Human & AI: Fool Us*

Crafting Synergistic Interactions with Insights from Magic

Tommaso Turchi[1]

[1]*Department of Computer Science, University of Pisa, Pisa, Italy*

Abstract

This position paper investigates how the art of magic, with its focus on audience perception and expectation, can inspire more effective Artificial Intelligence (AI) interactions, promoting a synergistic relationship between humans and AI. By adapting techniques from stage magic psychology, we explore ways to enhance transparency and engagement in AI systems, thereby improving user trust and engagement. Our approach focuses on developing adaptive interaction designs that encourage a genuine collaboration between humans and AI, leveraging their unique strengths for mutual augmentation. This contribution argues how magic can help AI transcend its conventional role, evolving into a co-decision ally that enriches human decision-making and creativity, paving the way for systems where humans and AI work together in truly flexible and complementary ways.

Keywords

Artificial Intelligence, Magic, Psychology, Human-AI Interaction, Synergistic Interaction

1. Introduction

Arthur C. Clarke once famously said, "Any sufficiently advanced technology is indistinguishable from magic" [1]. This idea highlights the wonder and mystery we often feel when faced with the latest advances in technology, and it's true especially for Artificial Intelligence (AI). Just like watching a magician perform on stage, when we see AI in action in decision-making areas, it can leave us amazed and sometimes baffled by what it does. An AI system, similar to a magician, works its black-box processes behind the scenes and then reveals results that can be impressive but can also leave us scratching our heads. A classic magician's prompt to "pick a card" or their confident "I will make a prediction" closely resembles the anticipatory and often unpredictable nature of AI, where outcomes are derived from often inaccessible data.

But there's a big difference between the mystery we enjoy in a magic show and what we need from AI systems. In areas where decisions really matter, it's crucial that AI systems are predictable, easy to understand, and trustworthy. This presents a challenge for those designing human-AI systems: how can we keep the advanced abilities of AI that astonish us without

making its workings as mysterious as a magic trick? A magician crafts illusions with the intent to deceive by leveraging human psychological biases. Conversely, in designing synergistic Human-AI interactions – aiming for both human-augmented AI and augmented human intelligence [2] – we can apply the principles of magic in a way that the system comprehensively understands its capabilities and limitations and addresses psychological biases not to deceive, but to enhance human cognitive capacities.

In this position paper, we argue that by looking into the art of magic, we can uncover important lessons for improving the interaction between humans and computers (HCI) and making AI more Human-Centered (HCAI). By analyzing how magicians engage their audiences – through controlling attention, telling stories, and forcing decisions – we can adapt these strategies to AI interaction design. Our goal is to ensure AI systems not only amaze us but also provide transparent insights into how they work, ensuring they are reliable and easier to understand. Thus, while the appeal of magic lies in not knowing how it's done, the future of AI should be about developing transparent, synergistic interactions that empower users with both the wonder of discovery and the clarity of understanding.

2. Background

Magic, with its long-standing tradition of creating illusions and engaging audiences, has been increasingly recognized for its potential to inform and enhance various disciplines, notably Human-Computer Interaction (HCI). Tognazzini's influential work on applying the principles, techniques, and ethics of stage magic to human interface design stands as a cornerstone in this domain [3]. By comparing strategies used by magicians to those in user interface design, he demonstrated the potential benefits of weaving elements of magic into HCI to foster more immersive and effective user interactions.

One of the most renowned methods in HCI that draws a direct line to the art of magic (through a famous novel and a later movie too) is the "Wizard of Oz" technique [4]. This methodology involves researchers secretly controlling parts of what users believe to be autonomous systems, allowing them to simulate and study how interfaces perform before the actual technology is fully operational. This approach has become a standard in experimental design, offering a pragmatic means to test hypotheses and refine user interactions with emerging technologies, mirroring the magician's control over their performance while remaining unseen by the audience.

In AI interaction design, the incorporation of magic principles presents a novel lens through which to enhance user experiences with intelligent systems. Research in Human-AI Interaction has highlighted the critical role of design guidelines and resources in creating effective AI products [5]. This research emphasizes the importance of integrating design principles, such as explainability, control, and feedback, into human-AI interactions to improve usability and user satisfaction.

For instance, the literature on AI interaction design highlights cognitive forcing as a technique to reduce users' over-reliance on AI decisions, inspired by similar strategies in magic that influence spectators' choices without their awareness. A study [6] demonstrated how cognitive forcing, compared to basic explainable AI methods, can significantly reduce this over-reliance, though it may also impact user satisfaction, especially among those more inclined to engage in

analytical thinking.

Additionally, Reeves et al. [7] explored how user experience design, analogous to creating a spectator experience in magic shows, can influence engagement and perception in interactive systems. Their study underlines the significance of considering the comprehensive user journey and the emotional impact of AI interactions.

The application of magic's principles extended also to video game design, as another study [8] drew parallels between magicians' illusions and the immersive worlds crafted by game designers. It argued that magic offers valuable design principles for creating believable and engaging game experiences, focusing on affording perceived causal relations and forcing a perceived-free choice.

Most recently, Lupetti and Murray-Rust [9] examined the enchantment factor within AI design, proposing a taxonomy of design approaches that modulate the perception of magic in AI technologies. Their work builds upon the dialogue surrounding recent AI advancements, focusing on how certain interaction qualities, such as algorithmic uncertainties, contribute to a sense of magic or disenchantment. They identified key principles that either enhance or diminish the magical experience, offering insights for design and HCI practitioners to navigate the delicate balance between enchantment and practicality in AI design.

To conclude, despite insightful discussions on using magic as inspiration for HCI and AI design, we found no studies that directly map some of the psychological insights exploited by magic to AI interaction design.

3. Principles of Magic for Human-AI Synergistic Interactions

Magic thrives on the magician's ability to guide audience perceptions and beliefs through misdirection, storytelling, and selective revelation [10]. Interestingly, these very principles could offer a blueprint for enhancing AI transparency and user trust. In the following, we will delve into these principles from magic and discuss their application to AI interaction design, highlighting how techniques developed for the stage can enhance the way AI systems engage and communicate with users.

Misdirection to Direction

Magicians use misdirection to draw the audience's attention away from the trick's method [11]. With AI, this principle can transform into guiding users' focus to critical elements of the AI's operation, making the process behind AI decisions more visible and less of a black box. Tools like feature importance scores and interactive visualizations can serve as the "spotlight," illuminating the path AI takes to arrive at conclusions. Integrating attention checks, as proposed in the context of Large Language Models (LLMs) [12], further exemplifies this.

Similarly, the principle of repetition [13], well-utilized in magic to condition the audience towards certain expectations, can find its place in AI by making certain interactions familiar and intuitive. By repetitively guiding users through the evaluation of AI outputs, we can foster a habit of critical engagement, ensuring users not only focus on but also critically assess the information presented by AI systems. This strategy could not only uncover the AI's operations

but also encourage a more thoughtful interaction between users and AI systems, enhancing the overall trustworthiness and reliability of AI interactions.

Storytelling for Engagement and Understanding

Magicians enhance the emotional impact and coherence of their performances through storytelling, skillfully navigating audience perceptions and expectations [14]. This art of narrative building resonates with the recent exploration into how the principles of magic can enlighten the advancement of Artificial Intelligence (AI) interaction design. By embedding AI decisions within narratives that elucidate the "why" and "how," we can explain the operations of AI, transforming abstract data and complex algorithms into relatable and understandable narratives [15]. This methodology not only makes AI more accessible but also augments the interaction with a sense of wonder and engagement, matching the magician's skill in crafting compelling stories that captivate their audiences.

Incorporating participatory design elements into AI storytelling reflects a shift towards a more inclusive and collaborative approach, where users actively contribute to shaping the AI narrative. This method, similar to myth-making in traditional storytelling, allows users to embed their own experiences and expectations into the AI development process, creating a shared narrative that enhances the system's relevance and user acceptance [16].

The notion that "Magic doesn't happen in the hand of the magician but in the mind of the spectator" [17] underscores the importance of active collaboration between the performer and the audience in creating the magic experience. Similarly, in AI, this collaboration can be mirrored in the dynamic between the system and its users, where the AI, prompted by human interaction, generates experiences that could be perceived as magical. This perspective shifts the role of AI from a passive tool to an active participant in a decision process, similar to a magician working alongside spectators to conjure moments of insight and discovery.

This approach shifts AI from a mere computational tool to a decision partner, echoing magic's interdisciplinary nature, which spans psychology, physics, and mathematics [14]. Through strategic storytelling, AI can guide users through its workings, not to obscure but to enlighten, mirroring the magician's narrative that leads to moments of insight and discovery. Thus, integrating storytelling from magic into AI design could foster a richer, more intuitive interaction, positioning AI as a companion in the journey towards understanding the digital and the magical.

Selective Revelation to Progressive Disclosure

In the realm of magic, only the outcome of a trick is revealed, with the method shrouded in secrecy. In contrast, AI systems aim for an open book approach, providing deep insights into their "method" – the algorithms and data underpinning decisions. This pursuit of transparency can be implemented through progressive disclosure, where explanation interfaces offer users varying levels of detail, from high-level summaries to intricate technical descriptions, tailored to their expertise and curiosity [18].

Expanding this notion, drawing back the curtain in AI means demystifying the system's inner workings comprehensively. It's about moving beyond merely showcasing outcomes to

communicating the decision-making journey in an accessible format. Employing techniques from the field of explainable AI (XAI) and adopting transparency-by-design principles are key to this effort. These strategies ensure that AI doesn't just present results but also invites users into the decision-making process, fostering a clear understanding of how conclusions are reached [19]. This approach not only aligns with the goal of making AI systems more intelligible and user-centric but could also reinforce the shift from the magician's secrecy to a paradigm of open exploration and knowledge sharing.

Countering Biases

Magic leverages psychological biases such as priming, stereotypical behaviour, and saliency to subtly direct audience decisions [20, 21]. These techniques manipulate spectators' choices by making specific outcomes more appealing or accessible, altering decisions in a seemingly natural way [22]. However, in the realm of AI, ethical interaction design aims to shift from exploiting these human cognitive biases to mitigating them, including addressing AI's own inherent biases derived from its training data [23].

This ethical approach in AI design involves both countering human psychological biases and diligently working to identify and correct biases within AI systems themselves. These inherent biases, often a reflection of skewed or unrepresentative training data, can lead to AI decisions that inadvertently perpetuate stereotypes or unfair outcomes. By implementing strategies that ensure balanced and diverse data sets, AI designers can reduce the impact of these biases, promoting fairness and neutrality in AI-generated options and suggestions.

The challenges of fully disclosing technological processes in AI systems are compounded by the socio-technical nature of these systems. As explored in [24], the diverse motivations and roles of stakeholders in even small systems indicate that transparency is not only a technical challenge but also a socio-technical one. Stakeholders vary in their support for the system and each other, affecting local to global scales. This underscores the need for responsible system design that considers these varied impacts, where the ethical design is foundational not only to functionality but also to identity and cultural integrity [25].

In parallel, AI systems can be designed to present information and choices in a way that encourages users to make informed, reflective decisions. This includes providing transparent explanations of how AI systems arrive at their conclusions and educating users about both their own cognitive biases and the potential biases within AI systems. Such transparency not only fosters trust but also empowers users to critically evaluate AI suggestions, leading to more informed choices.

Hence, transitioning from exploiting biases in magic to nurturing informed choices in AI design represents a profound shift towards ethical and responsible technology use. By addressing both human cognitive biases and the inherent biases within AI systems, designers can create AI interactions that are fair, transparent, and aligned with human values, supporting a more equitable and informed decision-making process.

Adapting to People and Outcomes

The anticipation of a single, predetermined ending is often superseded by the magician's ability to navigate through several potential conclusions. This flexibility ensures that the performance can adapt in real-time to the choices and reactions of the audience, embodying the principle of multiple outcomes as seen in mentalism and forcing [26]. Similarly, adaptive AI systems should be designed to not just react to user inputs but to anticipate and align with multiple potential user goals and scenarios, dynamically adjusting their outputs to fit the context of each interaction.

Moreover, the effectiveness of a magic performance is deeply influenced by the audience's beliefs and perceptions. Younger viewers, with more blurred boundaries between reality and fantasy, might experience a magic trick very differently from adults, whose understanding of the world is more rigidly defined [27]. This variance in audience reception highlights the need for AI systems to adapt not only to the explicit inputs provided by users but also to their underlying beliefs, expectations, and cognitive biases.

Incorporating these adaptive strategies into AI design involves creating systems that are not only responsive to direct input but are also sensitive to the broader context of user interactions. By integrating mechanisms similar to the magician's multiple outs and tailoring responses to individual user profiles, AI can achieve a higher degree of personalization. Such adaptive AI interactions promise a more engaging and intuitive user experience, related to the personalized engagement found in a magic performance, creating more meaningful and impactful interactions that resonate on a personal level.

By embracing the principles of adaptability in magic, AI systems can evolve beyond static algorithms to become dynamic entities capable of crafting bespoke experiences for each user [28]. This approach not only enhances the usability and effectiveness of AI but also enriches the relationship between humans and technology, fostering a collaborative and adaptive partnership that mirrors the dynamic interplay of a magician and their audience.

Informing AI interaction design through principles of magic invites a shift from opacity to clarity, from confusion to understanding. By directing attention, crafting narratives, and embracing transparency, AI can become not only more explainable but also more engaging and trustworthy. This approach does not strip AI of its wonder but rather opens a window for users into the "magic" of its technology, fostering an environment where advanced AI capabilities are met with informed surprise rather than mystified apprehension.

4. Discussion and Conclusion

In weaving together the threads of our discourse on infusing Artificial Intelligence (AI) interaction design with insights from the art of magic, a compelling narrative emerges. It challenges us to reimagine the relationship between humans and AI, evolving the latter into a partner that communicates, interacts, and enlightens through synergistic interactions. This paper argues for a shift from the mystique of magic to the clarity of understanding, advocating for AI systems that are as transparent in their workings as they are advanced in their capabilities.

The implications of this shift could be profound for the field of human-centered artificial

intelligence (HCAI). By drawing parallels with magic, we've highlighted a path towards AI systems that foster trust through transparency and engagement through explanation. Trust, in this context, emerges not from obscuring the complexity of AI but from demystifying it, enabling users to grasp the "how" and "why" behind AI decisions. This transparency is not a mere luxury but a necessity in building systems that are embraced by their users.

Moreover, our discussion underscores the importance of maintaining a sense of wonder about AI, similar to the awe inspired by a well-crafted magic trick. However, the wonder should stem from an appreciation of AI's capabilities and the elegance of its design, rather than from a lack of understanding. Such synergistic interactions between humans and AI promise to not only captivate but also empower users, enriching their experiences with both insight and a sense of discovery.

Ethical considerations are integral to achieving true human-AI synergy. As we draw inspiration from magicians who understand their audience to enhance the performance, so too must AI systems ethically gather user data to personalize experiences and enhance engagement. This process must be guided by principles that prioritize user privacy, consent, and autonomy, ensuring that personalization enhances the AI experience without compromising user rights [29].

Looking to the future, the journey toward synergistic human-AI interactions invites a multidisciplinary collaboration that bridges technologists, psychologists, narrative experts, and magicians. Such collaboration can lead to innovative approaches that make AI systems not only more understandable and engaging but also more integral and responsive to the human experience [19].

In conclusion, our exploration into magic-inspired guidelines for AI Interaction strives to foster synergistic interactions that augment human intelligence. By advocating for systems that are both transparent and trustworthy, we aim to facilitate a partnership where AI complements and enhances human capabilities, rather than merely serving as a tool. This approach not only satisfies our practical needs but also engages our innate curiosity, encouraging a collaborative journey with AI that deepens our understanding and enriches our experiences. Through this synergy, we envision a future where AI actively participates in our quest for knowledge and creativity, making every interaction an opportunity for growth.

Acknowledgments

References

[1] A. C. Clarke, Profiles of the Future: An Inquiry into the Limits of the Possible, [millennium ed.], 1. publ ed., Gollancz, London, 1999.

[2] M. H. Jarrahi, C. Lutz, G. Newlands, Artificial intelligence, human intelligence and hybrid

intelligence based on mutual augmentation, Big Data & Society 9 (2022) 205395172211428. doi:10.1177/20539517221142824.

[3] B. Tognazzini, Principles, techniques, and ethics of stage magic and their application to human interface design, in: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, CHI '93, Association for Computing Machinery, New York, NY, USA, 1993, pp. 355–362. doi:10.1145/169059.169284.

[4] A. Dix (Ed.), Human-Computer Interaction, 3rd print ed., Prentice Hall, New York, 1993.

[5] S. Amershi, D. Weld, M. Vorvoreanu, A. Fourney, B. Nushi, P. Collisson, J. Suh, S. Iqbal, P. N. Bennett, K. Inkpen, J. Teevan, R. Kikin-Gil, E. Horvitz, Guidelines for human-ai interaction, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 1–13. doi:10.1145/3290605.3300233.

[6] Z. Buçinca, M. B. Malaya, K. Z. Gajos, To trust or to think: Cognitive forcing functions can reduce overreliance on ai in ai-assisted decision-making, Proc. ACM Hum.-Comput. Interact. 5 (2021). doi:10.1145/3449287.

[7] S. Reeves, S. Benford, C. O'Malley, M. Fraser, Designing the spectator experience, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '05, Association for Computing Machinery, New York, NY, USA, 2005, p. 741–750. doi:10.1145/1054972.1055074.

[8] S. Kumari, S. Deterding, G. Kuhn, Why game designers should study magic, Proceedings of the 13th International Conference on the Foundations of Digital Games (2018). doi:10.1145/3235765.3235788.

[9] M. L. Lupetti, D. Murray-Rust, (un)making ai magic: a design taxonomy, in: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, CHI 2024, Honolulu, Hawaii, USA, May 11-16, 2024, ACM, 2024. doi:10.1145/3613904.3641954.

[10] G. Kuhn, A. A. Amlani, R. A. Rensink, Towards a science of magic, Trends in Cognitive Sciences 12 (2008) 349–354. doi:10.1016/j.tics.2008.05.008.

[11] G. Kuhn, H. A. Caffaratti, R. Teszka, R. A. Rensink, A psychologically-based taxonomy of misdirection, Frontiers in Psychology 5 (2014). doi:10.3389/fpsyg.2014.01392.

[12] S. J. Gould, D. P. Brumby, A. L. Cox, Chattl;dr – you really ought to check what the llm said on your behalf, in: Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24), May 11–16, 2024, ACM, Honolulu, HI, USA, 2024. doi:10.1145/3613905.3644062.

[13] G. Kuhn, P. Kingori, K. P. Grietens, Misdirection – magic, psychology and its application, Science & Technology Studies 35 (2022) 13–29. doi:10.23987/sts.112182.

[14] J. Steinmeyer, Hiding the Elephant: How Magicians Invented the Impossible and Learned to Disappear, 1st carroll & graf trade pbk. ed ed., Carroll & Graf Publishers, New York, 2003.

[15] J. Kim, S. Lee, Are two heads better than one?: the effect of student-ai collaboration on students' learning task performance, TechTrends 67 (2022) 365–375. doi:10.1007/s11528-022-00788-9.

[16] R. Jacobs, J. Spence, F. Abbott, A. Chamberlain, W. Heim, A. Yemaoua Dayo, D. Kemp, S. Benford, D. Price, R. Shackford, J. Robson, C. Locke, J. King, Future machine: Making myths & designing technology for a responsible future: Making myths and entan-

glement: Community engagement at the edge of participatory design and user experience, in: Proceedings of the 26th International Academic Mindtrek Conference, Mindtrek '23, Association for Computing Machinery, New York, NY, USA, 2023, p. 108–118. URL: https://doi.org/10.1145/3616961.3616979. doi:10.1145/3616961.3616979.

[17] A. Stone, Fooling Houdini: Magicians, Mentalists, Math Geeks, & the Hidden Powers of the Mind, Harper, New York, 2013.

[18] C. Panigutti, A. Beretta, D. Fadda, F. Giannotti, D. Pedreschi, A. Perotti, S. Rinzivillo, Co-design of human-centered, explainable ai for clinical decision support, ACM Trans. Interact. Intell. Syst. 13 (2023). doi:10.1145/3587271.

[19] A. Malizia, F. Paternò, Why is the current xai not meeting the expectations?, Communications of the ACM 66 (2023) 20–23. doi:10.1145/3588313.

[20] A. Pailhès, G. Kuhn, Influencing choices with conversational primes: How a magic trick unconsciously influences card choices, Proceedings of the National Academy of Sciences 117 (2020) 17675–17679. doi:10.1073/pnas.2000682117.

[21] Banachek, D. Dyment, S. R. Wells, Psychological Subtleties. 2, 2nd ed ed., Magic Inspirations, Houston, Tex., 2007.

[22] J. A. Olson, A. A. Amlani, R. A. Rensink, Perceptual and cognitive characteristics of common playing cards, Perception 41 (2012) 268–286. doi:10.1068/p7175.

[23] T. Turchi, A. Malizia, S. Borsci, Reflecting on algorithmic bias with design fiction: the minicode workshops, IEEE Intelligent Systems (2024) 1–13. doi:10.1109/MIS.2024.3352977.

[24] A. Crabtree, A. Chamberlain, Making it "pay a bit better": design challenges for micro rural enterprise, in: Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing, CSCW '14, Association for Computing Machinery, New York, NY, USA, 2014, p. 687–696. URL: https://doi.org/10.1145/2531602.2531618. doi:10.1145/2531602.2531618.

[25] A. M. Piskopani, A. Chamberlain, C. Ten Holter, Responsible ai and the arts: The ethical and legal implications of ai in the arts and creative industries, in: Proceedings of the First International Symposium on Trustworthy Autonomous Systems, TAS '23, Association for Computing Machinery, New York, NY, USA, 2023. URL: https://doi.org/10.1145/3597512.3597528. doi:10.1145/3597512.3597528.

[26] A. Pailhès, G. Kuhn, Mind control tricks: Magicians' forcing and free will, Trends in Cognitive Sciences 25 (2021) 338–341. doi:10.1016/j.tics.2021.02.001.

[27] J. A. Olson, I. Demacheva, A. Raz, Explanations of a magic trick across the life span, Frontiers in Psychology 6 (2015). doi:10.3389/fpsyg.2015.00219.

[28] T. Turchi, A. Malizia, F. Paternò, S. Borsci, A. Chamberlain, Adaptive xai: Towards intelligent interfaces for tailored ai explanations, in: 29th International Conference on Intelligent User Interfaces, IUI '24 Companion, Association for Computing Machinery, New York, NY, USA, 2024. doi:10.1145/3640544.3645253.

[29] T. Turchi, G. Prencipe, A. Malizia, S. Filogna, F. Latrofa, G. Sgandurra, Pathways to democratized healthcare: Envisioning human-centered AI-as-a-service for customized diagnosis and rehabilitation, Artificial Intelligence in Medicine 151 (2024) 102850. doi:10.1016/j.artmed.2024.102850.