After graduating Stanford University, I went to work as a software engineer in Silicon Valley. I quickly realized my passion was not for pure engineering, but for research. I was fortunate enough to be accepted to the Wharton School's PhD program in Statistics and there it was confirmed for me that research and teaching is my lifelong calling.

`GPT-4o` says my work "sits at the intersection of statistical methodology, software engineering, and applied social/biomedical science, with a strong commitment to reproducible open-source tools and educational impact". I would agree but further clarify: the common theme is innovating methods that solve real-world applied problems using statistics, mathematics and computer science. I then release their solutions via open-source software which I maintain continuously for a worldwide userbase. I believe the accessibility of my research explains why it is highly referenced: I have 4,200 citations on my 40+ publications, one patent, one manuscript in review and four nearing submission. I've published in Economics, Sociology, Linguistics, Educational Technology, Psychiatry, Opthamology, Oncology, Neurology, Chemistry and, of course, Statistics. I've received multiple grants, many of which are highly competitive. Herein, I'll outline three streams of research, their past (with a sample of published works cited) and their near future (in blue).

Experimental Design

My current focus is to understand the how the *degree of randomization* in randomized classic pill-placebo trials affects performance (joint with Abba M Krieger of the Wharton School of the University of Pennsylvania and David Azriel of the Technion). While a visiting scholar for a semester at the Technion six years ago, we were puzzled by two seemingly irreconcilable experimental design approaches: classicists say, "completely randomize subject assignments" but many today say, "use deterministic assignments from numerical optimization which minimizes covariate imbalance". We introduced a novel performance metric (risk measured by quantile of MSE) and demonstrated that the optimal design is somewhere in the middle of these two approaches without sacrificing the benefits of either. This translates to clinical trials running faster and cheaper with less risk of failure [1]. This result was so surprising that I was invited to present the research at five conferences, nine universities (including Harvard) and to Amazon, Inc.'s experimentation research team in Seattle, WA.

To dig deeper, we were awarded a prestigious US-Israel Binational Science Foundation grant (who only fund statistics proposals once every other year) for $167,000 through 2025. On this grant, we have explored many fruitful extensions. We proved that nearly optimal designs are almost as random as complete randomization releasing the `R` package `GreedyExperimentalDesign` which has ≈30K downloads [2], we optimized the rerandomization design (which dates to Student in the 1920's) [3], we explored the relationship between degree of randomization and power [4], we improved the pairwise matching design by further minimizing covariate imbalance [5], we showed that pairwise matching is optimal for incidence response [6] and recently showed that pairwise matching is optimal for general response types, i.e., incidence, count, proportion, and uncensored survival [7]. We are currently working on optimality properties of Fisher's blocking designs and showing that pairwise matching provides optimal power in the Cochran-Mantel Haenszel test.

A parallel line of work researches improvements to sequential (online) experimental designs. This began with my work in graduate school which matched subjects together on-the-fly while leaving idiosyncratic subjects in a "reservoir". At the end of the study, both types of subjects contribute to an estimate of the treatment effect. And this is shown to provide much more accurate estimation and higher power for continuous response [8]. A few years ago, we improved many of the shortcomings of this design by removing the need for strong assumptions and further allowed for optimized matching when some of the response values are known in advance [9]*. I am currently working on generalizing this latter work for all general response types, i.e., incidence, count, proportion and censored survival and upgrading `SeqExpMatch`, its `C++`-optimized `R` package which has ≈10K downloads.

Machine Learning and its Applications

My other main research stream is in machine learning / artificial intelligence (AI). Herein lies my most cited work on "explainable AI". This work tries to provide insight into the inner workings of all popular complicated black box models [10]. This publication, and its corresponding `R` package `ICEbox` has 1800+ citations, ≈90K downloads and was recently ported to `Python` as it is widely used in industry. In academic publications its methodology has been used to understand models for, inter alia, biomedical image recognition, driverless cars, forest populations, concrete foam compressive strength and the genomics of maize.

Modeling a phenomenon of interest with numeric measurements is considered by many to be *the* fundamental problem in science and it is a problem that deeply motivates me as well. My `R` package `bartMachine` has 450+ citations and ≈300K downloads [11]. It has been used by scientists to predict a wide variety of phenomena, inter alia, severity of depression, election results, storm outages in electric grids, appropriate dosing for pharmaceuticals, human fetal growth, lifespans of European aristocracy. And, of course, I "eat my own dog food"; using `bartMachine`, I teamed up with an opthamological surgeon in Wisconsin to build a model that optimizes the location of lenses in cataract surgery. Its accuracy is near optimal, i.e., at the lens manufacturers' tolerance [12]. Additionally, we were interested to understand how these models work by understanding and statistically testing covariates for usefulness [13], a work that is widely used and thus has 200+ citations.

In this stream, I currently switched to the newly-ubiquitous *deep learning*. I am using deep learning to classify aesthetic value of a wide variety of digital art genres.

Interdisciplinary Research

Working with passionate scholars in other fields is (I believe) the most rewarding part of being an applied statistician. I've always been playing in everyone's backyard.

In college, I developed software called `GemIdent` that used Random Forests (where features are built from various radii of rings of colored pixel scores) on microscopic images to localize different types cell nuclei in histological samples [14]. The output of this software was then used to understand how breast cancer spreads in lymph nodes [15].

Before graduate school, I worked on a vocabulary-learning program for high-school students by showing contextual snippet examples. I teamed up with reading specialists to create the website dictionarysquared which was awarded a $1.5M grant. We proved that this method is pedagogically effective [16] and then we also developed a Random-Forest-based model to identify the best contextual snippets [17]. I'm currently revamping this model to use deep learning's large language attention models. Preliminary results are astonishingly accurate.

In graduate school, I teamed up with a psychologist who studies depression, specifically why some succeed in therapy and some succeed with drugs. We developed methods for personalized treatments and wrapped it in the `R` package `PTE` [18] which has ≈35K downloads. This work was then used by a company called Aifred Health who validated it in clinical studies [19] and were rewarded a patent for the method of personalized treatment [20]. Using `PTE` to discover personalization models via analysis of new clinical datasets is ongoing.

As an assistant professor, I was once casually chatting with William Blanford, an Arizona-trained hydrologist, and realized estimating the slope-divided-by-the-intercept in OLS is fundamental to the measuring in Henry's Law constants. Working out the mathematics, we found an optimal design for the independent variable distribution [21]. We then released this methodology in our `R` package `optDesignSlope` which has ≈25K downloads.

Teaming up my friend from Dana Chandler, an MIT-trained economist, we experimentally demonstrated that making gig work more meaningful leads to more worker retention and people working longer [22]. This was one of the first empirical results on this topic (and thus became a canonical to cite having ≈600 and winning the journal's award for most cited work). Further, this study was one of the first to use Amazon's Mechanical Turk (MTurk) as a platform to run natural field experiments and was likely instrumental in its popularization.

Designing custom MTurk experiments using `Ruby on Rails` is an expertise of mine. Teaming up with Dana Weinberg, a Harvard-trained Sociologist, we used a natural field experiment to test gender discrimination in book publishing [23]. We then followed up this pilot study to test gender, age and race discrimination across a wide variety of book genres [24]*. This latter study may be the largest Internet-based social science natural field experiment in history with nearly 10,000 subjects assessed on almost 30,000 book covers.

Dana Weinberg and I are currently working together on a large Air Force Reseach Laboratory grant with a budget of $1.9M. We are co-investigating topics of disinformation and radicalization using another custom MTurk experiment.

Concluding Words

I believe to have demonstrated that research is what makes me tick. I would be honored to spend the rest of my productive life through publishing and disseminating original and novel science. And I am also excited to continue my passion for interdisciplinary research by teaming up with faculty from different departments.

Select Cited Works

[1] Adam Kapelner, Abba M Krieger, Michael Sklar, Uri Shalit, and David Azriel. Harmonizing optimized designs with classic randomization in experiments. *The American Statistician*, 75(2):195–206, 2021.

[2] Abba M Krieger, David Azriel, and Adam Kapelner. Nearly random designs with greatly improved balance. *Biometrika*, 106(3):695–701, 2019.

[3] Adam Kapelner, Abba M Krieger, Michael Sklar, and David Azriel. Optimal rerandomization designs via a criterion that provides insurance against failed experiments. *JSPI*, 219:63–84, 2022.

[4] Abba M Krieger, David Azriel, Michael Sklar, and Adam Kapelner. Design choices in randomization tests that affect power. *Communications in Statistics-Theory and Methods*, 53(9):3276–3291, 2024.

[5] Abba M Krieger, David A Azriel, and Adam Kapelner. Better experimental design by hybridizing binary matching with imbalance optimization. *Canadian Journal of Statistics*, 51(1):275–292, 2023.

[6] Adam Kapelner, Abba M Krieger, and David Azriel. The role of pairwise matching in experimental design for an incidence outcome. *Australian & New Zealand Journal of Statistics*, 65(4):379–393, 2023.

[7] Adam Kapelner, Abba M Krieger, and David Azriel. The pairwise matching design is optimal under extreme noise and extreme assignments. *Stat*, 14(2):e70058, 2025.

[8] Adam Kapelner and Abba Krieger. Matching on-the-fly: Sequential allocation with higher power and efficiency. *Biometrics*, 70(2):378–388, 2014.

[9] Adam Kapelner and Abba Krieger. A matching procedure for sequential experiments that iteratively learns which covariates improve power. *Biometrics*, 2020.

[10] Alex Goldstein, Adam Kapelner, Justin Bleich, and Emil Pitkin. Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *JCGS*, 24(1):44–65, 2015.

[11] Adam Kapelner and Justin Bleich. bartmachine: Machine learning with bayesian additive regression trees. *Journal of Statistical Software*, 70(4), 2016.

[12] Gerald P Clarke and Adam Kapelner. The bayesian additive regression trees formula for safe machine learning-based intraocular lens predictions. *Frontiers in big data*, 3:572134, 2020.

[13] Justin Bleich, Adam Kapelner, Edward I. George, and Shane T. Jensen. Variable selection for BART: An application to gene regulation. *The Annals of Applied Statistics*, 8(3):1750 – 1781, 2014.

[14] Susan Holmes, Adam Kapelner, and Peter P Lee. An interactive java statistical image segmentation system: Gemident. *Journal of Statistical Software*, 30(10), 2009.

[15] Andrew Y Chang, Nupur Bhattacharya, Jian Mu, A Francesca Setiadi, Valeria Carcamo-Cavazos, Gerald H Lee, Diana L Simons, Sina Yadegarynia, Kaveh Hemati, Adam Kapelner, et al. Spatial organization of dendritic cells within tumor draining lymph nodes impacts clinical outcome in breast cancer patients. *Journal of translational medicine*, 11:1–12, 2013.

[16] Suzanne M Adlof, Lauren S Baron, Joanna Scoggins, Adam Kapelner, Margaret G McKeown, Charles A Perfetti, Elaine Miller, Jeanine Soterwood, and Yaacov Petscher. Accelerating adolescent vocabulary growth: Development of an individualized, web-based, vocabulary instruction program. *Language, speech, and hearing services in schools*, 50(4):579–595, 2019.

[17] Adam Kapelner, Jeanine Soterwood, Shalev Nessaiver, and Suzanne Adlof. Predicting contextual informativeness for vocabulary learning. *IEEE Transactions on Learning Technologies*, 11(1):13–26, 2018.

[18] Adam Kapelner, Justin Bleich, Alina Levine, Zachary D Cohen, Robert J DeRubeis, and Richard Berk. Evaluating the effectiveness of personalized medicine with software. *Frontiers in Big Data*, 4:8, 2021.

[19] Akiva Kleinerman, Ariel Rosenfeld, David Benrimoh, Robert Fratila, Caitrin Armstrong, Joseph Mehltretter, Eliyahu Shneider, Amit Yaniv-Rosenfeld, Jordan Karp, Charles F Reynolds, et al. Treatment selection using prototyping in latent-space with application to depression treatment. *PloS one*, 16(11):e0258400, 2021.

[20] Caitrin Armstrong, David Benrimoh, Robert Fratila, Adam Kapelner, Akiva Kleinerman, Joseph Mehltretter, and Ariel Rosenfeld. Systems and methods for treatment selection, July 9 2024. US Patent 12,033,756.

[21] Adam Kapelner, Abba Krieger, and William J Blanford. Optimal experimental designs for estimating henry's law constants via the method of phase ratio variation. *Journal of Chromatography A*, 1468:183–191, 2016.

[22] Dana Chandler and Adam Kapelner. Breaking monotony with meaning: Motivation in crowdsourcing markets. *Journal of Economic Behavior & Organization*, 90:123–133, 2013.

[23] Adam Kapelner and Dana Weinberg. Do readers judge books by author gender? results from a randomized experiment. *Socius*, 5:2378023119868893, 2019.

[24] Dana B Weinberg and Adam Kapelner. Do book consumers discriminate against black, female, or young authors? *Plos one*, 17(6):e0267537, 2022.