Marking the Virtual: Exploring Visual Indicators for Distinguishing Physical and Virtual Objects in Augmented Reality

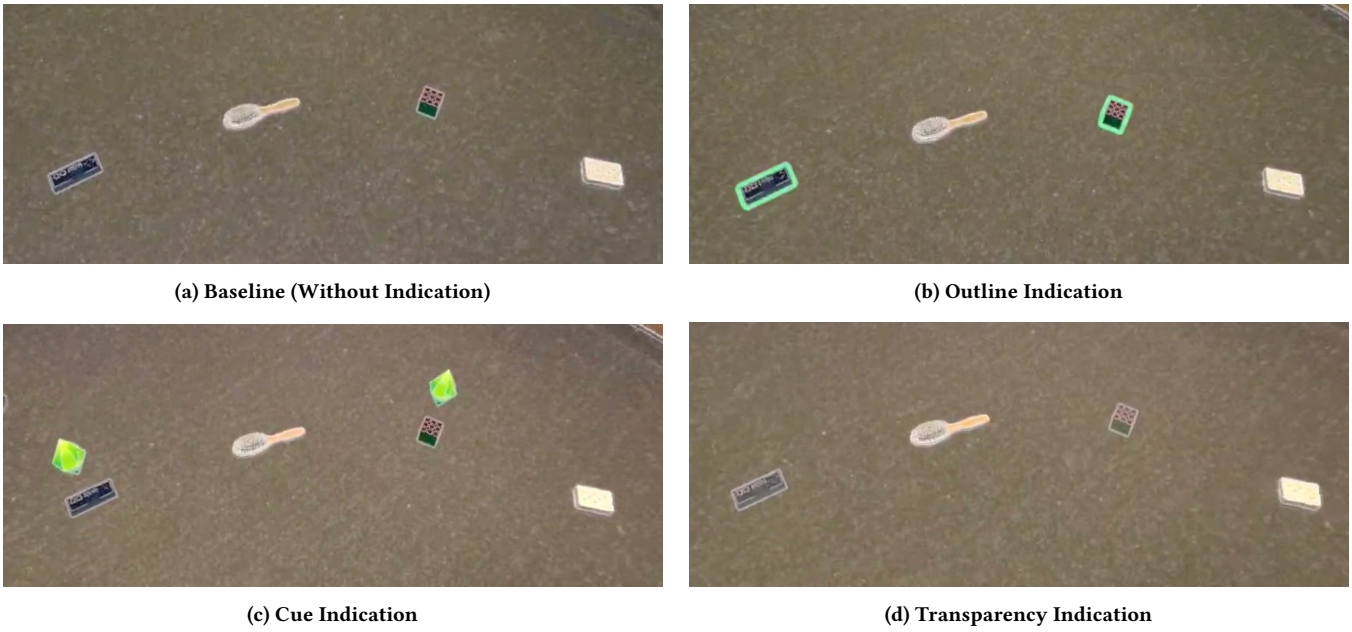
Christopher Katins
christopher.katins@hu-berlin.de
HU Berlin
Berlin, Germany

Keno Goertz
keno.goertz@tu-dresden.de
TU Dresden
Dresden, Germany

Ali Aslan
ali.aslan@student.hu-berlin.de
HU Berlin
Berlin, Germany

Christopher Lazik
christopher.lazik@hu-berlin.de
HU Berlin
Berlin, Germany

Thomas Kosch
thomas.kosch@hu-berlin.de
HU Berlin
Berlin, Germany



(a) Baseline (Without Indication)

(b) Outline Indication

(c) Cue Indication

(d) Transparency Indication

Figure 1: We investigated three different visual indicators to support users in distinguishing physical and virtual objects in Augmented Reality ((b) to (d)). Note that, even without a visual indicator (see (a)), the AR environment is stylized with a Sobel edge detection filter to make it harder to distinguish between physical and virtual objects. Objects from left to right are: NES controller (virtual), hair brush (physical), Rubik's cube (virtual), playing cards (physical).

Abstract

As Augmented Reality (AR) applications and their respective virtual components approach photorealism, users are challenged to distinguish physical objects from virtually rendered ones, which can undermine trust in the environment and lead to interaction errors. We address this challenge by introducing three indicators that mark virtual content to support reliable physical-virtual discrimination. Since photorealistic virtual content is still hard to achieve in real-time AR, we induce ambiguity instead by applying a matched Sobel edge-detection stylization to both the passthrough view and the virtual objects. We evaluated the indicators in a user study (N=20) where participants classified objects in an AR environment as either physical or virtual. Our results show that a salient hovering cue significantly improved confidence-weighted classification accuracy over the baseline and the transparency indicator, while it did not differ significantly from an outline. However, misclassifications still occurred for all indicators, suggesting that current designs are not fail-safe. These results give first indications on how users of AR applications could be informed about the physicality and virtuality of components once indistinguishable by the human eye alone.

This work is licensed under a Creative Commons Attribution 4.0 International License.
MUM '26, Glasgow, United Kingdom
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2791-7/2026/12
https://doi.org/10.1145/3861188.3861196

CCS Concepts

• **Human-centered computing → Empirical studies in HCI**; **Mixed / augmented reality**.

Keywords

Augmented Reality, Mixed Reality, Reality Design, Perceptual Manipulation, Visual Indicators, Photorealism

ACM Reference Format:

Christopher Katins, Keno Goertz, Ali Aslan, Christopher Lazik, and Thomas Kosch. 2026. Marking the Virtual: Exploring Visual Indicators for Distinguishing Physical and Virtual Objects in Augmented Reality. In *25th International Conference on Mobile and Ubiquitous Multimedia (MUM '26), December 01–04, 2026, Glasgow, United Kingdom.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3861188.3861196

1 Introduction and Background

Over recent years, AR, herein understood as a locally present experience that combines physical and virtual elements, following the definition by Rauschnabel et al. [15], has seen substantial advances in hardware. Together with ongoing improvements in real-time rendering, these developments have pushed AR experiences toward increasingly lifelike visual fidelity. Recent work showed that photorealistic rendering is possible, enabling switching between physical and virtual environments without users noticing [24].

However, photorealistic AR could introduce new critical challenges. If virtual objects become visually indistinguishable from physical ones, users may no longer be able to rely solely on their perception to distinguish between the two. Photorealistic rendering also lowers the barrier to perceptual manipulation attacks, which have been identified as a distinct challenge for AR [6]. For example, prior work showed attacks in AR that exploit users' multisensory perception (visual, auditory, and situational awareness) to bias judgments and actions, while user responses and defenses remain poorly understood [2]. Consequently, perceptual manipulation could be used to subtly steer users' physical movements and cognition, therefore posing an urgent challenge for protecting users from such attacks [22]. Slater et al. discuss the ethical implications of such realism, warning that when people cannot reliably separate physical from virtual, they may lose trust in their own experiences and memories, potentially weakening "the capacity to decouple sensory experience from ground truth," [19, p. 8] with broader societal consequences. Such concerns are also not limited to researchers, with a recent survey showing public concerns about the effects of AR and VR on health, social interactions, and overreliance on the technology [8].

Marking objects visually is a long-standing HCI problem. Work on immersion-preserving highlighting shows that subtle cues can direct attention without breaking the experience [10], and studies of distractors show that the surrounding scene strongly moderates whether a highlight is noticed at all [5]. Pakanen et al. found that, without visual indication, users of a virtual office environment were unsure which elements were interactive, and that they preferred a glow effect [13], a finding later confirmed for web-based 3D travel portals [12]. Games face the same problem, and Dillman et al. derived a framework of visual interaction cues from 49 video games and showed that it transfers to AR [4]. All of this work asks which elements of a clearly virtual scene are *interactive* or relevant, and it guides users toward known targets. We instead ask which elements of a mixed scene are *virtual* at all. In our setting, the scene itself is visually ambiguous, and the user does not know in advance which objects are marked or how many there are. Two works come closer to this question. In AR-based telepresence, participants preferred a see-through hologram avatar because it prevents mixing up the avatar with a real person in the room [11], and Pouke et al. deliberately blended a mirror-world copy of a physical library with fantasy layers in VR [14]. In both, the boundary between physical and virtual is a design choice made for a single, known entity or layer. We treat an unmarked boundary between physical and virtual objects as a risk and, to our knowledge, are the first to compare indicators whose sole purpose is to mark virtuality when physical and virtual objects are hard to tell apart.

To address this gap, we explored if and how indicators can effectively signal that an entity in an AR environment is virtual (see Figure 1). Our research was guided by the following research question:

RQ: How do different visual indicators affect users' ability to distinguish physical from virtual objects in an ambiguous AR environment?

We make two contributions. First, we contribute a method for inducing controlled physical-virtual ambiguity with today's hardware. Because photorealistic virtual content is very hard to achieve in real-time AR, we approach indistinguishability from the opposite direction. Rather than raising virtual fidelity to match the physical world, we apply a matched Sobel edge-detection stylization to both the passthrough view and the virtual objects, lowering both to a shared level of visual fidelity. This lets us study the consequences of physical-virtual ambiguity now, rather than waiting for rendering to catch up. Second, we contribute an exploratory study (N=20) of three visual indicators under this manipulation, addressing the research question above. We show that a salient cue significantly improves confidence-weighted classification accuracy over the baseline, that misclassifications persist under every indicator, and that an indicator sharing a visual dimension with the surrounding ambiguity may be at a disadvantage. We suggest that if photorealistic AR becomes commonplace, future indicator designs may need to operate robustly, potentially even at a more subconscious level, to ensure reliable discriminability. We therefore see this work as an initial step and call for further research in this direction.

2 Investigating Visual Indicators in an Induced-Ambiguity AR Environment

Fully photorealistic virtual objects remain difficult to realize with current real-time AR rendering pipelines, and waiting for photorealism before studying its consequences would leave the field reactive. We therefore decouple the research question from the rendering capability as our focus is not that virtual objects look photorealistic, but that users cannot tell the two apart. Rather than raising virtual fidelity to match the physical environment, we stylized the passthrough view and applied a matching stylization to virtual objects, creating controlled visual ambiguity while keeping all conditions technically feasible. We refer to this manipulation as *induced ambiguity*. We consider this a useful research instrument, though

it might be limited in usefulness once photorealism is even easier to reach with newly available AR devices. Induced ambiguity is not perceptually equivalent to photorealism. It removes surface, texture, and shading cues from physical and virtual content, whereas photorealism would produce confusion by making virtual content match physical appearance. Findings obtained this way are evidence about a state of low discriminability, not about photorealistic rendering as such. Practically, degrading the passthrough view degrades the user's view of the actual room. This confines the method to stationary, controlled tasks such as ours, where participants were guided to a fixed position and asked only to observe and rate. It is not suitable for studies involving locomotion, manipulation of physical objects, or any setting where reduced visibility of the physical environment raises safety concerns.

2.1 User Study

Having established a setting in which physical and virtual objects are hard to tell apart, we asked whether an explicit visual marking can restore that distinction. The study was conducted using a Meta Quest 3 head-mounted display, using Meta's passthrough API, which provides an option to apply an edge-detection filter to the passthrough view. We then implemented a Sobel edge-detection filter [20] for the virtual objects as well, tuning its parameters to closely match the passthrough filter. The edge-detection stylization was active in all experimental conditions. We conducted a within-subjects study (N=20) in which participants classified objects in an induced-ambiguity AR scene as physical or virtual, under four levels (Baseline, Outline, Cue, and Transparency) of a single independent variable, the indicator (see Figure 1). Because the *Baseline* condition contains no marking, it also serves as a check on the induced ambiguity introduced in Section 2.

2.1.1 Participants. We recruited 20 participants (12 identified as female, 5 as male, and 3 as non-binary) via personal networks and university group chats. Participants ranged from 21 to 54 years ($\bar{x} = 27.5$, $s = 7.8$) in age. Overall, participants reported little to no prior experience with AR. Only one participant indicated regular AR headset use[1]. Ten participants had never used an AR headset, and nine had only tried AR once.

2.1.2 Independent Variables. We define the Indicator as the single independent variable of our study, with four levels: *Baseline*, *Outline*, *Cue*, and *Transparency*. In *Baseline*, no visual indicator is displayed (see Figure 1a), while the three remaining levels each mark virtual objects in a different way. *Outline* draws a contour around them (see Figure 1b), *Cue* places a symbol above them (see Figure 1c), and *Transparency* reduces their opacity (see Figure 1d). Outlines and glows are the most common way to mark objects in games and 3D virtual environments [4, 12, 13], and see-through rendering has been used in AR to signal that content is not physically present [11], while also being a low-clutter alternative that preserves scene composition. For the cue, we used a "The Sims"-inspired bright green dodecahedron because it is a game-like marker that usually floats above an entity to signal selection or special status, making it a salient cue for marking virtual objects in AR. Objects were everyday, immediately nameable items of comparable size (see Figure 2).

[1]Approximately two to four times per month.

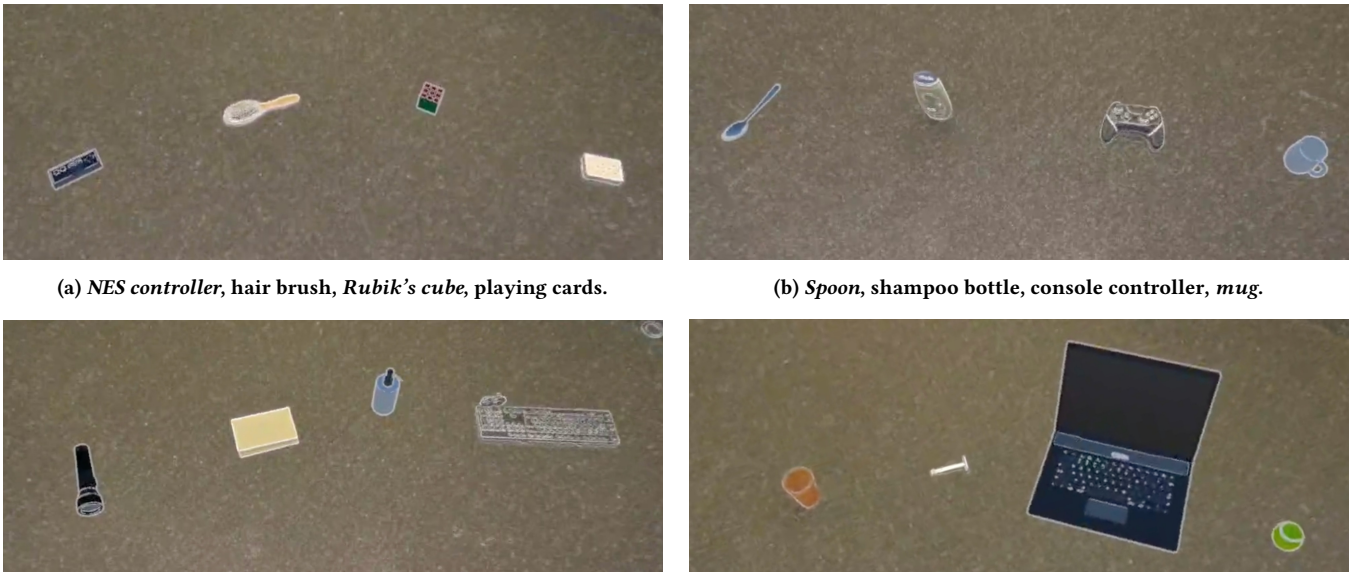
2.1.3 Dependent Variables. We measure the Task Completion Time (TCT), confidence-weighted accuracy, the Single Ease Question (SEQ) [16], and the Slater–Usoh–Steed Presence Questionnaire (SUS-P) [23]. We included presence because marking an object as virtual may also reduce how much virtual content feels like part of the physical room [21], which needs to be considered against improved discrimination. We measured TCT as the elapsed time from the AR scene's activation to the participant's final object rating. To measure confidence and accuracy, participants evaluated each object on an 11-point scale ranging from 0 ("I am certain the object is virtual") to 10 ("I am certain the object is physical"), with intermediate values reflecting graded certainty. We operationalize this confidence-weighted accuracy, where we averaged four values per set of objects:

$$\text{accuracy} = \begin{cases} \frac{\text{rating}}{10}, & \text{if the object is physical,} \\ 1 - \frac{\text{rating}}{10}, & \text{if the object is virtual.} \end{cases}$$

This yields a score in $[0, 1]$ that combines correctness and confidence: 1.0 requires both a correct classification and maximal certainty, and 0.5 corresponds to complete uncertainty. We combined correctness and confidence because binary correctness alone cannot distinguish a lucky guess from certain judgment and, with four objects per condition, can only take five values. A weighted score penalizes confident misclassification more than uncertain guesses. After completing the object ratings, participants answered the question "Overall, how easy or difficult did you find this task?" for each condition using the Single Ease Question (SEQ) [16] consisting of a single 7-point Likert item[2]. Participants completed the Slater–Usoh–Steed Presence Questionnaire (SUS-P) [23], which comprises 6 items rated on a 7-point Likert scale. Following Usoh et al. [23], the SUS-P score is the number of items rated 6 or 7, ranging from 0 to 6.

2.1.4 Procedure. At the start of each session, we explained the study and obtained informed consent. We then collected demographic information and participants' self-reported AR headset usage. Next, we ensured the Meta Quest 3 was properly adjusted for fit and visual clarity, and introduced AR by showing the Quest 3 settings menu floating in passthrough mode. Then, we explained that the study evaluates visual indicators intended to mark virtual objects and described the upcoming scene (i.e., objects placed on the floor) and the rating task. Participants were informed that some trials would include indicators that exclusively mark virtual objects, but we did not disclose which indicators would appear or the number of physical or virtual objects. We also described the stylized rendering as a filter applied to the environment to increase ambiguity, clarifying that it was not itself a visual cue. After participants confirmed their understanding, they left the room while we prepared the scene. The experiment used four object sets and assigned conditions using a balanced Latin square. Before each run, we turned off the headset display so the headset acted as a blindfold, guided participants to a position 1 meter from the objects, and turned on the passthrough remotely. Ratings were collected verbally to avoid requiring participants to learn to use the headset controller. After all ratings were provided, we turned off the

[2]Ranging from 1: "Very difficult" to 7: "Very easy"



(a) ***NES controller,*** **hair brush,** ***Rubik's cube,*** **playing cards.**

(b) ***Spoon,*** **shampoo bottle, console controller,** ***mug.***

(c) ***Flashlight,*** **book,** ***soap dispenser,*** **computer keyboard.**

(d) **Cup, razor,** ***laptop, tennis ball.***

Figure 2: Four distinct object sets were used to evaluate the four experimental conditions, as viewed through the Meta Quest 3. Virtual objects are shown in *italics,* and physical objects in bold. The experimental conditions were counterbalanced across the object sets. Within each set, the object order was identical for all participants. Each set contained two physical and two virtual objects.

display, removed the physical objects, and participants completed a questionnaire including the SEQ and the SUS-P questionnaire. This procedure was repeated for all conditions. We recorded participants' remarks during the task, but as our study approach was quantitative in nature, we report these as illustrative rather than systematic.

2.2 Results

2.2.1 Task Completion Time. We measured the TCT for each condition in seconds. A Shapiro-Wilk test revealed a violation of normality for all conditions, $p < .05$. A Friedman test showed that the TCT did not differ significantly across the four conditions, $\chi^2(3, N = 20) = 0.30$, $p = .960$. The *Outline* condition required the lowest TCT, $\bar{x} = 25.23$, $s = 14.65$, followed by *Cue*, $\bar{x} = 28.88$, $s = 22.23$, *Baseline*, $\bar{x} = 30.07$, $s = 17.81$, and *Transparency*, $\bar{x} = 31.39$, $s = 23.73$. Figure 3a shows the distribution for each condition.

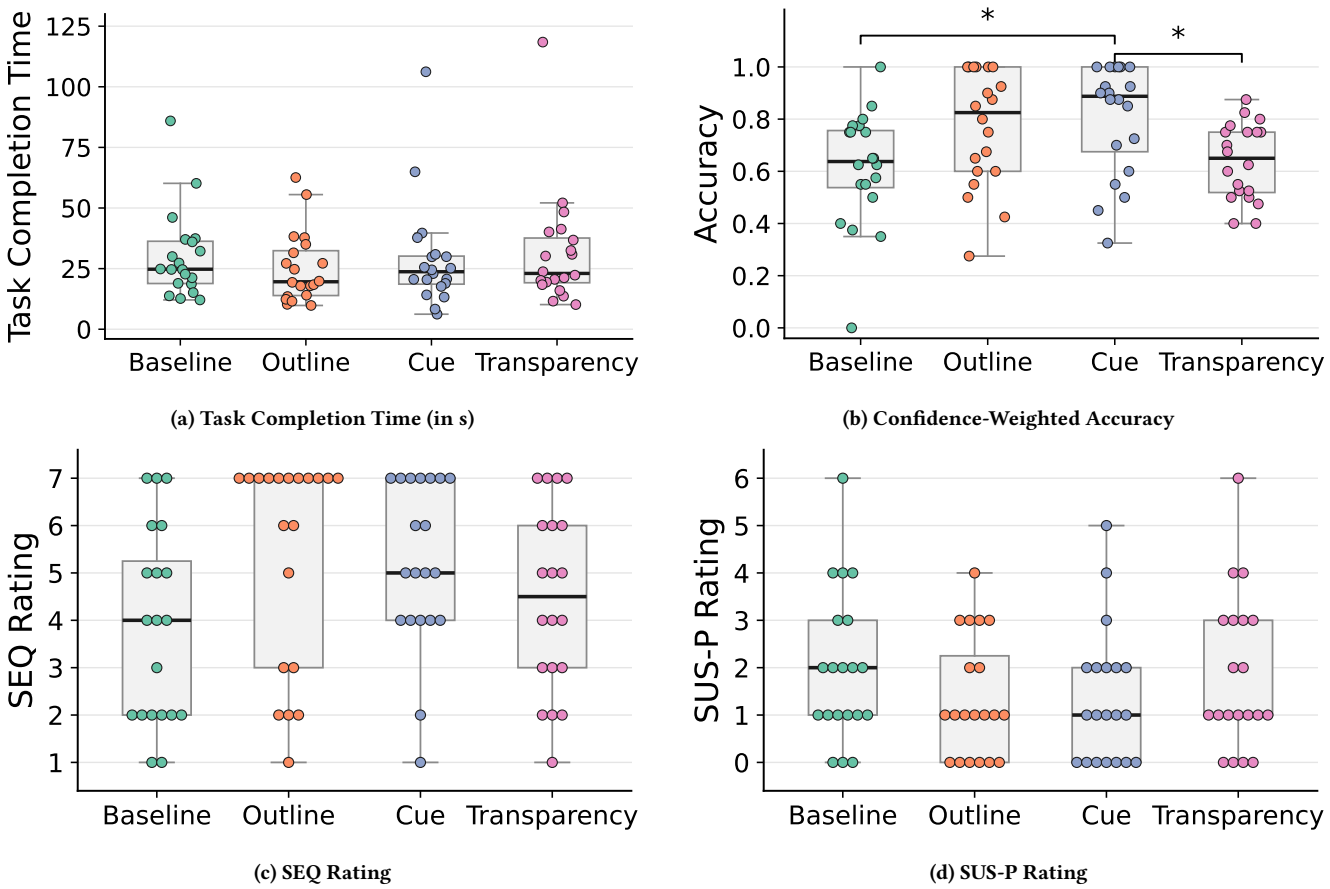
2.2.2 Confidence-Weighted Accuracy. Normality of the dependent measure was assessed separately for each condition using Shapiro-Wilk tests, $n = 20$ per condition. The results indicated deviations from normality for *Cue*, $W = 0.844$, $p = .004$, and *Outline*, $W = 0.898$, $p = .038$, while *Baseline*, $W = 0.935$, $p = .191$, and *Transparency*, $W = 0.921$, $p = .104$, did not significantly deviate from normality. Thus, we tested for main effects using a Friedman test. The analysis revealed a significant effect of condition, $\chi^2(3, N = 20) = 11.855$, $p = .008$. Then, we tested for pairwise differences using two-sided Wilcoxon signed-rank tests with Holm correction for multiple comparisons. Accuracy was significantly higher for the *Cue* condition compared to *Transparency*, $W = 26.5$, $p_{\text{Holm}} = .010$, and *Baseline*, $W = 32.0$, $p_{\text{Holm}} = .022$. All remaining comparisons were not significant after Holm correction: *Baseline* vs. *Outline*, $W = 42.0$, $p_{\text{Holm}} = .099$, *Outline* vs. *Transparency*, $W = 46.0$, $p_{\text{Holm}} = .194$, *Cue* vs. *Outline*, $W = 46.5$, $p_{\text{Holm}} = 1.000$, and *Baseline* vs. *Transparency*, $W = 70.5$, $p_{\text{Holm}} = 1.000$. The *Baseline* condition received the lowest classification accuracy, $\bar{x} = 0.61$, $s = 0.22$, followed by *Transparency*, $\bar{x} = 0.64$, $s = 0.14$, *Outline*, $\bar{x} = 0.77$, $s = 0.22$, and *Cue*, $\bar{x} = 0.81$, $s = 0.21$. Figure 3b shows the distribution for each condition. To check the induced ambiguity, we compared *Baseline* accuracy against chance using a one-sample t-test ($t(19) = 2.24$, $p = .037$).

2.2.3 SEQ. Participants provided the SEQ values after each condition. A Friedman test did not reveal a significant effect across the conditions, $\chi^2(3, N = 20) = 5.95$, $p = .114$. The *Baseline* condition received the lowest rating, $\bar{x} = 3.85$, $s = 2.06$, followed by *Transparency*, $\bar{x} = 4.45$, $s = 1.93$, *Cue*, $\bar{x} = 5.2$, $s = 1.77$, and *Outline*, $\bar{x} = 5.35$, $s = 2.23$. Figure 3c shows the distribution for each condition.

2.2.4 SUS-P. Participants filled in a SUS-P after each condition. A Friedman test did not reveal a significant effect across the conditions, $\chi^2(3, N = 20) = 5.57$, $p = .134$. The condition *Cue*, $\bar{x} = 1.35$, $s = 1.42$, and *Outline*, $\bar{x} = 1.35$, $s = 1.27$ received the lowest rating followed by *Transparency*, $\bar{x} = 1.85$, $s = 1.63$, and *Baseline*, $\bar{x} = 2.00$, $s = 1.59$. Figure 3d shows the distribution for each condition.

3 Discussion and Future Work

Within the ambiguity manipulation we tested, the choice of visual indicator affected discrimination accuracy. Post-hoc comparisons



(a) Task Completion Time (in s)

(b) Confidence-Weighted Accuracy

(c) SEQ Rating

(d) SUS-P Rating

Figure 3: Box plots of all dependent variables measured. While our results did not show a significant effect across conditions for TCT, SEQ, or SUS-P, we did find a significant effect on accuracy. An asterisk marks significant differences ($p < .05$ after Holm correction).

suggest the biggest differences involve *Cue*, whereas effects involving *Outline* provide weaker evidence after correction. *Transparency* showed limited practical benefit as several participants reported that they did not register the reduced opacity as an indicator at all, and its perceptual strength depended on object luminance and background. By design, it also made the marked object itself harder to see. This contrasts with AR-based telepresence, where a see-through avatar was valued because it signals virtuality [11]. A full-color avatar seen against an unaltered room may make reduced opacity far more noticeable than our edge-stylized objects did. Further, subjective preferences were polarized between *Cue* and *Outline*. Some participants overlooked the *Outline*. This however might not be a property of outlines as such, but of outlines under edge stylization as our filter rendered the entire scene as contours, so an added contour competes on a dimension already saturated with visual noise. In addition, our *Outline* was considerably less bright than the green *Cue*, which may explain part of the difference between the two. We therefore suggest that an indicator should not share its visual dimension with the source of ambiguity, and that outlines, the closest of our designs to the glow effects preferred in prior work [12, 13], may perform better under photorealistic ambiguity. Others found *Cue* visually detached and potentially cluttering. Overall, and most importantly, misclassifications persisted even with salient indicators, raising concerns about reliability in everyday AR where attention is divided and mistakes can be unsafe [17].

Prior work on the ethics of realism argues that the critical issue is not absolute photorealism but whether users can still reliably decouple immediate sensory experience from ground truth [19]. In practice, virtual elements only need to be convincing enough, or sufficiently well integrated into the surrounding environment, to remain unnoticed in the moment, as recent work has already shown through perceptual transitions between Virtual and Augmented Reality environments [24]. Our study cannot quantify how much realism is required, since we induced ambiguity from the opposite direction. What it does show is that once discriminability is low, indicators mitigate but do not eliminate misclassification. Consequently, indicators may not be a complete solution, as any robust marking system would require consistent implementation across AR platforms, applications, and content types while also managing

trade-offs around clutter, habituation, and trust. Principles for responsible innovation in AR and VR already call for such markers, for example "reality watermarks" that distinguish platform-generated, advertiser-placed, and user-created virtual content [3]. Our results suggest that such markers need empirical validation, as even our most salient indicator did not prevent all misclassifications. Nevertheless, the central implication is that users' immediate visual perception can no longer be treated as fully reliable evidence of what is physically present when using such technologies. We therefore urge the research community to investigate physical-virtual ambiguity more systematically and to iterate on the indicators presented here toward more salient, practical, and context-aware mechanisms for raising user awareness in everyday AR.

Yet, our results must be interpreted cautiously. Most importantly, induced ambiguity is a proxy in that it degrades the physical scene, so effects cannot be attributed to indicator design alone. Moreover, the induced ambiguity reduced but did not eliminate discriminability, as *Baseline* accuracy was above chance. Our findings therefore describe a state of low, not zero, discriminability, and indicators may matter even more when objects are truly indistinguishable. Our indicators were also not matched in color and luminance as the *Cue* was a brighter, more saturated green than the *Outline*, and *Transparency* reduced the visibility of the very object it marked. Differences between indicators may therefore partly reflect overall salience rather than the visual dimension each indicator uses. Future comparisons should vary one visual parameter at a time while keeping size, color, and brightness constant [11]. Further, our participants were mostly AR novices, who may be less familiar with how passthrough and virtual content usually look, which may have increased overall uncertainty. Experienced users might classify objects more accurately even without indicators, but might also habituate to indicators faster. Degrading the passthrough view also makes the physical world look more virtual, which may shift participants' priors toward classifying objects as virtual. Ratings were collected verbally by a non-blinded experimenter. TCT may be confounded by disorienting reorientation after blindfolding, and SEQ responses may vary due to differing participant interpretations. We found no significant effect of indicators on presence, in line with prior work reporting limited direct impacts of visual manipulation on presence [7, 21]. However, presence was descriptively lowest for *Cue* and *Outline*, so a small cost of salient indicators cannot be ruled out. Future work should evaluate indicator designs that balance unobtrusiveness with sufficient salience under divided attention, ideally in longer-term, realistic deployments that capture habituation and learning. Promising directions include conditionally activating cues (e.g., manual toggles or context-aware triggering via distance [1]) to reduce clutter while maintaining safety. Finally, research should more explicitly justify when photorealistic AR is desirable and map use-case-specific risks to appropriate safeguards and ethical considerations [18]. In multisensory, distributed mixed reality, harm can also arise from unwanted touch-like stimulation [9], so future indicators may need to mark virtuality beyond the visual channel.

4 Conclusion

In this paper, we investigated how AR environments can support users in reliably discerning physical from virtual objects as AR approaches photorealism. Using a controlled AR classification task with stylized passthrough and three visualization designs for *Outline*, *Cue*, and *Transparency*, we show that one of three tested indicators, a hovering cue, significantly improved the confidence-weighted accuracy relative to no indicator (*Baseline*), while outline and transparency did not reach significance after correction. Yet none are foolproof, and misclassifications occurred even with salient cues. These results reveal a core design tension in AR systems in that indicators must remain noticeable and trustworthy while avoiding excessive visual clutter, especially in everyday contexts.

Acknowledgments

This work is supported by the German Research Foundation (DFG), CRC 1404: "FONDA: Foundations of Workflows for Large-Scale Scientific Data Analysis" (Project-ID 414984028).

References

[1] Till Ballendat, Nicolai Marquardt, and Saul Greenberg. 2010. Proxemic Interaction: Designing for a Proximity and Orientation-Aware Environment. In *ACM International Conference on Interactive Tabletops and Surfaces (ITS '10)*. Association for Computing Machinery, New York, NY, USA, 121–130. doi:10.1145/1936652.1936676

[2] Kaiming Cheng, Jeffery F. Tian, Tadayoshi Kohno, and Franziska Roesner. 2023. Exploring User Reactions and Mental Models towards Perceptual Manipulation Attacks in Mixed Reality. In *Proceedings of the 32nd USENIX Conference on Security Symposium (SEC '23)*. USENIX Association, USA, 911–928.

[3] Alicia G. Cork, Mike Richardson, Heide K. Lukosch, Mohamed Khamis, Christopher Katins, Veronika Krauß, Lauren Ruffin, Sarah Papazoglakis, Victoria Sanchez, Xueni Pan, Michael J. Proulx, and Danaë Stanton Fraser. 2025. Shaping the Future: Principles for Policy Recommendations for Responsible Innovation in Virtual Worlds. *Frontiers in Virtual Reality* 6 (Oct. 2025). doi:10.3389/frvir.2025.1645330

[4] Kody R. Dillman, Terrance Tin Hoi Mok, Anthony Tang, Lora Oehlberg, and Alex Mitchell. 2018. A Visual Interaction Cue Framework from Video Game Environments for Augmented Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3173574.3173714

[5] Nina Doerr, Katrin Angerbauer, Melissa Reinelt, and Michael Sedlmair. 2023. Bees, Birds and Butterflies: Investigating the Influence of Distractors on Visual Attention Guidance Techniques. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, 1–7. doi:10.1145/3544549.3585816

[6] David Drascic and Paul Milgram. 1996. Perceptual Issues in Augmented Reality. In *Stereoscopic Displays and Virtual Reality Systems III*, Vol. 2653. SPIE, 123–134. doi:10.1117/12.237425

[7] Crescent Jicol, Christopher Clarke, Emilia Tor, Rebecca M Dakin, Tom Charlie Lancaster, Sze Tung Chang, Karin Petrini, Eamonn O'Neill, Michael J Proulx, and Christof Lutteroth. 2023. Realism and Field of View Affect Presence in VR but Not the Way You Think. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–17. doi:10.1145/3544548.3581448

[8] Christopher Katins, Katja Chen, Christopher Lazik, Pascal Knierim, Karola Marky, and Thomas Kosch. 2025. Public Perceptions on Psychological and Social Impacts of Virtual and Augmented Reality. In *Proceedings of the Mensch Und Computer 2025 (MuC '25)*. ACM, New York, NY, USA, 163–180. doi:10.1145/3743049.3743079

[9] Katja Krug, Wolfgang Büschel, and Marc Satkowski. 2024. Safeguarding Sensations: Harm Prevention in Multisensory Distributed Mixed Reality. In *2024 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, Bellevue, WA, USA, 6–8. doi:10.1109/ISMAR-Adjunct64951.2024.00008

[10] Daniel Lange, Tim Claudius Stratmann, Uwe Gruenefeld, and Susanne Boll. 2020. HiveFive: Immersion Preserving Attention Guidance in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3313831.3376803

[11] Minna Pakanen, Paula Alavesa, Niels van Berkel, Timo Koskela, and Timo Ojala. 2022. "Nice to See You Virtually": Thoughtful Design and Evaluation of Virtual

Avatar of the Other User in AR and VR Based Telexistence Systems. *Entertainment Computing* 40 (Jan. 2022), 100457. doi:10.1016/j.entcom.2021.100457

[12] Minna Pakanen and Leena Arhippainen. 2014. User Experiences with Web-Based 3D Virtual Travel Destination Marketing Portals: The Need for Visual Indication of Interactive 3D Elements. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design (OzCHI '14)*. Association for Computing Machinery, New York, NY, USA, 430–439. doi:10.1145/2686612.2686680

[13] Minna Pakanen, Leena Arhippainen, Seamus Hickey, and Antti Karhu. 2013. Visual Indication of Interactive 3D Elements in 3D Virtual Environments. In *Proceedings of International Conference on Making Sense of Converging Media (AcademicMindTrek '13)*. Association for Computing Machinery, New York, NY, USA, 1–4. doi:10.1145/2523429.2523452

[14] Matti Pouke, Johanna Ylipulli, Ilya Minyaev, Minna Pakanen, Paula Alavesa, Toni Alatalo, and Timo Ojala. 2018. Virtual Library: Blending Mirror and Fantasy Layers into a VR Interface for a Public Library. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia (MUM '18)*. Association for Computing Machinery, New York, NY, USA, 227–231. doi:10.1145/3282894.3282927

[15] Philipp A. Rauschnabel, Reto Felix, Chris Hinsch, Hamza Shahab, and Florian Alt. 2022. What Is XR? Towards a Framework for Augmented and Virtual Reality. *Computers in Human Behavior* 133 (Aug. 2022), 107289. doi:10.1016/j.chb.2022.107289

[16] Jeff Sauro and Joseph S. Dumas. 2009. Comparison of Three One-Question, Post-Task Usability Questionnaires. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1599–1608. doi:10.1145/1518701.1518946

[17] Susanne Schmidt, Oscar Javier Ariza Nunez, and Frank Steinicke. 2019. Blended Agents: Manipulation of Physical Objects within Mixed Reality Environments and Beyond. In *Symposium on Spatial User Interaction (SUI '19)*. Association for Computing Machinery, New York, NY, USA, 1–10. doi:10.1145/3357251.3357591

[18] Mel Slater. 2021. Beyond Speculation About the Ethics of Virtual Reality: The Need for Empirical Results. *Frontiers in Virtual Reality* 2 (Aug. 2021). doi:10.3389/frvir.2021.687609

[19] Mel Slater, Cristina Gonzalez-Liencres, Patrick Haggard, Charlotte Vinkers, Rebecca Gregory-Clarke, Steve Jelley, Zillah Watson, Graham Breen, Raz Schwarz, William Steptoe, Dalila Szostak, Shivashankar Halan, Deborah Fox, and Jeremy Silver. 2020. The Ethics of Realism in Virtual and Augmented Reality. *Frontiers in Virtual Reality* 1 (March 2020), 1. doi:10.3389/frvir.2020.00001

[20] Irwin Sobel and Gary Feldman. 1973. A 3x3 Isotropic Gradient Operator for Image Processing. In *Pattern Classification and Scene Analysis*, Richard O. Duda and Peter E. Hart (Eds.). Wiley, New York, 271–272.

[21] William Steptoe, Simon Julier, and Anthony Steed. 2014. Presence and Discernability in Conventional and Non-Photorealistic Immersive Augmented Reality. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 213–218. doi:10.1109/ISMAR.2014.6948430

[22] Wen-Jie Tseng, Elise Bonnail, Mark McGill, Mohamed Khamis, Eric Lecolinet, Samuel Huron, and Jan Gugenheimer. 2022. The Dark Side of Perceptual Manipulations in Virtual Reality. In *CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–15. doi:10.1145/3491102.3517728

[23] Martin Usoh, Ernest Catena, Sima Arman, and Mel Slater. 2000. Using Presence Questionnaires in Reality. *Presence* 9, 5 (Oct. 2000), 497–503. doi:10.1162/105474600566989

[24] Jonas Wombacher, Zhipeng Li, and Jan Gugenheimer. 2025. SwitchAR: Enabling Perceptual Manipulations in Augmented Reality Leveraging Change Blindness and Inattentional Blindness. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, 1–4. doi:10.1145/3706599.3721159