Personalized Recommendations in Mixed Reality Enhance Explanation Satisfaction and Hedonic User Experience in Board Game Learning

Sandra Dojcinovic
University St. Gallen
St.Gallen, Schweiz
sandra.dojcinovic@student.unisg.ch

Jannis Strecker-Bischoff
University St. Gallen
St. Gallen, Switzerland
jannis.strecker-bischoff@unisg.ch

Simon Mayer
University St. Gallen
St. Gallen, Switzerland
simon.mayer@unisg.ch

Kenan Bektaş
University St. Gallen
St. Gallen, Switzerland
HOCH Health Ostschweiz
St. Gallen, Switzerland
kenan.bektas@unisg.ch

Abstract

Board games often involve strategic decision making and procedural planning tasks. Such tasks require learners to make decisions based on dynamically evolving game state and changing information that is situated in a physical environment. Recommender systems can filter available information and provide learners with personalized and actionable suggestions that simplify their decision making while playing board games. Such recommendations can further be spatially aligned with relevant physical elements through Mixed Reality (MR). We present an MR system called GLAMRec for an engine-building strategy board game. GLAMRec provides personalized, transparent recommendations by integrating user data, real-time game state tracking, and ontology-based reasoning during a complex board game, which we use as a proxy environment for procedural learning tasks. We interviewed six board game designers to improve the GLAMRec and conducted a within-subjects design user study (N=32) to investigate how personalized explanations affect explanation satisfaction, user experience, and trust. We found that personalized recommendations significantly improve explanation satisfaction and hedonic user experience without affecting trust ratings, recommendation compliance, and game performance. These findings suggest that personalization primarily shaped perception of enjoyment rather than measurable learning outcomes or trust.

CCS Concepts

• **Information systems → Personalization**; **Recommender systems**; • **Human-centered computing → Mixed / augmented reality**; **Ubiquitous and mobile computing systems and tools**; • **Applied computing → Interactive learning environments**.

Keywords

personalized learning, immersive learning, decision-support systems, board games

ACM Reference Format:
Sandra Dojcinovic, Jannis Strecker-Bischoff, Simon Mayer, and Kenan Bektaş. 2026. Personalized Recommendations in Mixed Reality Enhance Explanation Satisfaction and Hedonic User Experience in Board Game Learning. In *31st International Conference on Intelligent User Interfaces (IUI '26), March 23–26, 2026, Paphos, Cyprus.* ACM, New York, NY, USA, 20 pages. https://doi.org/10.1145/3742413.3789129

IUI '26, Paphos, Cyprus

ACM ISBN 979-8-4007-1984-4/2026/03
https://doi.org/10.1145/3742413.3789129

1 Introduction

Personalized recommender systems are the subject of research in various domains, with the goal to help users navigate information overload by offering a subset of options that are individually tailored. By selecting content that best matches a user's profile, personalized recommendations can help users make better choices and potentially improve the learning experience over generic, one-size-fits-all approaches [26]. Personalization is hence generally valued for improving relevance and user satisfaction [40], better preference matching [72], and supporting diverse user needs and abilities [41]. Personalized recommender systems can provide substantial individual value also from the perspective of technology-enhanced learning [71], with promising findings in e-learning environments and beyond [11, 55, 75]. However, a lack of *transparency* in personalization has been linked to lower acceptance of recommendations [39]. In the context of intelligent systems, we define transparency as the user's appropriate understanding of how the system came to derive a particular decision or recommendation [63]. This is especially relevant in scenarios where users must learn from or make decisions based on the system's suggestions. For instance, studies have shown that offering "Why" or "How" explanation alongside recommendations can enhance users' perceived competence [15]. Accordingly, the availability of a structured, explainable model of system decisions (e.g., in an underlying ontology) can support interpretability of explanations [13].

In addition to transparency, previous research has shown that the choice of medium is also important in learning contexts. For example, various computer games incorporate scientifically validated learning principles, such as providing players with continuous feedback tailored to their immediate actions, as well as their abilities and learning styles [20]. Immersive virtual environments can support mental activities and reduce the cognitive demand to make inferences while solving problems [57]. Dalgarno and Lee argue that immersive virtual environments offer valuable benefits, as learners can better understand the representation of spatial knowledge, while immersion can enhance their engagement and motivation [16]. Furthermore, immersive technologies outperform traditional methods in procedural assembly tasks, improving accuracy and efficiency [10]. Mixed Reality (MR) technologies can enrich learning environments, allowing for tangible and spatial interaction with physical objects and novel input modalities such as user gaze to trigger specific content [14, 43, 79]. Specifically in board games, virtual overlays can guide players' attention without interrupting their learning flow [32] and MR can enhance immersion and engagement of players while assisting them in tracking the game state and decreasing their cognitive load [34, 46, 47, 70].

In this paper, we explore how personalized recommendations through MR can support learning experience. We present the *GLAMRec* ("Game Learning Assistance in MR through Recommendations") decision support system for board games. We use a complex physical board game as a proxy environment for procedural learning tasks in other domains such as industry, medicine, or agriculture. GLAMRec tracks the state of the board game using a head-worn MR device and overlays recommendations to supply users with real-time immersive guidance. We test two conditions where this guidance is either generic or personalized (i.e., recommendations are formulated with respect to user profile and real-time data) to investigate the influence of such personalization on *user learning experience*, *user explanation satisfaction*, and *trust*. In this context, we evaluate learning in terms of the perceived, subjective learning experience, how users perceive understanding, engagement and support while playing, rather than through objective, cognitive or procedural measures.

Our overall goal is to inform the design of recommender systems in learning-oriented contexts such that they remain effective and adaptive to the current user situation. In doing so, we contribute to the broader discourse on trustworthy and user-aware recommender systems that deliver explainable recommendations and support a satisfactory user and learning experience while respecting the user's agency in the decision-making process.

In summary, with this paper we contribute:

(1) A prototype system called GLAMRec for personalized, transparent recommendations in MR for learning a complex strategic engine-building board game: *It's a Wonderful World*.
(2) An empirical study to evaluate the effect of personalization on users' subjective learning experience, satisfaction and trust in the system, showing that personalized recommendations increase explanation satisfaction and hedonic user experience.
(3) Design suggestions for MR recommender systems in learning-oriented contexts that provide personalized, affordance-based explanations grounded in user data and real-time game state.

2 Related Work

We start with a presentation of the underlying assumptions, central themes, and most relevant research results that establish the baseline of our investigation, which informs the use of MR, recommender systems, and personalized learning support in an engine-building board game.

2.1 Technology-Supported Board Gaming

When approaching a new game, players are typically supported through user manuals, tutorials, or other onboarding mechanisms that explain rules and mechanics of the game. To reduce the initial learning barrier, printed materials are sometimes accompanied by digital tutorials such as *DIZED*[1]. Prior work has explored the integration of digital technologies with board games in various ways: through MR, embedding a fully digital board game into the player's physical environment or augmenting an existing physical board game to enhance immersion and engagement [46, 47]; through accompanying apps, where physical tabletop play is augmented by a digital app that manages gameplay functions such as information control and bookkeeping, reducing cognitive load and helping in narrative delivery [34]; and through physiological feedback, which integrates digital sensing technologies into multiplayer board games to transform gameplay mechanics [70]. MR augments physical objects and scenes with spatially aligned and interactive virtual objects [64]. Through spatial sound (and audio), the use of MR can thereby offer an interactive and immersive learning experience with interactive elements that go beyond the possibilities of traditional learning materials [27]. With respect to personalization, it is important to conceive affordances as action opportunities that exist *in relation to* the users' abilities as well as the current user-environment situation [22]. Previous research has used games as a learning environment, since they can showcase multiple affordances to the user in order to facilitate interaction [21, 68]. *Gameplay affordances* are defined as possible actions that improve players' performance such as score outcomes [29]. The support offered at the beginning of a game, often referred to as *tutorials*, can be implemented as cognitive and metacognitive scaffolding cues in MR that guide players toward appropriate actions without disrupting their flow state [32]. Incorporating user and game-state data and adding textual information to those cues may move such scaffolding toward personalized recommendations.

2.2 Recommender Systems and Transparency

Recommender systems are software tools or techniques that offer suggestions for items which are useful to the user [54]—for our investigation, these suggestions refer to actions and items that are relevant while playing an engine-building board game.

[1] https://dized.com/. Last accessed January 19, 2026.

Recommender systems are today widely used across industries such as e-commerce, retail, media, and entertainment to help users navigate the information overload and help them make a decision [23, 35, 58, 77]. To derive recommendations, systems may use different knowledge sources, including demographics, individual or group preferences, and needs; other knowledge sources include contextual and domain-specific knowledge as well as any inferences that could be made from these sources [18]. Hybrid recommender systems, in this context, integrate several knowledge sources to derive a recommendation which is tailored to the individual user and context [18].

With growing system complexity and reliance of users on recommendations, the need for *transparency* increases as well [48]. In the context of recommender systems, transparency refers to whether and how a system conveys *how it arrived* at a particular recommendation [63]. The provisioning of explanations has been shown to significantly increase user confidence when making decisions as well as user trust in the system [9, 39]. Maartje and Malle argue about the need for AI systems to provide explanations for their actions *in the same way as humans would* in order to become explainable [17]. While the topic of explainability has gained popularity in the field of recommender systems (e.g., [62, 69]), it remains underrepresented in educational recommender systems [4]. This is a missed opportunity, since in an educational context, *transparent* explanations could bring several benefits at once: Not only could they increase trust in the system, hence providing indirect learning support, but they could furthermore support the learning of a user *directly* by bringing system choices to their attention and thereby enhancing self-reflection.

2.3 Learning Support and Personalization

Explanations have been shown to be an essential part of an individuals' learning process, helping to recognize underlying patterns [38, 44]. Explanations provided by a system should hence be given in a way (format, language, visual appeal, etc.) that is familiar to the user, and the rationale behind actions should be communicated in a way that is intuitive for the user. Specifically, to motivate and justify possible actions, a system should clarify these actions by offering reasons that are driven by beliefs, goals, or duties [17]. Offering such transparency in an educational context to support the learning of specific content has been studied in only a few use cases unrelated to recommender systems (cf. [76, 78]). Research on *transparent* recommender systems in education is hence also rare; rather, the field has primarily focused on systems that recommend *what to learn* rather than the learning activity itself [3, 4]. Finally, explanations themselves might be *personalized*; however, explainable recommender systems are typically kept generic and do not consider individual users, leaving a large potential of using personalization to enhance user-centered explainability (cf. [25]).

According to Strecker et al., *personalization* describes a system-initiated adaption based on personal data for any delivery or process method [67]. In an educational context, Pérez-Ortiz et al. [52] takes user data such as interests and provides personalized recommendations on learning material based on the user's prior knowledge and interaction. Jiang et al. [37] also includes dynamic data such as time-aware interaction patterns to recommend individual learning paths, personalized to the user's needs. In e-learning settings, personalization has been shown to improve learning outcomes and enhance learners' autonomy [8, 26]. Researchers are further exploring the use of generative AI methods for personalized learning; for instance, Shu et al. [61] uses large language models (LLMs) to generate tailored learning paths based on user performance and interaction data for guitar learning.

3 GLAMRec: An MR System for Personalized Decision-support for a Board Game

To leverage these prior findings on improving effective learning through personalized [8, 52, 61] and transparent [76, 78] recommendations, we propose the GLAMRec system, which provides in-game recommendations for a board game to support learning how to play the game. The recommendations are based on the board game's rules, real-time game-state data and personal user data provided by the players. Previous research highlights that (complex) board games and their mechanics can help to facilitate learning various disciplines and skills, including computational thinking, teamwork, and creativity [5, 53]. Especially strategy-focused board games can be complex and cognitively demanding [45]. In our study, we therefore focus on personalized recommendations in a strategy game, which can be thought as a proxy for other learning processes in domains such as healthcare (e.g., following a sanitation procedure for medical instruments), hospitality (e.g., learning how to prepare roast beef), or craftsmanship (e.g., learning how create a sculpture).

GLAMRec provides recommendations on the learning activity itself (e.g., recommendations on how to play the game), rather than on the learning material as in prior recommender system implementations (e.g., recommendations on which practice content to select next for learning or which board game to play next) [3, 4, 33]. We intentionally limit the scope of recommendations to learning-oriented decision support. While other aspects of gameplay, such as social interaction and narrative strengthening, could also be supported, understanding game mechanics and early strategic choices represents a challenge when engaging with complex board games. Furthermore, the recommendations are shown in MR, with virtual content visible next to the physical board game, to allow for an immersive learning experience that complements rather than replaces the physical board game. In the following, we present our rationale for the specific board game that we selected for our study (see Section 3.1) and discuss the overall GLAMRec system architecture (see Section 3.2). We then detail how GLAMRec creates recommendations (see Section 3.3) and illustrate how users interact with the system (see Section 3.4). The subsequent Sections 4-7 introduce and discuss our evaluation of GLAMRec through expert interviews and a user study.

3.1 Board Game Selection

The board game genre "engine-building" is particularly well suited for our study because engine-building games typically induce a complex and incremental learning environment with high strategic depth. For game selection within this genre, we defined three further requirements:

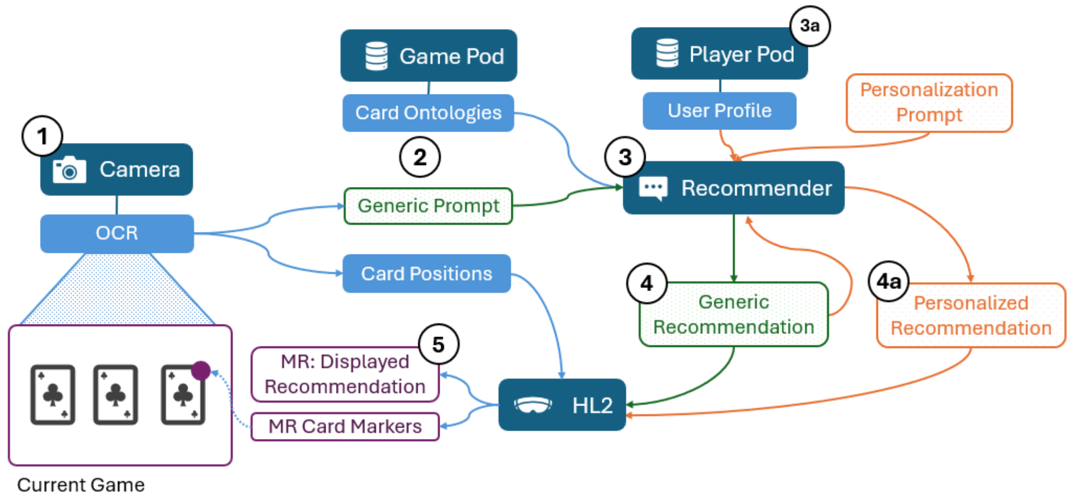


Figure 1: An overview of GLAMRec's components. (1) A camera is mounted above the game and uses OCR to recognize the text on individual cards and their positions. (2) The current recognized cards are included in a prompt that is sent to an LLM together with the cards' ontologies. (3) The LLM creates a generic recommendation (4) for which cards to choose in the current game phase. (4a) If a personalized recommendation should be created, the generic one is then sent again to the LLM, along with the player's user profile (3a) and another prompt that specifies how to personalized the generic recommendation. (5) The generated recommendation (generic or personalized) is displayed in MR. Additionally, card markers that indicate which ones to choose are visible when a user directs their gaze to the specific zone.

- The game needed to require low reliance on chance (e.g., dice-rolling) and dexterity (e.g., physical skills).
- The game needed to exhibit *sufficient complexity* for strategic reasoning, such that a recommender system may offer valuable assistance beyond trivial observations.
- The game needed to support *solo player mode* and have a rather *short playtime*; to ensure a feasible, controlled, and reproducible study setting.

Based on these requirements, we selected the board game *It's a Wonderful World (IAWW)*.[2] In IAWW, players build and manage an expanding empire by drafting and constructing cards to optimize resource generation. The aim is to develop the most prosperous civilization and earn most victory points after four rounds of strategic planning and resource production. The game ,[3] supports one to five players, and has an average playtime of 45 minutes per playthrough. In solo player mode, each round consists of two planning phases and one production phase. During the planning phase, the player selects cards to either construct, recycle, or discard, considering various factors like the card type, cost, production, conditional bonuses, and victory points. During the production phase the player can produce resources from the built deck and finish constructing cards. Throughout the game, the player has to make decisions in each draft, allocate resources, and continue to optimize their strategy. From a practical perspective, the game offers consistent card layouts and typography suitable for OCR.

3.2 GLAMRec System Architecture

We created GLAMRec to provide transparent and personalized decision support for players of the IAWW. The recommendations are visualized in MR including game-affordance-based explanations grounded in real-time game state as well as user data. GLAMRec consists of five main components (see Figure 1) that we introduce in the following.

Real-time game state recognition is implemented through an OCR-based camera-feed pipeline that recognizes game cards and tracks their positions across zones on the playing table (see (1) in Figure 1 and Figure 2). The text recognition was implemented with the open-source tool PaddleOCR [4]. To support accurate card identification and game state modeling, a full game ontology was created for the IAWW game using RDF [73] (see (2) in Figure 1). This ontology encodes all cards of the game and their associated attributes, costs, and conditions (see Figure 3). This approach relates to broader efforts that combine object recognition and identification with semantic grounding [65, 66]. The ontology is hosted on a Solid Pod[5] for decentralized, user-controlled data access (cf. [56]), and is loaded once at system initialization. When cards are recognized on the playing field, the game ontology is accessed to establish the current game state. This serves as input for the recommendation generation (see (3) in Figure 1), which can be run in one of two modes: *generic* (see (4)) and *personalized* (see (4a)). To personalize generic recommendations, the system makes use of user profiles with data collected through a questionnaire before the game (see Appendix D.2). The profiles are stored on Solid Pods and contain

[2]https://www.laboitedejeu.fr/en/its-a-wonderful-world/. Last accessed January 19, 2026.
[3]https://boardgamegeek.com/. Last accessed January 19, 2026.
[4]https://github.com/PaddlePaddle/PaddleOCR. Last accessed January 19, 2026.
[5]https://solidproject.org/. Last accessed January 19, 2026.

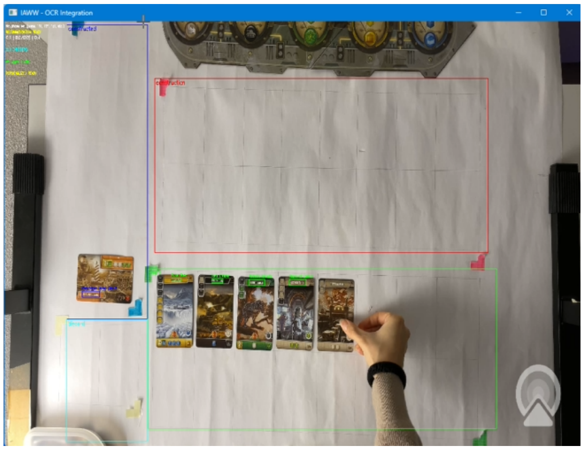


Figure 2: Real-time OCR integration to detect and classify cards that are placed in different defined zones while playing the *It's a Wonderful World.*

personal background information about the player (see (3a) in Figure 1). Finally, generic as well as personalized recommendations are passed to the MR interface which is implemented using a Microsoft HoloLens 2 device (see (5) in Figure 1). This interface contextually displays the recommendations and visual affordance cues. The cues displayed gaze-contingent, i.e., they appear in MR based on where the player directs their attention, which aligns with prior research that uses gaze to opportunistically adapt to user's context [7, 24].

3.3 Recommendation Generation

Recommendations in GLAMRec are created in two different modes: The *generic recommendations explain the game* independent of user background while the *personalized recommendations* tailor explanations based on user data and apply analogy-based reasoning. For the generation of recommendations, we used an LLM-based approach (cf. [42, 74]) and specifically GPT-4o-mini as our recommendation LLM based on early pilot tests across 74 recommendation generation instances in three game sessions.

3.3.1 Generation of Generic Recommendations. To generate generic recommendations, we used the current game state together with a tailored prompt containing the IAWW game rules as input to the model. Because the official game manual[6] emphasizes multiplayer gameplay with just a brief description of the adjustments to enable solo play, we created a summary of the instructions that focuses on solo mode. The generic recommendations are linked with three different zones featured in IAWW (i.e., the *Constructed Zone*, *Construction Zone*, and *Draft Zone*) and provide two layers of explanation: (1) *immediate tactical reasoning* (e.g., "construct this card for the energy recycling bonus which you can use"), and (2) *strategic context* (e.g., "this card synergizes with your current deck"). In early rounds of the game, GLAMRec provides detailed guidance while verbosity is gradually reduced in later rounds to encourage independent decision-making.

[6]See https://www.laboitedejeu.fr/wp-content/uploads/2021/01/ITS-EN-STD-rules-web.pdf. Last accessed January 19, 2026.

3.3.2 Generation of Personalized Recommendations. GLAMRec's *personalized* recommendations are based on a user profile, which is created through a short personalization form that takes approximately 5-10 minutes to complete (see Appendix D.2). This form captures several aspects of the user's background and preferences, which we derived through exploratory discussions, and were selected for specific reasons:

- **Known board and video games** indicate prior experience with strategic or digital game mechanics, enabling the recommender to relate to known game principles.
- **Academic and professional background** to draw on familiar disciplinary concepts, domains or work-related contexts in explanations.
- **Media preferences, passions, and hobbies** allow recommendations to be framed using analogies that resonate with the player's personal interest.
- **Preferred recommender tone** supports adjusting the communication style to user expectations.
- **Preferred language** ensures further accessibility and a more natural interaction with the system.

Making use of this additional information, personalized recommendations add a third layer on top of the *immediate tactical reasoning* and *strategic context* introduced above: (3) *analogy-based explanations* relate abstract strategies to familiar concepts (e.g., "try increasing your science production, similar to how you would expand a settlement in Catan to maximize resource generation" refers to a similar mechanism in the game *Settlers of Catan* that this user has prior experience with).[7] This design choice was later reinforced by the expert interviews (see Section 4), in which the designers of board games emphasized that they often relate back to known games or concepts when explaining how to play a new game to a player. The focus was set on experiential and contextual dimensions to obtain relatable explanations and simultaneously to avoid using demographic variables such as age or gender to reduce the risk of reinforcing stereotypes. Personalized recommendations furthermore adapt their *tonality* as prior research has shown that tone-aware explanations can increase perceived integrity, persuasiveness, transparency, and satisfaction with recommender systems [49]. A recent study by Okoso et al. revealed that explanation tone also significantly affects the user decision-making process [50]. Based on these prior findings, the four tone variants used in this paper (i.e., analytical/neutral, formal/professional, supportive/friendly and engaging/enthusiastic) cover a variety from restrained factual expression to more socially and emotionally engaging styles. Finally, to avoid boredom, personalized recommendations vary the referenced profile dimensions across turns by randomizing the chosen prompt. In the design of the GLAMRec, we employed a privacy-preserving approach. User profiles as well as the game's ontology are stored on a Solid Pod[8] to decouple our system from the data it uses (cf. [56]). Since data storage and access control are handled within the Solid ecosystem, a separate database is not required and users retain fine-grained control of what data they share and what applications they

[7]Examples of a generic and a personalized recommendation can be found in Appendix A.

[8]For this project, an instance of the open-source community Solid server was used, see https://github.com/CommunitySolidServer/CommunitySolidServer.


```xml
<iaww:Card rdf:about="http://example.org/iaww-ontology#ParallelDimension">
  <rdfs:label>Parallel Dimension</rdfs:label>
  <iaww:hasCardType rdf:resource="http://example.org/iaww-ontology#Discovery"/>
  <iaww:hasCopiesInDeck rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">1</iaww:hasCopiesInDeck>
  <iaww:hasResourceCost>
    <rdf:Description>
      <iaww:hasResourceType rdf:resource="http://example.org/iaww-ontology#Science"/>
      <iaww:hasQuantity rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">3</iaww:hasQuantity>
    </rdf:Description>
  </iaww:hasResourceCost>
  <iaww:hasResourceCost>
    <rdf:Description>
      <iaww:hasResourceType rdf:resource="http://example.org/iaww-ontology#Exploration"/>
      <iaww:hasQuantity rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">4</iaww:hasQuantity>
    </rdf:Description>
  </iaww:hasResourceCost>
  <iaww:hasResourceCost>
    <rdf:Description>
      <iaww:hasResourceType rdf:resource="http://example.org/iaww-ontology#General"/>
      <iaww:hasQuantity rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">1</iaww:hasQuantity>
    </rdf:Description>
  </iaww:hasResourceCost>
  <iaww:providesKrystallium>
    <rdf:Description>
      <iaww:hasResourceType rdf:resource="http://example.org/iaww-ontology#Krystallium"/>
      <iaww:hasQuantity rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">3</iaww:hasQuantity>
    </rdf:Description>
  </iaww:providesKrystallium>
  <iaww:hasSpecialAbility>
    <iaww:SpecialAbility>
      <iaww:abilityType>score_multiplier</iaww:abilityType>
      <iaww:abilityDescription>3 VP per Research card</iaww:abilityDescription>
      <iaww:affectsCardType rdf:resource="http://example.org/iaww-ontology#Research"/>
      <iaww:pointMultiplier rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">3</iaww:pointMultiplier>
    </iaww:SpecialAbility>
  </iaww:hasSpecialAbility>
  <iaww:recyclesTo rdf:resource="http://example.org/iaww-ontology#Exploration"/>
</iaww:Card>
```

Figure 3: The figure illustrates how the card "Parallel Dimension" is represented in the ontology in RDF format, including its card type, resource cost, construction bonus, scoring and recycling bonus.

share it with. The advantage of this approach is that data remains under user ownership and control while remaining accessible to applications in an interoperable format (cf. [6]). In Solid, applications' access rights are verified upon each request—meaning that our system's personalized recommender can gracefully degrade when the user chooses to share less data. We have consciously aligned GLAMRec with the interests of the Linked Data community, demonstrating how Solid-based data infrastructures can be leveraged for user-facing applications, in this case for personalized and transparent recommendations in MR environments.

3.3.3 Transparent Recommendations. Transparency in GLAMRec is implemented on two levels: (1) *explanation transparency* concerns how recommendations are communicated to the user while (2) *decision transparency* makes visible how the recommender's outputs relate to the current game state. By exposing both tactical and strategic recommendations, GLAMRec enables users to understand not only what to do but also why to do it. Previous as well as ongoing research has explored using semantic sources, such as ontologies and knowledge graphs, to be able to provide explainable recommendations in an educational setting [1, 2]. We support this approach and show how our game ontology can serve as the basis for providing transparent explanations based on the current game state. This layered approach of transparency helps complete novices as well as more experienced players to benefit from the generated recommendations. In the MR interface, gaze-enabled visual card markers are overlaid directly on the physical game cards that are being recommended for construction, recycling and discarding in the Draft Zone or that are recommended to be prioritized for building in the Construction Zone. This makes GLAMRec 's suggestions explicitly traceable to concrete game elements on the playing field, ensuring that users can always identify which part of the game state a recommendation refers to. By anchoring recommendations in the visual field, the system reduces ambiguity and may support greater transparency without overwhelming the player. This design empowers players to follow, question or ignore the system's reasoning by glancing at the playing field, thereby maintaining agency while interacting with the recommender. This layered design follows prior research on explanation transparency in recommender systems [63] and aligns with MR-focused proposals for increasing user agency and perceptual awareness [67].

3.4 User Interaction in MR

GLAMRec runs in real time on the HoloLens 2 device and was developed with the Unity Engine, using building blocks from the Mixed Reality Toolkit (MRTK)[9] and the Augmented Reality Eye Tracking Toolkit (ARETT)[10]. Eye-gaze tracking using ARETT ensures a fixed sampling rate (30Hz). The zones of the game's playing field are

[9]https://github.com/MixedRealityToolkit/MixedRealityToolkit-Unity. Last accessed January 19, 2026.
[10]https://github.com/AR-Eye-Tracking-Toolkit/ARETT. Last accessed January 19, 2026.

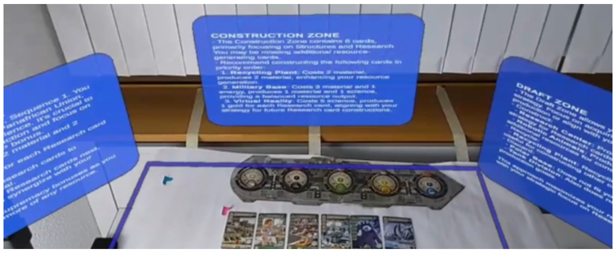


Figure 4: The MR interface shows three floating recommendations that relate to one of the three zones in the board game IAWW. The content updates each round to guide next actions. (Note: This image is taken from a recording; UI visibility is much better on the HL2 device.)

configured in MR to align with the given physical setup. Next, three floating text panes are positioned above the playing field (to the left, front, and right of the player's perspective, see Figure 4)—generated recommendations appear in these spaces. The recommendation system communicates with the MR interface using a lightweight UDP connection for real-time gameplay support. Detected card positions from the camera are mapped to the predefined board zones in MR using a 2x8 grid system, allowing individual card recommendations to be spatially anchored to the correct cards with 3D card markers. The user's gaze triggers the display of further contextual information. This is triggered based on gaze dwell time and leads to a highlighting of the corresponding zone borders and card markers. This affordance-based interaction ensures that relevant recommendations (e.g., for individual cards) appear when the user directs their attention to the respective area during gameplay.

In the Construction Zone, GLAMRec marks cards with a dot in a turquoise color to indicate construction prioritization (see Figure 5a). In the Draft Zone, GLAMRec marks cards that are recommended to construct in blue and cards that are recommended to be recycled or discarded in orange (see Figure 5b). These card markers are set based on the given recommendations and are gaze-contingent, appearing if the user is looking at the corresponding zone. The color selection for the card markers was based on the consideration to remain distinguishable under common forms of color blindness. Each time the game progresses to the next phase and the game state changes, the recommendation system can be triggered manually for the appropriate round/sequence or updates automatically when card movements during the discard/redraw step are detected. The MR interface updates accordingly, so that users always see relevant recommendations based on the current game state. The recommendations and affordance cues were integrated into the physical space, allowing users to perceive them while actively playing the board game.

(a)

(b)

Figure 5: Gaze-enabled card markers for each zone in MR to visually indicate the zone-specific recommendation (Highlighted with pink circles in these screenshots. The circles are not visible for players). Markers update in real time as recommendations appear or cards get removed. Image taken from a recording; UI visibility is better on the HL2. (a) Construction Zone. GLAMRec marks cards it recommends prioritizing for construction with light-blue markers. (b) Draft Zone. GLAMRec marks cards it recommends for constructing with dark blue markers and for recycling/discarding with orange markers.

4 Expert Interviews

In the course of the development of GLAMRec, we conducted six *expert interviews* with experienced board game designers to gain qualitative insights on a prototype version, towards improving the final version of GLAMRec. The aim of these expert interviews was to understand what domain experts considered helpful in an MR-based recommender system for board games, what they liked or disliked about our prototype, and what features they would envision particularly in MR.

4.1 Expert Selection and Interviews Method

We recruited six board game designers via the BoardGameGeek platform, where we only selected individuals with at least one published game with a minimum complexity rating of 2 and preferably similar mechanics to IAWW, such as closed drafting, end game bonuses, or deck/bag/pool building. The expert pool consists of five men and one woman, with the majority of experts aged around 55-64 years (see Table 1). We provide their names with their consent:

- *E1: Mac Gerdts* is the inventor of the rondel game mechanism and is internationally recognized for his contributions to the Eurogame genre. Games developed by him include *Imperial* and *Concordia*, which is rated as the 23rd-best among all strategy games (and 26th-best among all games) on *BoardGameGeek*.
- *E2: Thomas Sing* won the German *Kennerspiel des Jahres* ("expert game of the year") and the *Deutscher Spielepreis*

Table 1: Demographic and experience data of the interviewed board game designers.

ID	Name	Gender	Age Group	No. Games developed	Exp. (yrs)
E1	Mac Gerdts	Male	55–64	13	20
E2	Thomas Sing	Male	55–64	15	12
E3	Stefan Malz	Male	55–64	9	17
E4	Louis Malz	Male	25–34	5	13
E5	Geoffrey Engelstein	Male	55–64	25	20
E6	Rita Modl	Female	35–44	14	9

2020 ("German Game Prize 2020") for his board game *The Crew: The Quest for Planet Nine*. His games include various variation of *The Crew* and *The Key*.

- *E3: Stefan Malz* is part of the father–son duo *Malz Spiele*, organizes board game events across Germany, and collaborates with publishers on rulebooks, editing, and translations. Notable games include *Rococo*, which was nominated for several board game awards, and *Edo*.
- *E4: Louis Malz* is the other half of the father–son duo *Malz Spiele*, also organizes board game events across Germany, and collaborates with publishers on rulebooks, editing, and translations. Louis Malz has extensive knowledge of over 1000 board games.
- *E5: Geoffrey Engelstein* teaches board game design at NYU Game Center and is an author of books on tabletop gaming. He is also well known for his long-running "GameTek" segment on the podcast *The Dice Tower* and for his popular BoardGameGeek geeklists on the history and theory of games. Board games he developed include *The Expanse* and *Space Cadets*. In addition, he co-founded Mind Bullet Games, a small design studio responsible for *The Ares Project*.
- *E6: Rita Modl* is a member of the Executive Committee of the *Spiele-Autoren-Zunft* ("Game Authors' Guild") that is mainly based in Germany but internationally oriented. She is engaged in various associations, such as *Blick aufs Brett* and *Bayerisches Spiele-Archiv Haar e.V.*. With her debut game *Men at Work*, she won the Austrian Games Award "Spiele Hit mit Freunden" in 2018. She designed games such as *Kuhfstein* and *King of 12*.

All experts completed our system's personalization form (see Appendix B) prior to their interview. One expert was already familiar with the game. During the interviews, each expert was shown five personalized recommendations that were based their user profile (see Figure 6). The interviews were semi-structured and allowed for open comments by the experts (see Appendix C). We were especially interested to find out what they thought about the recommendations in terms of usefulness and accuracy, and how they thought that the recommendations would impact the game play.

4.2 Interview Results

The interview data preparation followed the thematic analysis approach by Braun and Clarke [12], which resulted in six main themes (see Table 2 for an overview).

Figure 6: An example of a recommendation as it was shown to E5 during the interview.

Table 2: Overview of the themes identified in the expert interviews (N=6).

Theme	Summary
Learning to Play a Game	Strategic cues and known-game analogies support new players.
Transparency	Explaining why actions are not recommended is helpful.
Personalization Dimensions	Responses to personalized analogies are highly individual.
Frequency of Personalization	Too many personalized references can feel artificial over time.
Simplicity	Concise summaries help reduce cognitive load during gameplay.
Visualization	Visual cues were preferred over text-heavy explanations for quick guidance.

Learning to Play a Game. When approaching a new game, the experts emphasized the importance of intuitive gameplay. Experienced players "often draw on the experience from other games they know, searching for analogies in their mind, often subconsciously, to derive an initial strategy." (E1); such analogies are hence deemed very helpful for learning. The experts "think it's great to give new players an overall strategy... there are just so many choices, so having an initial strategy is really helpful to narrow down what people have to consider." (E5), and E4 mentioned that "With complex games, it is difficult for many to make a decision in the beginning. They are aware that the decision doesn't have to be correct and that games are about fun, but still many players want to make the right decision at the start." Traditionally, this guidance is done by another person who knows the game well while this role might be taken by recommender systems such as GLAMRec. E6 agrees, stating that "That makes sense, because if I play a game for the first time and don't understand something, usually someone sitting next to me says, 'Take that card because next round you'll get that resource or these victory points.' The system basically does what an experienced player would do, giving me a little hint."

Transparency. For the system to be transparent in explaining why it would recommend certain choices—and, particularly, why it does not recommend others—is considered important since players can learn the most from such explanations: "I think it is even more important to know why I would not take the other cards... For the overall understanding of the whole game, it is important to know what these specific cards would have done for you, or rather what they would not have done for you." (E3). Overall, the majority of experts appreciated the generic explanation of the recommendation more than the personalized version, saying that it is either unnecessary to improve understanding or could be potentially disruptive.

Personalization Dimensions. Our prototype's personalized references were received mixed feedback from the experts. Some experts liked the analogies made to known games: "That is also what we often do. When someone asks what kind of game they should play, we ask which games they know." (E4) and "If somebody's played Dominion and we're sitting down to play another deck-building game, I'll say, hey, you're going to draw five cards… those steps are similar to Dominion, just to orient them." (E5). E6 further highlighted emotional connections: "I don't play it much, but I have a very positive connection to Catan. And when something is explained with reference to Catan, I immediately have a positive experience of the game." However, for others, references to other games represented more of a hindrance: "When I read through it quickly, I notice that I constantly have to switch mental tracks." (E3). Some experts appreciated the references to other dimensions of their user profile, such as job experience: "I had to smile a little, because I am effectively using and managing resources [on my job]." (E1). In contrast, E4 remarked that such analogies might become superficial: "It is difficult to capture what really fits when working with very rough concepts." With respect to further personalization dimensions, E2 suggested that "For me it would have been much more informative if they had asked what I have done throughout my life, what I have enjoyed...because that shaped me. Since I collected so much information from different experiences, I can now act intuitively."

Frequency of References to User Profile. Most experts found the frequency of references to user profile dimensions overwhelming: "From my personal feeling, I think there are too many personalized references. Overall it works, it gives connections, that you are familiar with." (E4). This was also observable over time: While E2's initial reaction to personalized references was "I believe the referral to my personal data does something to my motivation. I like it, I think it increases my motivation.", with more references E2 noted that "In the meantime, it now feels a bit artificial. I don't need it. It is too repetitive. I now get the impression, it has to give me these hints even though I don't need them anymore... Now it feels forced... I wouldn't call it manipulation, but it doesn't come across well."

Simplicity. All experts emphasized the value of simplicity when it comes to formulation and length of explanations, as well as gameplay in general: "For most people, they don't want to read the rules. I always say that, the rules are the original sin of board games. It's the root of many of the problems with board games. People get intimidated, people get nervous, people start to say 'it's going to take too much time' or get bored." (E3). On the length of our prototype's recommendations, E3 noted that "It says the right things, but it is expressed in a very cumbersome way." Experts further would welcome a summary feature for the produced resources: "This is something that often overwhelms players, all the small bureaucratic details... That's what many in our community appreciate in online games, that the system takes over the scoring and bookkeeping. It takes away a lot of work." (E4).

Figure 7: Study setup: A participant wearing the HoloLens 2 while playing the board game *It's a Wonderful World*. An overhead camera tracks the game state; predefined card decks are beside the play area.

Visualization. Connecting to the point of simplicity, visualization was seen as beneficial for giving players an overview. E6 suggested that "I see the card and it is outlined like a traffic light: green if it's very good, yellow if it's okay, red if it's not great. Then I immediately know which one to take." E2 emphasized the value of visuals over text: "Texts are always difficult... but symbols are great, they explain things quickly." Similarly, E3 highlighted the benefit of visualizing game constraints: "The cards I can build would be shown in one color and the ones I can't would be different. That would be a huge advantage, taking away the mechanical part of the game."

4.3 Implications for the GLAMRec system

Our findings from the expert interviews were used to inform and refine GLAMRec's design before conducting our main user study. Importantly, we adjusted the system to use personalization more sparingly and to introduce more variation, ensuring that recommendations did not repeatedly draw on the same personalization source (e.g., prior experience with board games). Because experts furthermore liked the idea of visual indications similar to those used in video games, we made refinements in regards to the visual presentation of the recommender systems' information by implementing card markers not only indicating the mentioned cards for faster orientation during gameplay but also directly indicating suggested actions.

5 GLAMRec Evaluation

We conducted a within-subject user study with N=32 participants. Figure 7 shows the experimental setup. Each participant experienced both recommender system types (i.e., generic and personalized) in MR under counterbalanced conditions.

5.1 Hypotheses

We compare generic and personalized recommendations with regards to perceived support for learning, and subjective learning experience. Since the level of perceived "learning support" and "learning experience" can be very subjective and difficult to measure directly, we use proxy and connected measures for these. Concretely, we investigate the overall user experience, as an important pre-requisite for an efficient learning process.

A good learning experience includes that using the system is enjoyable and that it provides a good user experience in general, where the user is engaged in the process. Thus, we measure the user experience using the short version of the *User Experience Questionnaire (UEQ-S)* [59], which consists of bipolar items rated on a 7-point Likert scale and captures pragmatic and hedonic qualities (see Appendix D.3.1).

When learning a new game, it is important that one can understand the instructions and is satisfied with the explanations one gets. We hypothesize that personalization will increase the satisfaction for the recommendations, as the personalized version will be more relatable for the participants. We measure explanation satisfaction using the *Explanation Satisfaction Questionnaire (ESQ)* [30] on a 5-point Likert scale (see Appendix D.3.2).

Additionally, we consider trust in the system an important prerequisite for an efficient learning process. As previous research suggests an understanding of how the system derived the recommendations, would lead to a potentially higher trust rating [9, 39]. We thus measure trust in the recommender system using the *Trust in Automation Questionnaire (TiAQ)* [36], which contains 12 items on 7-point Likert scales (see Appendix D.3.3).

The independent variable therefore was the type of recommender system, while the dependent measures, include the aforementioned *explanation satisfaction*, *user experience*, and *trust*.

Thus, we formulated our hypotheses as follows:

- H1 Participants report a better user experience (measured with the UEQ-S) in the personalized condition compared to the generic condition.
- H2 Participants report higher explanation satisfaction (measured via ESQ) in the personalized condition compared to the generic condition.
- H3 Participants report greater trust (measured via TiAQ) in the personalized condition compared to the generic condition.

Moreover, we collected further quantitative measures and treated them as observational; these include *game performance* (rounds reached, total points scored, number of cards built) and *recommendations compliance* (fraction of recommendations followed by the user). We explored whether participants would show greater adoption of the provided recommendations by analyzing the recommendation compliance for each move in each condition.[11] A recommendation was considered valid if it was feasible within the game rules and internally consistent (e.g., each card was associated with a single recommended action and discard actions included two cards). Otherwise, the recommendation was counted as invalid and excluded from the compliance counts. Each valid recommendation was counted either as followed, if the player executed the recommended action for that card, or as not followed otherwise (e.g., if a card was recommended for recycling and was instead constructed by the player). These compliance values were aggregated per participant and game round. Recommendation compliance was analyzed for the draft zone. Recommendations provided in the construction zone (e.g., prioritization of up to three cards for construction) were excluded from the compliance analysis, since resource allocation actions were not tracked and therefore could not be reliably matched to recommendation compliance. Additionally, we also considered game performance including total points scored, number of cards constructed and rounds completed as exploratory measure.

[11]To ensure validity, recommendations that were clearly erroneous (e.g., duplicated or inconsistent) were excluded from this analysis. We observed an error rate of around 7.5% in our setup.

5.2 Method

Prior to conducting the user study, we performed a sensitivity power analysis in G*Power for a two-tailed paired-samples t-test. The significance level was adjusted to $\alpha = .0125$ (Bonferroni–Holm correction for four comparisons: UEQ-Hedonic, UEQ-Pragmatic, ESQ, and TiAQ per condition). Based on Sawant et al. [51], effect sizes are often not reported for research papers in the human-computer interaction field; however, based on their meta-study of quantitative CHI papers, the thresholds for within-group effects considered small, medium, and large are 0.12, 0.32, and 0.72, respectively. Taking these field-specific recommendations and feasibility constrains (e.g., time, resources) into account we chose an expected medium-to-large effect size (Cohen's d = 0.6) with a power of 0.70, requiring us to recruit at least 29 participants.

5.2.1 Experimental Procedure. Our within-subject study took ca. 60 minutes per participant with two different primary phases. Prior to their study slot, participants were asked to complete two pre-surveys regarding their demographics (see Appendix D.1) and for the personalization (see Appendix D.2). After arriving in our laboratory, the participants were welcomed, received information about the user study, and signed a consent form regarding data collection and use. Each participant put on the HoloLens 2 and went through the built-in calibration procedure.

The first study phase was conducted outside of the MR environment, with the intention to mitigate any potential biases in ratings caused by participants' negative attitudes toward MR in general. Participants were shown two different versions of the instructions for the IAWW (i.e., generic and personalized) on a computer screen, with the order counterbalanced in each trial. The instructions were generated using prompts similar to those for the recommendations, with the personalized version based on the same user profile as the gameplay recommendations. After reading each version, the participants rated the instructions using the ESQ [30].

The second phase of the study consisted of two short gameplays of the IAWW game, with a duration of 15 minutes each, with the MR recommender, and counterbalanced between the generic and personalized recommender systems during each play. To mitigate learning effects and offer varied gameplay, two different, counterbalanced decks with different starting conditions were used (Deck A and B) in the two gameplays. Before starting the first game round, participants received a brief explanation of the playing field from the researcher, since the instructions before were in written form and did not include any visual references. After each gameplay, participants evaluated the recommendations via the UEQ-S and TiAQ. At the end of the study, participants were compensated with an equivalent of ca. USD 35 per hour for their time. The study was granted ethical approval according to our institutions' regulations.

5.2.2 Pilot Study. Before conducting the main user study, we carried out a pilot study with a colleague who self-reported a lot of experience with playing board games. The pilot study showed that a brief additional verbal explanation at the playing field was helpful to solidify participants' understanding prior to the gameplay and reduced the number of open questions during the session. Additional insights from the pilot study led to the provisioning of clearer indications of the current game phase in recommendations, and the enforcement of a more consistent recommendation structure.

5.3 Participants

The participants of our main study were required to not be familiar with the board game IAWW, while familiarity with board games in general was considered a plus. The participants were recruited through our University's platform and consisted of a total N=32 participants (17 men and 15 women), of which 12 (37.5%) were bachelor students, 15 (46.9%) master students, and 5 (15.6%) PhD candidates in various fields related to Business, Economics, Law, and Computer Science. Roughly half of the participants (17, 53.1%) were aged 18–24, and the other half (15, 46.9%) were aged 25–34. When asked to evaluate their familiarity with Augmented (AR) and Virtual Reality (VR)[12] on a 5-point Likert scale from 1="not at all familiar" to 5="extremely familiar", participants on average stated that their familiarity was 2.63 and 3.09, respectively, with a majority stating to have used AR or VR rarely or never; 3 participants out of 32 (9.4%) indicated using AR and VR a few times per month. With respect to familiarity with board games and video games on a 5-point Likert scale from 1="not at all familiar" to 5="extremely familiar", participants on average stated that their familiarity was 3.94 and 3.69 respectively. Out of 32 participants, 16 (50.0%) reported playing board games a few times per month, 9 (28.1%) rarely or never, 6 (18.8%) once per week, and 1 (3.1%) a few times per week. Regarding video games, 14 participants (43.8%) reported playing video games rarely or never, 11 (34.4%) a few times per month, 5 (15.6%) a few times per week, 1 (3.1%) once per week, and 1 (3.1%) every day. Regarding the tonality of the personalization 9 participants (28.1%) selected an analytical/neutral tone, 2 (6.3%) a formal/professional tone, 14 (43.8%) a supportive/friendly tone, and 7 (21.9%) an engaging/enthusiastic tone. Regarding recommendation language, 9 (28.1%) out of 32 participants indicated their preferred language to be German, while the remaining 23 (71.9%) preferred to receive recommendations in English.

5.4 Results

In this section, we present the main quantitative results of our user study as well as further exploratory observations we made while conducting the study.

5.4.1 Quantitative results.

Explanation Satisfaction. The ESQ results show significantly higher explanation satisfaction in the personalized condition (M = 3.88, SD = 0.70) compared to the generic condition (M = 3.55, SD = 0.70; t(31) = 2.66, p = .012, d = 0.47), which reflects a medium effect size based on HCI thresholds [51]. At item level, we observe significant differences in regard to the *availability of sufficient detail* (Mdiff = 0.47, p = .015, d = 0.46) and *completeness* (Mdiff = 0.38, p = .047, d = 0.36). Other items showed non-significant trends toward higher satisfaction in the personalized condition (all p > .05).

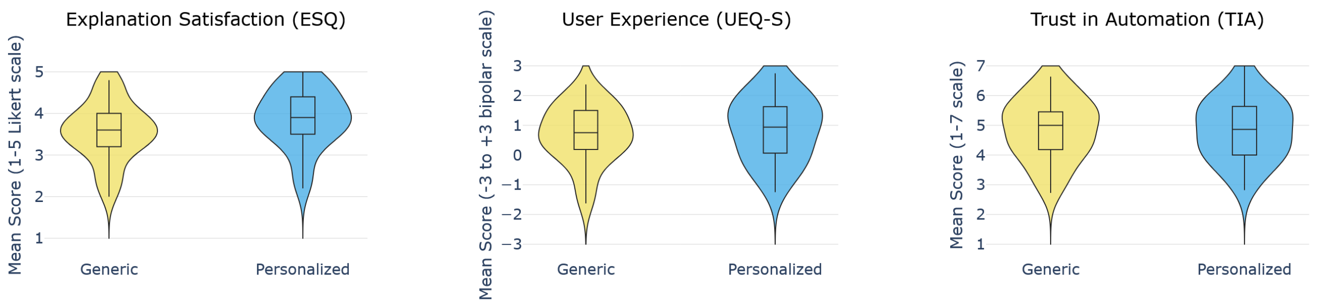
User Experience. The UEQ-S results show no significant difference for overall user experience between the personalized (M = 0.89, SD = 1.07) and generic (M = 0.69, SD = 0.97) conditions (t(31) = 1.40, p = .17, d = 0.25). As for the UEQ-S subscales, *hedonic* quality was rated significantly higher (t(31) = 2.65, p = .013, d = 0.47) in the personalized condition (M = 0.77, SD = 1.28) than in the generic condition (M = 0.20, SD = 1.45). However, *pragmatic* quality did not differ significantly between conditions (p = .33). Item-level analysis (see Figure 9) indicates a significant difference on the *usual/leading edge* item (Mdiff = 0.81, p = .002, d = 0.65), and non-significant trends on the items *boring/exciting* (towards the personalized condition) and *inefficient/efficient* (towards the generic condition). Descriptive patterns suggested that generic recommendations were perceived as somewhat easier, clearer, and more efficient, while personalized recommendations were perceived as more exciting, inventive, and leading-edge. These differences did not reach significance but align with qualitative interview insights.

Trust. The TiAQ results show no significant difference in overall trust between the personalized (M = 4.86, SD = 1.03) and generic (M = 4.86, SD = 1.03) conditions (t(31) = 0.02, p = .99, d < 0.01). Subscales for *positive trust* items (personalized: M = 4.50, SD = 1.09; generic: M = 4.53, SD = 1.17) and *distrust* items (reversed) (personalized: M = 5.30, SD = 1.25; generic: M = 5.26, SD = 1.16) also revealed no significant differences (all p > .80). Effect sizes were negligible. Small descriptive tendencies emerged at the item level: Generic recommendations were rated slightly higher on *perceived security*, *reliability*, and *reduced suspicion*, while personalized recommendations were seen as *less underhanded* and *less deceptive*. However, none of these item differences reached statistical significance (all p > .05).

Recommendation compliance. Exploratory analyses of the recommendation compliance showed no significant differences between conditions. Participants followed recommendations at comparable rates in the generic (M = 0.66, SD = 0.17) and personalized (M = 0.66, SD = 0.19) conditions (t(31) = 0.13, p = .90, d = 0.02). No significant change in following of recommendations across the two game sessions was observed (p = .078). Recommendation compliance also did not differ by board game experience level or gameplay order (all p > .20).

Game Performance. Game performance as exploratory measure did not significantly differ between recommender conditions across points scored ($M_{personalized}$ = 5.34, $M_{generic}$ = 5.44, p = .94), cards built ($M_{personalized}$ = 2.03, $M_{generic}$ = 2.16, p = .78) or rounds reached ($M_{personalized}$ = 4.97, $M_{generic}$ = 5.47, p = .21). It also did not significantly differ between Deck A and B for the recommender conditions (all p > .05). However, a strong learning effect was observed across sessions: participants scored more points (first: M = 2.94 vs. second: M = 7.84), built more cards (first: M = 1.06 vs.

[12] As the term Mixed Reality is not as prevalent in public discourse outside of academia, we used the terms AR and VR in the questionnaire to make them more relatable for the participants.



(a) (b) (c)

Figure 8: Overview of main results across recommender conditions. Violin plots show distribution and mean differences for (a) Explanation Satisfaction, (b) User Experience, and (c) Trust in Automation.

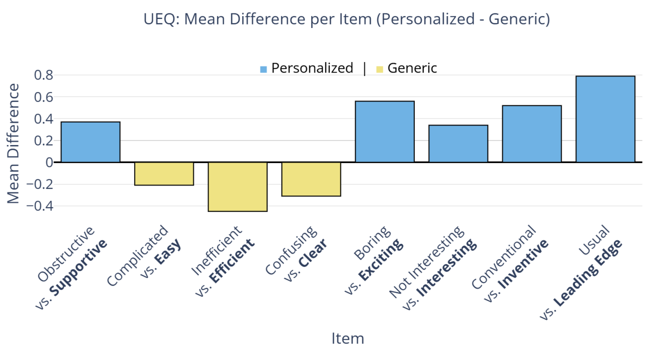


Figure 9: Item-level mean differences in the UEQ-S. Positive values indicate that participants preferred personalized recommendation. The difference in *usual/leading-edge* was significant.

second: M = 3.13), and reached later rounds (first: M = 4.31 vs. second: M = 6.13) in their second gameplay compared to the first (all $p < .001, d > 1.0$).

5.4.2 Further Observations. Many participants mentioned that the experience was enjoyable. Voiced opinions on personalized references differed, some described the recommendations as helpful for clarifying strategies, particularly when personalized references (e.g., cycling, investment) were perceived as relatable. Others expressed skepticism toward certain personalized references (e.g., when the system drew an analogy to the user's fondness of K-Dramas/Koreanovelas, where, according to GLAMRec, "narratives build up to a dramatic scene much like the construction of cards to a powerful engine"), which were considered either distracting or suspicious. A recurring pattern was that participants relied heavily on recommendations in the beginning but shifted toward trusting their own intuition and playing more independently with only doing cross-checks in the second round. In a few cases, participants reported unease when the recommendations seemed incorrect (e.g., missing to mention a card), whereas others were not bothered by this. They either played on their own or used the recommendations to verify their own intended moves. No clear differences in interaction behavior were observed between participants with lower vs. higher prior board game experience. While participants with higher board game familiarity sometimes appeared to gain confidence in the gameplay slightly earlier by stating that they had understood the game, both groups put in similar effort into understanding the game mechanics and engaging with the recommendations.

6 Discussion

H1—User Experience. In our study, personalization showed a significant impact on the *hedonic* quality of the user experience, indicating a more enjoyable experience due to the personalized condition. The hedonic effect reached significance at the conventional $\alpha = .05$ level, but falls just short of the threshold when applying a Bonferroni correction for multiple comparisons. We hence interpret with caution that personalization appears to enhance the user experience particularly in terms of the *perceived innovativeness* and *excitement*. However, the personalized condition did not result in improvements in the *pragmatic* quality of the user experience. Hypothesis H1, which states that participants report a better user experience in the personalized condition was therefore not supported as it did not improve aspects of pragmatic experience. The players also did not perform better across any game performance related measures such as score, built cards, or rounds reached, which would indicate better learning effects. This finding is contrary to prior findings of Shu et al., where personalized learning paths led to better performance outcomes [61]. This might be explained by the difference that participants in our study had to learn a completely new and unfamiliar board game and additionally mostly having little prior board game experience, whereas in the research before it was about participants improving their craft and being familiar with the domain. This could mean that in new learning experiences, personalization may not be as central and generic explainable recommendations may be sufficient for initial understanding. When comparing the first gameplay and second gameplay sessions, game performance improved. These improvements were consistent across both recommender conditions, indicating a practice effect associated with repeated gameplay rather than the influence of personalization.

H2—Explanation Satisfaction. Significantly higher explanation satisfaction was found in the personalized condition for the presented instructions. The effect was primarily driven by perceptions of sufficiency of detail and completeness, which suggests that richer contextualization enhances the perceived explanatory quality of recommendations. This aligns with previous qualitative findings, where personalized learning paths also led to higher satisfaction being expressed by the participants [61]. Although higher satisfaction was present in the personalized condition, our results show no significantly higher recommendation compliance, which we analyzed as an exploratory measure. There was also no indication that prior familiarity with board games affects participants' tendency to follow recommendations, which is consistent with the absence of moderation effects in the compliance analysis. The average recommendation compliance by users was nearly the same for the generic and personalized conditions across all games, which could potentially be related to the overall high trust in the system. Overall, hypothesis H2 was hence supported. Combining this result with our results on H1, we show a distinction between *perceived explanatory quality* and *actual effectiveness for learning as reflected by game performance*: While richer contextualization may make explanations *feel* more engaging and complete, effective learning outcomes seem to depend more strongly on clarity and generic explainability, as also suggested by the expert interviews and similar to our findings about game performance. These findings suggest that even though personalization increased satisfaction through more engaging explanations, it did not lead to greater behavioral reliance on the system, which is consistent with the findings regarding the game performance.

H3—Trust. No significant difference regarding trust between the personalized and generic condition was found. The hypothesis H3, stating that participants report higher trust in the personalized condition, is not supported. Since the trust ratings were high in both conditions, overall explainability seems to have established a sufficiently high baseline already, which aligns with previous research findings [9, 39]. Bernardo et al. [9] showed that trust is dependent on the perceived usefulness of a system, meaning correctly explaining a correct decision of the system—this would explain the similar ratings in our study, since we can assume that both conditions appropriately explained the action. Descriptive patterns indicate that generic recommendations are associated with slightly stronger perceptions of *security* and *reliability*, while personalized recommendations were viewed as marginally less *deceptive* and *underhanded.* This might point at greater relatability of the personalized condition, making it appear more trustworthy in these item measures. So, while personalization improved satisfaction, it did not significantly change the trust in the system.

6.1 Limitations

GLAMRec relies on PaddleOCR for text recognition, using an iPad Pro camera to record the game state in 4K resolution. While this setup provided overall reliable results, the system struggled with subtle variations in lighting and shadows. Despite studio lighting, some flickering effects remained and frame rate changes only partially mitigated these issues.

Furthermore, despite extensive prompting and refinement, the employed LLM occasionally produced errors, where it recommended multiple actions for an individual card or missed mentioning a card. The error rate was around 7.5%. This underlines that, even with careful design, it is not possible to eliminate all inaccuracies when using LLM-based recommendations within complex settings.

Personalization in GLAMRec combined multiple elements, including tone adaptation, personalized analogies, and contextual framing of explanations. These elements were applied jointly in the recommendations, which does not allow the present study to attribute observed benefits to individual personalization mechanisms. More controlled future studies could therefore systematically isolate specific personalization components to examine which primarily enhance enjoyment or which support understanding. GLAMRec personalization was also heavily dependent on the amount and specificity of the contextual data provided by the user. Furthermore, the LLM does not have exhaustive knowledge of all board games that users mentioned (or of other aspects of user profiles), which sometimes resulted in imprecise references. Given the diversity of games and players, there is a trade-off between giving broad personalized explanations, which may appear impersonal, and highly specific personalization, which may fail to resonate with a particular user. Both cases could potentially reduce the overall quality of the personalization.

GLAMRec focused on card-based game state tracking and did not capture every game element, such as the exact placement of the resource cubes. Tracking additional game assets could provide a more complete representation of the game state, but would potentially also require more hardware, more computational resources, and could hence lead to increase latency. Furthermore, since there were occasional errors with the current setting, adding another input layer to consider might lead to more errors.

Finally, a limitation of our user study was the sample size. Smaller effects with N=32 participants in our study might have gone undetected. Additionally, the two gameplays had a duration of 15 minutes each in the study, which aligns with the natural playtime of the chosen board game and was considered due to practical considerations for the user study. Further studies with a larger sample size or duration length could help to strengthen the findings; with respect to effects on long-term learning, a repeated study after several weeks or months might be relevant. Moreover, apart from game performance, this study did not include dedicated cognitive or procedural learning assessments; future work could incorporate such measures (e.g., knowledge tests or transfer tasks) to more directly capture learning outcomes beyond game performance.

6.2 Future Work

We argue that our proposed personalization approach can be translated to a variety of other settings, where several of the following conditions apply: (1) availability of some user data and/or context data (e.g., progress, state, environment), (2) a complex task space in which some actions are appropriate to take and others are inadvisable, and (3) need for domain-specific or procedural knowledge that can be integrated into the recommendation process. Since recommendations are grounded in ontological reasoning, this also supports responsible design: the system's decisions can be traced back and investigated by the user.

We specifically propose that GLAMRec's approach could be used in work or industry environments, such as an *assembly* scenario, where the system could track a worker's progress while also considering their training level (e.g., [28]). It could then support learning about the assembly process, suggesting the actions to take and which tools to use (or avoid). Similarly, in *maintenance* work, real-time sensor data from the machine as well as user data that includes the users' maintenance history and well-known other machines could be leveraged to advise on diagnostics, guiding the user on what to inspect and what issues can be ruled out—similar to GLAMRec by drawing analogies between other games the user knows well. The implicitness of the recommendations can be adjusted as well depending on the user scenario, to either provide full guidance or just subtly nudge towards the right direction. There is great potential for (automatic) adaptation of how the system supports learning and decision-making, what data can be applied, and for whom it is tailored, for instance in combination with ubiquitous task assistants, i.e., digital companions (cf. [19]). Such systems—and GLAMRec itself—would furthermore profit from the integration of user feedback, which we have not implemented for our implementation and study. Specifically, we believe that the integration of user feedback on recommendations could lead to stronger preference effects: the system could dynamically adjust the *level of personalization* based on learned user preferences—this is similar to the arguments on dynamic robot adaptation presented by Hostettler et al. [31]. We believe dynamic adaptation would address the mixed reception in H1 by giving frequent personalized references only to users who enjoy them.

In professional as well as home scenarios, building on the current implementation, GLAMRec could be extended to work for multi-person settings as well, where multiple people are playing, cooking, or assembling together (e.g., cf. [80] for collaborative assembly and cf. [32] for peer-scaffolding). The system state would then incorporate the actions of multiple individuals and reasoning would be extended to incorporate others' tasks and experiences as well, leading to a type of group adaptation that is personalized through individual interfaces. Past interaction patterns among the members of the group could also be taken into account, and the system could provide recommendations to help individuals make better strategic decisions as well as to learn collaboratively, while raising awareness of how their choices affect and are affected by others.

Finally, a fascinating option that we considered but chose to not implement for this study is what we refer to as *cross-learning* (drawing from a similar concept in machine learning, cf. [60]). GLAMRec might use user profile data to understand *what else* (other than IAWW gameplay) the user currently intends to learn about (e.g., from information about a student's upcoming exams). The system could then subtly integrate appropriate topics in gameplay explanations. While such integration clearly needs to be precisely calibrated, we believe that this idea might carry far and open fascinating potential for life-long cross-learning.

7 Conclusion

This paper presents GLAMRec, a recommender system that visualizes game-specific affordances based on the real-time game state and incorporates user data to provide personalized recommendations for supporting learning in a strategic engine-building board game. Our results show that, compared to generic recommendations, personalized recommendations lead to significantly higher explanation satisfaction; they further provide a significantly more enjoyable user experience by enhancing hedonic qualities such as novelty and stimulation during gameplay. However, personalization did not lead to greater trust in the system. These findings suggest that the value of personalization in recommender systems lies primarily in improving user engagement rather than in educational or trust-building outcomes; however, we believe that tuning the system's personalization level to individual users has the potential for pareto-improvement across both dimensions. We finally would like to highlight the modular and privacy-aware system design through Solid Pods, which can guide future implementations of adaptive recommender systems for learning environments and opens large potential for future research with respect to integration of user feedback, extension to further domains.

Acknowledgments

We thank Prof. Dr. Clemens Stachl for his advice on data processing and statistical analysis, and Alessandro Giugno for his advice on board game dynamics and the user study experience. This research is partially supported through a grant from the Research Committee of the HOCH Health Ostschweiz (Grant number 24/25).

References

[1] Neda Afreen. 2024. Explainable and Faithful Educational Recommendations through Causal Language Modelling via Knowledge Graphs. In *Proceedings of the 18th ACM Conference on Recommender Systems* (Bari, Italy) *(RecSys '24)*. Association for Computing Machinery, New York, NY, USA, 1358–1360. doi:10.1145/3640457.3688022

[2] Neda Afreen, Giacomo Balloccu, Ludovico Boratto, Gianni Fenu, Francesca Maridina Malloci, Mirko Marras, and Andrea Giovanni Martis. 2024. Learner-centered Ontology for Explainable Educational Recommendation. In *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization* (Cagliari, Italy) *(UMAP Adjunct '24)*. Association for Computing Machinery, New York, NY, USA, 567–575. doi:10.1145/3631700.3665226

[3] Jordan Barria-Pineda. 2020. Exploring the Need for Transparency in Educational Recommender Systems. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (Genoa, Italy) *(UMAP '20)*. Association for Computing Machinery, New York, NY, USA, 376–379. doi:10.1145/3340631.3398676

[4] Jordan Barria-Pineda, Kamil Akhuseyinoglu, Stefan Želem undefinedelap, Peter Brusilovsky, Aleksandra Klasnja Milicevic, and Mirjana Ivanovic. 2021. Explainable Recommendations in a Personalized Programming Practice System. In *Artificial Intelligence in Education: 22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14–18, 2021, Proceedings, Part I* (Utrecht, The Netherlands). Springer-Verlag, Berlin, Heidelberg, 64–76. doi:10.1007/978-3-030-78292-4_6

[5] Rebecca Yvonne Bayeck. 2020. Examining Board Gameplay and Learning: A Multidisciplinary Review of Recent Research. *Simulation & Gaming* 51, 4 (2020), 411–431. doi:10.1177/1046878119901286

[6] Kenan Bektaş, Jannis Strecker, Simon Mayer, Kimberly Garcia, Jonas Hermann, Kay Erik Jenß, Yasmine Sheila Antille, and Marc Solèr. 2023. GEAR: Gaze-enabled augmented reality for human activity recognition. In *Proceedings of the 2023 Symposium on Eye Tracking Research and Applications* (Tubingen, Germany) *(ETRA '23)*. Association for Computing Machinery, New York, NY, USA, Article 9, 9 pages. doi:10.1145/3588015.3588402

[7] Kenan Bektas. 2020. Toward A Pervasive Gaze-Contingent Assistance System: Attention and Context-Awareness in Augmented Reality. In *ACM Symposium on Eye Tracking Research and Applications* (Stuttgart, Germany) *(ETRA '20 Adjunct)*. Association for Computing Machinery, New York, NY, USA, Article 36, 3 pages. doi:10.1145/3379157.3391657

[8] Soulef Benhamdi, Abdesselam Babouri, and Raja Chiky. 2017. Personalized recommender system for e-Learning environment. *Education and Information Technologies* 22, 4 (2017), 1455–1477. doi:10.1007/s10639-016-9504-y

[9] Ezekiel L. Bernardo and Rosemary R. Seva. 2024. Exploration of Explainable AI for Trust Development on Human-AI Interaction. In *Proceedings of the 2023 6th Artificial Intelligence and Cloud Computing Conference* (Kyoto, Japan) *(AICCC '23)*. Association for Computing Machinery, New York, NY, USA, 238–246. doi:10.1145/3639592.3639625

[10] Bhaskar Bhattacharya and Eliot H. Winer. 2019. Augmented reality via expert demonstration authoring (AREDA). *Computers in Industry* 105 (2019), 61–79. doi:10.1016/j.compind.2018.04.021

[11] J. Bobadilla, F. Serradilla, and A. Hernando. 2009. Collaborative filtering adapted to recommender systems of e-learning. *Knowledge-Based Systems* 22, 4 (2009), 261–265. doi:10.1016/j.knosys.2009.01.008 Artificial Intelligence (AI) in Blended Learning.

[12] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. doi:10.1191/1478088706qp063oa

[13] Shruthi Chari, Oshani Seneviratne, Mohamed Ghalwash, Sola Shirai, Daniel M. Gruen, Pablo Meyer, Prithwish Chakraborty, and Deborah L. McGuinness. 2024. Explanation Ontology: A General-Purpose, Semantic Representation for Supporting User-Centered Explanations. *Semantic Web* 15, 4 (Oct. 2024), 959–989. doi:10.3233/SW-233282

[14] Jiahao Chen, D. Antony Chacon, Muhammad Bilal, Qiushi Zhou, and Wafa Johal. 2025. Mr.LfD: A Mixed Reality Interface for Robot Learning from Demonstration. In *Proceedings of the 36th Australasian Conference on Human-Computer Interaction (OzCHI '24)*. Association for Computing Machinery, New York, NY, USA, 275–285. doi:10.1145/3726986.3727004

[15] Henriette Cramer, Vanessa Evers, Satyan Ramlal, Maarten van Someren, Lloyd Rutledge, Natalia Stash, Lora Aroyo, and Bob Wielinga. 2008. The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling and User-Adapted Interaction* 18, 5 (2008), 455–496. doi:10.1007/s11257-008-9051-3

[16] Barney Dalgarno and Mark J. W. Lee. 2010. What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology* 41, 1 (2010), 10–32. doi:10.1111/j.1467-8535.2009.01038.x

[17] Maartje M.A. De Graaf and Bertram F. Malle. 2017. How people explain action (and autonomous intelligent systems should too). In *FS-17-01*, Vol. FS-17-01 - FS-17-05. AI Access Foundation, 19–26. 2017 AAAI Fall Symposium ; Conference date: 09-11-2017 Through 11-11-2017.

[18] A. Felfernig and R. Burke. 2008. Constraint-based recommender systems: technologies and research issues. In *Proceedings of the 10th International Conference on Electronic Commerce* (Innsbruck, Austria) *(ICEC '08)*. Association for Computing Machinery, New York, NY, USA, Article 3, 10 pages. doi:10.1145/1409540.1409544

[19] Kimberly Garcia, Jonathan Vontobel, and Simon Mayer. 2024. A Digital Companion Architecture for Ambient Intelligence. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 2, Article 66 (May 2024), 26 pages. doi:10.1145/3659610

[20] James Paul Gee. 2003. What video games have to teach us about learning and literacy. *Comput. Entertain.* 1, 1 (Oct. 2003), 20. doi:10.1145/950566.950595

[21] James Paul Gee. 2008. Video games and embodiment. *Games and culture* 3, 3-4 (2008), 253–263.

[22] James J. Gibson. 2014. *The Ecological Approach to Visual Perception: Classic Edition* (1st classic edition ed.). Psychology Press, New York, NY. doi:10.4324/9781315740218

[23] Carlos A. Gomez-Uribe and Neil Hunt. 2016. The Netflix Recommender System: Algorithms, Business Value, and Innovation. *ACM Trans. Manage. Inf. Syst.* 6, 4, Article 13 (Dec. 2016), 19 pages. doi:10.1145/2843948

[24] Jan Grau, Simon Mayer, Jannis Strecker, Kimberly Garcia, and Kenan Bektas. 2024. Gaze-based Opportunistic Privacy-preserving Human-Agent Collaboration. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 176, 6 pages. doi:10.1145/3613905.3651066

[25] Mouadh Guesmi, Mohamed Amine Chatti, Shoeb Joarder, Qurat Ul Ain, Rawaa Alatrash, Clara Siepmann, and Tannaz Vahidi. 2023. Interactive Explanation with Varying Level of Details in an Explainable Scientific Literature Recommender System. arXiv:2306.05809 [cs.IR] https://arxiv.org/abs/2306.05809

[26] Xue Guo, Xiangchun He, and Zhuoyun Pei. 2024. Data-driven Personalized Learning. In *Proceedings of the 2023 6th International Conference on Educational Technology Management* (Guangzhou, China) *(ICETM '23)*. Association for Computing Machinery, New York, NY, USA, 49–54. doi:10.1145/3637907.3637988

[27] Muhammad Naufal Sinai Harjana, Hanley Yunanda Saputra, and Cuk Tho. 2023. A Review of the Potential Use of Mixed Reality Learning Methods in Comparison to Traditional Learning Methods. *Procedia Computer Science* 227 (2023), 734–742. doi:10.1016/j.procs.2023.10.578 8th International Conference on Computer Science and Computational Intelligence (ICCSCI 2023).

[28] Michael Haslgrübler, Benedikt Gollan, and Alois Ferscha. 2018. A Cognitive Assistance Framework for Supporting Human Workers in Industrial Tasks. *IT Professional* 20, 5 (2018), 48–56. doi:10.1109/MITP.2018.053891337

[29] Drew Hicks, Zhongxiu Liu, Michael Eagle, and Tiffany Barnes. 2016. Measuring Gameplay Affordances of User-Generated Content in an Educational Game. In *Proceedings of the 9th International Conference on Educational Data Mining (EDM 2016)*. International Educational Data Mining Society, 78–85.

[30] Robert R. Hoffman, Shane T. Mueller, Gary Klein, and Jordan Litman. 2023. Measures for explainable AI: Explanation goodness, user satisfaction, mental models, curiosity, trust, and human-AI performance. *Frontiers in Computer Science* 5 - 2023 (2023). doi:10.3389/fcomp.2023.1096257

[31] Damian Hostettler, Simon Mayer, and Christian Hildebrand. 2022. Human-Like Movements of Industrial Robots Positively Impact Observer Perception. *International Journal of Social Robotics* (Dec 2022). doi:10.1007/s12369-022-00954-2

[32] Huei-Tse Hou, Ying-Sang Fang, and Joni Tzuchen Tang. 2023. Designing an alternate reality board game with augmented reality and multi-dimensional scaffolding for promoting spatial and logical ability. *Interactive Learning Environments* 31, 7 (2023), 4346–4366. doi:10.1080/10494820.2021.1961810

[33] Michael Ion, Dimitris Sacharidis, and Hannes Werthner. 2020. Designing a recommender system for board games. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing* (Brno, Czech Republic) *(SAC '20)*. Association for Computing Machinery, New York, NY, USA, 1465–1467. doi:10.1145/3341105.3375780

[34] Melissa J. Rogerson, Lucy A. Sparrow, and Martin R. Gibbs. 2021. Unpacking "Boardgames with Apps": The Hybrid Digital Boardgame Model. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 111, 17 pages. doi:10.1145/3411764.3445077

[35] Dietmar Jannach and Kolja Hegelich. 2009. A case study on the effectiveness of recommendations in the mobile internet. In *Proceedings of the Third ACM Conference on Recommender Systems* (New York, New York, USA) *(RecSys '09)*. Association for Computing Machinery, New York, NY, USA, 205–208. doi:10.1145/1639714.1639749

[36] Jiun-Yin Jian, Ann Bisantz, and Colin Drury. 2000. Foundations for an Empirically Determined Scale of Trust in Automated Systems. *International Journal of Cognitive Ergonomics* 4 (03 2000), 53–71. doi:10.1207/S15327566IJCE0401_04

[37] Shantao Jiang, Yiping Wen, Jun Shen, Gaoxian Peng, Guosheng Kang, and Jianxun Liu. 2025. Personalized Learning Path Recommendation with Time-Aware Attention-Based Reinforcement Learning. *ACM Trans. Intell. Syst. Technol.* 16, 5, Article 112 (Sept. 2025), 24 pages. doi:10.1145/3747594

[38] Frank C. Keil. 2006. Explanation and understanding. *Annual Review of Psychology* 57 (2006), 227–254. doi:10.1146/annurev.psych.57.102904.190100

[39] René F. Kizilcec. 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* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 2390–2395. doi:10.1145/2858036.2858402

[40] Bart P. Knijnenburg, Martijn C. Willemsen, Zeno Gantner, Hakan Soncu, and Chris Newell. 2012. Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction* 22, 4 (2012), 441–504. doi:10.1007/s11257-011-9118-4

[41] Xiangdong Li, Yunzhan Zhou, Wenqian Chen, Preben Hansen, Weidong Geng, and Lingyun Sun. 2019. 8. Towards personalized virtual reality touring through cross-object user interfaces. In *Personalized Human-Computer Interaction*, Mirjam Augstein, Eelco Herder, and Wolfgang Wörndl (Eds.). De Gruyter, Berlin, Boston.

[42] Jianghao Lin, Xinyi Dai, Yunjia Xi, Weiwen Liu, Bo Chen, Hao Zhang, Yong Liu, Chuhan Wu, Xiangyang Li, Chenxu Zhu, Huifeng Guo, Yong Yu, Ruiming Tang, and Weinan Zhang. 2025. How Can Recommender Systems Benefit from Large Language Models: A Survey. *ACM Trans. Inf. Syst.* 43, 2, Article 28 (Jan. 2025), 47 pages. doi:10.1145/3678004

[43] Shi Liu, Peyman Toreini, and Alexander Maedche. 2022. Designing Gaze-Aware Attention Feedback for Learning in Mixed Reality. In *Proceedings of Mensch Und Computer 2022* (Darmstadt, Germany) *(MuC '22)*. Association for Computing Machinery, New York, NY, USA, 503–508. doi:10.1145/3543758.3547565

[44] Tania Lombrozo. 2006. The structure and function of explanations. *Trends in Cognitive Sciences* 10, 10 (2006), 464–470. doi:10.1016/j.tics.2006.08.004

[45] Brian Mayer and Christopher Harris. 2010. *Libraries Got Game: Aligned Learning Through Modern Board Games*. American Library Association, Chicago. p. cm. pages.

[46] Eray Molla and Vincent Lepetit. 2010. Augmented reality for board games. In *2010 IEEE International Symposium on Mixed and Augmented Reality*. 253–254. doi:10.1109/ISMAR.2010.5643593

[47] Victor Emanuel Montes Moreira, Lucas Gregory Gomes Almeida, Márcio Fontana Catapan, Daniella Rosito Michelena Munhoz, André Leonardo Demaison Medeiros Maia, and Ingrid Winkler. 2024. Exploring Mixed Reality in Digital Board Games: A Comparative Analysis of Cooperative and Competitive Modes. In *Proceedings of the 26th Symposium on Virtual and Augmented Reality* (Manaus, Brazil) *(SVR '24)*. Association for Computing Machinery, New York, NY, USA, 71–79. doi:10.1145/3691573.3691583

[48] Mohammad Naiseh, Deniz Cemiloglu, Dena Al Thani, Nan Jiang, and Raian Ali. 2021. Explainable Recommendations and Calibrated Trust: Two Systematic User Errors. *Computer* 54, 10 (2021), 28–37. doi:10.1109/MC.2021.3076131

[49] Ayano Okoso, Keisuke Otaki, Satoshi Koide, and Yukino Baba. 2025. Impact of Tone-Aware Explanations in Recommender Systems. *ACM Trans. Recomm. Syst.* 3, 4, Article 55 (April 2025), 34 pages. doi:10.1145/3718101

[50] Ayano Okoso, Mingzhe Yang, and Yukino Baba. 2025. Do Expressions Change Decisions? Exploring the Impact of AI's Explanation Tone on Decision-Making. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 824, 22 pages. doi:10.1145/3706598.3713744

[51] Anna-Marie Ortloff, Florin Martius, Mischa Meier, Theo Raimbault, Lisa Geierhaas, and Matthew Smith. 2025. Small, Medium, Large? A Meta-Study of Effect Sizes at CHI to Aid Interpretation of Effect Sizes and Power Calculation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 483, 28 pages. doi:10.1145/3706598.3713671

[52] Maria Perez-Ortiz, Claire Dormann, Yvonne Rogers, Sahan Bulathwela, Stefan Kreitmayer, Emine Yilmaz, Richard Noss, and John Shawe-Taylor. 2021. X5Learn: A Personalised Learning Companion at the Intersection of AI and HCI. In *Companion Proceedings of the 26th International Conference on Intelligent User Interfaces* (College Station, TX, USA) *(IUI '21 Companion)*. Association for Computing Machinery, New York, NY, USA, 70–74. doi:10.1145/3397482.3450721

[53] Ruth Pinedo, Noelia García-Martín, Débora Rascón, César Caballero-San José, and Manuel Cañas. 2022. Reasoning and learning with board game-based learning: A case study. *Current Psychology* 41, 3 (2022), 1603–1617. doi:10.1007/s12144-021-01744-1

[54] Francesco Ricci, Lior Rokach, and Bracha Shapira. 2011. Introduction to Recommender Systems Handbook. In *Recommender Systems Handbook*, Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor (Eds.). Springer, New York, NY, 1–35. doi:10.1007/978-0-387-85820-3_1

[55] Mojtaba Salehi and Isa Nakhai Kmalabadi. 2012. A Hybrid Attribute–based Recommender System for E–learning Material Recommendation. *IERI Procedia* 2 (2012), 565–570. doi:10.1016/j.ieri.2012.06.135 International Conference on Future Computer Supported Education, August 22- 23, 2012,Fraser Place Central - Seoul.

[56] Andrei Vlad Sambra, Essam Mansour, Sandro Hawke, Maged Zereba, Nicola Greco, Abdurrahman Ghanem, Dmitrij Zagidulin, Ashraf Aboulnaga, and Tim Berners-Lee. 2016. Solid : A Platform for Decentralized Social Applications Based on Linked Data. https://api.semanticscholar.org/CorpusID:49564404

[57] Mike Scaife and Yvonne Rogers. 1996. External cognition: how do graphical representations work? *International Journal of Human-Computer Studies* 45, 2 (Aug. 1996), 185–213. doi:10.1006/ijhc.1996.0048

[58] J. Ben Schafer, Joseph Konstan, and John Riedl. 1999. Recommender systems in e-commerce. In *Proceedings of the 1st ACM Conference on Electronic Commerce* (Denver, Colorado, USA) *(EC '99)*. Association for Computing Machinery, New York, NY, USA, 158–166. doi:10.1145/336992.337035

[59] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2017. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence* 4 (01 2017), 103. doi:10.9781/ijimai.2017.09.001

[60] Artemios-Anargyros Semenoglou, Evangelos Spiliotis, Spyros Makridakis, and Vassilios Assimakopoulos. 2021. Investigating the accuracy of cross-learning time series forecasting methods. *International Journal of Forecasting* 37, 3 (2021), 1072–1084. doi:10.1016/j.ijforecast.2020.11.009

[61] Xin Shu, Lei Shi, Jiacheng Cheng, Lingling Ouyang, Mengdi Chu, and Xinhuan Shu. 2025. FretMate: ChatGPT-Powered Adaptive Guitar Learning Assistant. In *Proceedings of the 30th International Conference on Intelligent User Interfaces (IUI '25)*. Association for Computing Machinery, New York, NY, USA, 715–726. doi:10.1145/3708359.3712080

[62] Ítallo Silva, Leandro Marinho, Alan Said, and Martijn C. Willemsen. 2024. Leveraging ChatGPT for Automated Human-centered Explanations in Recommender Systems. In *Proceedings of the 29th International Conference on Intelligent User Interfaces* (Greenville, SC, USA) *(IUI '24)*. Association for Computing Machinery, New York, NY, USA, 597–608. doi:10.1145/3640543.3645171

[63] Rashmi Sinha and Kirsten Swearingen. 2002. The role of transparency in recommender systems. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems* (Minneapolis, Minnesota, USA) *(CHI EA '02)*. Association for Computing Machinery, New York, NY, USA, 830–831. doi:10.1145/506443.506619

[64] Maximilian Speicher, Brian D. Hall, and Michael Nebeling. 2019. What is Mixed Reality?. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3290605.3300767

[65] Jannis Strecker, Khakim Akhunov, Federico Carbone, Kimberly García, Kenan Bektaş, Andres Gomez, Simon Mayer, and Kasim Sinan Yildirim. 2023. MR Object Identification and Interaction: Fusing Object Situation Information from Heterogeneous Sources. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 3, Article 124 (Sept. 2023), 26 pages. doi:10.1145/3610879

[66] Jannis Strecker, Kimberly García, Kenan Bektaş, Simon Mayer, and Ganesh Ramanathan. 2023. SOCRAR: Semantic OCR through Augmented Reality. In *Proceedings of the 12th International Conference on the Internet of Things* (Delft, Netherlands) *(IoT '22)*. Association for Computing Machinery, New York, NY, USA, 25–32. doi:10.1145/3567445.3567453

[67] Jannis Strecker, Simon Mayer, and Kenan Bektaş. 2025. Towards Societally Beneficial Personalized Realities: A Conceptual Foundation for Responsible Ubiquitous Personalization Systems. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference (DIS '25)*. Association for Computing Machinery, New York, NY, USA, 1792–1814. doi:10.1145/3715336.3735709

[68] Matilda Ståhl, Katri Hansell, Sandra Bäck, and Mattias Wingren. 2025. Affordances for In-Game Interaction and Language Learning Through Children's Collaborative Play in Minecraft. *International Journal of Game-Based Learning* 15, 1 (2025). doi:10.4018/IJGBL.370559

[69] Chun-Hua Tsai and Peter Brusilovsky. 2019. Explaining recommendations in an interactive hybrid social recommender. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (Marina del Ray, California) *(IUI '19)*. Association for Computing Machinery, New York, NY, USA, 391–396. doi:10.1145/3301275.3302318

[70] Joseph Tu, Eugene Kukshinov, Reza Hadi Mogavi, Derrick M. Wang, and Lennart E. Nacke. 2025. Designing Biofeedback Board Games: The Impact of Heart Rate on Player Experience. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 466, 21 pages. doi:10.1145/3706598.3713543

[71] Katrien Verbert, Nikos Manouselis, Xavier Ochoa, Martin Wolpers, Hendrik Drachsler, Ivana Bosnic, and Erik Duval. 2012. Context-Aware Recommender Systems for Learning: A Survey and Future Challenges. *IEEE Transactions on Learning Technologies* 5, 4 (2012), 318–335. doi:10.1109/TLT.2012.11

[72] Jari Vesanen. 2007. What is personalization? A conceptual framework. *European Journal of Marketing* 41, 5/6 (June 2007). doi:10.1108/03090560710737534

[73] W3C. 2014. RDF 1.1 Turtle - Terse RDF Triple Language. https://www.w3.org/TR/turtle/.

[74] Hao Wang, Wei Guo, Luankang Zhang, Jin Yao Chin, Yufei Ye, Huifeng Guo, Yong Liu, Defu Lian, Ruiming Tang, and Enhong Chen. 2025. Generative Large Recommendation Models: Emerging Trends in LLMs for Recommendation. In *Companion Proceedings of the ACM on Web Conference 2025* (Sydney NSW, Australia) *(WWW '25)*. Association for Computing Machinery, New York, NY, USA, 49–52. doi:10.1145/3701716.3715865

[75] Shu-Lin Wang and Chun-Yi Wu. 2011. Application of context-aware and personalized recommendation to implement an adaptive ubiquitous learning system. *Expert Systems with Applications* 38, 9 (2011), 10831–10838. doi:10.1016/j.eswa.2011.02.083

[76] Jessica Wen, Angela Zavaleta Bernuy, Naaz Sibia, Andrew Petersen, and Michael Liut. 2025. Enhancing Self-Explanation in Student Learning Through Large Language Models. In *Proceedings of the 30th ACM Conference on Innovation and Technology in Computer Science Education V. 2* (Nijmegen, Netherlands) *(ITiCSE 2025)*. Association for Computing Machinery, New York, NY, USA, 762. doi:10.1145/3724389.3730790

[77] Chuhan Wu, Fangzhao Wu, Yongfeng Huang, and Xing Xie. 2023. Personalized News Recommendation: Methods and Challenges. *ACM Trans. Inf. Syst.* 41, 1, Article 24 (Jan. 2023), 50 pages. doi:10.1145/3530257

[78] Yuansong Xu, Yuheng Shao, Jiahe Dong, Shaohan Shi, Chang Jiang, and Quan Li. 2025. Advancing Problem-Based Learning with Clinical Reasoning for Improved Differential Diagnosis in Medical Education. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 1022, 32 pages. doi:10.1145/3706598.3713772

[79] Chenggong Zhai, Yan Li, Xinguang Yuan, Xinrong Cao, Xiaodong Li, and Yaoqi Zhang. 2024. On the Reform of Education Methods under the Background of Mixed Reality Technology. In *Proceedings of the 2024 2nd Asia Conference on Computer Vision, Image Processing and Pattern Recognition* (Xiamen, China) *(CVIPPR '24)*. Association for Computing Machinery, New York, NY, USA, Article 37, 6 pages. doi:10.1145/3663976.3664022

[80] Yan Xiang Zhang, Jia Qi Cheng, Jia Yu Wang, and Lei Zhao. 2021. Co-assemble: A collaborative AR cross-devices teaching system for assemble practice courses. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 512–513. doi:10.1109/VRW52623.2021.00138

A Recommendation Example

Table 3: Example of a generic and personalized recommendation (personalized for a participant passionate about investing with finance background; tone: analytical/neutral). Differences are marked in *italics*.

Zone	Generic	Personalized
Constructed Zone	The Constructed Zone holds fully constructed cards that generate resources. The goal is to maximize Victory Points (VP) through efficient resource production and scoring. **Current Round:** 1, **Sequence:** 2. **Status:** 1 Empire card constructed. Produces a total of 1 material and 2 gold. No conditional bonuses currently active. Focus on constructing Project cards for the 2× VP bonus. Prioritize constructing more yellow Project cards to synergize with the current Empire card. No supremacy bonuses are active since production is below 5 for any resource.	The Constructed Zone holds fully constructed cards that generate resources. *Given your background in economics and financial services, think of this zone as your investment portfolio where the goal is to maximize your returns—Victory Points (VP)—through efficient resource production and scoring.* **Current Round:** 1, **Sequence:** 2. **Status:** 1 Empire card constructed. Produces a total of 1 material and 2 gold. No conditional bonuses currently active. Focus on constructing Project cards for the 2× VP bonus. Prioritize constructing more yellow Project cards to synergize with the current Empire card. No supremacy bonuses are active since production is below 5 for any resource.
Construction Zone	The Construction Zone contains cards selected for construction that need to be built as soon as possible. Focus on building cards that enhance resource production. Recommended to construct: **Magnetic Train** — Requires 1 energy, 1 science, and 3 gold. Produces 1 gold and synergizes with Structures for additional gold production. **Propaganda Center** — Requires 3 gold. Produces 1 gold and synergizes with Project cards for additional gold production. Prioritize these cards due to their resource generation potential and synergy with the constructed Empire card.	The Construction Zone contains cards selected for construction that need to be built as soon as possible. *Consider this zone your strategic planning phase, akin to analyzing investment opportunities in your financial internships. Focus on building cards that yield resources.* Recommended to construct: **Magnetic Train** — Requires 1 energy, 1 science, and 3 gold. Produces 1 gold and synergizes with Structures for additional gold production. **Propaganda Center** — Requires 3 gold. Produces 1 gold and synergizes with Project cards for additional gold production. Prioritize these cards due to their resource generation potential and synergy with the constructed Empire card.
Draft Zone	The Draft Zone contains randomly drawn cards that can be constructed, recycled, or discarded. Focus on selecting cards that will enhance your strategy. Construct: **Casino City** — Produces 2 gold and fits the Project strategy for 2× VP. Recycle: **Unknown Technology** — Recycling provides science, which is needed for constructing other cards. **Airborne Laboratory** — Recycling provides science, which is also beneficial for future constructions. Discard: **Mega-Drill** **Tank Division** — Both cards do not align with the current strategy and can be discarded to draw new options.	The Draft Zone contains randomly drawn cards that can be constructed, recycled, or discarded. *This is similar to diversifying your investment portfolio; focus on selecting cards that will enhance your strategy by either having valuable resources or scoring potential.* Construct: **Casino City** — Produces 2 gold and fits the Project strategy for 2× VP. Recycle: **Unknown Technology** — Recycling provides science, which is needed for constructing other cards. **Airborne Laboratory** — Recycling provides science, which is also beneficial for future constructions. Discard: **Mega-Drill** **Tank Division** — Both cards do not align with the current strategy and can be discarded to draw new options. *This strategy will optimize your resource production and set up for future rounds effectively, much like building a well-rounded investment strategy.*

B Personalization Questionnaire for the Expert Interviews

This is the English version of the Personalization Questionnaire for the Experts. If they wished, they could also choose to fill in a German version.

(1) **Demographics. This section won't be used for personalization.**
- What is your name?
- What is your gender? Male \| Female \| Non-binary \| Prefer not to say
- What is your age? Under 18 \| 18–24 \| 25–34 \| 35–44 \| 45–54 \| 55–64 \| 65–74 \| 75+
- How many and which games have you designed (published)?
- How many years have you been designing board games?

(2) **Personalization Section The following section will be used to personalize the gameplay prototype for your interview.**

(2.1) **Games You Know. Knowledge about the known games, helps to reference known mechanics and co.**
- Which board games have you played before and are familiar with? Please list up to 10 board games you know well enough to understand the basic rules or strategies.
- Which video games (if any) have you played before and are familiar with?Please list up to 10 video games you've played before (on PC, console, or any other platform).

(2.2) **Academic Background. Understanding your academic background allows the system to draw references from familiar concepts and domains.**
- What is the highest level of education you have completed? Please select the option that best describes your highest completed level of education. No education beyond compulsory schooling | Upper secondary: Vocational education (e.g. apprenticeship) | Upper secondary: General education (e.g. high school diploma) | Tertiary: Professional college / advanced vocational education | Tertiary: University / academic higher education
- Any further details you'd like to share? (e.g., your specific specialization, degree program, research focus)

(2.3) **Professional Experience. Your professional experience helps the system relate recommendations to relevant fields or work-related contexts.**
- What are your previous work experiences/current background? Choose one or more sectors that best describe your work experience. Agriculture, Forestry, and Fishing | Mining, Stones and Earth | Food, Beverages, and Tobacco | Textiles and Clothing | Leather, Leather Goods, and Shoes | Wood, Wicker, Basket, and Cork Products | Paper and Printing Industry | Chemical Industry, Petroleum Processing | Rubber and Plastic Products | Glass, Ceramics, Cement Products | Metal Production and Metal Products | Electrical Engineering, Electronics, Watches, Optics | Mechanical Engineering | Vehicle Manufacturing | Furniture; Repair of Machinery | Energy Supply | Recycling; Water Supply Construction Industry | Trade; Motor Vehicle Trade and Repair | Motor Vehicle Trade and Repair | Wholesale | Retail | Transport and Logistics | Hospitality Industry | Information and Communication | Computer Science; IT Services | Financial and Insurance Services | Banks | Insurance | Real Estate | Professional, Technical and Scientific Services | Research and Development | Other Economic Services | Public Administration, Social Insurance | Education | Health and Social Services | Arts, Entertainment and Recreation | Other Services | Private Households | Other:
- Would you like to add more detail about your professional experience? (e.g., job titles, relevant projects, industry specializations)

(2.4) **Media Preferences Media preferences can be used to reference familiar themes or narratives in the recommendations.**
- Are there any specific media or entertainment you enjoy?(e.g., movies, genres, TV shows, books, comics). Feel free to list as many as you like and names — the more detailed, the better!

(2.5) **Passions. If you're passionate about certain topics, hobbies or interests, the system can incorporate them into the experience.**
- What are you passionate about? (e.g., topics, hobbies or interests)

(2.6) **Recommender Tone. This helps personalize how the system communicates its suggestions in a tone that feels natural or effective to you.**
- Do you have a tone preference for the recommender system's communication? coach/motivational | analytical/neutral
- What language of the recommendations do you prefer? English| German

C Interview Guideline for the Expert Interviews

Before showing the system:
- Are you already familiar with the game *It's a Wonderful World*?
- How do you typically approach board games yourself?
  → Do you read the rulebook and try it out right away?
  → Do you watch reviews or gameplay videos beforehand?
  → Or do you just start playing without preparation?

While showing the system (if the expert comments spontaneously):
- What is your impression?
- What do you think about the system's generated recommendation and personalization?
- Does the reasoning make sense to you?
- At first glance, does the recommendation seem understandable to you?
- To what extent do you think this recommendation would genuinely help a player make a decision, or would it rather cause confusion?

After showing the system:
Perception
- How did you generally perceive the system?
- Was it clear to you how the recommendations were generated?
- How did you find the transparency of the decision-making logic?

Personalization
- How appropriate or helpful did you find the personalized recommendations in relation to your background?
- Were there recommendations that became clearer or more compelling to you because of the personalized explanation?
- Were there situations in which you found the personalization excessive, unnecessary, or
- Did you feel "represented" in the system's language and reasoning? (e.g., through profession, interests, etc.)
- Do you think personalization is generally useful?

Gameplay Flow
- Did you feel that the system supported the flow of the game, or did it rather disrupt the experience?
- How would you assess the usefulness of such a system for players who don't know the game yet?

- Do you think such a system would help a complete board game beginner?

Assessment

- From your perspective as a game developer, how do you view the idea of giving players personalized and transparent recommendations?
- Did you feel the system helped you understand the game and its mechanics (better)?
- Do you think such a system could be especially useful for more complex games, as orientation or for learning?
- Do you see any risks or opportunities for the player experience — for example, in terms of fun, challenge, or learning?
- What would you improve or do differently in this approach?
- Do you think players feel more engaged or understand what to do faster if analogies are used (e.g., economic terms for business students)?
- What types of user data would you consider useful? Are there data you would add, remove, or prioritize?
- How important do you think it is that a recommendation system explains why it makes certain suggestions?

Reflection

- Is there anything that stood out to you, positively or negatively, that we haven't talked about yet?
- Would you do anything differently when developing a game if you knew such a system was going to be used?

D Questionnaires for the User Study

D.1 Demographic Questionnaire

(1) Please indicate the date and time of your appointment for this experiment.
(2) What nickname will you use? Please choose a nickname and use it the same way in all questionnaires and throughout the user study. This helps us match your responses correctly.
(3) What is your gender? Male | Female | Non-binary | Prefer not to say
(4) What is your age? Under 18 | 18–24 | 25–34 | 35–44 | 45–54 | 55–64 | 65–74 | 75+
(5) What is your current level of study or academic position? Bachelor student | Master student | PhD student | Other
(6) What is your field of study? (B=Bachelor's, M=Master's) Business Administration (B) | Economics (B) | Law (B) | Law & Economics (B) | International Affairs (B) | Computer Science (B) | Strategy and International Management (M) | Banking & Finance (M) | Business Innovation (M) | General Management (M) | Marketing Management (M) | Accounting & Corporate Finance (M) | International Affairs & Governance (M) | Economics (M) | Quantitative Economics & Finance (M) | Computer Science (M) | Other
(7) How familiar are you with Augmented Reality? 5-Point Likert scale: Not at all familiar | Slightly familiar | Somewhat familiar |Fairly Familiar | Extremely Familiar
(8) How familiar are you with Virtual Reality? 5-Point Likert scale: Not at all familiar | Slightly familiar | Somewhat familiar |Fairly Familiar | Extremely Familiar
(9) How often do you use Augmented Reality? 5-Point Likert scale: Every Day | A Few Times a Week | Once a Week | A Few Times a Month |Rarely or Never
(10) How often do you use Virtual Reality? 5-Point Likert scale: Every Day | A Few Times a Week | Once a Week | A Few Times a Month |Rarely or Never
(11) How familiar are you with Board Games? 5-Point Likert scale: Not at all familiar | Slightly familiar | Somewhat familiar |Fairly Familiar | Extremely Familiar
(12) How familiar are you with Video Games? 5-Point Likert scale: Not at all familiar | Slightly familiar | Somewhat familiar |Fairly Familiar | Extremely Familiar
(13) How often do you play Board Games? 5-Point Likert scale: Every Day | A Few Times a Week | Once a Week | A Few Times a Month |Rarely or Never
(14) How often do you play Video Games? 5-Point Likert scale: Every Day | A Few Times a Week | Once a Week | A Few Times a Month |Rarely or Never

D.2 Personalization Questionnaire

- What is your nickname?
  Games You Know. Knowledge about the known games, helps to reference known mechanics and co.
- Which board games have you played before and are familiar with? Please list up to 10 board games you know well enough to understand the basic rules or strategies.
- Which video games (if any) have you played before and are familiar with?Please list up to 10 video games you've played before (on PC, console, or any other platform).
  Academic Background. Understanding your academic background allows the system to draw references from familiar concepts and domains.
- What is the highest level of education you have completed? Please select the option that best describes your highest completed level of education. No education beyond compulsory schooling | Upper secondary: Vocational education (e.g. apprenticeship) | Upper secondary: General education (e.g. high school diploma) | Tertiary: Professional college / advanced vocational education | Tertiary: University / academic higher education
- Please share some further details. (e.g., your specific specialization, degree program, research focus)

(2.3) **Professional Experience. Your professional experience helps the system relate recommendations to relevant fields or work-related contexts.**

- What are your previous work experiences/current background? Choose one or more sectors that best describe your work experience. Agriculture, Forestry, and Fishing | Mining, Stones and Earth | Food, Beverages, and Tobacco | Textiles and Clothing | Leather, Leather Goods, and Shoes | Wood, Wicker, Basket, and Cork Products | Paper and Printing Industry | Chemical Industry, Petroleum Processing | Rubber and Plastic Products | Glass, Ceramics, Cement Products | Metal Production and Metal Products | Electrical Engineering, Electronics, Watches, Optics | Mechanical Engineering | Vehicle Manufacturing | Furniture; Repair of Machinery |

Energy Supply | Recycling; Water Supply Construction Industry | Trade; Motor Vehicle Trade and Repair | Motor Vehicle Trade and Repair | Wholesale | Retail | Transport and Logistics | Hospitality Industry | Information and Communication | Computer Science; IT Services | Financial and Insurance Services | Banks | Insurance | Real Estate | Professional, Technical and Scientific Services | Research and Development | Other Economic Services | Public Administration, Social Insurance | Education | Health and Social Services | Arts, Entertainment and Recreation | Other Services | Private Households | Other:

- Please share some more detail about your professional experience. (e.g., job titles, relevant projects, industry specializations)

(2.4) **Media Preferences Media preferences can be used to reference familiar themes or narratives in the recommendations.**

- Are there any specific media or entertainment you enjoy?(e.g., movies, genres, TV shows, books, comics). Feel free to list as many as you like and names — the more detailed, the better!

(2.5) **Passions. If you're passionate about certain topics, hobbies or interests, the system can incorporate them into the experience.**

- What are you passionate about?

(2.6) **Recommender Tone. This helps personalize how the system communicates its suggestions in a tone that feels natural or effective to you.**

- Do you have a tone preference for the recommender system's communication? Analytical/Neutral | Formal/Professional | Supportive/Friendly | Engaging/Enthusiastic
- What language of the recommendations do you prefer? English| German

D.3 Additional Questionnaires used in the User Study

D.3.1 UEQ-Short. For the short version of the User Experience Questionnaire (UEQ-S) [59], for each pair of opposite terms, participants select the point on a 7-point Likert scale that best reflects their impression (1 =left term, 7 = right term).

- Obstructive vs. Supportive
- Complicated vs. Easy
- Inefficient vs. Efficient
- Confusing vs. Clear
- Boring vs. Exciting
- Not interesting vs. Interesting
- Conventional vs. Inventive
- Usual vs. Leading edge

D.3.2 Explanation Satisfaction. For the Explanation Satisfaction questionnaire [30] participants state their agreement on a 5-Likert scale, ranging from 1=Strongly Disagree to 5=Strongly Agree.

- From the explanation, I know how the game system works.
- This explanation of how the game system works is satisfying.
- This explanation of how the game system works has sufficient detail.
- This explanation of how the game system works seems complete.
- This explanation of how the game system works tells me how to use it.
- This explanation of how the game system works is useful to my goals.
- This explanation of the game system shows me how accurate the game system is.

D.3.3 Trust in Automation. For the Trust in Automation questionnaire [36], the participants indicate their agreement on a 7-point Likert scale, ranging from 1=Not at all to 7=Extremely.

- The system is deceptive.
- The system behaves in an underhanded manner.
- I am suspicious of the system's intent, action, or outputs.
- I am wary of the system.
- The system's actions will have a harmful or injurious outcome.
- I am confident in the system.
- The system provides security.
- The system has integrity.
- The system is dependable.
- The system is reliable.
- I can trust the system.
- I am familiar with the system.