CLINICAL TRIAL REGISTRATION

Registration date
August 6, 2021

Approval and release date
August 6, 2021

PRS review and public post date
August 12, 2021

NCT number
NCT05004649

Website
https://register.clinicaltrials.gov

Unique Protocol Identification Number
305435

Brief Title
Studying the Effects of Natural Visual Scene Changes on Typical Adult Visual Perception

Official Title
The Effects of Stimulus Variability in Natural Visual Scenes

Secondary IDs
K99NS118117

Study Type
Observational

Record Verification Date
August 2021

Overall Recruitment Status
Recruiting

Study Start Date
August 9, 2021 (Actual)

Primary Completion Date
January 31, 2022 (Anticipated)

Study Completion Date
January 31, 2022 (Anticipated)

Responsible Party, by Official Title
Sponsor

Name of the Sponsor
University of Pennsylvania

Collaborators:
National Institute of Neurological Disorders and Stroke (NINDS)

Studies a U.S. FDA-regulated Drug Product
No

Studies a U.S. FDA-regulated Device Product
No

Studies a U.S. FDA IND/IDE
No

Human Subjects Protection Review Board Status
Submitted, approved

Data Monitoring Committee
No

FDA Regulated Intervention
No

Brief Summary

The natural visual environment is complex and rich with different stimuli and features. The visual system must constantly extract behaviorally relevant visual information from an abundance of irrelevant information in the visual scene. To complicate matters further, the visual feature or stimulus that is most relevant at any given moment can change quickly and frequently in realistic visual environments. The mechanisms by which task-relevant information guides perceptual behavior are not fully understood.

In this study, psychophysical experiments will be used to measure participants' ability to discriminate the horizontal position of a central object within a complex, natural visual scene, as well as to measure how that ability is affected by within-trial variability in the features of background objects in the scene.

The goal of this study is to investigate the overarching prediction that the visual system extracts task-relevant information in a manner that reflects realistically complex visual environments in which the stimuli change quickly and frequently. Specifically, this study will test the hypothesis that task-irrelevant variability in the scene affects participants' ability to discriminate the visual feature that is relevant to the task at hand.

Detailed Description

Background

The natural visual environment is often complex, with a single visual scene containing a wide variety of stimuli. These stimuli can change quickly and frequently. To further complicate matters, the behavioral relevance of any particular stimulus or visual feature can also change at any given moment. Visual perception in realistically complex visual environments requires the constant extraction of task-relevant stimulus information from an abundance of irrelevant information. Understanding how task-relevant information is used to guide behavior in the context of constantly changing, feature-rich visual environments is a key component of understanding perception. This may be particularly true if the visual system is evolutionarily optimized to perform in realistically complex environments in which the behaviorally relevant stimulus or feature can change quickly and frequently.

The goal of this study is to investigate the overarching prediction that the visual system is optimized to extract visual information in a generalized manner that is flexible to the wide variety of constantly changing visual features encountered in natural environments. Specifically, this study will test the hypothesis that, in the context of realistically complex natural scenes, task-irrelevant visual feature variability negatively affects participants' ability to discriminate the visual feature that is relevant to the task at hand.

Methods

Participants

The experimental protocols are approved by the University of Pennsylvania Institutional Review Board. Participants will be invited to volunteer to participate in this study. Participants will provide informed consent. To ensure that the participants meet the eligibility criteria, prior to the experiment, they will be interviewed and they will fill out a survey. Also prior to the experiment, they will be screened for visual acuity using a Snellen eye chart and for color deficiencies using the Ishihara plate test. They will be excluded prior to the experiment if their best-corrected visual acuity is worse than 20/40 in either eye or if they make any errors on the Ishihara plate test.

For enrolled participants, their threshold for horizontal position discrimination (in a control condition without any task-irrelevant variability; see *Session organization* below for details) will be calculated based on their performance on the experimental task during their first session. Participants will be excluded after the conclusion of their first session if their horizontal position discrimination threshold in the control condition is higher than a maximum value of 0.6 degrees of visual angle, and participants excluded at this point will not participate in any further experimental sessions. If very few enrolled participants satisfy this criterion, then this maximum threshold value for participant inclusion will be increased. In the case that the maximum threshold value for inclusion is increased, participants that had been previously excluded will not be re-included in the experiment posthoc.

Apparatus

A calibrated LCD color monitor (27-inch NEC MultiSync PA271Q QHD Color Critical Desktop W-LED Monitor with SpectraView Engine; NEC Display Solutions) will be used to display the stimuli in an otherwise dark room, after participants have dark-adapted in the experimental room for a minimum of 5 minutes. The monitor will be driven at a pixel resolution of 1920 x 1080, with a refresh rate of 60 Hz and with 8-bit resolution for each RGB channel. The host computer for this monitor will be an Apple Macintosh with an Intel Core i7 processor. The head position of each participant will be stabilized using a chin cup (Headspot, UHCOTech, Houston, TX). The participant's eyes will be centered horizontally and vertically with respect to the monitor. The distance between the participant's eyes and the monitor will be 75 cm. The participant will input their responses using a Logitech F310 gamepad controller.

Stimuli

All of the stimuli are variants of the same natural visual scene: a square image (subtending 8 degrees of visual angle in both width and height), in which a central object (a banana, subtending approximately 4 degrees of visual angle in height) is presented on an approximately circular array of overlapping background objects (subtending approximately 5 degrees of visual angle and made up of overlapping branches and leaves). The central object (the banana) and/or what will be referred to as the "background objects" (the branches and leaves) change in horizontal position, rotation, and/or depth across different stimuli. The central object and background objects are presented in the context of other objects that do not ever move across different stimuli (a rock ledge, a skyline, and three moss-covered stumps).

The natural visual scene was created using Blender, an open-source 3D creation suite (https://www.blender.org, Version 2.81a). The central object and/or the background objects were moved in horizontal position, rotation, and/or depth to create different stimuli using iset3d, an open-source software package (https://github.com/ISET/iset3d) that works with a modified version of pbrt (https://github.com/mmp/pbrt-v3). The

images were created using iset3d at a resolution of 1920 x 1920 with 100 samples per pixel, at 31 equally spaced wavelengths between 400 nm and 700 nm.

The images created using iset3d were converted to RGB images using custom software (Natural Image Thresholds; https://github.com/AmyMNi/NaturalImageThresholds) written using MATLAB (MathWorks; Natick, MA) and based on the software package Virtual World Color Constancy (github.com/BrainardLab/VirtualWorldColorConstancy). Natural Image Thresholds is dependent on routines from the Psychophysics Toolbox (http://psychtoolbox.org), iset3d (https://github.com/ISET/iset3d), and isetbio (http://psychtoolbox.org). To convert a hyperspectral image created using iset3d to an RGB image for presentation on the calibrated monitor, the hyperspectral image data were first used to compute LMS cone excitations. Then, the LMS cone excitations were converted to a metameric rendered image in the RGB color space of the monitor, based on the monitor calibration data. The RGB image was gamma corrected using a common scaling that brought all of the RGB images in the stimulus set into the display gamut of the monitor.

The stimuli will be presented on the calibrated monitor in the context of the psychophysical task. The stimuli will be presented on a uniform gray background (~100 cd/m^2), which will be presented on the monitor for the duration of the experimental session.

Psychophysical task

A psychophysical task will be used to measure participants' ability to discriminate the horizontal position of the central object that is presented within the context of background objects in a natural visual scene. The task will be a two-interval forced choice task that presents one stimulus per interval. Each interval will have a duration of 250 ms. Each stimulus will be presented at the center of the monitor. Between the two stimulus intervals, two masks will be shown in succession at the center of the monitor. Each mask will be presented for a duration of 400 ms, for a total interstimulus interval of 800 ms (see *Session organization* below for mask details).

The task of the participant will be to determine whether, compared to the central object presented in the first interval, the central object presented in the second interval is to the left or to the right. Following the two intervals, the participant will have an unlimited amount of time to press one of two response buttons on the gamepad (the upper left trigger to indicate that the central object in the second interval was to the left, the upper right trigger to indicate that it was to the right). One of two feedback tones will be presented after the response is entered, indicating whether the participant was correct or incorrect. For trials in which there is no difference in the position of the central object between the two intervals, the response that will receive the correct feedback tone will be randomly selected per trial. The trials will be separated by an intertrial interval of approximately 1 second.

The experimental programs can be found in the custom software package Natural Image Thresholds (https://github.com/AmyMNi/NaturalImageThresholds). They were written in MATLAB (MathWorks; Natick, MA) and were based on the software package Virtual World Color Constancy (github.com/BrainardLab/VirtualWorldColorConstancy). They rely on routines from the Psychophysics Toolbox (http://psychtoolbox.org) and mgl (http://justingardner.net/doku.php/mgl/overview).

Session organization

It is expected that participants will complete this pilot experiment in six sessions. The first session will include participant enrollment procedures (informed consent, vision tests, etc.; see *Participants* above for details) as well as familiarization trials (see next paragraph) and will last approximately one and a half hours. The second through sixth sessions will last approximately one hour each.

In the first session, prior to beginning the task, the participant will be provided with task instructions and will be given the opportunity to practice pressing the response buttons. For the first session only, the participant will begin with 30 familiarization trials. The familiarization trials will comprise, in order: 10 randomly selected easy trials (the largest position-change comparisons), 10 randomly selected medium-difficulty trials (the 4th and 5th largest position-change comparisons), and 10 randomly selected trials from all possible position-change comparisons. The familiarization trials will not include any task-irrelevant variability. Data from the familiarization trials will not be saved. The familiarization trials will be followed by a break, during which the participant will be given the opportunity to ask any questions they may have. The break will end when the participant indicates that they are ready (using a button press), and the experiment will begin.

In each session, there will be two conditions: “condition” refers to the reference position of the central object. For each reference position, there will be 11 “comparison” positions for the central object: five comparison positions in the positive horizontal direction, five comparison positions in the negative horizontal direction, and a comparison position of 0 indicating no change. In each trial, one interval will contain a reference stimulus and the other interval will contain one of that reference stimulus’s comparison stimuli. The order in which these two stimuli are presented within a trial will be selected randomly per trial.

A “block” of trials will consist of 2 conditions and 11 comparisons per condition, for a total of 22 trials. The trials within a block will be run in randomized order. A block will be completed before the next block of trials begins. There will be 14 iterations of a block, for a total of 308 trials.

This set of 308 trials will make up a single “noise level”. A single session will consist of three noise levels: Noise Level 0, Noise Level 1, and Noise Level 2. The trials for each noise level will be divided into two “runs” (154 trials per run). Thus, each run will comprise a single noise level. The six runs will be run in random order per session. Each run will be separated by a break that lasts at least one minute and during which the participant will be instructed to stand or stretch as needed. Each break will end when the participant indicates that they are ready (using a button press).

Across all six runs, there will be a total of 924 trials. Additionally, each session will begin with four practice trials (including the first experimental session, which will be preceded by familiarization trials as described above). Each session will also include one practice trial following each of the five breaks. Each practice trial will be randomly selected from the set of easy trials (described above) and will not include any task-irrelevant variability. The data from the practice trials will not be saved. Including the nine practice trials, there will be a total of 933 trials per session.

For Noise Level 0, there will not be any changes to the background objects (the branches and leaves). Noise Level 0 will be the control condition and will be used to determine the participant’s threshold for discriminating the horizontal position of the central object without any task-irrelevant stimulus noise. Noise Levels 1 and 2 will be used to determine the participant’s threshold for discriminating the horizontal position of the central object in the presence of task-irrelevant stimulus noise.

Noise Level 1 will consist of task-irrelevant noise in a single task-irrelevant feature: rotation. A task-irrelevant rotation amount will be applied to each stimulus separately. For each stimulus, a single rotation amount will be drawn randomly from a pool of 51 rotation amounts, and the background objects in the stimulus will all be rotated by that rotation amount. The rotation amount will be drawn separately (randomly with replacement) for each of the two stimuli presented in a trial (the reference position stimulus, and the comparison position stimulus). The pool of 51 rotation amounts will comprise: one rotation amount of zero (no change to the background objects), 25 equally spaced rotation amounts in the clockwise direction, and 25 equally spaced rotation amounts in the counterclockwise direction.

Noise Level 2 will consist of task-irrelevant noise in two task-irrelevant features: rotation and depth. For Noise Level 2, there will also be a pool of 51 noise amounts, but each noise amount in the pool will consist of both a rotation amount (the same 51 rotation amounts as in Noise Level 1) and a depth amount. There will be 51 possible depth amounts (one depth amount of zero, 25 equally spaced depth amounts in the position direction, and 25 equally spaced depth amounts in the negative direction). For the Noise Level 2 pool of 51 noise amounts, one of the noise amounts will consist of a rotation amount of zero and a depth amount of zero. For the remaining 50 noise amounts in the pool, each of the remaining 50 rotation amounts will be randomly assigned (without replacement) to one of the remaining 50 depth amounts. From this Noise Level 2 pool of

51 noise amounts, a single noise amount will be randomly drawn (with replacement) for each of the two stimuli in a trial separately.

Finally, as noted above (see *Psychophysical task*), two masks will be shown per trial during the interstimulus interval. All masks across all noise levels will be drawn from the same distribution of stimuli (stimuli with Noise Level 0, thus containing no task-irrelevant noise). To create each of the two masks in a trial: first, the central object positions in the first and second intervals of the trial will be determined. The two stimuli with Noise Level 0 that match the central object positions in the first and second intervals will be used to create the trial masks. For each of these two stimuli, the average intensity will be calculated per RGB channel per 16 x 16 block of the stimulus. Next, each 16 x 16 block of a mask will be randomly drawn from either one or the other stimulus, and will be made up of the uniform average intensity calculated per RGB channel for that block of the selected stimulus. Thus, the two masks shown per trial will each consist of a different random draw per 16 x 16 block.

Data Analysis

Per session, the participant's threshold for discriminating object position will be measured for each noise level. First, for each comparison position, the proportion of trials on which the participant responded that the comparison stimulus was located to the right of the reference stimulus will be calculated. Next, the proportion the comparison was chosen as rightwards will be fit with a cumulative normal function using the Palamedes Toolbox (http://www.palamedestoolbox.org). To estimate all four parameters of the psychometric function (threshold, slope, lapse rate, and guess rate), the lapse rate will be set equal to the guess rate and will be forced to be in the range [0, 0.05], and the model will be fit to the data using the maximum likelihood method. The threshold will be calculated as the difference between the stimulus levels at performances (proportion the comparison was chosen as rightwards) equal to 0.7602 and 0.5 as determined by the cumulative normal fit.

Primary Disease or Condition Being Studied in the Trial, or the Focus of the Study
Visual Perception

Keywords
Perception, Sensory Processing, Vision

Observational Study Design
Observational Study Model: Case-Only
Time Perspective: Prospective
Biospecimen Retention: None Retained
Enrollment: 10 anticipated enrolled participants
Number of Groups/Cohorts: 1

Primary Outcome Measure Information
Title: Psychophysical threshold measurements
Description: Behavioral assays of psychophysical threshold measurements for discriminations of object position
Time Frame: Approximately 3 weeks

Eligibility Sex
All

Age Limits Minimum Age
18

Age Limits Unit of Time
Years

Age Limits Maximum Age/Unitof Time
N/A (No limit)

Accepts Healthy Volunteers
Yes

Eligibility Criteria
Inclusion Criteria:
- Normal visual acuity
- Capable of giving informed consent
- Fully vaccinated against COVID-19

Exclusion Criteria:
- Known color deficiencies
- Diagnosis of retinal disease or inherited retinal disease from family history
- A psychophysical threshold for horizontal position discrimination that is greater than 0.6 degrees of visual angle (to be determined during the first experimental session)

Study Population Description
The subject population for this study will be drawn primarily from the University of Pennsylvania community and the surrounding Philadelphia community.

Sampling Method
Non-Probability Sample

Central Contact Person
Amy M. Ni, Ph.D.
412-268-3922
amyni@sas.upenn.edu

Overall Study Officials

1) Amy M. Ni, Ph.D., University of Pennsylvania, Study Principal Investigator
2) David H. Brainard, Ph.D., University of Pennsylvania, Study Director

Facility Information

University of Pennsylvania, Philadelphia PA, 19104 USA

Individual Site Status

Recruiting

Facility Contact

Amy M. Ni, Ph.D.
412-268-3922
amyni@sas.upenn.edu

Citations

1) Brainard DH. The Psychophysics Toolbox. Spat Vis. 1997;10(4):433-6. Epub 1997/01/01. PubMed PMID: 9176952.
2) Cottaris NP, Jiang H, Ding X, Wandell BA, Brainard DH. A computational-observer model of spatial contrast sensitivity: Effects of wave-front-based optics, cone-mosaic structure, and inference engine. J Vis. 2019;19(4):8. Epub 2019/04/04. doi: 10.1167/19.4.8. PubMed PMID: 30943530.
3) Cottaris NP, Wandell BA, Rieke F, Brainard DH. A computational observer model of spatial contrast sensitivity: Effects of photocurrent encoding, fixational eye movements, and inference engine. J Vis. 2020;20(7):17. Epub 2020/07/22. doi: 10.1167/jov.20.7.17. PubMed PMID: 32692826; PubMed Central PMCID: PMCPMC7424933.
4) DiCarlo JJ, Cox DD. Untangling invariant object recognition. Trends Cogn Sci. 2007;11(8):333-41. Epub 2007/07/17. doi: 10.1016/j.tics.2007.06.010. PubMed PMID: 17631409.
5) DiCarlo JJ, Zoccolan D, Rust NC. How does the brain solve visual object recognition? Neuron. 2012;73(3):415-34. Epub 2012/02/14. doi: 10.1016/j.neuron.2012.01.010. PubMed PMID: 22325196; PubMed Central PMCID: PMCPMC3306444.
6) Gauthier I, Tarr MJ. Visual Object Recognition: Do We (Finally) Know More Now Than We Did? Annu Rev Vis Sci. 2016;2:377-96. Epub 2017/05/24. doi: 10.1146/annurev-vision-111815-114621. PubMed PMID: 28532357.
7) Heasly BS, Cottaris NP, Lichtman DP, Xiao B, Brainard DH. RenderToolbox3: MATLAB tools that facilitate physically based stimulus rendering for vision research. J Vis. 2014;14(2). Epub 2014/02/11. doi: 10.1167/14.2.6. PubMed PMID: 24511145; PubMed Central PMCID: PMCPMC3919102.
8) Ni AM, Huang C, Doiron B, Cohen MR. A general decoding strategy explains the relationship between behavior and correlated variability. bioRxiv 2020.10.08.331850. doi: https://doi.org/10.1101/2020.10.08.331850.
9) Ni AM, Ruff DA, Alberts JJ, Symmonds J, Cohen MR. Learning and attention reveal a general relationship between population activity and behavior. Science. 2018;359(6374):463-5.

Epub 2018/01/27. doi: 10.1126/science.aao0284. PubMed PMID: 29371470; PubMed Central PMCID: PMCPMC6571104.
10) Prins N, Kingdom FAA. Applying the Model-Comparison Approach to Test Specific Research Hypotheses in Psychophysical Research Using the Palamedes Toolbox. Front Psychol. 2018;9:1250. Epub 2018/08/08. doi: 10.3389/fpsyg.2018.01250. PubMed PMID: 30083122; PubMed Central PMCID: PMCPMC6064978.
11) Ruff DA, Ni AM, Cohen MR. Cognition as a Window into Neuronal Population Space. Annu Rev Neurosci. 2018;41:77-97. Epub 2018/05/26. doi: 10.1146/annurev-neuro-080317-061936. PubMed PMID: 29799773; PubMed Central PMCID: PMCPMC6571103.
12) Singh V, Cottaris NP, Heasly BS, Brainard DH, Burge J. Computational luminance constancy from naturalistic images. J Vis. 2018;18(13):19. Epub 2018/12/29. doi: 10.1167/18.13.19. PubMed PMID: 30593061; PubMed Central PMCID: PMCPMC6314111.
13) Singh V, Burge J, Brainard DH. Equivalent noise characterization of human lightness constancy. bioRxiv 2021.06.04.447171. doi: https://doi.org/10.1101/2021.06.04.447171.

Document Section
Uploaded PDF: Informed Consent Form