ARL-TN-1313 ● JUNE 2026

PARADIGMS: Platform for Accelerating Research And Development that is Integrative and Generative over Massive Scales – Final Report

by Stephen R. Mitroff, Dwight J. Kravitz, and Kelvin S. Oie

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

ARL-TN-1313 ● JUNE 2026

PARADIGMS: Platform for Accelerating Research And Development that is Integrative and Generative over Massive Scales – Final Report

Stephen R. Mitroff and Dwight J. Kravitz
The George Washington University and The National Science Foundation

Kelvin S. Oie
DEVCOM Army Research Laboratory

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
		START DATE	END DATE
June 2026	Technical Note	9/29/2023	9/28/2025

4. TITLE AND SUBTITLE
PARADIGMS: Platform for Accelerating Research And Development that is Integrative and Generative over Massive Scales – Final Report

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
W911NF-23-2-0097		
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**

6. AUTHOR(S)
Stephen R. Mitroff, Dwight J. Kravitz, and Kelvin S. Oie

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
DEVCOM Army Research Laboratory ATTN: FCDD-RLA-FA Aberdeen Proving Ground, MD 21005	ARL-TN-1313

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES
ORCIDs: Stephen R. Mitroff, 0000-0003-3524-6396; Dwight J. Kravitz, 0000-0002-8108-1329; Kelvin S. Oie, 0000-0002-0311-800X.
This technical note represents the final report for cooperative agreement W911NF-23-2-0097 between the DEVCOM Army Research Laboratory and The George Washington University.

14. ABSTRACT
Meeting the challenges of adaptability and variability in human behavior requires more thorough and extensive quantifications of behavior with timelines that can also keep up with the increasing pace of the required innovation. To address this need, this 2-year effort developed a proof-of-concept implementation of a modular, scalable platform known as PARADIGMS (Platform for Accelerating Research And Development that is Integrative and Generative over Massive Scales). This delivered capability provides an end-to-end research-grade solution that integrates experimental tasks, preregistration, and automated sampling/resampling and analyses to drastically increase sample sizes and shorten experimentation timelines. The project successfully delivered an initial operating capability and a proof-of-concept demonstration experiment, which recruited and obtained data from 936 participants in just 1 month. The PARADIGMS platform demonstrates the utility of crowdsourced experimentation for advancing the understanding of human cognitive mechanisms, generally, and U.S. Army priorities, specifically.

15. SUBJECT TERMS
crowdsourced experimentation, research methods, automated data collection, behavioral research, Humans in Complex Systems

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT	b. ABSTRACT	C. THIS PAGE		
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UU	19

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER *(Include area code)*
Kelvin S. Oie	(520) 691-5705

STANDARD FORM 298 (REV. 5/2020)
Prescribed by ANSI Std. Z39.18

Contents

List of Tables

1. Introduction and Proposed Approach

The primary goal of this cooperative agreement was to deliver an initial operating capability (IOC) for a research-grade platform that can improve the efficiency and efficacy of U.S. Army S&T involving human participants by 1) drastically shortening experimentation timelines; 2) decreasing cost and resourcing needs; and 3) improving the transparency, repeatability, and quality of research results.

Human behavior is remarkable for its adaptability and variability among individuals. While these attributes can lead to behaviors that clearly benefit individuals and society, they present challenges for U.S. Army decision-making as new systems (e.g., intelligent agents) are developed. Even when these systems are only enhancing or extending existing capabilities, they may cause unforeseen changes in Soldier behaviors, let alone when an entirely new system is deployed. Further, each Soldier is different, making it crucial to quantify how the efficacy of a new system varies across personnel and whether that efficacy changes as the personnel does. Meeting these challenges requires more thorough and extensive quantifications of behavior with timelines that can also keep up with the increasing pace of the required innovation.

This 2-year effort developed a proof-of-concept implementation of a modular, scalable platform: PARADIGMS (Platform for Accelerating Research And Development that is Integrative and Generative over Massive Scales). The delivered capability matured existing capabilities developed by The George Washington University (GWU) and implemented new capabilities to provide an end-to-end research-grade solution that integrates experimental tasks, preregistration, and automated sampling/resampling and analyses, in a modular system to improve the efficiency, transparency, and rigor of future research efforts.

The following subsections outline the main objectives of Tasks 1–3 for the PARADIGMS project.

1.1 Task 1: PARADIGMS System Development

The efforts executed to accomplish Task 1 of the project were to mature and integrate existing software components developed by GWU, implement new components as needed, and test component and system performance to deliver an IOC of the PARADIGMS system. Additionally, year 2 efforts focused on updates to component capabilities in the IOC implementation based on identified needs arising from system deployments in support of research data collection efforts.

1.2 Task 2: Proof-of-Concept Demonstration Experiment

Task 2 focused on efforts to design, develop, implement, and execute an experiment utilizing the full capabilities of the PARADIGMS platform to provide a proof-of-concept demonstration. Project objectives aimed for an experiment that would align with (threshold) or directly address (objective) Army-identified scientific objectives or questions within current Army S&T programs and priorities.

1.3 Task 3: System Documentation

Task 3 for the project was to document the software and provide a manual to use the PARADIGMS system, as well as to deliver any accompanying materials, as applicable—including updates necessitated by changes to the IOC software code base and system capabilities.

2. Project Accomplishments

Over the course of this award's period of performance, GWU and U.S. Army Combat Capabilities Development Command (DEVCOM) Army Research Laboratory (ARL) research teams worked together to address the goals of Tasks 1–3. This resulted in multiple accomplishments, including successful completion of all project deliverables and scientific advances, which have been reported in numerous knowledge products. (These are itemized in the Appendix.) At a high level, this award contributed to 7 manuscripts published or in-process, 4 conference talks, 10 conference poster presentations, and 2 invited talks. Eight of the works produced under this project were jointly authored by GWU and ARL collaborators. Ongoing efforts that extend from the research results and advances made within this project will expand these numbers further.

2.1 Deliverables and Notable Advances

2.1.1 Task 1: PARADIGMS System Development

Task 1 focused on the creation and initial optimization of the PARADIGMS system. Fundamentally, PARADIGMS is an integration between experimental code hosted on self-contained servers and crowd-sourcing platforms (e.g., Amazon Mechanical Turk, Prolific), toward the greatest possible automation in the recruitment, replacement, and counterbalancing of large samples of human participants.

As such, it required the maturation and integration of existing software components, and the implementation of new components, as needed. The platform was developed to have the following modules/capabilities:

- *Modular Operation*. PARADIGMS was implemented so that experiment presentation, data collection, and analyses are largely independent of any specific online crowdsourcing platform (e.g., Amazon Mechanical Turk, Prolific). This allows PARADIGMS to leverage most, if not all, of these platforms—or any new platforms that may be developed (internal or external to the Army), without being dependent on any of them.
- *Automated Data Collection*. Preregistration requires an a priori specification of the number of participants in each condition, as well as any exclusion criteria. In many cases that creates a need to run basic analyses and solicit participants multiple times as random participants are excluded. PARADIGMS automates this process, reducing cost while increasing transparency and reliability. The platform automatically assigns participants to the right version of the experiment, runs preregistered exclusion checks, and creates new requests to replace the excluded participants.
- "*One-click" Replication*. For most of the possible research questions, all experiments, in their totality (e.g., presentation, data collection, exclusion checks, preregistered analyses), can be captured within a server image that can be included in all preregistrations. These images can be updated to include experiment pilot data and any exploratory analysis code, experimental data after data collection, and be made available concomitant with publication. Anyone with these images will be able to spin-up new versions of these servers capable of immediately replicating the experiment all the way to figure creation on new participants.
- *Modular Code Base*. To an extent, each experiment is unique, requiring at least different stimulus presentation. However, there are several aspects that are generally common (e.g., demographic questions, data logging, reading counterbalancing files). PARADIGMS experiments can be built with predefined classes/functions for these common operations, reducing the burden of making alterations to existing server images to extend/modify prior experiments. Documentation can also be provided within each experiment server image explaining how to leverage these classes/functions.

2.1.2 Task 2: Proof-of-Concept Demonstration Experiment

Task 2 centered on the design, development, implementation, and execution of an experiment utilizing the full capabilities of the PARADIGMS platform to provide proof-of-concept demonstration. While we have run several experiments using the platform's capabilities, here we highlight a particularly extensive project that resulted in publication in a high-tier psychology journal (Mahableshwarkar et al. 2025).

The project was focused on the cognitive mechanisms underlying the ability to make spatial judgements about environments depicted in 2D images, a common aspect of many Army tasks and of increasing importance with the expansion of the remote operation of mission assets. In brief, participants were presented with 156 unique images of unfamiliar scenes taken with a camera positioned like a human observer. Across two tasks (Numerical Estimate, Discrimination) and three stimulus durations (125, 250, 1000 ms), they were asked to judge the distance between the observer and a target embedded in the scene. In the Numerical Estimate task, participants gave an explicit estimate of the distance, whereas in the Discrimination task, they judged which of two targets was farther away.

A total of 156 observers participated in each of the 6 experiments, such that the order of the stimuli fell into a Latin square design, with each stimulus occurring in each trial slot once. The stimulus orders were matched between the Numerical Estimate and Discrimination tasks. In total, it only took approximately 1 month to recruit and obtain data from all 936 included participants. This controlled and efficient counterbalancing, enabled by the PARADIGMS platform, revealed a remarkable mechanism that humans use to extract depth information from images.

We observed the following:

1) Humans are extremely accurate at these judgements;
2) They form such perceptions in as little as 125 ms; and
3) They make systematic errors in such perceptions that generalize across observers, tasks, and stimulus durations.

The observed data were so systematic that the stimulus unique perception of one group of participants in a different task with a different stimulus duration predicted the perceptions in the other conditions. Thus, there had to be stimulus features responsible for the generalizable perception, and the quality of the data allowed us to demonstrate two such features: ground plane size and familiar size of objects. Follow-up experiments are currently under review that further explore the impact of these features on depth perception and the deployment of attention in 2D images depicting 3D scenes.

Overall, this project and the other experiments our lab has performed and is continuing to perform using the PARADIGMS platform demonstrate the utility of crowdsourced experimentation for advancing the understanding of human cognitive mechanisms, generally, and Army priorities, specifically.

2.1.3 Task 3: System Documentation

As part of the development of the PARADIGMS platform and our general commitment to good Open Science practice, we directly provide commented code for all of the experimentation and analyses underlying our published work. Each individual project is fully preregistered and contains a link to an Open Science Framework repository[*] containing all of the code, stimuli, and data (e.g., Mahableshwarkar et al. [2025]). Moreover, because the PARADIGMS system is designed to function on cloud-based resources, each project also has an image of an Amazon Web Services instance that can be shared and spun-up to immediately replicate the project, including the mechanics of the PARADIGMS system. Finally, the lab maintains an overall GitHub repository,[†] which shares all relevant code for PARADIGMS projects.

[*] https://osf.io/28vjd
[†] https://github.com/orgs/GWVisCogLab/repositories

3. References

Mahableshwarkar P, Houck L, Philbeck J, Kravitz DJ. Distance perception in natural scene images generalizes across individuals and viewing time. J Exp Psychol Gen. 2025;154(6):1545–1570.

Appendix. Project Knowledge Product Metrics

A.1 Research Team

Table A-1. The George Washington University (GWU) researchers.

Co-Principal Investigator:	Stephen R. Mitroff
Co-Principal Investigator:	Dwight J. Kravitz
Postdoctoral Fellow:	Audrey Siqi-Liu
Ph.D. Graduate Student:	Samoni Nag
Ph.D. Graduate Student:	Emma Siritzky
Ph.D. Graduate Student:	Justin Grady
Ph.D. Graduate Student:	Sarah Malykke
Undergraduate Researcher:	Sarah Hodgdon
Post-Baccalaureate Researcher:	Victor Platzer
Programmer:	Adam Hereth

Table A-2. U.S. Army Combat Capabilities Development Command (DEVCOM) Army Research Laboratory (ARL) researchers.

Co-Principal Investigator:	Kelvin S. Oie
Army Educational Outreach Program (AEOP) Postdoctoral Fellow:	Ellie Robbins
Oak Ridge Associated Universities (ORAU) Postdoctoral Fellow:	Melanie Bancilhon

A.2 Accomplishments by Category*

Publications and Manuscripts

Callahan-Flintoft C, Cox PH, Siritzky EM, Mitroff SR, **Oie KS**, Kravitz DJ. Repeated exposure to task-relevant and task-irrelevant information—and their interaction—affect visual search performance. Atten Percept Psychophys. 2026;88(3):91. https://doi.org/10.3758/s13414-025-03217-0

Mahableshwarkar P, Houck L, Philbeck J, Kravitz DJ. Distance perception in natural scene images generalizes across individuals and viewing time. J Exp Psychol Gen. 2025;154(6):1545–1570. https://doi.org/10.1037/xge0001741

Ritchie JB, Wardle S, Vaziri-Pashkam M, Kravitz DJ, Baker CI. Rethinking category-selectivity in human visual cortex. Cogn Neurosci. 2026;17(2):1–28. https://doi.org/10.1080/17588928.2025.2543890

* ARL authors in **bold**.

Baker CI, Kravitz DJ. Insights from the evolving model of two cortical visual pathways. J Cog Neurosci. 2024;36(12):2568–2579. https://doi.org/0.1162/jocn_a_02192

Kravitz DJ, Mitroff SR, **Callahan-Flintoft C**, **Oie KS**. Crowdsourced data collection opens new avenues for the behavioral sciences to impact real-world applications. Policy Insights Behav Brain Sci. 2024;11(2):123–131. https://doi.org/10.1177/23727322241274745

Nau M, Schmid A, Kaplan S, Baker CI, Kravitz DJ. Towards a task-centered framework for cognitive neuroscience. Nat Neurosci. Manuscript in press.

Tkacz-Domb, S, Shomstein S, Kravitz DJ. Greater interference between perception and working memory across the lifespan: feature- and sex-dependent effects. Manuscript under review.

Conference Podium Presentations

Kravitz DJ, **Robbins E**, Mahableshwarkar P, Paranjape S, Hereth A, **Oie KS**, Mitroff SR. PARADIGMS: a new system for precise, accurate, and efficient measurement of human behavior. University of Colorado Human-Machine Teaming Paradigms Meeting; 2024 Mar 15–16; Boulder, CO.

Tkacz-Domb S, Shomstein S, Kravitz DJ. Greater interference between ongoing perception and working memory across the lifespan for both high- and low-level visual features. Psychonomic Society 65th Annual Meeting; 2024 Nov 21–24; New York, NY.

Cox PH, **Callahan-Flintoft C**, Kramer MR, Mitroff SR, Kravitz DJ. An extremely high-fidelity quantification of the time course of learning consolidation from a massive behavioral dataset. Psychonomic Society 64th Annual Meeting; 2023 Nov 16–19; San Francisco, CA.

Paranjape S, Shomstein S, Kravitz DJ. Faces in working memory cause racial biases in subsequent trustworthiness judgments. OPAM 31: Object Perception, Attention, and Memory Workshop; 2023 Nov 16; San Francisco, CA.

Conference and Workshop Poster Presentations

Bancilhon M, Siqi-Liu A, Platzer V, Kravitz DJ, **Oie KS**, Mitroff SR. Exploring individual differences in reliance on AI-assisted visual search. VSS 2025: 25th Annual Meeting of the Vision Sciences Society; 2025 May 16–20; St Pete Beach, FL.

Dochnal M, Siqi-Liu A, Malykke SB, Mitroff SR. Using mouse tracking data to examine adaptation of visual search behaviors to transitions in task complexity. GWU CCAS Research Showcase; 2025 Apr 2; Washington, DC.

Grady JN, Siritzky EM, Kravitz DJ, Mitroff SR. Characterizing the independent and joint impacts of previous category evidence and visual and semantic similarity on visual attention. VSS 2025: 25th Annual Meeting of the Vision Sciences Society; 2025 May 16–20; St Pete Beach, FL.

Malykke SB, Siqi-Liu A, **Oie KS**, Kravitz DJ, Mitroff SR. The time course of an error: Benefits of examining pre-error speeding and post-error slowing simultaneously. VSS 2025: 25th Annual Meeting of the Vision Sciences Society; 2025 May 16–20; St Pete Beach, FL.

Paranjape S, Shomstein S, Kravitz DJ. Interactions between visual working memory and face perception. J Vision. 2024 Sep;24(10):950.

Robbins E, Kravitz DJ, Hereth A, **Oie KS**, Mitroff SR. PARADIGMS: a new system for precise, accurate, and efficient measurement of human behavior: collaboration opportunities. University of Colorado Human-Machine Teaming Paradigms Meeting; 2024 Mar 15–16; Boulder, CO.

Tkacz-Domb S, Shomstein S, Kravitz DJ. The strength of visual working memory representation deteriorates with age. Society for Neuroscience, 2024 Oct 5–9; Chicago, IL.

Tkacz-Domb S, Shomstein S, Kravitz DJ. The influence of visual perception on working memory throughout the adult lifespan. VSS 2024: 24th Annual Meeting of the Vision Sciences Society; 2024 May 17–22; St Pete Beach, FL.

Mahableshwarkar P, Philbeck J, Kravitz DJ. Establishing task-general representations and the cues which drive them in spatial cognition. Society for Neuroscience; 2023 Nov 11–14; Washington, DC.

Tkacz-Domb S., Shomstein, S, Kravitz, DJ. The strength of visual working memory degrades with age. OPAM 31: Object Perception, Attention, and Memory Workshop. 2023 Nov 16; San Francisco, CA.

Other Invited Talks

Kravitz DJ. Embracing complexity in theory and practice. Laboratory of Brain and Cognition Lab Meeting, National Institutes of Mental Health (NIMH); 2024; Bethesda, MD.

Mitroff SR. Using "theory-driven big data" to advance psychology research. National Academies, Board on Behavioral, Cognitive, and Sensory Sciences. National Academies of Science and Engineering (US); 2025; Washington, DC.

Brownbag Talks

Bancilhon M. The effect of individual differences and presentation formats on decision-making. The George Washington University (US), Department of Psychological & Brain Sciences, Cognitive Neuroscience Brownbag Series; 2024.

Paranjape S. Interactions between working memory and face perception. The George Washington University (US), Department of Psychological & Brain Sciences, Cognitive Neuroscience Brownbag Series; 2023.

Tkacz-Domb S. The interaction of visual working memory and perception across the adult lifespan. The George Washington University (US), Department of Psychological & Brain Sciences, Cognitive Neuroscience Brownbag Series; 2024.

Awards

Mahableshwarkar P. MR People's Choice Award; 2nd Place, GWU 3-Minute Thesis (3MT) Competition.

Degrees and Promotions

- Hodgdon SV. Accepted to Lehigh University Ph.D. program, 2025 June
- Siritzky EM. Comprehensive examinations, 2025 June
- Malykke SB. Comprehensive examinations, 2025 Apr
- Malykke SB. Master's degree, 2024 Dec
- Siritzky EM. Master's degree, 2024 Dec

List of Symbols, Abbreviations, and Acronyms

2/3D	Two-/Three-Dimensional
AEOP	Army Educational Outreach Program
ARL	Army Research Laboratory
DEVCOM	U.S. Army Combat Capabilities Development Command
GWU	The George Washington University
IOC	Initial Operating Capability
ORAU	Oak Ridge Associated Universities
PARADIGMS	Platform for Accelerating Research And Development that is Integrative and Generative over Massive Scales
S&T	Science and Technology

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
1 (PDF)	DEVCOM ARL FCDD RLB CB TECH LIB