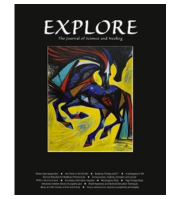
Explore 22 (2026) 103506

Contents lists available at ScienceDirect

EXPLORE

journal homepage: www.elsevier.com/locate/jsch



Original Research

Assessing correspondence between subjective reports and visual targets during attempted out-of-body experiences

Marina Weiler[a,*], Damon Abraham[b], David J.P. Acunzo[a], Niffe Hermansson[c]

[a] Department of Psychiatry and Neurobehavioral Sciences, University of Virginia Health System, Charlottesville, VA, USA
[b] Entangled Labs, Irvine, CA, USA
[c] Department of Statistics, Faculty of Science, University of Auckland, New Zealand

ARTICLE INFO

Keywords:
veridical perception
extrasensory perception
anomalous cognition
remote viewing
semantic similarity
text embeddings

ABSTRACT

Claims of veridical perception often accompany out-of-body experiences (OBEs), but systematic evaluation under controlled conditions remains difficult. We examined whether descriptions produced after attempted OBEs corresponded to concealed visual targets beyond chance using computational semantic similarity. Twenty-one individuals affiliated with an institute specializing in OBE training attempted to perceive a randomly selected target displayed in a separate, inaccessible room. Targets were drawn from a standardized database and remained unknown to participants and researchers. Ten participants reported an OBE, and 13 provided target descriptions for analysis. Participant descriptions and target captions were converted into high-dimensional vector embeddings using a pretrained language model. For each trial, cosine similarity between the participant's description and each of 100 candidate targets was computed; the correct target was evaluated against the within-trial decoy distribution using a rank-based metric (our primary analysis) and standardized z-scores (a secondary, descriptive analysis). Control analyses confirmed that the pipeline identified correct targets when semantic correspondence was present, though these controls do not establish sensitivity to naturalistic human language. No individual trial showed statistically significant above-chance correspondence, and a combined rank-based test across all 13 trials likewise showed no evidence of above-chance performance. Two participants spontaneously reported accurate details about experimenters' locations and activities in an adjacent room. Although the target-identification results did not support anomalous cognition, the spontaneous environmental reports may have methodological value and warrant systematic investigation.

Introduction

Out-of-body experiences (OBEs) are characterized by the subjective feeling of disembodiment — the sense that one's consciousness exists apart from the physical body — regardless of other accompanying phenomenological features[1–3] This distinguishes OBEs from related but phenomenologically distinct states, such as autoscopic hallucinations or heautoscopy of neurological origin, in which an individual perceives a visual double of the body without the accompanying sense of detachment from it.[4] Historically, they have been associated with claims of acquiring information inaccessible through ordinary sensory channels.[5] Reports of veridical perception during OBEs have appeared in spontaneous case collections,[6,7] and survey research suggests that approximately one-third of spontaneous OBE narratives include claims of extrasensory perception (ESP).[a,8] While such reports do not constitute controlled evidence, they motivate empirical investigation under laboratory conditions.

Experimental attempts to test ESP[a] during OBEs began in the mid-20th century. Early work by Charles Tart examined hypnosis-induced OBEs in controlled settings.[9] In one frequently cited case ("Miss Z"), a participant correctly reported a randomly generated five-digit number during an alleged OBE, a statistically significant result given the

* Corresponding author at: Department of Psychiatry and Neurobehavioral Sciences, University of Virginia School of Medicine, P. O. Box 800152, Charlottesville, VA 22908, USA.

E-mail address: zzs2jq@uvahealth.org (M. Weiler).

[a] We use the term extrasensory perception (ESP) to refer to the putative acquisition of information about external events, objects, or locations through means not mediated by the known senses (Cardeña & Alvarado, 2014). Cardeña E, Alvarado CS. Anomalous self and identity experiences. In: Cardeña E, Lynn SJ, Krippner S, eds. Varieties of Anomalous Experience: Examining the Scientific Evidence, 2nd ed. American Psychological Association; 2014:175-212

https://doi.org/10.1016/j.explore.2026.103506
Received 23 June 2026; Received in revised form 24 July 2026; Accepted 7 August 2026
Available online 12 August 2026

1-in-100,000 chance of this occurring. Subsequent experiments with Robert Monroe and other participants yielded mixed results, offering limited support for ESP during OBEs.[9]

John Palmer's research in the 1970s investigated whether OBEs or hypnagogic states could enhance ESP abilities. Despite a substantial number of participants reporting OBEs, the results showed no significant correlation between OBEs and ESP performance.[10–12] This suggested that OBEs may be a byproduct of entering a deep hypnagogic state rather than directly contributing to ESP. Smith and Irwin's (1981)[13] study found a correlation between a participant's 'OBEness' score and ESP performance, although methodological concerns regarding the use of identical targets raised questions about bias. In recent years, research has continued to explore the possibility of non-local perception during OBEs. For instance, Del Prete and Tressoldi (2005)[14] conducted an experiment where participants exhibited significantly higher hit rates during hypnotically induced OBEs compared to self-relaxation conditions, suggesting that OBEs may influence ESP performance.

Taken together, the literature suggests that claims of veridical perception during OBEs have persisted for decades, yet robust experimental support remains inconclusive. A central challenge has been methodological: small sample sizes, limited blinding, concerns about target leakage, subjective judging procedures, and analytic flexibility have complicated interpretation. Advances in experimental design, stimulus control, and computational text analysis now provide opportunities to reassess such claims under more stringent and transparent conditions. We note that laboratory paradigms of this kind have themselves been the subject of methodological critique, on the grounds that sterile, fixed-target designs may misalign with the attentional and motivational dynamics of the OBE state.[15] We return to this critique in the Discussion, situating the present findings within it.

The present investigation addresses a focused empirical question: whether semantic correspondence between participant descriptions and visually presented targets exceeds chance expectation under controlled conditions. By integrating rigorous blinding procedures, standardized image stimuli, and computational text-embedding methods for quantitative similarity assessment, this study applies contemporary analytic tools to a longstanding experimental paradigm. In doing so, it aims to provide transparent, methodologically robust data to inform ongoing discussions about perception during OBEs. Beyond its empirical status, the question of whether OBEs are accompanied by genuine ESP also carries theoretical significance: if such perception occurs and cannot be accounted for by known sensory or inferential mechanisms, it would suggest that some aspect of mind or consciousness operates beyond the explanatory repertoire of conventional brain-based accounts of perception.[5]

Methods

Participants

Forty individuals affiliated with the *International Institute of Projectiology and Conscientiology* (IIPC, Brazil) were initially screened. The IIPC is a Brazilian institution whose training programs emphasize meditation, visualization, and techniques described by participants as involving energetic self-regulation, to facilitate OBEs. Participants were asked: "If I asked you to induce an out-of-body experience now, what do you think your likelihood of success would be?" Individuals were eligible for inclusion if they estimated a success likelihood greater than 40%. Based on this criterion, 21 participants were included in the study. Of these, 11 were female, and the mean age was 47 years (SD = 10). The study was approved by the University's Institutional Review Board, and all participants provided informed consent before participation.

Experimental setup

Over three weeks, a series of experimental sessions was conducted at the IIPC headquarters in Foz do Iguaçu, Brazil. The facility consisted of a small building containing three enclosed rooms (Rooms 1–3). Participants were scheduled for either morning or afternoon sessions and completed the protocol individually. Room 1, designated for the attempted induction of OBEs, was furnished with two beds, an armchair, and blankets, allowing participants to choose between a supine or seated position. Participants were also permitted to adjust ambient temperature and lighting to create conditions they perceived as conducive to the intended experiential state. During all sessions, electroencephalography (EEG) data were also recorded as part of a broader investigation of the neurophysiological correlates of these experiences. However, analysis of the EEG data is beyond the scope of the present paper and will be reported separately.

Once participants were comfortable and the EEG signal was stable, they were instructed to attempt to perceive a randomly selected target displayed in Room 3. They were given up to one hour to induce their OBEs (except for two participants who requested and were granted 90 min). The door from Room 1 was left ajar for participant monitoring, and the door from Room 3 remained closed during the session. As soon as participants were told they could begin using their preferred techniques to induce an OBE (for OBE induction techniques used by the participants, see,[16] the researchers moved to Room 2, where they remained throughout the session to monitor the physiological recordings. No interaction occurred between the researchers and participants during the experiment.

Target presentation and blinding

Room 3, located approximately 7 m from Room 1, contained a small desk with a laptop that displayed the target image. A custom MATLAB script randomly selected one image from a pool of 100 standardized stimuli, selected from the BOSS database,[17] to be displayed on the laptop screen. Each image contained an object against a uniform white background. The laptop screen was oriented to face the wall to prevent any inadvertent visual access if someone opened the door of Room 3. The 100 target images were hand-selected from the larger BOSS database by the first author to maximize visual and semantic diversity across the pool (e.g., object category, color palette, and structural complexity), rather than drawn at random or restricted to a pre-normed BOSS subset. Images were not specifically screened for cultural specificity or familiarity relative to the Brazilian sample. The randomization script (MATLAB) used to select the session-specific target from this pool is available in the study's OSF repository (10.17,605/OSF.IO/86UBZ).

While one researcher prepared the participant in Room 1, a second researcher initiated the MATLAB script in Room 3 by pressing the "Enter" key. The target image was displayed with a delay of 10 s, during which the experimenter turned the laptop so that the screen faced the wall, exited the room, and closed the door. Neither researcher viewed the target image at any point. The identity of the target remained unknown to both participants and researchers until the data analysis phase, which took place several months later.

In retrospect, we note that BOSS images — discrete objects isolated against a uniform white background — are visually and semantically decontextualized relative to the kind of content typically reported in spontaneous OBEs, which often involves emotionally salient, socially meaningful, or environmentally embedded scenes rather than isolated objects. This limitation was not anticipated at the time of study design; recognizing it in the present data aligns with a subsequent methodological analysis,[15] which we draw on further in the Discussion.

Session termination and interview procedure

Once the allotted time to induce the OBE had elapsed (60–90 min), a member of the research team returned to Room 1 to inform the participant that the session had ended. Simultaneously, a second researcher entered Room 3, terminated the image display by pressing the "Esc" key, while the laptop was still facing the wall, and transferred the laptop to the researcher's room (Room 2). After Room 3 was confirmed to be clear, participants were escorted into that room for a post-session interview.

Interviews were conducted to document participants' experiences. Participants were asked whether they believed they had experienced an OBE, to describe the phenomenology of their experience, and to report any perceptions they believed were related to the target image. Responses were provided in free-form narrative format. Participants were also asked to rate their confidence in any reported perceptions on a scale from 0 (no confidence) to 5 (maximum confidence).

Target caption generation

Text captions were generated for each image in the target pool (N = 100) using Google's Gemini 2.5 Flash multimodal model. Images were submitted programmatically through a Python pipeline, together with a standardized prompt instructing the model to identify the main objects, shapes, colors, and textures/materials of the subject in the photograph. These captions were generated once and stored in a CSV file for all subsequent analyses.

The prompt used for literal caption generation was: 'Identify and literally describe the main objects, shapes, colors, and textures/materials of the subject in the photograph. Do not describe the background. The description should be a single, concise paragraph.' The prompt used for conceptual caption generation was: 'Identify and conceptually describe the higher-level scene category, functional context, and abstract or thematic associations suggested by the subject in the photograph. Do not describe specific visual details such as individual objects, shapes, colors, or textures. The description should be a single, concise paragraph.'

Text embedding and preprocessing

Participant descriptions and target captions were converted into vector embeddings using OpenAI's text-embedding-3-large model. Embeddings were generated using a Python pipeline that loaded caption text from CSV files, removed line breaks, and submitted each caption programmatically to OpenAI via its application programming interface (API) to obtain high-dimensional embedding vectors. The resulting vectors were stored in JSON format, allowing all subsequent analyses to be conducted offline. The final dataset consisted of 100 target caption embeddings and 13 participant trial captions with associated embeddings. All embeddings were generated once using a fixed model version and identical preprocessing parameters, stored locally, and reused for all analyses to ensure reproducibility.

Semantic similarity analysis

We computed two complementary metrics of semantic correspondence for each trial: a rank-based metric, which we treat as our primary inferential statistic given that it does not depend on the shape of the decoy similarity distribution, and standardized z-scores, which we report as a secondary, descriptive analysis.

For each participant trial, cosine similarity was computed between the participant's caption embedding and each of the other 100 possible target caption embeddings using an embedding-based semantic similarity approach similar to that described previously.[18] The 99 non-selected target captions served as decoys and were used to define the within-trial null distribution of semantic similarity (i.e., similarity to randomly selected non-target stimuli). The cosine similarity value for the true target was then compared against this distribution and converted into a z-score using the sample mean and standard deviation, indicating how closely the participant's description matched the actual target relative to the distribution of non-selected targets. Each trial's z-score, therefore, reflected the degree of semantic correspondence between the transcript and the true target relative to chance expectation. Positive z-scores indicate that the true target is more semantically similar to the participant's response than expected relative to the decoy set, whereas negative z-scores indicate lower-than-expected similarity.

In addition to z-score standardization, we computed a rank-based metric for each trial. The cosine similarity values were ordered from highest to lowest, and the rank of the true target was defined as its ordinal position within this distribution, with rank = 1 indicating the highest semantic similarity and rank = 100 indicating the lowest. This rank-based measure provided a nonparametric index of relative correspondence, free from distributional assumptions, and allowed straightforward interpretation of how the true target performed relative to all alternative candidates within a given trial.

This procedure situates the target similarity within the empirical distribution defined by the decoys. The decoy set serves as an appropriate within-trial null model because it represents the distribution of semantic similarities expected under random matching between the participant's caption and non-target captions drawn from the same stimulus pool, thereby controlling for baseline similarity structure in the embedding space. Finally, two-tailed p-values were computed from these z-scores to quantify the extremity of the observed target similarity under the within-trial null distribution. Pooled across all 13 trials, the decoy cosine-similarity distribution ($n = 1287$ comparisons) was right-skewed (skewness = 0.71) with excess kurtosis (1.52) and departed from normality (Shapiro-Wilk $p < .001$), which we take into account in our treatment of the rank-based metric as the primary inferential statistic (see Results).

To further contextualize performance, we computed the same metrics (cosine similarity, z-scores based on decoy mean and decoy std, and rank among all 100 candidates) under two control conditions derived from the stimulus set itself. In the first control condition, each image was paired with its own literal caption, consisting of a direct, object-focused description of the image's visible elements (e.g., primary objects, shapes, colors, and textures). In the second control condition, each image was paired with a more conceptual caption, designed to capture higher-level semantic or thematic features (e.g., scene category, functional context, or abstract associations), which may be more appropriate for evaluating descriptions that are impressionistic, symbolic, or non-literal in character.

For both control conditions, each caption was treated analogously to a participant description and compared against all 100 target captions. Cosine similarity distributions, decoy mean, decoy std, z-scores, and rank of the correct image were computed using the identical analytic pipeline applied to participant data. These control analyses served two purposes. First, they provided a benchmark for the discriminative sensitivity of the embedding-based similarity method, quantifying the extent to which the analytic pipeline can successfully identify the correct image when given an accurate description. Second, they allowed us to evaluate whether the method is differentially sensitive to literal versus conceptual forms of correspondence. This distinction is particularly relevant in ESP-related paradigms, where reported perceptions are often described as fragmentary, symbolic, or semantically abstract rather than strictly perceptual replicas of the stimulus.

It should be noted that both control conditions used captions generated by the same AI model applied to the stimulus set, and therefore assess the pipeline's sensitivity within an AI-generated linguistic register. They do not directly establish how the pipeline performs when applied to naturalistic, first-person human language, which may differ systematically from AI-generated captions in structure and content.

Results

Participants in this study reported a mean age of 13 ± 10 years for their first OBE, with an average of 19 ± 9 years of practice in techniques aimed at inducing them. Of the 21 participants, 10 reported experiencing a state they identified as an OBE during the experimental session.

This state was characterized by a subjective sense of disembodiment or perceiving the environment from a vantage point outside the physical body. Among this group, eight participants claimed to have “visually” perceived the target stimulus. The other two participants reported having had OBEs, but not “seeing” anything related to the target.

Interestingly, however, a broader group of 13 participants—including but not limited to the seven who reported OBEs—described perceiving elements of the target stimulus. Notably, five of these 13 did not report having had an OBE. Instead, they described the perceptual content as appearing internally, often referring to a “mental screen” or “inner vision.” These individuals stated that while they did not feel as though they had left their bodies, they nonetheless received or formed impressions of the target during the experimental period. Of the 10 participants who reported experiencing an OBE during the session, eight volunteered content they associated with the target and were included in the semantic similarity analyses. The remaining 2 (P6 and P9) did not provide any target-related content; in both cases, their reported OBEs involved attention directed toward unrelated content rather than the target (Fig. 1). Out of the 13 participants who provided descriptions of the target, nine provided a self-rated confidence score (five of them had an OBE, and four did not). Among the 9 who provided, one rated their perception as 0, two as 1, two as 2, one as 3, one as 4, and two as 5 (Table 1).

Participants’ descriptions of the target stimuli varied in content and complexity, but typically included references to colors, general shapes, or broad categories such as a landscape or objects. Across the 13 analyzed trials, within-trial cosine similarity distributions showed substantial variability, ranging from 0.105 to 0.233, and were generally close to the corresponding decoy means (range: 0.148–0.269). Decoy-pool cosine similarities (i.e., the null distribution against which each trial’s target similarity was standardized) had a pooled mean of 0.210 ($SD = 0.067$) across all 13 trials ($n = 1287$ decoy comparisons total), with per-trial decoy standard deviations ranging from 0.035 to 0.069 (Table 2). As a secondary, descriptive analysis, standardized effects (z-scores) ranged from −1.240 to 0.851 (mean $z \approx -0.42$), with 9 of 13 trials negative — a pattern better described as skewed toward below-chance performance than as centered or distributed around zero. No individual trial approached statistical significance (all p-values $\geq .215$). A one-sample *t*-test on the pooled z-scores was marginally significant in the negative direction ($t = -2.28$, $p = .042$); we do not treat this as a robust conclusion, however, as the small sample of 13 trial-level z-scores provides limited power to assess the adequacy of the normal approximation (see Methods), making this parametric result assumption-dependent rather than robust (Table 2).

Table 1
Participant characteristics and session-level self-reports.

Participant	Age at first OBE	Years inducing OBEs	Had an OBE?	Provided descriptions of the target?	Self-rated confidence score
1	7	6	No	No	-
2	12	12	Yes	**Yes**	2
3	8	28	No	**Yes**	1
4	25	11	No	**Yes**	-
5	5	7	No	No	-
6	8	20	Yes	No	-
7	13	10	Yes	**Yes**	-
8	3	8	Yes	**Yes**	3
9	7	36	Yes	No	-
10	7	15	Yes	**Yes**	-
11	20	15	No	No	-
12	14	12	No	**Yes**	1
13	7	25	Yes	**Yes**	-
14	13	26	Yes	**Yes**	0
15	8	19	No	**Yes**	2
16	7	13	Yes	**Yes**	5
17	46	25	No	No	-
18	37	15	Yes	**Yes**	5
19	13	17	No	**Yes**	4
20	10	33	No	No	-
21	6	36	No	No	-

Note. Table 1 summarizes demographic and experiential characteristics of the 21 participants included in the experimental sessions. Participants reported first experiencing OBEs at a mean age of 13 years (*SD* = 11; range = 3–46) and had practiced OBE-induction techniques for an average of 19 years (*SD* = 9; range = 6–36). During the experimental session, ten participants reported having had an OBE. However, 13 participants provided descriptions of perceived target-related content and were therefore included in the semantic similarity analyses. Among those who provided target descriptions, nine participants also gave a self-rated confidence score, ranging from 0 to 5. A dash indicates that no confidence rating was provided.

Rank-based analyses, which we treat as our primary inferential approach, yielded a similar pattern. The true target ranked between 16th and 92nd out of 100 possible images (median rank near the midpoint), corresponding to empirical percentiles ranging from 8.081 to 84.848. Only one trial (P10) placed the target within the top 20 (rank = 16), and none ranked the target first. A combined test across all 13 trials (mean rank = 61.5 against a null expectation of 50.5) did not reach significance (Wilcoxon signed-rank test, $p = .147$). Overall, we find no evidence of above-chance semantic correspondence between participant

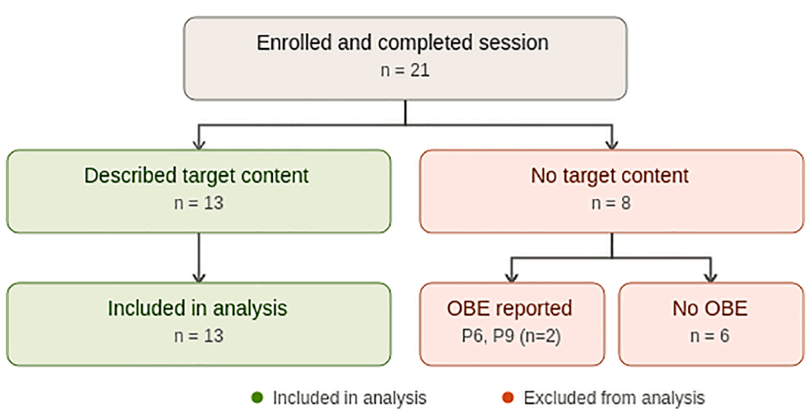


Fig. 1. Participant selection flow diagram.
Flow diagram of participant selection from enrollment to inclusion in the semantic similarity analyses. Of 21 enrolled participants who completed a session, 13 volunteered content during the post-session interview that they associated with the concealed target and were included in the primary analyses; 8 did not offer any target-related content and were excluded. Of these 8, 2 participants (P6 and P9) had reported experiencing an OBE during the session but described attention directed toward unrelated content (e.g., one participant reported an OBE in which she perceived herself in a war zone assisting children) rather than the target; the remaining 6 did not report an OBE. Inclusion judgments were made by the first author under full blinding: the identity of the target image was not known to any member of the research team until the data analysis phase, conducted several months after data collection concluded.

Table 2
Semantic similarity between participant descriptions and concealed visual targets.

Participant	Decoy Mean	Decoy Std	Target Cosine Similarity	Target z-score	Target p-value	Target rank	Target empirical percentile
P02	0.249	0.055	0.227	−0.388	0.698	63	37
P03	0.177	0.035	0.142	−0.996	0.319	84	16
P04	0.220	0.047	0.210	−0.208	0.835	56	44
P07	0.242	0.052	0.233	−0.175	0.861	57	43
P08	0.196	0.052	0.139	−1.090	0.276	92	8
P10	0.182	0.052	0.226	0.851	0.395	16	85
P12	0.174	0.056	0.114	−1.073	0.283	90	10
P13	0.266	0.052	0.222	−0.849	0.396	79	21
P14	0.169	0.068	0.200	0.453	0.651	21	80
P15	0.226	0.068	0.215	−0.150	0.880	50	51
P16	0.209	0.069	0.227	0.256	0.798	24	77
P18	0.269	0.049	0.208	−1.240	0.215	89	11
P19	0.148	0.052	0.105	−0.824	0.410	78	22

Note. Table 2 presents the semantic similarity results for the 13 participants who provided descriptions of perceived target-related content. For each participant, the target cosine similarity was compared with the distribution of similarities between that participant's description and the 99 decoy images. Target z-scores ranged from −1.240 to 0.851, indicating that the correct target was generally no more semantically similar to the participant's description than were the decoy images. No trial reached statistical significance, with all target p-values $\geq$ .215. Target ranks ranged from 16th to 92nd out of 100 candidate images, and empirical percentiles ranged from 8.081 to 84.848. Overall, these results show no evidence of above-chance performance, with no evidence of systematic above-chance semantic correspondence between participant descriptions and the concealed targets.

descriptions and the correct target images during OBE trials (Table 2).

Because the 13 analyzed trials included eight participants who reported an OBE with target-related perception and five who reported internal imagery without an OBE, we additionally examined these two subgroups separately. Using our primary, rank-based approach, neither subgroup differed significantly from chance (OBE-confirmed: mean rank = 55.1, Wilcoxon signed-rank test, p = .641; imagery-only: mean rank = 71.6, p = .125), nor did the two subgroups differ significantly from one another (Mann-Whitney U test, p = .622). We note descriptively that participants in the imagery-only group produced descriptions that were, on average, less semantically similar to the AI-generated description of the target than would be expected by chance alone. We interpret this cautiously: it derives from a post-hoc subdivision of an already small sample, and our semantic similarity design does not offer as direct an interpretation of a below-chance deviation as would a forced-choice paradigm. We therefore report it as a weak, exploratory observation rather than a confirmed finding.

The full results for both control conditions are provided in the **Supplementary Materials**. In the literal-description control condition, the analytic pipeline demonstrated near-ceiling discriminative performance. Across all 13 trials, the correct target image ranked 1st among all 100 candidate images (100th empirical percentile in every case). Target cosine similarity values ranged from 0.714 to 0.853, yielding extremely large standardized effects (z-scores from 7.736 to 13.534), with all associated *p*-values $<$ .001. These results confirm that when the input description accurately and directly corresponds to the visual content of the image, the embedding-based similarity procedure reliably and unambiguously identifies the correct target (**Supplementary Table 1**).

In the conceptual-description control condition, performance remained robust but showed greater variability, as expected given the more abstract nature of the captions. In 10 of 13 trials, the correct target ranked 1st (100th percentile), and in one additional trial, it ranked 2nd (98.99th percentile). Two trials yielded lower ranks (7th and 12th–15th range), corresponding to non-significant z-scores (p = .086, .172, and .238). Target cosine similarity values ranged from 0.301 to 0.563, with z-scores spanning 1.179 to 6.082. Overall, these findings indicate that the analytic framework is sensitive not only to literal object-level correspondence but also, in most cases, to higher-level semantic or conceptual similarity, thereby establishing an empirical benchmark for interpreting participant performance (**Supplementary Table 2**).

Spontaneous accounts

In addition to the quantitative analyses, two participants reported spontaneous observations that did not pertain to the designated visual target but are noteworthy. These reports were not part of the predefined outcome measures and were volunteered during post-session interviews.

Participant 7 stated that during her OBE, she perceived the two experimenters in the adjacent monitoring room. She described Researcher 1 seated to the left, focused on a computer screen, and Researcher 2 seated farther back in a corner, reading a book. According to contemporaneous session records, this description corresponded to the actual configuration at that specific time—namely, while the participant was in the induction room with the doors closed. Immediately after the session ended and the participant was escorted out, the seating arrangement changed. Importantly, neither researcher had adopted that specific configuration nor engaged in reading books during prior sessions.

Participant 13 independently reported perceiving the experimenters in the adjacent room while attempting to view the target. She described Researcher 1 seated on the right, taking notes, and Researcher 2 positioned on the left at the computer. This configuration was also accurate for that session and occurred only while the participant was in the other room connected to the EEG equipment, with the doors closed. As with Participant 7, this arrangement differed from the seating positions typically used across sessions.

These observations are presented for completeness, to illustrate findings that motivate further investigation. They were anecdotal, not systematically elicited, and were not subjected to formal statistical analysis. No other participants provided spontaneous environmental descriptions.

Discussion

In the present study, we examined whether verbal reports produced following attempted OBEs would show greater semantic correspondence to concealed visual targets than expected by chance. Using a fully blinded design, standardized stimuli, and embedding-based computational similarity analysis, we quantified the correspondence between participant descriptions and all 100 candidate target images within each trial. Across 13 analyzed sessions, similarity scores for the true target did not reliably exceed the within-trial null distribution derived from decoy images. Standardized effects were small and centered around zero, and rank-based analyses showed that correct targets were broadly distributed across possible positions rather than clustering near the top of the candidate set. Under the present experimental conditions, therefore, participant reports did not demonstrate systematic above-chance semantic alignment with the concealed targets.

Importantly, control analyses confirmed that the analytic pipeline was capable of detecting genuine semantic correspondence when it existed. When literal, object-focused captions were compared against the target set, the correct image ranked first in every case with extremely large z-scores. Conceptual captions, designed to capture more abstract, thematic properties, also performed strongly in most trials, albeit with greater variability. These findings establish that the embedding-based method is sensitive to both literal and higher-level semantic similarity when the input text is itself AI-generated. However, because both control conditions relied on captions generated by the same multimodal AI model used to caption the target pool, they demonstrate within-register validity rather than cross-register sensitivity: they do not establish how naturalistic, first-person human descriptions — of the kind produced by participants — would compare to AI-generated captions in this embedding space. Given that conceptual (more abstract) AI captions already showed a marked decline in discriminative performance relative to literal AI captions, a comparison involving human phenomenological language, which is likely to differ further in register, vocabulary, and structure from AI-generated captions, could plausibly show an even larger decline. We therefore cannot rule out that part of the null result in the OBE trials reflects this register mismatch rather than the absence of genuine target-related correspondence, and we treat this as an important limitation of the present design rather than as fully resolved by the control analyses.

A notable phenomenological finding was that over a third of the participants who reported perceiving aspects of the target did not report having experienced an OBE. Instead, they described receiving impressions on a "mental screen" or through inner imagery, without a sense of spatial disembodiment. This dissociation between the subjective experience of leaving the body and the experience of target-related imagery is theoretically relevant: the present data suggest that target-related impressions—veridical or not—may occur independently of the full phenomenological profile of an OBE. This aligns with earlier laboratory work suggesting that ESP attempted during OBEs may be associated with hypnagogic or imagery-prone states[10–,12] and that disembodiment per se may not be the operative variable in anomalous cognition paradigms.[1,19]

The broader pattern of null target-identification results is consistent with much of the prior experimental OBE literature[9,11–13,20], and others generally failed to demonstrate reliable above-chance ESP performance during OBEs. However, Tart's early investigations produced mixed and interesting findings, and of particular relevance is the case of participant Robert Monroe. Although Monroe did not correctly identify the concealed target in that session, he reportedly provided accurate, detailed descriptions of the technician's adjacent room, including contextual details that were not part of the formal target. This pattern—failure on predefined symbolic targets but apparent success in describing aspects of the immediate environment—bears resemblance to the spontaneous observations reported in the present study. Two participants described the researchers' seating positions and activities in the adjacent room with accuracy, despite those details not being part of the experimental outcome measures. Future studies could strengthen the evidentiary value of such spontaneous environmental reports by incorporating a fully blinded, prearranged target embedded in the researchers' own appearance or environment during the session (for example, an unusual or distinctive item of clothing unknown to participants in advance). A spontaneous, specific description of such a detail, given that participants would have no ordinary means of anticipating or guessing it, would constitute stronger evidence than can be claimed for incidentally accurate reports of researcher positioning or activity, as observed in the present study.

These observations parallel a broader literature on veridical perception occurring in near-death experiences (NDEs) during cardiac arrest, a type of OBE that occurs under conditions of severely compromised or absent normal sensory function. In one frequently cited case, a comatose, cyanotic patient undergoing resuscitation later accurately described where a nurse had placed his dentures during intubation — including the specific drawer and its contents — despite being deeply unresponsive and requiring ventilation throughout the resuscitation.[21] The AWARE and AWARE-II studies[22,23] represent the most direct methodological precedent to the present design, having prospectively incorporated concealed visual and auditory targets in hospital resuscitation areas, alongside continuous EEG and cerebral oxygenation monitoring, to test for awareness and veridical perception during cardiac arrest. However, due to low survival rates and the small number of patients who both survived and could be interviewed with sufficient recall, neither study yielded a confirmed instance of accurate identification of the concealed targets; AWARE-II instead focused primarily on documenting electrocortical biomarkers of consciousness emerging during CPR. This underscores the same fundamental challenge illustrated by our own findings: rare, spontaneous perceptual events of this kind are exceptionally difficult to capture and verify under planned, hidden-target conditions. As in the NDE literature, our own spontaneous reports were not tied to a predetermined, deliberately concealed target, and therefore cannot be assigned the same evidentiary weight as our primary, fully blinded target-identification analysis.

While the anecdotal observations reported here were not formally evaluated and cannot be treated as confirmatory evidence, their convergence with historical reports raises methodological questions. Emotionally salient, socially meaningful, or contextually rich elements of the environment may be more likely to be reported than static, decontextualized objects displayed on a laptop screen. In many remote viewing paradigms, targets are information-rich locations or dynamic scenes.[24] Some evidence suggests that target characteristics may influence performance. For example, recent research found that targets later judged as more interesting were more likely to be correctly predicted in precognitive remote viewing tasks,[25] and participants have higher hit rates with "numinous" or emotionally evocative images compared to low-engagement stimuli.[26] Ganzfeld research also suggested differences between static and dynamic targets.[27] While these findings remain debated, they raise the possibility that target salience, complexity, or emotional resonance may modulate performance and may have affected our results.

More broadly, it has been argued that psi phenomena (including ESP), if they occur, may be weak, context-sensitive, and difficult to elicit on demand under tightly controlled laboratory conditions. Some authors have characterized psi as fundamentally spontaneous, suggesting that it cannot be reliably summoned at will within formal experimental settings.[28] Discussions of repeated failed replications have further proposed that the replication problem may be systematic rather than incidental, potentially reflecting a mismatch between the nature of the phenomenon and the constraints of standard laboratory methodology.[29] Reviews of meta-analytic findings have emphasized the importance of ecologically valid designs and the need to examine session-level variables associated with successful outcomes.[30] It has also been suggested that reliance on isolated, tightly controlled experiments evaluated primarily through statistical significance testing may be insufficient to detect subtle anomalous effects, and that diversified methodological strategies may be required.[31] Taken together, these perspectives converge on the view that conventional laboratory paradigms—while indispensable for ensuring rigor—may not always provide optimal conditions for investigating phenomena described as spontaneous or context-dependent. In this context, it is notable that some participants reported discomfort with the EEG equipment, which may have interfered with their ability to induce the intended experiential states.

Therefore, the longstanding difficulty in replicating veridical perception during OBEs may stem less from the absence of the phenomenon and more from a methodological mismatch between laboratory paradigms and the lived dynamics of the OBE state. After reviewing historical case reports and experimental studies, Weiler et al. (2026)[15] identified five determinants that may constrain outcomes in conventional designs: the participant's mental and emotional state, the participant–researcher relationship, researcher intention, variability in participant skill, and the intrinsic purpose of the experience.[15] It proposes that sterile, target-identification tasks using arbitrary stimuli may fail to capture the attentional, motivational, and relational conditions under which OBEs are reported to occur spontaneously. To address this, the authors advocate for greater ecological validity, integration of well-documented spontaneous cases using explicit evidentiary criteria, and experimental designs that preserve falsifiability while better aligning with the phenomenology of the state.

Considered against this framework, some aspects of the present design likely mitigated certain concerns previously raised,[15] while others may still have constrained the likelihood of observing veridical perception. Regarding participant state and setting, sessions were conducted at the participants' own training institute (IIPC headquarters), rather than in an unfamiliar laboratory environment, allowing participants to use their own preferred induction techniques and adjust the room to their liking; however, some participants reported that the concurrent EEG equipment was physically uncomfortable, which may have interfered with induction. Regarding the participant–researcher relationship, care was taken to establish trust and rapport with participants before and throughout the study, which we believe helped support engagement with the protocol despite the necessary constraints of blinding. Regarding participant skill, although our sample consisted of experienced practitioners, induction success and target-related reporting still varied considerably across participants. Finally, regarding the intrinsic purpose of the experience, participants were asked to attend to a fixed, arbitrary, and impersonal stimulus (an object photographed on a white background) rather than a personally meaningful or ecologically salient target, which suggests it may be poorly matched to the motivational structure of spontaneous OBEs. Illustrating this potential mismatch, one of the excluded participants, rather than attending to the designated target, described an OBE in which she traveled to a war zone and perceived herself helping children — an experience that, while not counted toward the experimental outcome measure, exemplifies the kind of personally or emotionally significant content that participants may be spontaneously drawn toward during OBEs, even under explicit task instructions. We do not offer these considerations as post hoc excuses for a null result, but as concrete, addressable design parameters for future studies attempting to bridge rigorous blinding with greater ecological validity.

Several limitations warrant consideration. The sample size was modest, and although participants were experienced practitioners, induction success varied. The target pool consisted of static, object-focused images that may not have been optimally engaging. Moreover, while embedding-based semantic similarity provides an objective and reproducible metric, it remains a proxy for human-level meaning and may not fully capture idiosyncratic symbolic correspondences. Future studies could systematically manipulate target characteristics (e.g., emotional salience, dynamism, personal relevance), incorporate pre-registered analytic plans with larger samples, and examine individual-difference variables associated with performance. We also note that the present study was likely underpowered to detect anything short of a large effect. At $n = 13$, a one-sample test would require an effect size of approximately $d = 0.85$ to achieve 80% power at $\alpha = .05$ (two-tailed); our observed effect size ($d \approx 0.63$, based on the secondary parametric analysis reported above) corresponds to post-hoc power of approximately 55%. A large effect size of this kind would be an unlikely magnitude for a psi-type effect. Null or non-robust results at this sample size are therefore also consistent with a real but modest effect that this study was not adequately powered to detect, and should be interpreted with this limitation in mind.

In conclusion, this study applied contemporary computational semantic methods to a longstanding experimental question: whether descriptions produced during attempted OBEs correspond to concealed targets beyond chance expectation. Under rigorous blinding and quantitative evaluation, we found no evidence of systematic above-chance correspondence. At the same time, the phenomenological and anecdotal patterns observed underscore the complexity of studying extraordinary experiences in laboratory settings. Continued investigation will benefit from integrating methodological stringency with careful attention to experiential context, target properties, and state characterization, thereby advancing a more nuanced empirical understanding of OBEs and related anomalous cognition claims.

Funding

This work was supported by a small philanthropic gift that funded data collection.

CRediT authorship contribution statement

Marina Weiler: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Damon Abraham:** Writing – review & editing, Software, Resources, Methodology, Investigation, Formal analysis, Data curation. **David J.P. Acunzo:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Niffe Hermansson:** Writing – review & editing, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation.

Declaration of competing interest

The authors declare no conflict of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.explore.2026.103506.

References

1. Weiler M, Moreira-Almeida A, Monti MM. Distinguishing out-of-body experiences from lucid dreaming: a phenomenological analysis. *Rev Front Psychol*. 2025;16-2025.
2. Campillo-Ferrer T, Alcaraz-Sánchez A, Demšar E, et al. Out-of-body experiences in relation to lucid dreaming and sleep paralysis: a theoretical review and conceptual model. *Neurosci Biobehav Rev*. 2024;163, 105770.
3. Carruthers G. Confabulation or experience? Implications of out-of-body experiences for theories of consciousness. *Theory Psychol*. 2018;28(1):122–140.
4. Blanke O, Mohr C. Out-of-body experience, heautoscopy, and autoscopic hallucination of neurological origin: implications for neurocognitive mechanisms of corporeal awareness and self-consciousness. *Brain Res Rev*. 2005;50(1):184–199.
5. Weiler M, Acunzo DJ. What out-of-body experiences may tell us about the mind beyond the brain. *Int Rev Psychiatry*. 2024:1–10.
6. Green C. Out-of-the-body experiences. In: *Proceedings. Institute of Psychophysical Research*. 1968.
7. Ring K, Cooper S. Near-death and out-of-body experiences in the blind: a study of apparent eyeless vision. *J Near-Death Stud*. 1997;16(2):101–147.
8. Alvarado CS. ESP during spontaneous out-of-body experiences: a research and methodological note. *J Soc Psych Res*. 1986;53(804):393–397. Society for Psychical Research.
9. Tart C. Six studies of out-of-body experiences. *J Near-Death Stud*. 1998;17(2):73–99.
10. Palmer J, Vassar C. ESP and out-of-the-body experiences: an exploratory study. *J Am Soc Psych Res*. 1974.
11. Palmer J, Lieberman R. ESP and out-of-body experiences: a further study. *Res Parapsychol*. 1975:102–106.
12. Palmer J, Lieberman R. The influence of psychological set on ESP and out-of-body experiences. *J Am Soc Psych Res*. 1975;69(3):193–213.
13. Smith P, Irwin H. Out-of-body experiences, needs and the experimental approach: a laboratory study. *Parapsychol Rev*. 1981.
14. Del Prete G, Tressoldi PE. Anomalous cognition in hypnagogic state with OBE induction: an experimental study. *J Parapsychol*. 2005;69(2):329.
15. Weiler M, Pehlivanova M, Cozzolino P. From laboratory targets to experiential authenticity: reframing the study of veridicality in OBEs. *Front Psychol*. 2026;17, 1720998.
16. Weiler M, Casseb RF, Acunzo D. Using eye movement to signal the onset of self-induced out-of-body experiences. *J Conscious Stud*. 2025;32:135–149. https://doi.org/10.53765/20512201.32.3.135.
17. Brodeur MB, Guérard K, Bouras M. Bank of Standardized Stimuli (BOSS) phase II: 930 new normative photos. *PloS one*. 2014;9(9), e106953.

18. Mossbridge J, Green D, French CC, Pickering A, Abraham D. Future dreams of electric sheep: case study of a possibly precognitive lucid dreamer with AI scoring. *Int J Dream Res*. 2025;18(2):151–168.
19. Ballati A, Prati E, Pederzoli L, Tressoldi P. Remote viewing with and without controlled out-of-body consciousness. *Adv Res Psychol*. 2020. https://doi.org/10.46412/001c.14154.
20. Osis K. Perceptual experiments on out-of-body experiences. *Res parapsychol*. 1974:53–55.
21. Van Lommel P, Van Wees R, Meyers V, Elfferich I. Near-death experience in survivors of cardiac arrest: a prospective study in the Netherlands. *Lancet*. 2001;358(9298):2039–2045.
22. Parnia S, Shirazi TK, Patel J, et al. AWAreness during REsuscitation-II: a multi-center study of consciousness and awareness in cardiac arrest. *Resuscitation*. 2023;191, 109903.
23. Parnia S, Spearpoint K, De Vos G, et al. AWARE—AWAreness during REsuscitation—a prospective study. *Resuscitation*. 2014;85(12):1799–1805.
24. Targ R. Remote-viewing replication: evaluated by concept analysis. *J Parapsychol*. 1994;58:271.
25. Mossbridge J, Cameron K, State BM. Trait, and target parameters associated with accuracy in two online tests of precognitive remote viewing. *J Anomalous Exp Cogn*. 2024;4:88–121. https://doi.org/10.31156/jaex.24743.
26. Krippner S, Saunders DT, Morgan A, Quan A. Remote viewing of concealed target pictures under light and dark conditions. *Explore*. 2019;15(1):27–37.
27. L. Storm, P. Tressoldi Meta-analysis of free-response studies 2009-2018: assessing the noise-reduction model ten years on. 2020;doi:10.31234/osf.io/3d7at.
28. Marks DF. *Take-home message: psi is a spontaneous process that cannot be summoned at will in a laboratory experiment. Psychology and the Paranormal: Exploring Anomalous Experience*. SAGE Publications Ltd; 2020.
29. Walach H, Kirmse KA, Sedlmeier P, Vogt H, Hinterberger T, von Lucadou W. Nailing jelly: the replication problem seems to be unsurmountable—two failed replications of the Matrix experiment. *J Sci Explor*. 2021;35(4):788–828.
30. Cardeña E. The experimental evidence for parapsychological phenomena: a review. *Am Psychol*. 2018;73(5):663–677. https://doi.org/10.1037/amp0000236.
31. Utts J. Replication and meta-analysis in parapsychology. *Stat Sci*. 1991;6:363–378. https://doi.org/10.1214/ss/1177011577.