Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Testing an Anomalous Photograph Against Ordinary Image-Forming Mechanisms: A Null Hypothesis Framework and a Single-Case Demonstration from Battle Mountain, Virginia, USA

Thomas Alan Woolman

twoolman@ontargettek.com

On Target Technologies, Inc. https://orcid.org/0000-0002-4120-2157

Research Article

Keywords: anomalous photography, null hypothesis testing, look-elsewhere effect, extreme value theory, veiling glare, lens flare

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11707102/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Testing an Anomalous Photograph Against Ordinary Image-Forming Mechanisms: A Null Hypothesis Framework and a Single-Case Demonstration from Battle Mountain, Virginia, USA

Thomas A. Woolman

On Target Technologies, Inc., Ridgeley, WV, USA

ORCID: 0000-0002-4120-2157

Abstract

Photographs offered as evidence of anomalous phenomena are usually evaluated qualitatively, with little attempt to state which ordinary explanations are excluded or how improbable the image is under each. This study developed a reproducible null-hypothesis testing framework for a single anomalous photograph and demonstrated it upon an unedited 2015 smartphone image showing two pale, blue-white forms within the shaded tree line at Battle Mountain, Rappahannock County, Virginia, USA, site of an 1863 Confederate infantry assault. The global null hypothesis, that the forms are ordinary lighting artifacts, was decomposed into component tests for sensor noise, lens-flare ghosting and scene lighting, the last employing a scan statistic with a generalized Pareto tail correcting for the look-elsewhere effect after veiling glare had been removed. Sensor noise ($p \leq 0.00001$) and lens flare ($p = 4.2 \times 10^{-6}$) were rejected against an alpha of 0.001, while scene lighting was rejected against patches at the forms' own illumination level ($p < 2.2 \times 10^{-16}$) but not against the full glare-corrected tree line ($p = 0.95$ and 1.0). The author concluded that the outcome depends upon which comparison set is physically appropriate, and that although the analysis excludes the tested mechanisms, it cannot establish the cause.

Keywords: anomalous photography, null hypothesis testing, look-elsewhere effect, extreme value theory, veiling glare, lens flare

INTRODUCTION

Photographs are among the most frequently offered forms of evidence for anomalous phenomena, and yet they are among the least rigorously examined, as a typical analysis lists a series of possible ordinary causes such as dust, lens flare, reflections, long exposures, and pareidolia and then dismisses or accepts each of them by visual inspection alone. In general, three weaknesses recur throughout this type of analysis, each of which limits the ability of the analyst to make a falsifiable statement about the photograph in question.

First, the null hypothesis is rarely stated, so that an analyst who concludes that a feature "cannot be explained" has usually not specified which explanations were tested, what each of those explanations would predict about the image, or how far the image departs from those predictions. Second, anomalous features are found by looking, which means that the feature that draws attention is by construction the most unusual feature in the frame, and comparing it with a typical patch of the same image therefore overstates its significance; this problem is known in physics as the look-elsewhere effect (Gross & Vitells, 2010), and in spatial statistics it provides the motivation for the scan statistic (Kulldorff, 1997).

Third, images taken toward a bright light source carry veiling glare, which is a low-frequency haze produced by light that has scattered inside the lens (McCann & Rizzi, 2007). Because glare lifts the darkest shadows unevenly across the frame, brightness measured in one part of such an image is not directly comparable with brightness measured in another part, and any analysis that compares features across the frame without accounting for glare is likely to reach misleading conclusions.

This study addresses all three of these weaknesses, in that the author stated a global null hypothesis, decomposed it into component tests that each correspond to a distinct physical

mechanism, compared the anomaly with the maximum of the ordinary features in the image rather than with a typical one, and removed veiling glare before any brightness comparisons were made. The resulting framework was demonstrated on a single photograph for which the original camera file, with its full metadata record, is available.

The contribution of this study is therefore methodological in nature, and the demonstration is intended to show both what such a framework can establish and where its reach ends, since two of the ordinary mechanisms were rejected decisively while the outcome for the third depended upon how the comparison set was defined, and that dependence is reported here in full rather than being resolved by the selection of the most favorable setting. To the knowledge of the author, and on the basis of the review of prior work that follows, this study is the first to subject an individual anomalous photograph of the kind examined in psychical research to formal significance testing against explicit, physically derived component null hypotheses for sensor noise, lens-flare geometry, and scene illumination, with a correction for the look-elsewhere effect and a Bonferroni-controlled alpha test statistic.

Prior Work on Anomalous Photographs

The literature on anomalous photographs is extensive, but it is overwhelmingly descriptive or qualitative in nature. Historical surveys of photography in psychical research, from the spirit photographs of the nineteenth century to the contested cases of the twentieth, treat each image through provenance, context, and expert opinion rather than through quantified inference (Willin, n.d.), and the methods chapter of the principal German handbook of scientific anomalistics sets out a four-step protocol of photo analysis, re-enactment, context, and digital forensics that deliberately ends in judgments of plausibility rather than in statistical tests (Mayer, 2015). The closest single-case precedent in this literature, an analysis of mysterious objects

recorded by a wildlife camera, combined image enhancement, depth-of-field and exposure reasoning, and observer perception, but it did not apply statistical testing to the image (Mayer & Kornmeier, 2014).

Where statistics have been applied to anomalous photographs, they have been applied to collections of images rather than to an individual image. Lange and Houran (1997) cross-tabulated published anomalous photographs by type of anomaly and recording medium and concluded that the medium largely determines the effect, while Terhune et al. (2007) compared the incidence of photographic anomalies at an alleged haunt and a control site across 338 prints and several film and sensor types, using analyses of variance, nonparametric tests, and a Bonferroni-corrected alpha of 0.005. Schwartz and Creath (2005) explained so-called orb photographs through the optics of near-lens particles illuminated by the flash, which anticipates the lens-flare component of the present study, although their analysis did not take the form of hypothesis tests, and the thoughtography experiments of the 1960s were likewise reported as counts of outcomes across series of trials rather than as tests of individual images (Stevenson & Pratt, 1968).

The nearest quantitative analyses of single images lie outside of psychical research, in the study of claimed anomalies in planetary, aerial, and astronomical imagery. Carlotto (1997) assessed features in Viking images of Mars through fractal model fits combined into Bayesian odds under an assumed prior, rather than through a null distribution and p-value, and a technical analysis of the 1950 McMinnville photographs reported a detection statistic of approximately 2.4 to 2.5 standard deviations for a suspension thread after the tilt and position of the search had been tuned interactively, without a stated null hypothesis or a correction for the number of positions examined (Cousyn et al., 2013). The closest formal analogue is a recent analysis of

transient sources on mid-century astronomical survey plates, which tested each plate against a Monte Carlo null of random source placement and applied a Bonferroni correction across the plates examined (Doherty, 2026); that analysis, however, concerns astronomical plates under a single random-placement null and does not address the optical and illumination mechanisms that dominate terrestrial photographs.

The statistical machinery of the present study also has precedent in computer vision, where a contrario detection methods control the expected number of false detections across all of the positions and scales examined within a single image (Desolneux et al., 2000), and where the detectability of spots against textured backgrounds has been formalized in the same manner (Grosjean & Moisan, 2009), which is closely analogous to the scene lighting test used here. What has not previously been done, to the knowledge of the author, is to combine these elements into a test of an individual anomalous photograph, with separate null hypotheses for each ordinary image-forming mechanism, a look-elsewhere correction for the illumination test, and a stated alpha test statistic, and it is that combination that constitutes the contribution of this study.

Site and Historical Context

Battle Mountain

Battle Mountain is an isolated summit near Amissville, Rappahannock County, Virginia, USA, within the eastern Blue Ridge province, and it is composed of the Battle Mountain Alkali Feldspar Granite and associated felsite of the Late Neoproterozoic Robertson River Igneous Suite, for which U-Pb zircon ages of 705 ± 2 Ma for the granite and approximately 702 Ma for the felsite have been reported (U.S. Geological Survey, n.d.). Earlier gallium geochemistry by the author distinguished a plutonic western slope from a rhyolitic eastern slope and summit, and

on that basis interpreted the edifice as an extinct volcanic center with a lateral eruption directed toward the east (Woolman, 2016).

The author owned the property on the mountain from which the photograph was taken and therefore knew the ground firsthand, although the property was sold in late 2021 and is no longer accessible for the purposes of further study.

The Engagement of July 24, 1863

The largest single military engagement in Rappahannock County took place at and around Battle Mountain on the morning of July 24, 1863, during the Confederate withdrawal from Gettysburg (Virginia Civil War Trails, n.d.). Brigadier General George A. Custer, who was then commanding the Third Division of the Union Cavalry Corps in addition to his own Michigan brigade, advanced from Amissville toward Newby's Crossroads with a force of five regiments and two batteries (U.S. War Department, 1889a, pp. 997, 1002).

At the crossroads, Custer struck the Confederate infantry that was covering the march toward Culpeper. His report identified the enemy as Longstreet's corps, whereas the Civil War Trails marker describes the rear of Longstreet's corps together with the head of A. P. Hill's corps (U.S. War Department, 1889a, p. 1002; Virginia Civil War Trails, n.d.). The principal Confederate force engaged was the brigade of Brigadier General Henry L. Benning of Hood's division, which was then under the command of Law and was reinforced by the 4th and 15th Alabama regiments, and Benning had been ordered to hold the road until he could be relieved by Hill's corps, which was close behind (U.S. War Department, 1889b, pp. 418 to 419). That march followed the road named in the contemporary reports as the Richmond road, which survives in the modern secondary system as State Route 729, carrying the name Richmond Road through Rappahannock County and the name Eggbornsville Road where it continues south across the

county line into Culpeper County, so that the line of the Confederate column below the mountain can still be followed on a modern map, and the town of Culpeper lies approximately thirteen miles to the south-southeast of the summit, or approximately sixteen miles by that road.

Custer found the Confederates strongly posted on a ridge and opened fire with his artillery from the shoulder of Battle Mountain, after which he withdrew part of his force and left the 5th and 6th Michigan regiments, together with two guns of Battery M of the 2nd U.S. Artillery, as a rear guard. Benning, who judged that the Union guns occupied "a mountain position difficult to be carried by a front attack," sent a column on a flanking march of approximately four hours and four miles, while Colonel William C. Oates's 15th Alabama scouted along the mountainside (U.S. War Department, 1889b, pp. 418 to 419; Virginia Civil War Trails, n.d.).

The advanced Union force was nearly cut off, and Custer described it as "the most critical [position] I was ever in," with his men escaping only by cutting a road through the dense woods (U.S. War Department, 1889a, p. 1003). Lieutenant Carle A. Woodruff and Captain Smith H. Hastings later received the Medal of Honor for refusing orders to abandon the guns during this action.

Losses were light by the standards of the campaign, as Custer reported a loss not exceeding fifteen men, while Benning reported only that his men took a few prisoners and killed and wounded more (U.S. War Department, 1889a, p. 1003; 1889b, p. 419). No burial record at or near the site was located during the course of this study. Contrary to the common assumption that the mountain was named for this engagement, local history attributes the name to the Bataille family, French Huguenots who settled near Amissville and changed their name to Battle upon arriving in America (Hurley & Stanton, 2022), which indicates that the name predates the

events of 1863. The earliest primary attestation of the name located during the course of this study is Custer's own report, which referred to Battery M as being "engaged with the enemy at Battle Mountain, near Newby's Cross-Roads" (U.S. War Department, 1889a, p. 1001), so that the name was demonstrably in use by September of 1863, although no earlier deed, map, or survey bearing the name was located.

Relevance of the History to This Study

The photograph was taken at the base of the eastern slope of the mountain, and from his local knowledge of the ground the author understands this position to lie near the edge of the Confederate right during the engagement. The published sources reviewed for this study place Benning's flanking approach to the east and southeast of the crossroads, but none of them locates a Confederate line on the eastern slope specifically, and for that reason the placement is reported here as the author's understanding rather than as a documented fact.

The historical record is relevant to this study in two ways, the first being that it motivates a provenance hypothesis, which holds that the forms in the photograph, if they are not explained by ordinary mechanisms, are associated with the events of 1863, and the second being that it bears on interpretation, because knowledge of the history of a site has been shown to shape what observers report there (Wiseman et al., 2003).

The analysis was designed so that the second of these effects could not influence the statistical results, since the tests operate on pixel values and do not depend upon what the forms are taken to represent. The provenance hypothesis, however, lies outside of what any analysis of a single photograph can test, and the Discussion returns to what the results do and do not imply for that hypothesis.

The Photograph

Provenance and Capture

The file analyzed in this study is 20150718_184215.jpg, which was recovered as an attachment to an email that the photographer sent on February 14, 2016. It is the full-resolution camera output with its EXIF record intact, and it contains no editing-software signature, an embedded camera thumbnail, and a file name that matches the recorded capture time. A later copy of the image carries an added content-credential metadata block, and that copy was not used in the analysis. The capture record and the computed position of the sun are shown in Table 1.

Table 1

Capture Record (EXIF) and Computed Sun Position

Property	Value
Camera	Samsung Galaxy S6 (SM-G920V); firmware G920VVRU2AOF1
Captured	July 18, 2015, 6:42:14 PM EDT
Image size	5,312 × 2,988 pixels
Exposure	f/1.9, 1/1200 s, ISO 40; no flash; automatic exposure
Lens	4.3 mm (28 mm equivalent); horizontal field of view approximately 68°
Sun	Elevation 19.8°, azimuth 281° (west-northwest); sunset at 8:35 PM
Subject distance	Tree line at the base of the eastern slope, approximately 120 to 150 m

The camera faced the tree line with the sun in the upper left of the frame, partly occluded by the canopy at the crest of the ridge, and according to the photographer's observation the sun had dropped behind the shoulder of the mountain relative to the slope, leaving the tree line in diffuse light and shadow. The exposure of 1/1200 s at ISO 40 is a bright-daylight setting that the camera selected because the sun was within the frame, and at that setting the shaded tree line was recorded as nearly black, with a median brightness of approximately 14 on the 0 to 255 scale in the area around the forms.

The photographer took the frame on an unexplained impulse, saw nothing unusual at the time of capture, and noticed the forms only upon later review of the image, as shown in Figure 1.

Figure 1

Full Frame of the Photograph, Downscaled



Note. The red box marks the two forms. The yellow circle marks the position at which a lens-flare ghost of the sun is predicted to fall (see Lens-Flare Ghost Test).

The Forms

Two pale forms sit approximately 126 pixels apart in the shaded tree line at the lower right of the frame (Figure 2), and at the estimated distance one pixel spans approximately 3 to 4 cm. The dimensions and brightness of the forms are summarized in Table 2, and no pixel in either form is clipped.

Figure 2

The Two Forms at Eight Times Native Resolution



Note. The crop is enlarged eight times by nearest-neighbor replication, so that no interpolated detail is introduced, and the image was brightened by a uniform linear factor of 2.1 across all three channels for visibility, which preserves the relative luminance on which the tests are based. The scale bar indicates the length subtended by 1 m at an assumed subject distance of 135 m, which is an informed estimate rather than a measurement; the estimated range is 120 to 150 m.

Table 2

Dimensions and Brightness of the Forms

Feature	Core size (px)	Size at 120 to 150 m	Peak / local shade
Right form	40 wide × 43 tall	1.2 to 1.5 m wide, 1.3 to 1.6 m tall	118.7 / 13.7 (8.7×)
Left form	21 wide × 23 tall	0.6 to 0.8 m wide, 0.7 to 0.9 m tall	96.7 / 14.0 (6.9×)
Separation of centers	126	3.8 to 4.8 m	

Within the larger, right-hand pixel cluster, the photographer perceives two potentially human outlines, in the manner of the well-documented perception of faces and figures within ambiguous visual stimuli (Liu et al., 2014). On the left side of that cluster, an upright form is read by the photographer as a figure kneeling behind cover while holding a weapon, and to its right, a horizontal form that rises at its far end is read as a second figure lying in a prone position with the torso elevated against cover. The photographer further reads the smaller diagonal streak

to the left, approximately 4 to 5 m away at the estimated distance, as a possible third individual who is partially obscured by foliage.

The overall dimensions of the right-hand cluster, at approximately 1.2 to 1.5 m wide and 1.3 to 1.6 m tall, are consistent with that reading, as is the spacing between the two clusters. However, the image does not contain enough detail to resolve anatomy, and the perception of human figures within ambiguous shapes is precisely what face and figure pareidolia would predict, so the interpretation is reported here as that of the photographer and it plays no part in the statistical tests of this study, although the Recommendations for Future Research propose a formal test of it.

The forms are blue-white in color, with the blue share of their brightest pixels, expressed as B/(R + G + B), measuring 0.43 to 0.44, whereas the backlit foliage and the lawn in the same frame fall within a range of 0.33 to 0.37. The color of the forms therefore matches that of surfaces lit by open skylight rather than by the sun, which is consistent with the photographer's account that no direct sunlight reached the slope at the time of capture. Because color alone cannot distinguish reflected light from emitted light, no claim of self-luminosity is made in this study.

Explanations Excluded by Site Knowledge

Several ordinary explanations cannot be tested from pixel values and instead rest upon the photographer's knowledge of the site, and they are listed separately in Table 3 so that the reader can weigh them on that basis.

Table 3

Candidate Explanations Excluded by Site Knowledge Rather Than by Statistical Test

Candidate	Basis for exclusion
Water reflection	Dry July conditions on sloping, well-drained ground with no standing water.

Candidate	Basis for exclusion
Artificial light	No artificial light source was present at the site.
Deer	Deer came onto the open lawn at that hour rather than the tree edge, and neither the color nor the shape of the forms matches.
Pale bark or dead snags	The trees were oaks and tulip trees with no pale-barked species, and no pale trunks or snags stood at that spot; the forms are also brighter than any trunk in the frame (approximately 71 to 75) despite sitting in deeper shade.
Pale object (fabric, bag, debris)	The site was walked during the following week, and only leaf litter was present.

METHOD

Research Questions and Hypotheses

The global null hypothesis for this study states that the anomaly group of pixels is consistent with normally expected lighting artifacts within this photograph, including sensor noise, lens-flare ghosting, and the bright patches that ordinary light produces within a shaded tree line, such as sky gaps, sunflecks, backlit foliage, and bark. The alternative hypothesis states that the anomaly group is not explained by those mechanisms. It should be emphasized that rejecting the global null hypothesis supports only the statement that the forms are not explained by the tested mechanisms, and it does not identify what the forms are.

The global null hypothesis was decomposed into three research questions, each of which corresponds to one physical mechanism, and the global null hypothesis is rejected only if all three of the component null hypotheses are rejected. Consistent with the author's preferred practice for quantitative research into physical phenomena, where computational resources permit the necessary number of simulations, the alpha test statistic for the global null hypothesis was set at 0.001 rather than the conventional 0.05, and each component was tested against an alpha of 0.001 / 3 = 0.00033, applying a Bonferroni correction for the three tests. Following the recommendation of the American Psychological Association that effect sizes accompany

significance tests, each component test is reported with the effect size natural to its statistic, namely a standardized difference for the sensor-noise and flare-position tests and a ratio of the observed statistic to the strongest null patch, together with the percentile of the observed value within its comparison set, for the scene-lighting test. The substantive research questions for this study, together with their corresponding null and alternative hypothesis statements, are therefore as follows:

RQ1: Can the anomaly group of pixels be explained by random sensor noise at the local shade level?

$H1_o$: The integrated excess luminance of the right form is consistent with a patch of the same area drawn from sensor noise at the local shade level.

$H1_a$: The integrated excess luminance of the right form exceeds what sensor noise at the local shade level would produce.

RQ2: Can the anomaly group of pixels be explained as a lens-flare ghost of the in-frame sun?

$H2_o$: The position of the right form is consistent with the position predicted for a lens-flare ghost of the sun.

$H2_a$: The right form lies farther from the predicted ghost position than lens-flare scatter would allow.

RQ3: Can the anomaly group of pixels be explained by the ordinary lighting of the scene?

$H3_o$: The right form is no stronger than the strongest of the ordinary bright patches within the comparison set.

$H3_a$: The right form is stronger than ordinary scene lighting would produce, after correction for the look-elsewhere effect.

Luminance and Veiling Glare Correction

Luminance was taken as the mean of the red, green, and blue channels on the 0 to 255 scale. Because the sun is within the frame, veiling glare lifts the darkest shadows from approximately 5 at points far from the sun to approximately 65 at points near it, while the local contrast of the scene above that floor remains roughly constant, and since glare is additive in nature it was estimated and then subtracted from the image.

To accomplish this, the frame was divided into blocks of 64 pixels, and within each block the 2nd percentile of luminance was taken as the glare floor, after which the block values were smoothed with a 5 × 5 running median and interpolated bilinearly to pixel resolution. The corrected luminance was then computed as Lc = max(L − F, 0), where F is the smoothed glare floor, and the local scene level S was computed as the block median of Lc, smoothed and interpolated in the same manner, with a minimum value of 3.

Regions and Comparison Sets

The tree line was defined as the region between the skyline, which was taken as the first block in each column darker than a luminance of 120, and the edge of the lawn, which was traced on the photograph. Two foreground objects were excluded from the analysis, namely the sapling in mid-frame and the canopy of a foreground tree at the far left, through whose leaves open sky is visible. A box of 293 × 156 pixels around the forms defined the anomaly region, and patches inside that box were not used as null observations. Three comparison sets, or strata, were then defined for use in the scene lighting test as follows:

1. The lower tree line, consisting of the tree line at least 400 pixels below the local skyline after glare correction, which represents 17.2% of the frame.
2. The whole tree line, consisting of the full tree line including the upper canopy where sky gaps occur, after glare correction, which represents 28.6% of the frame.
3. The deep-shade band, consisting of those tree-line blocks whose raw median luminance lies between 8 and 25, which is the band that contains the forms, using raw luminance without glare correction, and which represents 3.8% of the frame.

Patch Detection and Test Statistic

A pixel was treated as a candidate if $Lc > 2S$ and $Lc > S + 8$, and candidate pixels were grouped into 8-connected patches, with patches smaller than 4 pixels being discarded. For each patch, the local scene level S was recomputed as the median of Lc within a ring from 8 to 30 pixels outside the bounding box of the patch, excluding any other patches. The test statistic was the normalized integrated excess luminance (NIEL), computed as follows:

$$NIEL = \Sigma \max(Lc / S - 1, 0), \text{ summed over the pixels of the patch}$$

The NIEL statistic represents the total excess brightness of a patch in units of its own local scene level, and it is therefore invariant to the level of illumination, since reflected light scales with the light falling upon the scene, while it increases with both the brightness and the size of the patch. Within the deep-shade stratum the same statistic was computed on raw luminance, and the right form, which is the strongest patch within the anomaly region, was used as the test case for all three research questions.

Sensor Noise Test

For the first research question, the noise standard deviation was estimated robustly from the pixel-to-pixel differences in the dark shade beside the forms, using the median absolute

deviation of the first differences divided by √2. A patch equal in area to the right form was then drawn 100,000 times from pure Gaussian noise, so that the smallest reportable p-value (0.00001) lies well below the corrected alpha, and its integrated excess luminance was recorded for each draw, with the p-value computed as the share of draws that reached the observed value, applying the conventional correction of adding one to both the numerator and the denominator.

Lens-Flare Ghost Test

For the second research question, the test relied upon the fact that internal reflections form ghost images of a bright source on the line that runs from the source through the optical center of the lens, on the opposite side of the frame. The sun disk was located automatically as the compact, fully saturated spot with the darkest surroundings, using the saturation fill within a 41 × 41 pixel core minus the fill within the surrounding 161 × 161 pixel box, and the predicted position of the ghost was computed as g = 2c − s, where c is the center of the frame and s is the position of the sun.

Because ghosts from different pairs of lens surfaces fall at slightly different points, and because the optical center of a phone lens need not coincide exactly with the center of the frame, the ghost position was modeled generously as isotropic Gaussian scatter about g with a standard deviation equal to 5% of the frame diagonal, or 305 pixels. The p-value was then computed as the Rayleigh tail probability exp(−d² / 2σ²) for the observed distance d between the right form and the predicted ghost.

Scene Lighting Test

For the third research question, because the forms were found by looking, they were compared with the maximum of the N ordinary patches within the stratum rather than with a typical patch. The upper tail of the null NIEL distribution was modeled with a generalized Pareto

distribution (GPD) fitted by maximum likelihood above the 90th percentile u of the null patches (Coles, 2001; Pickands, 1975), so that the probability that a single patch reaches the observed value x is given by the following expression:

$$p_1 = \zeta_u \, [1 + \xi \, (x - u) / \sigma]^{(-1/\xi)}$$

In this expression, ζ_u is the share of null patches above u, while σ and ξ are the fitted scale and shape parameters respectively. The look-elsewhere-corrected p-value is then the probability that the largest of the N patches reaches the observed value x:

$$p = 1 - (1 - p_1)^N$$

Uncertainty in the tail fit was assessed through a nonparametric bootstrap of the tail exceedances using 500 resamples (Efron & Tibshirani, 1993), and a component was classified as robust if it was rejected even at the upper end of the 95% range of the bootstrap. Where the fitted shape parameter is negative the tail has a finite endpoint, and an observed value beyond that endpoint produces a p-value below machine precision, which is reported here as $p < 2.2 \times 10^{-16}$.

Sensitivity Analysis

To assess the stability of the findings, the scene lighting test was repeated for each stratum across nine settings, consisting of detection thresholds of 1.75, 2.0, and 2.5 times the scene level, crossed with GPD thresholds at the 85th, 90th, and 95th percentiles, with the patches being re-detected for each detection threshold.

Software and Reproducibility

All analyses were conducted in R version 4.6.1 (R Core Team, 2026) under Windows 11, using the jpeg 0.1-11 (Urbanek, 2026), evd 2.3-7.1 (Stephenson, 2002), and spatstat.geom 3.8-3 (Baddeley et al., 2015) packages, with a fixed random seed of 20150718 for the purpose of reproducibility. The analysis script and the original image file have been deposited in a public

archive, which is identified in the Data and Code Availability statement below, so that the analysis reported here can be reproduced in full rather than merely described.

A-Priori Belief Disclosure

In accordance with the disclosure policy of this journal, the author records his prior expectation of the outcome. The hypotheses tested here concern ordinary image-forming mechanisms rather than psi, so that the rejection of any one of them establishes only that the mechanism in question does not account for the feature, which is a conclusion that neither supports nor requires a psi hypothesis. On the five-point scale of the journal concerning the belief that a psi hypothesis would be supported, upon which 1 indicates strong disbelief and 5 indicates strong belief, the prior position of the author was 3, which is to say agnostic. The author took the photograph and therefore held an interest in the outcome, and it is for that reason that the alpha was specified in advance at the stringent level of 0.001, that the analysis was decomposed into component tests each corresponding to a distinct physical mechanism, and that all three comparison strata are reported here rather than only the stratum within which the null hypothesis is rejected.

RESULTS

The results of the three component tests are summarized in Table 4, and each research question is then addressed in turn.

Table 4

Summary of Component Tests

Test	Component null hypothesis	p	H_0 rejected at $\alpha / 3 = 0.001 / 3 = 0.00033$
RQ1 Sensor noise	The forms are random noise at the local shade level	≤ 0.00001	Yes
RQ2 Lens flare	The forms are a flare ghost of the sun	4.2×10^{-6}	Yes

Test	Component null hypothesis	p	H_0 rejected at $\alpha / 3 = 0.001 / 3 = 0.00033$
RQ3 Scene lighting, lower tree line	The forms are no stronger than the strongest ordinary patch	0.95	No
RQ3 Scene lighting, whole tree line	Same as above	1.0	No
RQ3 Scene lighting, deep-shade band	Same, among patches at the forms' illumination level	$< 2.2 \times 10^{-16}$	Yes

Note. Each research question tests one component of the global null hypothesis, so the alpha test statistic of 0.001 is divided by the three tests (Bonferroni correction), giving $\alpha / 3 = 0.00033$ for each component. The global null hypothesis is rejected only if all three components are rejected at that level.

Research Question 1 Results

The estimated noise standard deviation was 2.45 on the 0 to 255 scale, and the integrated excess luminance of the right form was 50,843, against a pure-noise mean of 1,457 with a standard deviation of 55, which represents a standardized effect of $z = 893$, the observed value being 34.9 times the pure-noise mean. None of the 100,000 draws approached the observed value, so that $p \leq 0.00001$, which is the floor set by the number of draws and lies well below the corrected alpha of 0.00033. Thus, the null hypothesis statement that "the integrated excess luminance of the right form is consistent with a patch of the same area drawn from sensor noise at the local shade level" is rejected.

Research Question 2 Results

The sun disk was located at pixel (1827, 1065), with a core saturation fill of 0.96 against a fill of 0.11 in its surroundings, and the predicted ghost position falls at pixel (3486, 1924), which lies inside the foreground sapling (Figure 1). The right form lies 1,516 pixels from that point, which represents 24.9% of the frame diagonal and a standardized radial deviation of 4.97 relative to the allowed scatter, yielding a p-value of 4.2×10^{-6}. Thus, the null hypothesis statement that "the position of the right form is consistent with the position predicted for a lens-flare ghost of the sun" is rejected.

Research Question 3 Results

Within the deep-shade band, the stratum contained 98 null patches (N = 122 including the anomaly region), and the NIEL of the right form, at 3,450, was 23.7 times that of the strongest null patch, at 145.6, and exceeded every null patch in the stratum, placing it at the 100th percentile of that comparison set. The fitted tail had a shape parameter near zero ($\xi = -0.04$), and the corrected p-value fell below machine precision, so that the null hypothesis was rejected against the corrected alpha of 0.00033. The upper bound of the 95% bootstrap interval for the tail fit was 0.0069, however, which lies above that alpha, so the rejection is not robust to the uncertainty in the tail estimate. Furthermore, the null hypothesis was rejected in six of the nine sensitivity settings, with p-values ranging from 0.000095 down to below machine precision, while the three exceptions returned p-values between 0.00091 and 0.0034, which lie above the corrected alpha although still below the conventional 0.01, and two of those three occur at the loosest detection threshold, where the null set expands to 281 patches and the tail is accordingly estimated from a larger but shallower set of exceedances.

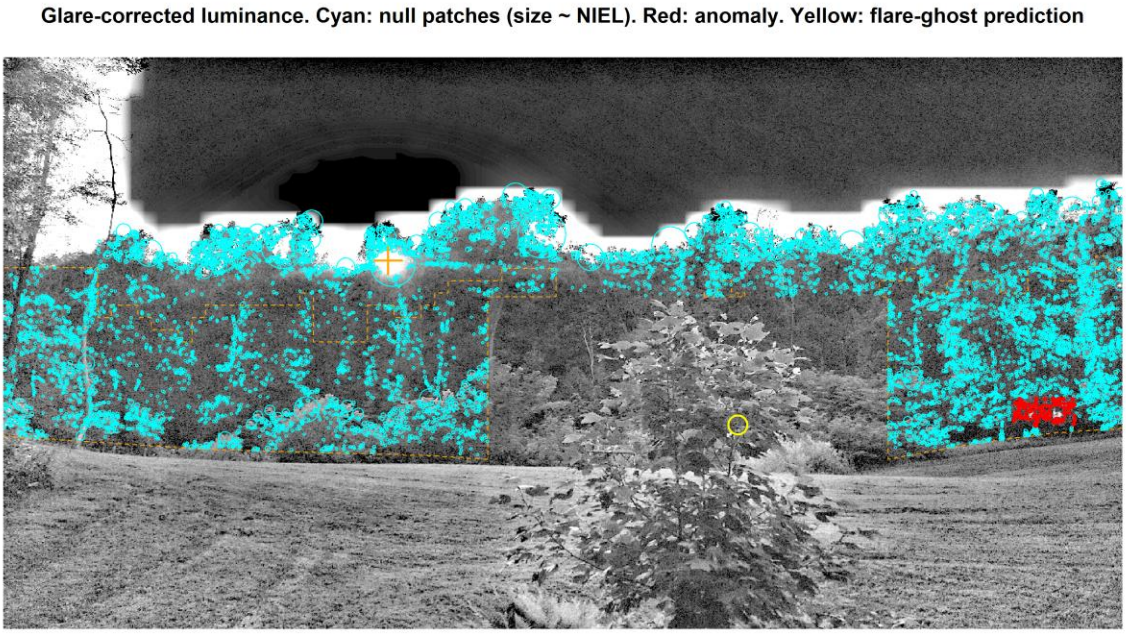
Within the lower tree line, the glare-corrected stratum contained 4,881 null patches (N = 4,977), and although the NIEL of the right form was 9,723, two patches exceeded it, namely backlit shrub clusters nearer to the sun with values of 21,598 and 14,086, so that the form reached 0.45 times the strongest null patch and sat at the 99.96th percentile of the 4,881 null patches. The corrected p-value was 0.95, with a 95% bootstrap interval of 0.77 to 0.99, and the null hypothesis was not rejected in any of the nine sensitivity settings, where p-values ranged from 0.53 to 0.98.

Within the whole tree line, the inclusion of the upper canopy added saturated sky gaps along the skyline, the largest of which had a NIEL of 84,404, so that the form reached only 0.12

times the strongest null patch and the corrected p-value was 1.0 and the null hypothesis was not rejected in any setting. The spatial distribution of the patches is shown in Figure 3, and the distribution of the NIEL statistic is shown in Figure 4.

Figure 3

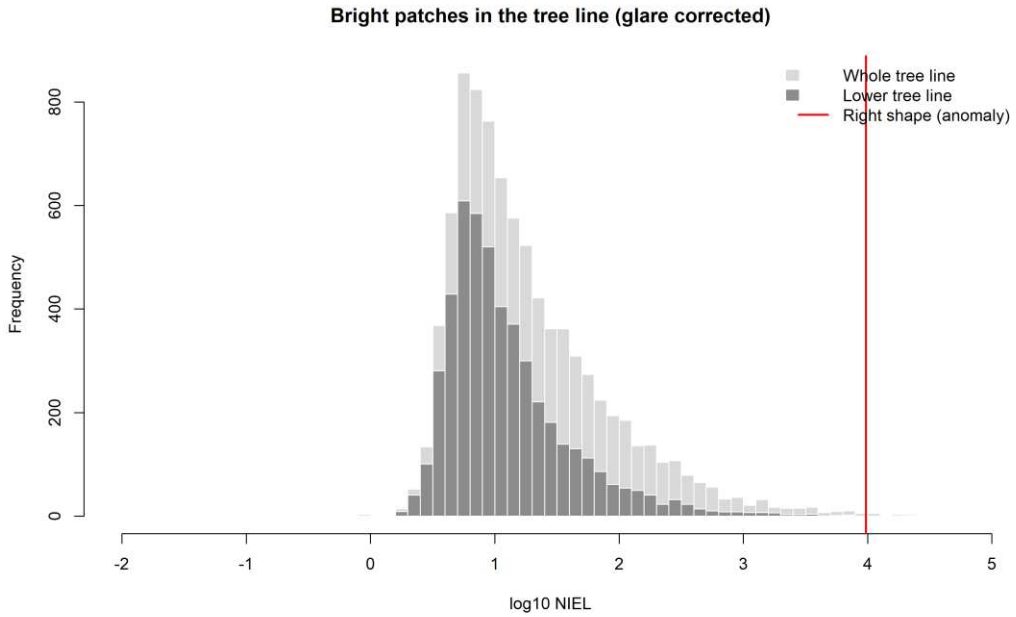
Glare-Corrected Luminance With Detected Patches



Note. Cyan circles mark null patches, sized by NIEL, and red circles mark patches within the anomaly region. The orange cross marks the sun disk, the yellow circle marks the predicted flare ghost, and the dashed orange line marks the boundary of the lower tree line. The foreground sapling and the canopy at the far left are excluded.

Figure 4

Distribution of NIEL for Null Patches in the Whole and Lower Tree Line



Note. The horizontal axis is on a log scale, and the red line marks the right form.

The patches that outscored the forms in the lower and whole tree line strata differ from them physically in several respects, since they are four to eleven times larger in area (6,000 to 16,000 pixels against 1,493), they sit where the glare floor is five to nine times higher (approximately 44 to 78 against 8), and they are yellow-green or sky-colored, with a blue share of 0.33 to 0.37, rather than blue-white at 0.43.

DISCUSSION

Discussion of Research Questions 1 and 2 Findings

Two of the three component null hypotheses were rejected decisively and without any dependence upon the choices made by the analyst. The forms are not sensor noise, and they are not a flare ghost of the in-frame sun, which would have fallen within the foreground sapling at a distance of a quarter of the frame diagonal from the forms.

Discussion of Research Question 3 Findings

The third component, ordinary scene lighting, is where the result depends upon a choice made by the analyst. Among patches at the forms' own illumination level, the forms are extreme by every measure used in this study, and the rejection survives six of the nine sensitivity settings at the stringent alpha of 0.001, although not the upper end of the uncertainty in the tail fit, whereas against the whole tree line, once glare has been removed, the forms are not the strongest bright feature in the image.

The author argues that the deep-shade comparison is the physically appropriate one, for the following three reasons:

1. The illumination mechanism differs, since the patches that outscore the forms are lit by the sun, either directly through backlit foliage near the azimuth of the sun or as open sky seen through the canopy, whereas the forms sit in shade and are lit by skylight alone, as their blue-white color indicates, and a comparison set should share the illumination that the null hypothesis attributes to the anomaly.
2. Glare correction is imperfect where glare dominates, since near the sun the corrected scene level is small relative to the glare floor that was removed, so that the ratios there are inflated by any error in the estimate of the floor, and the deep-shade band avoids this problem by comparing raw luminance where the glare floor is uniformly low.
3. The competing patches are a different kind of object, since they are broad masses of lit foliage and sky rather than compact bright forms, and because the NIEL statistic rewards area, a large and moderately bright patch can outscore a small and very

bright one, while the per-pixel contrast of the forms, at a mean of 7.5 times the local scene level after glare correction, is among the highest in the frame.

The counterargument, however, deserves equal weight, since all three strata were defined after the image had been examined, and the deep-shade band contains only 98 null patches, so that its tail was estimated from only ten exceedances, and a skeptical reader could therefore fairly treat the lower and whole tree line strata as the pre-specified tests and the deep-shade band as a post hoc refinement. On that reading, the global null hypothesis is not rejected, and for this reason the author has reported all three results so that the reader can make that judgment with the numbers in view.

Implications for the Provenance Hypothesis

If the deep-shade comparison is accepted, then the forms are not explained by sensor noise, lens flare, or the ordinary lighting of comparable shade. This result is necessary for any non-ordinary interpretation, but it is far from sufficient, since it leaves open other ordinary causes that pixel statistics cannot address, including those listed in Table 3, whose exclusion rests upon the photographer's knowledge of the site.

The photograph was taken on ground where Union and Confederate forces fought on July 24, 1863, and the measured sizes and spacing of the forms are consistent with the photographer's reading of a kneeling figure and a prone figure within the right-hand cluster, together with a possible third individual partially obscured by foliage in the smaller form to the left. These facts make the provenance hypothesis a natural one to state, and they are the reason that the history of the engagement has been presented in this study.

These facts are not, however, evidence for the provenance hypothesis, because nothing in the image connects the forms to a particular time or cause, and the reading of the forms as human

figures is interpretive in nature. The appropriate conclusion is that the photograph is consistent with the provenance hypothesis in the weak sense that it does not contradict it, and that testing that hypothesis would require evidence of a different kind than a single image is capable of supplying.

Limitations

The primary limitation of this study is that it relies upon a single frame, which makes it an ex-post facto analysis in which persistence, motion, and change over time cannot be assessed. In addition, the strata, thresholds, and test statistic were chosen after the image had been examined, and while the sensitivity analysis shows that the conclusions are largely stable within each stratum, with only three of the twenty-seven settings changing the outcome at the stringent alpha of 0.001, it cannot remove this limitation.

Likewise, the null set within the deep-shade band is small, so that its tail was estimated from ten exceedances, and the upper bound of its bootstrap range (0.0069) lies above the corrected alpha of 0.00033, so that the scene lighting rejection, while strong on its point estimate, is not robust to the uncertainty in the tail estimate. The site is also no longer accessible, since the property was sold in 2021, and as a result the scene cannot be re-photographed under the same conditions or searched again. The author further notes that the sensitivity settings nearest the corrected alpha are numerically unstable, in that re-running the identical analysis code on the identical camera file under a later build of the image-decoding library shifted the decoded luminance values by a few parts in ten thousand, and that shift moved the p-values of three of the nine deep-shade settings by as much as four orders of magnitude, from 0.00000034 to 0.0034 in one instance, so that the number of settings in which the null hypothesis is rejected fell from eight to six. The six settings that do reject fell below machine precision in both runs, and the

lower and whole tree line strata returned the same corrected p-values of 0.95 and 1.0 and the same absence of rejection at every setting, their patch counts and strongest null values shifting by one percent or less, so that the instability is confined to the settings whose tails are estimated from the smallest sets of exceedances. This is reported because it bears directly upon how much weight the sensitivity count can carry, and the figures given throughout this paper are those of the later run.

Furthermore, the sharpening and tone mapping performed by the camera's processing can make fragments appear more angular than they are, and although such processing cannot create brightness where none was recorded, it does affect the shape of the patches. The sizes of the forms also depend upon the estimated distance of 120 to 150 m, so that a 10% error in distance produces a 10% error in size, and finally, the position of the Confederate right relative to the photograph rests upon the author's local knowledge rather than upon a located primary source.

Recommendations for Future Research

The framework can be made confirmatory without returning to the site, and the requirement for doing so is an independent reference set, collected after the method has been fixed, of photographs that reproduce the imaging conditions of the original. Specifically, such a reference set would consist of a shaded, leaf-on tree line at a distance of 100 to 150 m, photographed toward a low summer sun at the crest of the canopy approximately two hours before sunset, ideally using the same camera model (Samsung Galaxy S6), or with multi-frame processing such as HDR, scene optimization, and night modes disabled on any substitute device, and comprising 30 to 50 frames across at least five scenes and evenings, together with control frames in which the sun is outside of the frame.

With the strata, thresholds, and test statistic fixed in advance, the null distribution of the maximum patch statistic would then be built from those reference frames rather than from the anomaly image itself, which would remove both the post hoc concern and the small-sample limitation of the deep-shade band that were identified in this study.

A second component of future research would address the photographer's interpretation of the forms directly, which the present study reports but does not test. For this purpose, the author proposes the use of simulation code to generate articulated human forms as parametric stick-figure outlines, varying posture (standing, kneeling, prone with the torso elevated, and partially occluded variants of each), orientation, limb angles, and body proportions across the range of adult human dimensions, and rendering each permutation at the scale of 3 to 4 cm per pixel that applies at the estimated distance, with the blur, noise, and compression characteristics of the original camera applied so that the simulated forms are degraded in the same manner as the photograph.

The degree to which the best-matching humanoid permutation fits each observed pixel cluster would then be measured with a shape-matching statistic, such as intersection over union or a boundary distance, and compared with the null distribution of the same statistic obtained by fitting the same library of humanoid templates to random shapes of comparable size, brightness, and texture, including the ordinary bright patches of the tree line identified in this study. Because the best match is selected from a large number of permutations, the null distribution would be built from the maximum match score over the entire template library, in the same manner as the look-elsewhere correction used in this study, and the resulting future research question and its hypothesis statements would be as follows:

RQ4: Do the anomalous pixel clusters match articulated human forms more closely than would be expected by random chance for shapes of comparable size and brightness?

$H4_0$: The maximum match score between the anomalous pixel clusters and the library of simulated human forms is consistent with the distribution of maximum match scores obtained from random shapes of comparable size and brightness.

$H4_a$: The maximum match score between the anomalous pixel clusters and the library of simulated human forms exceeds what random shapes of comparable size and brightness would be expected to produce.

Such a test would convert the question of pareidolia from a matter of interpretation into a falsifiable statistical question, since a rejection of $H4_0$ would indicate that the clusters resemble human forms more closely than chance alone would predict, whereas a failure to reject it would indicate that the perceived figures are no more humanlike than the shapes that ordinary scene lighting produces.

Conclusion

This study demonstrated that a single anomalous photograph can be tested against ordinary explanations through the use of stated hypotheses, a correction for having found the anomaly by looking, and a correction for veiling glare. Applied to the Battle Mountain photograph, the framework rejected sensor noise and lens flare decisively, and it rejected ordinary scene lighting only when the comparison was restricted to patches at the forms' own illumination level, which is a restriction that the author argues is physically appropriate but which was chosen after the image had been examined. To the knowledge of the author, this represents the first application of formal significance testing with a look-elsewhere correction to an individual anomalous photograph of the kind examined in psychical research.

The historical engagement of July 24, 1863 gives the site its significance and motivates the question that the photograph raises, and while the statistical analysis can narrow the range of ordinary explanations, it cannot answer that question on its own. The confirmatory design described in the Recommendations for Future Research would allow the scene lighting result to be tested on independent data, and the author believes that this framework represents a useful and reproducible precursor assessment for any photograph that is offered as evidence of an anomalous phenomenon.

Data and Code Availability

The original image file, the analysis scripts, and the supporting citation metadata have been deposited in a public archive at Zenodo under the concept digital object identifier 10.5281/zenodo.24139413, which resolves to the most recent version of the deposit, while the particular version archived at the time of writing carries the digital object identifier 10.5281/zenodo.24139414. The deposit was taken from a public repository maintained at https://github.com/twoolman/battle-mountain-anomaly-nulltest, where the development history of the analysis remains visible. The current script is bm_null_test_v3.R, which differs from the earlier bm_null_test_v2.R only in that it writes a record of the software environment at the end of every run. The numerical results reported above are therefore unaffected by the choice between the two. The source code is released under the MIT license, whereas copyright in the photograph is retained by the author and the photograph is released under the Creative Commons Attribution 4.0 International license. The random seed is fixed at 20150718, so that a run of the archived script upon the archived image reproduces the reported values exactly.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors, and it was conducted by the author without institutional support.

Conflict of Interest

The author took the photograph and owned the property on which it was taken, and the author declares no financial conflict of interest.

Declaration of Generative AI and AI-assisted technologies in the writing process

Claude Opus 5 (Anthropic; used from September through November 2026) was used collaboratively with the author in developing the analysis code, validating the statistical diagnostics, retrieving historical sources, and editing the manuscript text, with the author contributing directly to the code and the prose throughout rather than reviewing material produced independently. The study design, the specification of the hypotheses, the choice of the comparison strata, and the interpretation of the findings are the author's own, and the author directed and verified every analysis and takes full responsibility for the content of this manuscript.

References

Baddeley, A., Rubak, E., & Turner, R. (2015). *Spatial point patterns: Methodology and applications with R*. Chapman and Hall/CRC.

Carlotto, M. J. (1997). Evidence in support of the hypothesis that certain objects on Mars are artificial in origin. *Journal of Scientific Exploration, 11*(2), 123–145.

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer. https://doi.org/10.1007/978-1-4471-3675-0

Cousyn, A., Louange, F., & Quick, G. (2013). *The McMinnville pictures* [Technical report]. IPACO. https://www.nicap.org/reports/500511mcminnville_analysis_new.pdf

Desolneux, A., Moisan, L., & Morel, J.-M. (2000). Meaningful alignments. International Journal of Computer Vision, 40(1), 7–23. https://doi.org/10.1023/A:1026593302236

Doherty, B. (2026). *Statistically significant linear alignments among high-confidence transient candidates on POSS-I photographic plates* (arXiv:2605.01190). arXiv. https://arxiv.org/abs/2605.01190

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman & Hall.

Grosjean, B., & Moisan, L. (2009). A-contrario detectability of spots in textured backgrounds. *Journal of Mathematical Imaging and Vision, 33*(3), 313–337. https://doi.org/10.1007/s10851-008-0111-4

Gross, E., & Vitells, O. (2010). Trial factors for the look elsewhere effect in high energy physics. *The European Physical Journal C, 70*(1–2), 525–530. https://doi.org/10.1140/epjc/s10052-010-1470-8

Hurley, B., & Stanton, L. (2022, June 30). A journey through Amissville's history [Graphic feature produced for Foothills Forum]. *Rappahannock News*, 15.

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics: Theory and Methods, 26*(6), 1481–1496. https://doi.org/10.1080/03610929708831995

Lange, R., & Houran, J. (1997). Fortean phenomena on film: Evidence or artifact? *Journal of Scientific Exploration, 11*(1), 41–46.

Liu, J., Li, J., Feng, L., Li, L., Tian, J., & Lee, K. (2014). Seeing Jesus in toast: Neural and behavioral correlates of face pareidolia. *Cortex, 53*, 60–77. https://doi.org/10.1016/j.cortex.2014.01.013

Mayer, G. (2015). Fotografien in der Anomalistik. In G. Mayer, M. Schetsche, I. Schmied-Knittel, & D. Vaitl (Eds.), *An den Grenzen der Erkenntnis: Handbuch der wissenschaftlichen Anomalistik* (chap. 34). Schattauer.

Mayer, G., & Kornmeier, J. (2014). Rätselhafte Objekte auf den Bildern einer Wildkamera oder: Die Tücken der Wahrnehmung. *Zeitschrift für Anomalistik, 14*, 7–24.

McCann, J. J., & Rizzi, A. (2007). Veiling glare: The dynamic range limit of HDR images. In *Proceedings of SPIE 6492, Human Vision and Electronic Imaging XII*, 649213. https://doi.org/10.1117/12.703042

Pickands, J., III. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics, 3*(1), 119–131. https://doi.org/10.1214/aos/1176343003

R Core Team. (2026). *R: A language and environment for statistical computing* (Version 4.6.1). R Foundation for Statistical Computing.

Schwartz, G. E., & Creath, K. (2005). Anomalous orbic "spirit" photographs? A conventional optical explanation. *Journal of Scientific Exploration, 19*(3), 343–358.

Stephenson, A. G. (2002). evd: Extreme value distributions. *R News, 2*(2), 31–32.

Stevenson, I., & Pratt, J. G. (1968). Exploratory investigations of the psychic photography of Ted Serios. *Journal of the American Society for Psychical Research, 62*, 103–129.

Terhune, D. B., Ventola, A., & Houran, J. (2007). An analysis of contextual variables and the incidence of photographic anomalies at an alleged haunt and a control site. *Journal of Scientific Exploration, 21*(1), 99–120.

Urbanek, S. (2026). jpeg: Read and write JPEG images (Version 0.1-11) [Computer software]. Comprehensive R Archive Network.

U.S. Geological Survey. (n.d.). *Geolex: Battle Mountain Alkali Feldspar Granite* [Unit references]. National Geologic Map Database. https://ngmdb.usgs.gov/Geolex/UnitRefs/BattleMountainRefs_11312.html

U.S. War Department. (1889a). *The war of the rebellion: A compilation of the official records of the Union and Confederate armies* (Series I, Vol. 27, Part 1). Government Printing Office.

U.S. War Department. (1889b). *The war of the rebellion: A compilation of the official records of the Union and Confederate armies* (Series I, Vol. 27, Part 2). Government Printing Office.

Virginia Civil War Trails. (n.d.). *Battle Mountain* [Historical marker]. Newby's Crossroads, Rappahannock County, VA. Transcribed at the Historical Marker Database, https://www.hmdb.org/m.asp?m=50140

Willin, M. (n.d.). Photography and the paranormal. In *Psi Encyclopedia*. Society for Psychical Research. https://psi-encyclopedia.spr.ac.uk/articles/photography-and-paranormal

Wiseman, R., Watt, C., Stevens, P., Greening, E., & O'Keeffe, C. (2003). An investigation into alleged "hauntings." *British Journal of Psychology, 94*(2), 195–211. https://doi.org/10.1348/000712603321661886

Woolman, T. A. (2016). A gallium anomaly utilized in palaeogeographic reconstruction of Battle Mountain, Rappahannock County, Virginia. Transactions of the Kansas Academy of Science, 119(2), 155–172.

Woolman, T. A. (2026). *Analysis code for null-hypothesis tests of an anomalous pixel group in a single photograph* (Version 3.0.0) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.24139413