Citation Evidence Report

EB-1B Petition — Outstanding Professor or Researcher

8 CFR § 204.5(i)(3) · Authorship + Original Contributions

Tianyu Yang

University of Notre Dame

Google Scholar profile

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to the 8 CFR § 204.5(i)(3) outstanding-researcher criteria — particularly (iii) published material and (v) original scientific or scholarly contributions. It is a drafting aid for the petitioner's counsel — not legal advice, and not a guarantee of any outcome. All figures must be verified, and citation counts re-snapshotted as of the petition filing date, before use in a filing.

A. Overview & Filtering Statement

142	149	23	5
Citing papers mapped	Citation edges	Home papers mapped	h-index (GS)

Filtering statement — methodology & limits

Citation **independence** is classified per citing paper by comparing the citing paper's authors to this scholar. *Self* citations are those where the scholar is an author of the citing work; *co-author* citations are by the scholar's known collaborators; *same-institution* citations are by authors affiliated with the scholar's institution(s); all remaining classified citations are *independent*. Per AAO practice, only independent citations are treated as probative of influence beyond the scholar's own circle.

Known limitations — counsel must verify. (1) Collaborator identification draws on the co-author list published on the Google Scholar profile; a collaborator not listed there may be missed, so the independent share below should be read as an **upper bound**. (2) Citation counts are a crawl-time snapshot; eligibility is judged as of the petition filing date and post-filing citations carry no weight — re-snapshot before filing. (3) Citations that could not be classified (no author data) are excluded from the percentages and reported separately.

B. Citation Independence

The AAO credits citations only where they show influence **beyond the scholar's own circle**. Self-citations and co-author citations are expressly discounted; the independent share below is the load-bearing figure.

85.3% independent of 102 classified citing papers

Citation type	Count
Independent	87
Self-citation	5
Co-author	9
Same-institution	1

40 citing papers could not be classified (no author data) and are excluded from the percentages above.

C. Significant Contributions & Their Citation Evidence

Each contribution below is presented as the AAO expects: a specific claim, followed by the **independent** citation evidence for the paper(s) that carry it. Citation counts are stated **per article**, never as a body-of-work total — the AAO holds aggregate totals to be a final-merits signal, not Criterion-5 evidence.

Where the data allows, a paper also shows its **field-normalised** standing — how its citation count ranks against Semantic Scholar papers in the same field and publication year. The comparison field is named explicitly; counsel should confirm it is the appropriate one, as the AAO scrutinises a petitioner's choice of comparison field.

Contribution 1

Claim — Contribution 1

The researcher established a foundational multimodal mathematical reasoning framework, subsequently synthesizing the field's evolution from perception to alignment in a comprehensive survey.

The researcher's contribution centers on the development of Unimath, a foundational and multimodal mathematical reasoner introduced in 2023. This core work serves as the anchor for a broader research line that includes a 2026 survey examining the progression of multimodal mathematical reasoning from perception and alignment to reasoning.

This line of work appears to address the challenge of integrating multimodal inputs into robust mathematical reasoning systems. By first proposing a foundational model and later surveying the field's trajectory, the researcher demonstrates a sustained effort to define and structure this emerging domain, moving from specific architectural contributions to broader conceptual synthesis.

The significance of this work is evidenced by its uptake in the research community. The core paper has accumulated 45 citations, with 87.3% of the scholar's total citing papers originating from independent researchers. This high degree of independent citation suggests that the foundational framework has been widely adopted and built upon by the broader academic community, indicating substantial impact beyond the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 25 · **2 flagged influential by Semantic Scholar**

CORE PAPER

Unimath: A foundational and multimodal mathematical reasoner

2023 · EMNLP 2023, 7126-7133, 2023 · 45 citations (GS)

Field-normalised: 38 Semantic Scholar citations place it in the top 1% of Mathematics papers from 2023 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	A Comprehensive Survey of Scientific Large Language Models and Their Applications in Scientific Discovery	Texas A&M University, University of California, Los Angeles, University of Illinois at Urbana-Champaign	United States	—
2	MM-Verify: Enhancing Multimodal Reasoning with Chain-of-Thought Verification	—	—	—
3	Insight-V++: Towards Advanced Long-Chain Visual Reasoning with Multimodal Large Language Models	Nanyang Technological University, Tencent Hunyuan, Tsinghua University	China, Singapore	—
4	Insight-v: Exploring long-chain visual reasoning with multimodal large language models	Nanjing University, Nanyang Technological University, Tencent	China, Singapore	—
5	Mavis: Mathematical visual instruction tuning with an automatic data engine	ByteDance, CUHK, Oracle	China, Switzerland	—
6	Autoformalizing euclidean geometry	Caltech, University of Toronto	Canada, United States	—
7	Valley2: Exploring multimodal models with scalable vision-language design	ByteDance	China	—
8	Geox: Geometric problem solving through unified formalized vision-language pre-training	Fudan University, Shanghai AI Laboratory, Shanghai Artificial Intelligence Laboratory	China	—

No.	Citing paper	Citing institution(s)	Country	S2
9	A survey of mathematical reasoning in the era of multimodal large language model: Benchmark, method & challenges	Nanyang Technological University, Squirrel Ai Learning, The Hong Kong University of Science and Technology (Guangzhou)	China, Singapore	—
10	Eagle: Elevating geometric reasoning through llm-empowered visual instruction tuning	East China Normal University, Meituan Inc., Tsinghua University	Canada, China	—
11	Enhancing the geometric problem-solving ability of multimodal llms via symbolic-neural integration	iFLYTEK, University of Science and Technology of China	China	—
12	Solidgeo: Measuring multimodal spatial math reasoning in solid geometry	Chinese Academy of Sciences, TAL, University of Electronic Science and Technology of China	China	—
13	Engibench: A benchmark for evaluating large language models on engineering problem solving	Hong Kong Polytechnic University, Nanyang Technological University, Sofia University "St. Kliment Ohridski"	Bulgaria, China, Hong Kong	—
14	Open eyes, then reason: Fine-grained visual mathematical understanding in mllms	CSIRO, Georgia Institute of Technology, Nanjing University of Science and Technology	Australia, China, United Kingdom	—
15	Geocoder: Solving geometry problems by generating modular code through vision-language models	Google DeepMind, Polytechnique Montréal, Université de Montréal	Canada, United Kingdom	—
16	Geo-llava: A large multi-modal model for solving geometry math problems with meta in-context learning	Huawei	Singapore	—
17	Geodano: Geometric vlm with domain agnostic vision encoder	Australian National University, POSTECH	Australia, South Korea	Influential
18	Plane geometry problem solving with multi-modal reasoning: A survey	Australian National University, POSTECH	Australia, South Korea	—
19	Mathglm-vision: solving mathematical problems with multi-modal large language model	Beihang University, Tsinghua University, Zhipu.AI	China	—
20	Unlocking multimodal mathematical reasoning via process reward model	ByteDance, Ping An Technology Co., Ltd., Tsinghua University	China	—
21	Hologram reasoning for solving algebra problems with geometry diagrams	Central China Normal University	China	—
22	Theorem-validated reverse chain-of-thought problem generation for geometric reasoning	Baidu Inc, Huazhong University of Science and Technology	China	Influential
23	How Does a Virtual Agent Decide Where to Look? Symbolic Cognitive Reasoning for Embodied Head Rotation	Korea University, Kyung Hee University	South Korea	—

No.	Citing paper	Citing institution(s)	Country	S2
24	TrustGeoGen: Formal-Verified Data Engine for Trustworthy Multi-modal Geometric Problem Solving	Shanghai Artificial Intelligence Laboratory, Shanghai Jiao Tong University, The Chinese University of Hong Kong, Shenzhen	China	—
25	Enhancing Geometric Perception in VLMs via Translator-Guided Reinforcement Learning	Guangdong Laboratory of Artificial Intelligence and Digital Economy, Tsinghua University	China	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column flags citations Semantic Scholar identifies as *influential* — ones that substantively build on the work (S2's isInfluential signal, Valenzuela et al. 2015) — the "built on / relied upon" pattern the AAO credits. Counsel should quote the citing text for the strongest of these.

FOLLOW-UP WORK

A Survey of Multimodal Mathematical Reasoning: From Perception, Alignment to Reasoning

2026 · arXiv preprint arXiv:2603.08291, 2026 · 0 citations (GS)

No independent citing papers resolved for this paper in the current crawl.

Contribution 2

Claim — Contribution 2

The researcher established a multimodal benchmark for scientific college entrance exams and subsequently surveyed self-improvement mechanisms in multimodal large language models.

The researcher's contribution centers on advancing multimodal artificial intelligence for scientific education and model optimization. This line of work is anchored by the 2024 paper 'Scemqa: A scientific college entrance level multimodal question answering benchmark,' which appears to introduce a specialized evaluation framework for assessing model performance on complex, science-based entrance examination questions. The titles suggest a focus on bridging the gap between general multimodal capabilities and the rigorous demands of standardized scientific testing.

Originality in this trajectory is indicated by the progression from establishing a specific benchmark to examining broader model evolution. The follow-up 2025 paper, 'Self-Improvement in Multimodal Large Language Models: A Survey,' suggests the researcher expanded their scope to analyze how these models can autonomously enhance their capabilities. This chronological shift implies a comprehensive approach, moving from defining performance standards to investigating the mechanisms of continuous model improvement within the multimodal domain.

The significance of this work is evidenced by its reception in the academic community. The core benchmark paper has accumulated 42 citations, indicating it has become a recognized reference point for researchers in this niche. Furthermore, citation analysis reveals that 87.3% of the 102 classified citations originate from independent researchers, demonstrating that the work has influenced scholars outside the researcher's immediate institution and collaboration network. This high degree of independent uptake underscores the broad relevance and utility of the proposed benchmark and subsequent survey.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 18 · **1 flagged influential by Semantic Scholar**

CORE PAPER

Scemqa: A scientific college entrance level multimodal question answering benchmark

2024 · ACL 2024, 2024 · 42 citations (GS)

Field-normalised: 38 Semantic Scholar citations place it in the top 5% of Computer Science papers from 2024 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Vision-R1: Incentivizing Reasoning Capability in Multimodal Large Language Models	East China Normal University, Xiaohongshu Inc.	China	—
2	StructVRM: Aligning Multimodal Reasoning with Structured and Verifiable Reward Models	—	—	—
3	A survey on benchmarks of multimodal large language models	Peking University, Tencent	China	—
4	Vision language models are blind	Auburn University, University of Alberta	Canada, United States	—
5	A survey on evaluation of multimodal large language models	Nanyang Technological University	Singapore	—
6	Vrbench: A benchmark for multi-step reasoning in long narrative videos	Nanjing University, Shanghai Artificial Intelligence Laboratory, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences	China	—
7	Sciverse: Unveiling the knowledge comprehension and visual reasoning of lmms on multi-modal scientific problems	The Chinese University of Hong Kong	China	—
8	A survey on multimodal benchmarks: In the era of large ai models	Hong Kong University of Science and Technology, Zhejiang University	China, Hong Kong	—
9	A causality-aware paradigm for evaluating creativity of multimodal large language models	Harvard University, Singapore Management University, Sun Yat-sen University	China, Singapore, United States	**Influential**
10	Adacurl: Adaptive curriculum reinforcement learning with invalid sample mitigation and historical revisiting	Alibaba Group	China	—
11	Qwen look again: Guiding vision-language reasoning models to re-attention visual information	Baidu Inc., Peking University	China	—
12	Visco: Benchmarking fine-grained critique and correction towards self-improvement in visual reasoning	Stanford, University of California, Los Angeles	United States	—
13	Observe-r1: Unlocking reasoning abilities of mllms with dynamic progressive reinforcement learning	Zhejiang University	China	—
14	Argus inspection: do multimodal large language models possess the eye of panoptes?	Fudan University, Shanghai Artificial Intelligence Laboratory, Shanghai Jiao Tong University	China, Hong Kong	—
15	Eee-bench: A comprehensive multimodal electrical and electronics engineering benchmark	Boston University, Emory University, The University of Tokyo	Japan, United States	—

No.	Citing paper	Citing institution(s)	Country	S2
16	Towards llm agents for earth observation	Columbia University, Cornell University	United States	—
17	Qcbench: Evaluating large language models on domain-specific quantitative chemistry	Nanjing University, Shanghai AI Laboratory, Shanghai Artificial Intelligence Laboratory	China	—
18	Apo: Enhancing reasoning ability of mllms via asymmetric policy optimization	Zhejiang University	China	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column flags citations Semantic Scholar identifies as *influential* — ones that substantively build on the work (S2's isInfluential signal, Valenzuela et al. 2015) — the "built on / relied upon" pattern the AAO credits. Counsel should quote the citing text for the strongest of these.

FOLLOW-UP WORK

Self-Improvement in Multimodal Large Language Models: A Survey

2025 · EMNLP 2025, 2025 · 5 citations (GS)

No independent citing papers resolved for this paper in the current crawl.

Contribution 3

Claim — Contribution 3

The researcher established a benchmarking framework for evaluating foundation models on university-level physics problem solving, providing a critical standard for assessing AI capabilities in complex scientific reasoning.

CLAIM: The researcher's contribution centers on the 2025 paper 'PHYSICS: Benchmarking Foundation Models on University-Level Physics Problem Solving,' which appears to introduce a standardized method for assessing how well large language models handle advanced physics problems. This work serves as the foundational reference for this specific line of inquiry.

ORIGINALITY: The title suggests the researcher addressed a gap in evaluating AI performance on rigorous, university-level scientific tasks rather than simpler benchmarks. By focusing on physics problem solving, the work likely provided a novel dataset or evaluation protocol that challenges models to demonstrate deeper reasoning capabilities, distinguishing it from general-purpose language model assessments.

SIGNIFICANCE: The work has garnered 29 citations, indicating rapid uptake within the field. Notably, 87.3% of the 102 citing papers classified for this scholar originate from independent researchers, suggesting that the benchmark has been widely adopted by the broader scientific community as a reliable standard for evaluating foundation models in physics.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 21 **· 4 flagged influential by Semantic Scholar**

CORE PAPER

PHYSICS: Benchmarking Foundation Models on University-Level Physics Problem Solving

2025 · ACL 2025, 2025 · 29 citations (GS)

Field-normalised: 29 Semantic Scholar citations place it in the top 1% of Physics papers from 2025 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	VLMEvalKit: An Open-Source ToolKit for Evaluating Large Multi-Modality Models	Nanjing University, Southeast University, The Chinese University of Hong Kong	China	—

No.	Citing paper	Citing institution(s)	Country	S2
2	A Survey of Scientific Large Language Models: From Data Foundations to Agent Frontiers	Alibaba Group, Beijing Institute of Technology, Beijing Jiaotong University	Australia, China, Hong Kong	—
3	Sdar: A synergistic diffusion-autoregression paradigm for scalable sequence generation	Shanghai AI Laboratory, Shanghai Jiao Tong University, Tsinghua University	China, United States	—
4	Evaluating GPT-and reasoning-based large language models on Physics Olympiad problems: Surpassing human performance and implications for educational …	Free University of Berlin, Leibniz Institute for Science and Mathematics Education, Ludwigsburg University of Education	Germany	Influential
5	SeePhys: Does Seeing Help Thinking?-- Benchmarking Vision-Based Physics Reasoning	ETH Zurich, Huawei, Sun Yat-sen University	China, Hong Kong, Switzerland	Influential
6	Lexam: Benchmarking legal reasoning on 340 law exams	ETH Zurich, ETH Zurich, University of Lausanne, Max Planck Institute for Research on Collective Goods, Max Planck Institute for Research on Collective Goods	Germany, Switzerland	—
7	AtmosSci-Bench: evaluating the recent advance of large language model for atmospheric science	Hong Kong University of Science and Technology	Hong Kong	—
8	A survey on large language model benchmarks	Harbin Institute of Technology, Shenzhen, Institute of Software, Chinese Academy of Sciences, Shanghai AI Lab	China	—
9	MatSciBench: Benchmarking the Reasoning Ability of Large Language Models in Materials Science	Genentech/Roche, Princeton University, University of California, Los Angeles	United States	—
10	Reasoning over mathematical objects: on-policy reward modeling and test time aggregation	Carnegie Mellon University, Meta, Meta AI	United States	—
11	Classroom Final Exam: An Instructor-Tested Reasoning Benchmark	Analogy AI, Inc., Duke University, Northwestern University	United Kingdom, United States	—
12	QuantumQA: Enhancing Scientific Reasoning via Physics-Consistent Dataset and Verification-Aware Reinforcement Learning	Anhui University, Hefei Comprehensive National Science Center, National University of Singapore	China, Singapore	—
13	Fine-Tuning Small Reasoning Models for Quantum Field Theory	Perimeter Institute for Theoretical Physics, University of Wisconsin-Madison	Canada, United States	—
14	LLM Swiss Round: Aggregating Multi-Benchmark Performance via Competitive Swiss-System Dynamics	ByteDance, Carnegie Mellon University, Columbia University	China, United States	—
15	S2SServiceBench: A Multimodal Benchmark for Last-Mile S2S Climate Services	Beijing Normal University, Nanjing University of Infor-	China	—

No.	Citing paper	Citing institution(s)	Country	S2
		mation Science and Technology, The Hong Kong University of Science and Technology		
16	Evaluating NLP Embedding Models for Handling Science-Specific Symbolic Expressions in Student Texts	Leibniz Institute for Science and Mathematics Education, Leibniz University Hannover	Germany	—
17	LOCA: Logical Chain Augmentation for Scientific Corpus Cleaning	Peking University	China	**Influential**
18	DataChef: Cooking Up Optimal Data Recipes for LLM Adaptation via Reinforcement Learning	Fudan University, Shanghai AI Laboratory	China	—
19	S1-VL: Scientific Multimodal Reasoning Model with Thinking-with-Images	ScienceOne AI	—	**Influential**
20	Multi-Physics: A Comprehensive Benchmark for Multimodal LLMs Reasoning on Chinese Multi-Subject Physics Problems	The Chinese University of Hong Kong, Shenzhen	China	—
21	PhysUniBench: A Multi-Modal Physics Reasoning Benchmark at Undergraduate Level	Beihang University, Fudan University, Michigan State University	Australia, China, United States	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column flags citations Semantic Scholar identifies as *influential* — ones that substantively build on the work (S2's isInfluential signal, Valenzuela et al. 2015) — the "built on / relied upon" pattern the AAO credits. Counsel should quote the citing text for the strongest of these.

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Tsinghua University	China	SCImago #8 · THE 12 · QS =17	11
University of Notre Dame	United States	SCImago #1036 · THE 194 · QS =294	10
Zhejiang University	China	SCImago #6 · THE 39 · QS 49	8
Nanyang Technological University	Singapore	SCImago #137	6
The University of Hong Kong	Hong Kong	SCImago #195 · THE 33 · QS 11	6
Shanghai AI Laboratory	China	—	6
Shanghai Artificial Intelligence Laboratory	China	SCImago #563	6
Fudan University	China	SCImago #46 · THE 36 · QS 30	5
Peking University	China	SCImago #11 · THE 13 · QS 14	5
The Chinese University of Hong Kong	China	SCImago #163 · THE =41 · QS =32	5
Tencent	United States	—	5
University of Science and Technology of China	China	SCImago #77 · THE 51 · QS =132	5

Institution	Country	World ranking	Citing papers
Shanghai Jiao Tong University	China	SCImago #10 · THE 40 · QS =47	5
Alibaba Group	China	SCImago #226	4
Nanjing University	China	SCImago #178 · THE =62 · QS =103	4

Geographic distribution of citing authors

Country	Citing papers
China	61
United States	34
Singapore	11
Australia	8
Canada	8
Hong Kong	8
United Kingdom	4
South Korea	3
Germany	3
Switzerland	3
France	2
Denmark	2

Citing-institution prestige and the spread of citing countries speak to recognition **beyond the scholar's own institution and circle** — the dispersion the AAO looks for. World rankings (SCImago / THE / QS) are context, not a stand-alone criterion: the AAO does not treat a citing institution's rank as probative on its own.

F. AAO Precedent Considerations

Pre-filing self-check (AAO denial patterns)

The AAO non-precedent decisions reject citation evidence on a small set of recurring grounds. Confirm the petition addresses each before filing:

- ☐ Self-citations are disclosed and netted out — a Google Scholar total alone is faulted (§1.1).
- ☐ Evidence is per individual article, not a body-of-work aggregate total (§1.2).
- ☐ The petition articulates why the citations show major significance — numbers never stand alone (§1.5).
- ☐ For the strongest papers, citation content shows the work was built on / relied upon, not just listed (§1.6, §2.2).
- ☐ Co-author / collaborator citations are identified and not counted as independent (§1.7).
- ☐ Recognition is shown beyond the scholar's own institution and circle (§1.8).
- ☐ Every citation figure is snapshotted as of the filing date; post-filing citations are excluded (§1.9).
- ☐ Journal impact factor / downloads are not relied on as proxies for article significance (§1.10, §1.12).
- ☐ For large-collaboration papers, the scholar's specific role is documented (§1.13).
- ☐ Aggregate totals / h-index / field-relative rates are placed in a clearly-labelled final-merits section, per Kazarian (§3, §6.1.7).

Disclaimer

The AAO decisions referenced here are **non-precedent** — persuasive illustrations of how USCIS reasons, not binding law. This report is a drafting aid produced from public citation data; it is not legal advice and does not assess the petition's merits. All analysis must be reviewed by qualified immigration counsel.

G. Citation Evidence Index

Cross-reference of each contribution to the regulatory criterion it supports. Counsel should map these to the petition's exhibit numbers.

Contribution	Core paper	Indep. cites	Supports
Contribution 1	Unimath: A foundational and multimodal mathematical reasoner	25	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 2	Scemqa: A scientific college entrance level multimodal question answering benchmark	18	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 3	PHYSICS: Benchmarking Foundation Models on University-Level Physics Problem Solving	21	8 CFR 204.5(i)(3) — Outstanding Researcher