Yunyan Duan
https://yyd27.github.io
Updated February 2025
Email : yyduan.pku@gmail.com

Work Experience

Technische Universität Darmstadt Darmstadt, Hessen, Germany
Post-doctoral fellow & Research software engineer *Jul. 2022 – present*

- Conducted visual perception research using psychophysics methods to investigate how humans integrate multiple visual cues in visual processing. Utilized both behavioral experiments and computational modeling approaches to study the mechanism of vision.
- Managed and maintained lab hardware and software system to ensure the precise spatial and temporal display of visual stimuli. Provided technical support to lab users from diverse backgrounds and developed effective lab usage policies to ensure device availability and improve user experience.
- Taught undergraduate-level courses, and supervised Bachelor and Master's theses, mentoring students through research conceptualization and thesis writing.

Tencent Shenzhen, Guangdong, China
Data Scientist *Jul. 2021 – Jul. 2022*

- Designed, implemented, and monitored ETL pipelines in a data warehouse, supporting 10+ daily and real-time dashboards for WeChat's live streaming subsystem, which serves 1M+ DAUs.
- Conducted causal inference to develop explainable and automated solutions for evidence-based metric design in building a metrics monitoring and alerting system, validating findings through A/B testing.
- Regularly communicated data analysis insights to engineering and product teams via dashboards, reports, and documentation, supporting the development and iteration of the live streaming recommendation system.

Education

Northwestern University Evanston, IL, United States
Ph.D., ***Linguistics*** *2021*

- Dissertation: Word identification and eye movement control in reading as rational decision making
- Cognitive Science Specialization

Peking University Beijing, China
Bachelor of Science, ***Statistics & Psychology*** *(Double major)* *2013*

Online Learning

Coursera: Google IT Support Professional Certificate Sep. 2023

Coursera: Reinforcement Learning Specialization Mar. 2020

Coursera: Deep Learning Specialization Nov. 2018

Stanford Online: Mining Massive Datasets Jul. 2017

Academic Journal Publications

1. Chang, W., **Duan, Y.**, Qian, J., Wu, F., Jiang, X., & Zhou, X. (2020). Gender interference in processing Chinese compound reflexive: Evidence from reading eye-tracking. *Language, Cognition and Neuroscience*. 1-16.
2. **Duan, Y.**, & Bicknell, K. (2019). A rational model of word skipping in reading: Ideal integration of visual and linguistic information. In *Proceedings of the 41th Annual Conference of the Cognitive Science Society*: 275-281. ***Winner of best Computational Modeling paper in Perception & Action.***
3. **Duan, Y.**, & Bicknell, K. (2017). Refixations gather new visual information rationally. In *Proceedings of the 39th Annual Conference of the Cognitive Science Society*: 301-306.
4. Yu, H., **Duan, Y.**, & Zhou, X. (2017). Guilt in the eyes: Eye movement and physiological evidence for guilt-induced social avoidance. *Journal of Experimental Social Psychology, 71*, 128-137.
5. ***Duan, Y.**, & *Wu, O. (2016). Learning with auxiliary less-noisy labels. *IEEE Transactions on Neural Networks and Learning System, 28(7)*, 1716-1721. (* indicates equal contributions.)
6. Luo, Y., **Duan, Y.**, & Zhou, X. (2015). Processing rhythmic pattern during Chinese sentence reading: an eye movement study. *Frontiers in Psychology, 6*, 1881.
7. Wang, L., **Duan, Y.**, Theeuwes, J., & Zhou, X. (2014). Reward breaks through the inhibitory region around attentional focus. *Journal of Vision, 14(12):2*, 1-7.

Conference Presentations (excluding those with proceedings)

1. **Duan, Y.**, Mahncke, S., & Wallis, T. Combining surface reflectance and motion cues in peripheral target detection. Poster presentation at the 24th Annual Vision Sciences Society Meeting, St. Pete Beach, Florida, 17-22 May 2024.
2. Eicke-Kanani, L., **Duan, Y.**, & Wallis, T. From visual features of moving objects to subjective impressions of causality. Poster presentation at the 24th Annual Vision Sciences Society Meeting, St. Pete Beach, Florida, 17-22 May 2024.
3. Sheng, Y., **Duan, Y.**, & Wu, F. Corpus-based analysis of complement coercion in Mandarin Chinese. Poster presentation at the 25th annual conference on Architecture and Mechanisms for Language Processing (AMLaP), Moscow, Russia, 6-8 September 2019.
4. **Duan, Y.**, Berzak, Y., Bicknell, K., & Levy, R. Inferring sentence comprehension from eye movements in reading. Poster presentation at the 32nd annual CUNY Conference on Human Sentence Processing, University of Colorado, Boulder, Colorado, 29-31 March 2019.
5. **Duan, Y.**, & Bicknell, K. (2019). A rational model of word skipping in reading: ideal integration of visual and linguistic information. Poster presentation at the 32nd annual CUNY Conference on Human Sentence Processing, University of Colorado, Boulder, Colorado, 29-31 March 2019.
6. **Duan, Y.**, & Bicknell, K. (2016). Word identification in reading is constructive: Refixations seek new visual information. Poster presentation at the 22nd annual conference on Architecture and Mechanisms for Language Processing (AMLaP), Bilbao, Spain, 1-3 September 2016.
7. Hu, J., Liu, J., **Duan, Y.**, Zhao, C., Gong, X., Xiang, Y., Jiang, C., & Zhou, X. (2014). Resting-state functional connectivity indexes emotion recognition bias. Poster presentation at the 20th Annual Meeting of the Organization for Human Brain Mapping (OHBM), Hamburg, Germany, 8-12 June 2014.
8. **Duan, Y.**, Yu, H., & Zhou, X. (2014). Avoiding eyes reveals guilty heart: An eye movement study on interpersonal guilt. Poster presentation at the 6th Chinese International Conference on Eye Movements (CICEM), Beijing, China, 5-9 May 2014.
9. *Feng, W., ***Duan, Y.**, Luo, Y., & Zhou, X. (2013). When language hurts you: Aggression provoked by rhetorical questions. Poster presentation at the 1st Brain Research Symposium by PKU-IDG/McGovern Institute, Beijing, China, 20-21 August 2013.

Internship Experience

Tencent Guangzhou, China

Data Scientist (Internship) *Jul. – Aug. 2020*

- Evaluated and improved the quality of Mandarin speech-to-text subtitles to improve user engagement in an audiobook app
- Measured word error rate by annotating representative samples from a corpus of 1.8k documents with 8 million tokens, identifying a specific genre for targeted improvement through root cause analysis of word errors
- Implemented and evaluated text error correction pipelines integrating various components, including language models, neural networks, and NER models, selecting the best model based on evaluation results
- Incorporated the offline evaluation model into the subtitle feature, yielding reduced word error rate and ensuring satisfactory quality of subtitles

Google Los Angeles, CA

Analytical Linguist (Internship) *Jun. – Sep. 2019*

- Proposed and developed a feature to inspect user journeys as they navigated pilot online tasks/questionnaires on a crowd-sourcing platform, such that task designers could evaluate the quality of the results, identify ways to improve the task design, and decide whether to proceed with a full-scale deployment
- Analyzed logs from 100+ tasks, each consisting of 10-200 questions and responses from 100-500 raters, and extracted metrics (including clicks and time spent) that predicted result quality and flagged puzzling questions, in spite of high task variability
- Developed an R package and a Python module that implemented the feature, which automatically queried data from database, extracted predictive metrics, visualized typical user journey trajectories, and generated an analytical report for a standardized pilot task
- Presented analyses and findings to the team, delivering slides, code, and documentation to support further iteration

Research Projects

Rational models of eye movements in reading

Research Assistant at Department of Linguistics, Northwestern University, Evanston, IL *Sep. 2015 – Jun.2021*

- To understand how eye movement control and word identification work in reading, we use rational analysis with computational modeling to study how different sources of information (visual, linguistic, and contextual) act interactively to identify words and influence eye movement decisions.
- Techniques: *Bayesian inference, language models, reinforcement learning*

Inferring sentence comprehension from eye movements in reading

Visiting Student at Department of Brain and Cognitive Sciences, MIT, Cambridge, MA *Aug. 2018 – Dec. 2018*

- Although psycholinguistic experiments suggest a strong eye-mind link, the evidence is mostly based on aggregated data. We predict readers' comprehension of individual sentences from their eye movements using machine learning and neural networks, and evaluate how well the predictive models generalize to new readers and new sentences.
- Techniques: *Machine learning, convolutional neural network (CNN), language models, model selection*

Natural language processing for predicting readmission in pediatric ICU

Research Assistant at Feinberg School of Medicine, Northwestern University, Chicago, IL *Sep. 2016 – Aug. 2017*

- Useful information for predicting readmission after pediatric ICU hospitalization can come from electronic health records, especially social workers' notes. We formalize the idea by extracting text features from this unstructured data and implementing classifiers to predict readmission probability.
- Techniques: *Classification with unbalanced classes, natural language processing (bag-of-words model, sentiment analysis)*

Learning with auxiliary less-noisy labels

Research Intern at Institute of Automation, Chinese Academy of Sciences, Beijing, China *Apr. – Sep. 2014*

- We propose a learning method that considers both noisy labels and auxiliary less-noisy labels available in a small portion of the training data. Our method outperforms traditional classifiers that do not explicitly consider the auxiliary less-noisy labels.
- Techniques: *Matlab, Expectation–Maximization algorithm, logistic regression, crowd-sourcing*

Teaching

- **Experimentalpsychologisches Praktikum**: Bachelor-level practical course for psychology students

  Topics: Cue combination in continuous psychophysics tracking tasks (Winter 2024); Combination of color and motion cues in peripheral target detection (Winter 2023); Perceptual decision making and confidence report in tilt illusion (Winter 2022)
- **Experimental psychology lab project**: Bachelor-level practical course for cognitive science students

  Topics: Prior knowledge of color temperature and its influence on peripheral target detection (Summer 2024); Develop dead-leaves stimuli with a non-semantic proto-object defined by bottom-up visual cues (Summer 2023)
- **Cognitive Science I: Perception**: Bachelor-level seminar (Summer 2023 & 2024)

Skills

Programming language: Python (5+ years), R (5+ years), SQL (2+ years), MATLAB (1+ years), HTML/CSS/JavaScript (occasionally)
Data science: statistical inference, machine learning, causal inference
Natural language processing: Text analysis, sentiment analysis, neural network models
Other computer skills: Git, Linux/Unix, AWS (EC2)
Cognitive science: Eye-tracking, psycholinguistics, experiment design, psychophysics, computational cognitive models, decision making, scene perception
Statistics: Linear-mixed models, logistic regression, cluster analysis, Bayesian inference, etc.
Python Packages & tools: Scikit-learn, Numpy, Pandas, Stanford CoreNLP, Tensorflow
Natural language: English (C1), German (B1), Mandarin Chinese (C2)