Alyssa Hwang

My Experience in Research and Motivation for Grad School

People often ask me why I pursue research or want to go to grad school. This is my best (but still in-comprehensive) attempt at unravelling four long years of thinking, learning, and growing.

When I first stepped foot onto Columbia's campus in the fall of 2016, I knew one thing: I wanted to pursue research. Research in *what*? I didn't know yet. I just had a strong feeling that research would give me the greatest opportunities to strengthen my skills and learn from established professionals in my future field. From the beginning, research fueled my studies in computer science and became a significant part of my undergraduate experience. I sat in front of my computer with earnest and Googled *Columbia Engineering*. I made a list of all the majors available to me and scratched off the ones I didn't like. From there, I looked at every research lab available to each of those majors and eventually zeroed in on the *Columbia Natural Language Text Processing Lab.* "Our research combines linguistic insights into the phenomena of interest with rigorous, cutting edge methods in machine learning and other computational approaches," the website proclaimed. The projects displayed on the homepage ranged from automatic news colaters to understanding stories to generating emails that encourage eco-friendly habits. It was a fascinating combination of computer science and linguistics. It was complex problem-solving centered around humans. It was the one lab I wanted to join. I didn't have a concrete plan, but I had an instinctual feeling that this was the place I needed to be. I followed it.

I started off by cold-emailing the director of the NLP Lab. Professors, I have come to learn, have eternally overflowing inboxes. When I didn't receive a response from the original email or a followup, I asked my advisor for help. I was accepted to Columbia as an Egleston Scholar, a special scholars program for the "top 1% of Columbia Engineering applicants," and it has been one of the greatest privileges and advantages I have been blessed with throughout college. I don't know what made me part of this pretty 1%; I don't consider myself some kind of genius. I don't understand most (any?) complex subjects the first time around, I can't compute large calculations in my head, and I don't remember every single fact I've ever learned. Eventually, I would realize that there is kind of no such thing as genius. An expert in *anything* is just someone with adequate time, practice, and resources. Some of us start off with more resources, but all of us need to put in the time to practice. This is an important lesson I learned throughout my undergraduate experience. As an Egleston Scholar, I started off with incredible resources—the advisor of the program contacted the professor and helped introduce me. Soon, I was set up with a PhD student as a mentor, and I was excited to get started.

For those without this sort of resource, I want to share another important resource: knowledge. At the time, I thought that asking a professor was my only way to get started in

research. If I hadn't had extra help, I would have been stuck. One great tip I can give all of you is to try reaching out to PhD students. You will probably be directly mentored by a PhD student anyway, and PhD students have more time and attention to spare for replying to emails. There are multiple ways in to research. Sometimes, you have to keep trying until something works out.

Pursuing research is a beast. More specifically, pursuing research is like wrestling an invisible monster. I spent the first two years of my research life deeply scared: I was scared I didn't know enough, scared I was wasting my mentor's time and resources, scared that I wasn't smart enough to even try, scared that I would be kicked out at any second. I couldn't understand how everyone around me knew *so much*. I was stuck in a static mindset. I, for some reason, believed that there were two types of people: Smart Enough and Not Smart Enough. Smart Enough people just know what to do. They don't get confused. They don't struggle. Whatever they try just works. I was confused. I struggled a lot (I barely knew how to *print("hello world")* in Python or open my command line). Most everything I tried broke. I was Not Smart Enough and the Smart Enough people were far away from me, and I was loath to bother them with my hopeless questions.

Sound familiar? You might call it imposter syndrome or fixed mindset or pessimism or anxiety. Whatever *it* is, it turns out that plenty of people feel this way. A lot of the time, no person, place, or thing *makes* you feel this way; you just *do.* It happens because of a lack of connection to other people. It's easy to assume no one else struggles when all you see is their results out of context from their journeys. It's easy to isolate yourself because you feel out of place as a newcomer. Because my experience was the only one I truly saw all of, I started to believe that the City of Not Smart Enough had exactly one citizen: me. I went through all of this and figured a lot out the hard way, so I would love to give you a shortcut: there is no fundamental, biological, permanent difference between Smart Enough and Not Smart Enough. The fairer distinction would be Know Enough and Not Know Enough or Practiced Enough and Not Practiced Enough. No one pops into the world knowing how to solve computer science (which is what I somehow believed at first); they read and learn and listen and struggle through it. Everyone is capable of learning. Everyone is capable of growing. I see that now. I want everyone to see that. With enough time, practice, and resources, *anyone* has a place in computer science. That's why I care so much about sharing knowledge. I started school off with a special resource—the Egleston Scholars Program—and now I want to be a resource for everyone else. This is one of my main sources of motivation for grad school: I want disseminate knowledge and mentor young computer scientists as a professor of computer science.

The other half of my motivation comes from my research, of course. A PhD is a degree in research. I wouldn't be able to even consider a PhD if I didn't have some commitment to my research, even if I treated it as a means to an end to be a professor. Grad school represents dedicated time to expanding knowledge in a specific area I care about, and this has immense personal value to me. It's worth more than the opportunity cost of not going straight to industry, but I didn't find this particular niche right away. I started off in research clinging to my mentor's ideas. He was interested in persuasive speech—figuring out what makes a successful argument, predicting the success of a persuasive text, generating opinionated text,

and so on—so I began by annotating and analyzing data for him. All of our data was a collection of conversations from the Change My View Subreddit, and I ended up annotating and analyzing the different types of argumentation styles, cited sources, and other textual details. Along the way, my mentor submitted a workshop paper with me as a co-author because I helped design the annotation schema—my first publication! I submitted to undergraduate research conferences and presented my work—I thought my work was insignificant and disappointing at the time, but now I'm proud of how far I've come since I started. I can remember the exact turning point in my research, the moment I transitioned from a scared, codependent research assistant to a less scared, more independent researcher.

Some time in the beginning of junior year, my mentor was explaining his latest project to me. It was about collecting examples of summaries. For example, people will often summarize their opinion by saying, "tldr, <insert summary here>." They might also say "imo," "imho," or "tbh." The idea was to take the sentence that comes after this short abbreviation and use it to summarize an entire post or comment.

"Did you consider idioms?" I asked offhandedly. My mentor immediately perked up.

"Are you interested in idioms?" He replied. My mentor usually has a gentle nature, but his question was almost a demand. For some reason, he sounded very excited about my potential interest. I didn't know what to make of it.

"Maybe," I answered nervously. "I'm open to anything." This was my secret way of saying, *Eep! I don't even know what I'm interested in yet!* At this point, I was still blindly pursuing research because some instinct felt right, but I still didn't have a concrete motivation. Yet.

"I'll send you a paper," he promised. He did. It changed my career.

An idiom is a non-compositional multiword expression. In regular-person-speak, an idiom is a special collection of words that form a phrase whose meaning cannot be guessed based on the meanings of the individual words it is made up of. *It's raining cats and dogs,* for example, has nothing to do with animals. *Break a leg* basically means the opposite of what it says. Much of NLP research has transitioned to throwing text at a black box (another idiom!) machine learning algorithm and waiting for results to be chucked back at you. I was fascinated by idioms because they are anchored in linguistics. I want to understand language. I want to understand how language represents people. Once again, I was drawn to the human-centered approach to computer science research.

Throughout the semester, I kept working on my own idioms project. An idiom is defined as non-compositional for semantics; the *meaning* of an idiom cannot be predicted from the meanings of individual words. I wanted to show that this non-compositionality extended to *sentiment*. I had results in the middle of my junior spring and discovered the ACL Joint Workshop on Multiword Expressions and WordNet the same day. It was the perfect alignment of research I had done purely because I found it interesting. I had to submit something. The deadline was a week away. Oh my god.

I asked my parents a million times if I should submit to this workshop. They said a million and one times: yes, go for it, you should. In truth, I wanted them to say no. I wanted them to give me an excuse to not bother because it's impossible to fail if you don't even try. The imposter syndrome/fixed mindset/pessimism/anxiety I thought I had gotten over started to

creep back in. *Just do it*, my parents urged me. Fine. It's time to jump at all the opportunities I have in front of me. I locked myself in my room. I taped papers to the wall. I went absolutely crazy. I wrote an entire workshop short paper in 36 hours. I was accepted. I would be presenting my work in Florence, Italy. *O mio dio.*

This whole experience entrenched my commitment to computer science research. I was suddenly exposed to all the parts that slip through the cracks of NLP research. Idioms are one of them—I now assume that my mentor was so excited about my interest in idioms because not many people study them, which leaves the field wide open for me to play around with. I soon became interested in slang and the evolution language on social media—how can we teach machines to understand language as fast as we change it? From there I started learning about non-standard dialects of English, specifically African American (Vernacular) English. As NLP researchers, we love clean, predictable sources of data, but these datapoints don't reflect all English speakers. We need too include different dialects, slang, and ways of speaking. We make up weird new slang words everyday and don't have great ways of automatically learning their meanings; this is what I wrote my undergraduate thesis about. I am excited to pursue human-centered research in sociolinguistics and computer science.

I am grateful for the opportunities that Columbia has given me and excited for the opportunities that will come my way at UPenn. Throughout my time at Columbia, I have learned several important lessons. First, I need to make the most of every opportunity I am given, and sometimes I have to search them out. Who knows how my life would be different if I hadn't written that paper? What if I had given up on research the first hundred times I struggled? Second, sometimes, I need to trust my instincts. Some strong feeling drew me to research. Some strong feeling urged me to stay for four school years and two summers (I was a software engineering intern at Google during my third summer, but that deserves a separate story). I didn't have a clear future goal when I started in freshman year; I didn't really have a clear future goal when I was applying to grad school either. I realized *after* I submitted my applications that my PhD would allow me to have high impact as a professor; I nearly burst into tears of appreciation when my thesis advisor casually mentioned that I have a "good personality" for teaching (I love teaching. I TAed for a summer and three semesters. I am already dreaming up a specific philosophy for teaching, but that also deserves a separate story). Third, some of the most important things that separate beginners and experts are knowledge and resources. I hope I can be a source of both.