P7	92.50
G1-P8	72.50	–	–	–	–	G4-P8	52.50
–	–	–	–	–	–	G4-P9	67.50

Table B.3: Stage 2 Group 1 follow-up reflection responses.

Follow-up reflection item	N	Mean	Median	SD
Understood how PickNPlay used group preferences to produce an allocation	6	3.17	3.00	0.75
Would trust PickNPlay more if it showed why one allocation was chosen over alternatives	6	3.83	4.00	0.75
Would be useful when the group is large or has many possible games	6	4.67	5.00	0.52
Could imagine using PickNPlay without the researcher present	6	4.67	5.00	0.82
Some parts still felt unclear when thinking back to the study	6	2.17	2.00	1.17

Table B.4: Stage 3 Boardot participant SUS scores.

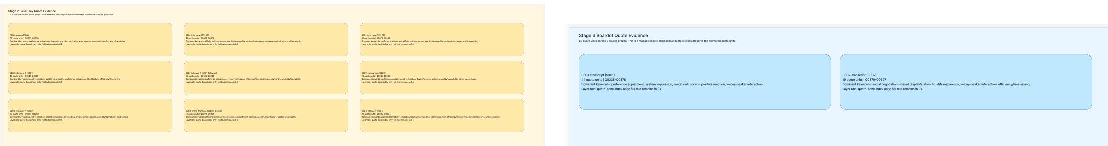
Group 1	SUS	Group 2	SUS
G1-P1	80.00	G2-P1	75.00
G1-P2	80.00	G2-P2	80.00
G1-P3	92.50	G2-P3	72.50
G1-P4	90.00	G2-P4	82.50
–	–	G2-P5	67.50
–	–	G2-P6	80.00
–	–	G2-P7	67.50
–	–	G2-P8	75.00
–	–	G2-P9	92.50

B.2 Thematic Analysis Artefacts

The thematic analysis was conducted through a staged artefact chain from corpus inventory to conceptual synthesis. The working quote bank is retained as a traceability source, but is not reproduced in full because it contains extensive participant speech.

Table B.5: Thematic-analysis artefact chain.

Artefact	Appendix use
Corpus inventory	Corpus boundary, stage separation, and source inventory.
Ethics and data-use notes	Ethics and data-use constraints applied during analysis.
Transcript status record	Derived transcript preparation status.
Analysis protocol	Coding protocol, label schema, and quality-audit checklist.
Full quote bank	Internal traceability source for quote units; not reproduced in full.
Keyword map	Quote-to-keyword mapping.
Initial codebook	Initial coding artefact.
Refined codebook	Refined codebook and consolidation decisions.
Sub-theme map	Sub-theme mapping.
Theme map	Theme mapping used to support Chapter 5.
FigJam board map	FigJam board map and verification record.
Conceptual synthesis	Conceptual synthesis used to organise the final qualitative findings.



(a) Stage 2 PickNPlay. (b) Stage 3 Boardot.

Figure B.1: Quote-extraction audit boards.

B.3 Prototype and System Links

B.4 Source Code and Deployment References

Table B.7 lists the main implementation components corresponding to the system design described in Chapter 4.

(a) Stage 2 PickNPlay. (b) Stage 3 Boardot.

Figure B.2: Keyword-selection audit boards.

(a) Stage 2 PickNPlay. (b) Stage 3 Boardot.

Figure B.3: Initial-code audit boards.

Table B.6: Prototype, design, and deployment links.

Artefact	Reference
PickNPlay live deployment	PickNPlay deployment
PickNPlay source repository	GitHub repository
Interactive PickNPlay wireframe	interactive wireframe
Thematic analysis FigJam board	FigJam board

Table B.7: Implementation source-code references.

Component	Role in the implemented system
Project overview and setup notes	Project overview, setup, verification commands, and deployment reference.
Frontend build configuration	Frontend dependencies, build scripts, and test scripts.
PickNPlay room workflow	Phone and room workflow: room creation, lobby, voting, waiting, and results.
Boardot station workflow	Station interface, shared display, speech, marker, and station-result behaviour.
Resolution integration module	Client-side resolution data structures and resolver integration support.
Preference-tier module	Tiered preference representation used by the voting interface.
Firebase integration module	Firebase room-state, persistence, and backend integration functions.
Cloud Functions backend	Python Firebase Cloud Functions, including resolver and generated-dialogue endpoints.
Firebase deployment configuration	Firebase hosting, emulator, and function configuration.
Netlify routing configuration	Netlify routing support for the single-page app.

Bibliography

Abdul, A.; Vermeulen, J.; Wang, D.; Lim, B. Y.; and Kankanhalli, M., 2018. Trends and trajectories for explainable, accountable and intelligible systems: An HCI research agenda. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–18. Association for Computing Machinery, New York, NY, USA. doi:10.1145/3173574.3174156. https://doi.org/10.1145/3173574.3174156. [Cited on pages 16, 55, and 58.]

Abdulkadiroğlu, A. and Sönmez, T., 1998. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, 66, 3 (1998), 689–702. doi:10.2307/2998580. [Cited on page 8.]

Adida, B., 2008. Helios: Web-based open-audit voting. In *Proceedings of the 17th USENIX Security Symposium*, 335–348. USENIX Association, San Jose, CA. https://www.usenix.org/conference/17th-usenix-security-symposium/helios-web-based-open-audit-voting. [Cited on page 10.]

Ágoston, K. C.; Biró, P.; and McBride, I., 2016. Integer programming methods for special college admissions problems. *Journal of Combinatorial Optimization*, 32, 4 (2016), 1371–1399. doi:10.1007/s10878-016-0085-x. https://doi.org/10.1007/s10878-016-0085-x. [Cited on page 8.]

Arrow, K. J., 1963. *Social Choice and Individual Values*. Yale University Press, New Haven, CT, 2 edn. [Cited on page 5.]

Barrett, C.; Sebastiani, R.; Seshia, S. A.; and Tinelli, C., 2021. Satisfiability modulo theories. In *Handbook of Satisfiability* (Eds. A. Biere; M. Heule; H. van Maaren; and T. Walsh), vol. 336 of *Frontiers in Artificial Intelligence and Applications*, 1267–1329. IOS Press, Amsterdam, 2 edn. doi:10.3233/FAIA201017. https://doi.org/10.3233/FAIA201017. [Cited on pages 8 and 52.]

Biró, P.; Fleiner, T.; Irving, R. W.; and Manlove, D. F., 2010. The college admissions problem with lower and common quotas. *Theoretical Computer Science*, 411, 34–36 (2010), 3136–3153. doi:10.1016/j.tcs.2010.05.005. [Cited on page 7.]

Bjørner, N.; Phan, A.-D.; and Fleckenstein, L., 2015. νZ – an optimizing SMT solver. In *Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2015)*, vol. 9035 of *Lecture Notes in Computer Science*, 194–199. Springer, Berlin. doi: 10.1007/978-3-662-46681-0_14. https://doi.org/10.1007/978-3-662-46681-0_14. [Cited on pages 8 and 52.]

Board Game Oracle, 2026. Game to pick a game: The gateway edition. https://www.boardgameoracle.com/boardgame/price/E0OJOBqrzt/game-to-pick-a-game-the-gateway-edition. Listing for a 2018 Chip Theory Games title about selecting a game to play, designed by Adam Carlson and Josh J. Carlson. Accessed 2026-04-28. [Cited on page 21.]

Board Game Pick, 2026. Board game pick. https://www.boardgamepick.com/. Random board game selector that reads a BoardGameGeek collection and proposes games using criteria such as player count and playing time. Accessed 2026-04-28. [Cited on pages 2, 20, and 21.]

BoardGameGeek, 2026. Collection. https://boardgamegeek.com/wiki/page/Collection. BoardGameGeek wiki page explaining that a user's collection can include games beyond those physically owned, such as played, rated, commented, or tracked titles. Accessed 2026-04-28. [Cited on page 20.]

BoardGameGeek forum community, 2026. How to select games to play in a large group? https://boardgamegeek.com/thread/3662600/how-to-select-games-to-play-in-a-large-group. Public recommendations forum thread. Accessed 2026-05-18. [Cited on pages 1 and 31.]

BoardGameGeek wiki contributors, n.d.a. Frequent suggestions. https://boardgamegeek.com/wiki/page/Frequent_Suggestions. BoardGameGeek wiki page, including discussion of large-player-count recommendation requests. Accessed 2026-05-18. [Cited on page 31.]

BoardGameGeek wiki contributors, n.d.b. How to host a game day. https://boardgamegeek.com/wiki/page/How_To_Host_a_Game_Day. BoardGameGeek wiki page. Accessed 2026-05-18. [Cited on page 31.]

Boneff-Peng, R., 2023. Board game chooser. https://boardgamechooser.com/. BoardGameGeek collection tool for combining collections and filtering games by category, status, recommended player counts, and related criteria. Accessed 2026-04-28. [Cited on page 20.]

Bouveret, S.; Chevaleyre, Y.; and Maudet, N., 2016. Fair allocation of indivisible goods. In *Handbook of Computational Social Choice* (Eds. F. Brandt; V. Conitzer; U. Endriss; J. Lang; and A. D. Procaccia). Cambridge University Press, Cambridge. doi:10.1017/CBO9781107446984.013. https://doi.org/10.1017/CBO9781107446984.013. [Cited on page 6.]

Brandt, F.; Conitzer, V.; Endriss, U.; Lang, J.; and Procaccia, A. D. (Eds.), 2016. *Handbook of Computational Social Choice*. Cambridge University Press. ISBN 9781107060432. doi:10.1017/CBO9781107446984. https://www.cambridge.org/core/books/handbook-of-computational-social-choice/8AF63E87F76A5FC974D5E73536C52BD6. [Cited on pages 1, 6, and 22.]

Braun, V. and Clarke, V., 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 2 (2006), 77–101. doi:10.1191/1478088706qp063oa. https://doi.org/10.1191/1478088706qp063oa. [Cited on pages 13 and 72.]

Braun, V. and Clarke, V., 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11, 4 (2019), 589–597. doi:10.1080/2159676X.2019.1628806. https://doi.org/10.1080/2159676X.2019.1628806. [Cited on pages 13 and 72.]

Brooke, J., 1996. SUS: A quick and dirty usability scale. In *Usability Evaluation in Industry* (Eds. P. W. Jordan; B. Thomas; B. A. Weerdmeester; and I. L. McClelland), 189–194. Taylor & Francis, London. doi:10.1201/9781498710411-35. https://doi.org/10.1201/9781498710411-35. [Cited on pages 13 and 66.]

Budish, E.; Cachon, G. P.; Kessler, J. B.; and Othman, A., 2017. Course match: A large-scale implementation of approximate competitive equilibrium from equal incomes for combinatorial allocation. *Operations Research*, 65, 2 (2017), 314–336. doi:10.1287/opre.2016.1544. https://doi.org/10.1287/opre.2016.1544. [Cited on page 9.]

Caine, K., 2016. Local standards for sample size at CHI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 981–992. Association for Computing Machinery, New York, NY, USA. doi:10.1145/2858036.2858498. https://doi.org/10.1145/2858036.2858498. [Cited on page 62.]

Carroll, J. M., 2000. Making use: scenarios and scenario-based design. In *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, DIS '00 (New York City, New York, USA, 2000), 4. Association for Computing Machinery, New York, NY, USA. doi:10.1145/347642.347652. https://doi.org/10.1145/347642.347652. [Cited on page 12.]

Carroll, J. M., 2003. *HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. ISBN 9780080491417. [Cited on page 11.]

Chevaleyre, Y.; Dunne, P. E.; Endriss, U.; Lang, J.; Lemaître, M.; Maudet, N.; Padget, J.; Phelps, S.; Rodríguez-Aguilar, J. A.; and Sousa, P., 2006. Issues in multiagent resource allocation. *Informatica*, 30, 1 (2006), 3–31. [Cited on page 6.]

Cohn, M., 2004. *User Stories Applied: For Agile Software Development*. Addison-Wesley Professional, Boston. ISBN 9780321205681. [Cited on page 12.]

Cooper, A., 1999. *The Inmates Are Running the Asylum: Why High-Tech Products Drive Us Crazy and How to Restore the Sanity*. Sams, Indianapolis, IN. ISBN 9780672316494. [Cited on page 12.]

de Moura, L. and Bjørner, N., 2008. Z3: An efficient SMT solver. In *Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2008)*, vol. 4963 of *Lecture Notes in Computer Science*, 337–340. Springer, Berlin. doi:10.1007/978-3-540-78800-3_24. https://doi.org/10.1007/978-3-540-78800-3_24. [Cited on pages 8 and 52.]

DeSanctis, G. and Gallupe, R. B., 1987. A foundation for the study of group decision support systems. *Management Science*, 33, 5 (1987), 589–609. doi:10.1287/mnsc.33.5.589. https://doi.org/10.1287/mnsc.33.5.589. [Cited on pages 14 and 15.]

Design Council, n.d. History of the double diamond. https://www.designcouncil.org.uk/our-resources/the-double-diamond/history-of-the-double-diamond/. Accessed 4 May 2026. [Cited on page 12.]

Desmet, P. M. A. and Hekkert, P., 2007. Framework of product experience. *International Journal of Design*, 1, 1 (04 2007), 57–66. [Cited on page 11.]

DiceDecider, 2026. Dicedecider. https://dicedecider.com/. Board game recommender and picker using criteria such as player count, play time, complexity, category, and BoardGameGeek collection data. Accessed 2026-04-28. [Cited on page 20.]

Dix, A.; Finlay, J. E.; Abowd, G. D.; and Beale, R., 2004. *Human-Computer Interaction*. Pearson Prentice Hall, 3rd edn. [Cited on page 11.]

Elections ACT, 2024. Electronic voting and counting. https://www.elections.act.gov.au/elections/our-electoral-system/elections-in-the-act/technology-assisted-voting-and-counting/electronic-voting. Accessed 3 May 2026. [Cited on page 10.]

Expand Your Game, 2019. The game to pick a game: Review. https://expandyourgame.blogspot.com/2019/11/game-to-pick-game-review.html. Review with photographs of the game's nomination, boost, and bidding components. Accessed 2026-05-17. [Cited on page 21.]

Furlong, M., 2016. Board game selection tool. https://github.com/mfurlong64/board_game_selector. Web app that filters a BGG collection by number of players and playing time categories, and can select a random game from the filtered list. [Cited on page 20.]

Game Night Picks, 2026. Game night picks. https://gamenightpicks.com/. Board game night website for game-night setup, invitations, voting, game suggestions, player discovery, and game libraries. Accessed 2026-05-28. [Cited on page 20.]

GameShelf.io, 2026. Gameshelf.io. https://gameshelf.io/. BoardGameGeek collection viewer and filtering tool. Accessed 2026-05-28. [Cited on page 20.]

Goldman, J. and Procaccia, A. D., 2015. Spliddit: Unleashing fair division algorithms. *SIGecom Exchanges*, 13, 2 (2015), 41–46. doi:10.1145/2728732.2728738. https://doi.org/10.1145/2728732.2728738. [Cited on pages 8 and 9.]

Google AI for Developers, 2025. Gemini api reference. https://ai.google.dev/api. Accessed 17 May 2026. [Cited on page 48.]

Google Firebase, 2026. Firebase developer documentation. https://firebase.google.com/docs. Accessed 17 May 2026. [Cited on page 48.]

Gotel, O. C. Z. and Finkelstein, A., 1994. An analysis of the requirements traceability problem. In *Proceedings of the First IEEE International Conference on Requirements Engineering*, 94–101. IEEE Computer Society Press. doi:10.1109/ICRE.1994.292398. https://doi.org/10.1109/ICRE.1994.292398. [Cited on page 29.]

Hamada, K.; Iwama, K.; and Miyazaki, S., 2011. The hospitals/residents problem with quota lower bounds. In *Algorithms – ESA 2011*, vol. 6942 of *Lecture Notes in Computer Science*, 180–191. Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-23719-5_16. https://doi.org/10.1007/978-3-642-23719-5_16. [Cited on page 7.]

Hassenzahl, M.; Diefenbach, S.; and Göritz, A., 2010. Needs, affect, and interactive products - facets of user experience. *Interact. Comput.*, 22, 5 (Sep. 2010), 353–362. doi:10.1016/j.intcom.2010.04.002. https://doi.org/10.1016/j.intcom.2010.04.002. [Cited on page 11.]

Heer, J., 2019. Agency plus automation: Designing artificial intelligence into interactive systems. *Proceedings of the National Academy of Sciences*, 116, 6 (2019), 1844–1850. doi:10.1073/pnas.1807184115. https://www.pnas.org/doi/10.1073/pnas.1807184115. [Cited on pages 1, 16, and 59.]

Hisfantor, 2023. Bgg best games for player count. https://github.com/Hisfantor/BGG-Best-Games-for-Player-Count. Google Sheets script that imports a user's BoardGameGeek collection and sorts games by suitability for a specified player count, with filters for rating, playtime, and complexity. [Cited on page 20.]

Horvitz, E., 1999. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 159–166. Association for Computing Machinery, New York, NY, USA. doi:10.1145/302979.303030. https://doi.org/10.1145/302979.303030. [Cited on pages 15 and 59.]

HSIEH, H.-F. AND SHANNON, S. E., 2005. Three approaches to qualitative content analysis. *Qualitative Health Research*, 15, 9 (2005), 1277–1288. doi:10.1177/1049732305276687. https://doi.org/10.1177/1049732305276687. [Cited on page 12.]

ION, M. A., 2018. *Designing and Evaluating a Board Game Recommender System*. Master's thesis, Technische Universität Wien. https://repositum.tuwien.at/bitstream/20.500.12708/1444/2/Ion%20Michael%20Antonius%20-%202018%20-%20Designing%20and%20evaluating%20a%20board%20game%20recommender...pdf. Diploma thesis. Accessed 2026-04-28. [Cited on page 22.]

KARAPANOS, E.; ZIMMERMAN, J.; FORLIZZI, J.; AND MARTENS, J.-B., 2009. User experience over time: An initial framework. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 729–738. Association for Computing Machinery, New York, NY, USA. doi:10.1145/1518701.1518814. https://doi.org/10.1145/1518701.1518814. [Cited on page 13.]

KELLY, D. AND TEEVAN, J., 2003. Implicit feedback for inferring user preference: A bibliography. *SIGIR Forum*, 37, 2 (2003), 18–28. doi:10.1145/959258.959260. https://doi.org/10.1145/959258.959260. [Cited on page 15.]

KIZILCEC, R. F., 2016. How much information?: Effects of transparency on trust in an algorithmic interface. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2390–2395. Association for Computing Machinery, New York, NY, USA. doi:10.1145/2858036.2858402. https://doi.org/10.1145/2858036.2858402. [Cited on pages 15, 55, and 58.]

KONSTAN, J. A. AND RIEDL, J., 2012. Recommender systems: From algorithms to user experience. *User Modeling and User-Adapted Interaction*, 22, 1–2 (2012), 101–123. doi:10.1007/s11257-011-9112-x. https://doi.org/10.1007/s11257-011-9112-x. [Cited on page 13.]

KRISTOFFERSEN, P., 2018. Cardboard butler. https://cardboardbutler.blob.core.windows.net/cardboardbutler/about.html. Web page that helps decide the next board game to play using BoardGameGeek data. Accessed 2026-05-28. [Cited on page 20.]

KUJALA, S.; ROTO, V.; VÄÄNÄNEN-VAINIO-MATTILA, K.; KARAPANOS, E.; AND SINNELÄ, A., 2011. UX curve: A method for evaluating long-term user experience. *Interacting with Computers*, 23, 5 (2011), 473–483. doi:10.1016/j.intcom.2011.06.005. https://doi.org/10.1016/j.intcom.2011.06.005. [Cited on page 13.]

LIM, Y.-K.; STOLTERMAN, E.; AND TENENBERG, J., 2008. The anatomy of prototypes: Prototypes as filters, prototypes as manifestations of design ideas. *ACM Transactions on Computer-Human Interaction*, 15, 2 (2008), 1–27. doi:10.1145/1375761.1375762. https://doi.org/10.1145/1375761.1375762. [Cited on page 13.]

LUDOPEDIA, 2026. I don't know, what do you want to play? https://ludopedia.com.br/jogo/i-don-t-know-what-do-you-want-to-play. Listing for a 2007 web-published card game designed by Tom Kiehl, with voting listed as a mechanic. Accessed 2026-05-03. [Cited on page 21.]

MANN, H. B. AND WHITNEY, D. R., 1947. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18, 1 (1947), 50–60. doi:10.1214/aoms/1177730491. https://doi.org/10.1214/aoms/1177730491. [Cited on page 13.]

MCI4ME, 2020. How to vote using the electronic voting system. https://vote.mci4me.at/ajuda/. Accessed 3 May 2026. [Cited on page 10.]

META OPEN SOURCE, 2026. React documentation. https://react.dev/. Accessed 17 May 2026. [Cited on page 48.]

NADAL, C.; MCCULLY, S.; DOHERTY, K.; SAS, C.; AND DOHERTY, G., 2022. The TAC toolkit: Supporting design for user acceptance of health technologies from a macro-temporal perspective. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22. Association for Computing Machinery, New York, NY, USA. doi:10.1145/3491102.3502039. https://doi.org/10.1145/3491102.3502039. [Cited on page 12.]

NAEEM, M.; OZUEM, W.; HOWELL, K.; AND RANFAGNI, S., 2023. A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. *International Journal of Qualitative Methods*, 22 (2023). doi:10.1177/16094069231205789. https://doi.org/10.1177/16094069231205789. [Cited on pages 13 and 72.]

NETLIFY, 2026. Netlify documentation. https://docs.netlify.com/. Accessed 17 May 2026. [Cited on page 48.]

NIELSEN, J., 1993. *Usability Engineering*. Academic Press, Boston. ISBN 9780125184069. [Cited on page 13.]

NISSENBAUM, H., 2004. Privacy as contextual integrity. *Washington Law Review*, 79, 1 (2004), 119–158. https://digitalcommons.law.uw.edu/wlr/vol79/iss1/10/. Accessed 2026-04-28. [Cited on page 23.]

NORMAN, D. A., 2004. *Emotional Design: Why We Love (or Hate) Everyday Things*. Basic Books, New York. [Cited on page 11.]

O'CATHAIN, A. AND THOMAS, K. J., 2004. "Any Other Comments?" open questions on questionnaires: A bane or a bonus to research? *BMC Medical Research Methodology*, 4, 25 (2004), 1–7. doi:10.1186/1471-2288-4-25. https://doi.org/10.1186/1471-2288-4-25. [Cited on page 12.]

PRUITT, J. AND ADLIN, T., 2006. *The Persona Lifecycle: Keeping People in Mind Throughout Product Design*. Morgan Kaufmann, Burlington, MA. ISBN 9780125662512. [Cited on page 12.]

REDDIT R/BOARDGAMES COMMUNITY, 2023. Tool to help a group decide on which game to play? https://www.reddit.com/r/boardgames/comments/17bhc8m/tool_to_help_a_group_decide_on_which_game_to_play/. Public forum discussion. Accessed 2026-05-18. [Cited on pages 1 and 31.]

RETTIG, M., 1994. Prototyping for tiny fingers. *Communications of the ACM*, 37, 4 (1994), 21–27. doi:10.1145/175276.175288. https://doi.org/10.1145/175276.175288. [Cited on page 13.]

RIEMAN, J., 1993. The diary study: A workplace-oriented research tool to guide laboratory efforts. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, 321–326. Association for Computing Machinery, New York, NY, USA. doi:10.1145/169059.169255. https://doi.org/10.1145/169059.169255. [Cited on page 13.]

ROCKET POWER SOFTWARE, LLC, 2025. Boardy: Ai board game night organizer. https://boardyboard.com/. SMS-native AI board game night assistant with game recommendations, scheduling, RSVPs, availability polls, BoardGameGeek-backed catalogue data, collection management, and play tracking. Accessed 2026-04-28. [Cited on pages 2, 20, and 21.]

ROSSI, F.; VAN BEEK, P.; AND WALSH, T. (Eds.), 2006. *Handbook of Constraint Programming*. Elsevier, Amsterdam. ISBN 9780444527264. [Cited on pages 8 and 52.]

RUDD, J.; STERN, K.; AND ISENSEE, S., 1996. Low vs. high-fidelity prototyping debate. *Interactions*, 3, 1 (1996), 76–85. doi:10.1145/223500.223514. https://doi.org/10.1145/223500.223514. [Cited on page 13.]

SHAPLEY, L. S. AND SCARF, H. E., 1974. On cores and indivisibility. *Journal of Mathematical Economics*, 1, 1 (1974), 23–37. doi:10.1016/0304-4068(74)90033-0. [Cited on page 8.]

SOFTWARE IMPROVEMENTS, 2005. eVACS and ACT legislative assembly elections. https://www.softimp.com.au/Common%20content/White%20Papers/eVACS_Development2.pdf. Accessed 3 May 2026. [Cited on page 10.]

SUCHMAN, L., 1987. *Plans and Situated Actions: The Problem of Human–Machine Interaction*. Cambridge University Press, Cambridge. [Cited on page 11.]

SWELLER, J., 1988. Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 2 (1988), 257–285. doi:10.1207/s15516709cog1202_4. https://doi.org/10.1207/s15516709cog1202_4. [Cited on pages 15 and 58.]

Tintarev, N. and Masthoff, J., 2012. Evaluating the effectiveness of explanations for recommender systems. *User Modeling and User-Adapted Interaction*, 22, 4–5 (2012), 399–439. doi:10.1007/s11257-011-9117-5. https://doi.org/10.1007/s11257-011-9117-5. [Cited on page 13.]

Triay, M. and Wood, A., 2021. Using board games as interfaces for social interactions. https://www.bbc.co.uk/rd/blog/2021-08-board-games-digital-storytelling-participation-experiences. BBC Research & Development. Accessed 2026-03-02. [Cited on page 2.]

van Lamsweerde, A., 2001. Goal-oriented requirements engineering: A guided tour. In *Proceedings of the Fifth IEEE International Symposium on Requirements Engineering*, 249–262. IEEE Computer Society. doi:10.1109/ISRE.2001.948567. https://doi.org/10.1109/ISRE.2001.948567. [Cited on page 29.]

Verrell, K. F., 2025. Boardroom: A board game voting system. COMP4560 Project Report, Australian National University. Canberra, Australia. [Cited on pages 1, 2, 7, 8, 16, 17, 18, 19, 48, and 50.]

Vite Contributors, 2026. Vite documentation. https://vite.dev/guide/. Accessed 17 May 2026. [Cited on page 48.]

What2Play, 2026a. Board game voting app for faster decisions. https://what2play.games/board-game-voting-app. Guide describing ranked voting, host controls, voting records, fallback options, and multi-table seat optimisation for board game events. Accessed 2026-04-28. [Cited on page 20.]

What2Play, 2026b. What2play: Board game night planner and voting app. https://what2play.games/. Board game night planning application with BoardGameGeek sync, event setup, AI-assisted ranking, voting, feedback, and multi-table seating features. Accessed 2026-04-28. [Cited on pages 2, 20, and 21.]

Yuan, Y.; Cao, J.; Wang, R.; and Yarosh, S., 2021. Tabletop games in the age of remote collaboration: Design opportunities for a socially connected game experience. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA. doi:10.1145/3411764.3445512. https://doi.org/10.1145/3411764.3445512. [Cited on page 2.]

Zalewski, J.; Ganzha, M.; and Paprzycki, M., 2019. Recommender system for board games. In *2019 23rd International Conference on System Theory, Control and Computing (ICSTCC)*, 249–254. IEEE. doi:10.1109/ICSTCC.2019.8885455. https://doi.org/10.1109/ICSTCC.2019.8885455. [Cited on pages 7 and 22.]

Zimmer, S., 2026. Board game picker. https://tools.sylvainzimmer.com/boardgames/picker/. CSV-based board game picker that filters a BoardGameGeek collection export and can choose a random game from the filtered set. Accessed 2026-04-28. [Cited on page 20.]