Understanding Post-Baccalaureate Cultural Gaps: Building Equitable Ecosystems for AI Research and What We Can Learn from Federal TRIO Programs

Anonymous Author(s)
Affiliation
Address
email

Abstract

This paper aims to survey the problem space around cultural barriers in research collaboration, specifically for Machine Learning (ML). We review (1) unequal representation in ML/AI and STEM, (2) socioeconomic influences on retention of scientists and researchers, and (3) existing educational opportunity programs for people from disadvantaged backgrounds, with emphasis on Post-Baccalaureate support. We provide evidence that students from disadvantaged backgrounds not only experience barriers to gaining intellectual and technical expertise, they often experience cultural gaps that impede their participation in graduate programs and inclusion in research collaborations. We discuss relevant research on culture differences and the ways that some U.S. Federal TRIO programs explicitly address them, highlighting standardization as one means of demystifying academic and research cultures. We conclude with recommendations toward understanding post-education culture gaps with the goal of finding better solutions for increasing diversity in research collaborations.

1 Introduction

> Jonathan's first day at Wheaton, he looked up his course syllabi and panicked. He couldn't afford the books. He also did not tell anyone he couldn't afford the books, he just never got them. [Joffe-Walt and Glass, 2015]

1.1 Identifying culture gaps from early education to early career scientists

We survey literature on culture gaps that we believe is relevant in designing interventions for ML career success. The inequalities that inhibit research collaboration range from a multifaceted history of research from historically unequal resource distribution [Seth, 2009, Roy, 2018] to the inequitable access to technology education today [Muro et al., 2018, Bell et al., 2019]. Additionally, studies have shown that skill gaps do not fully account for the lack of representation in AI and STEM, as evidenced by the attrition of people from underrepresented and disadvantaged communities after working in research and engineering fields [Palmer et al., 2011]. Researchers have found that before any skill gap is a culture gap that is challenging to short-circuit. Below are results from a cultural identity study of incoming freshmen at a private university, where students attended an hour-long student discussion about adjusting to college. The study reviewed the first year performance of first-generation students in correlation to the type of discussion they attended at the beginning of the year [Stephens et al., 2014].

> One group attended a session in which panelists talked about their social class background, and how it affected their transition to college. This was called "difference

> education." Another group attended a session in which social class backgrounds were not highlighted. Among the students who were in the standard session that did not highlight social class, first-generation students had significantly worse GPAs. But among those who were in the difference education sessions, first-generation students had pretty much the same GPA as continuing generation students. [Chang, 2018]

STEM disciplines, like ML and AI, have field-specific challenges as well [Freire et al., 2021]. In addition to (skill-based) coding and mathematical literacy, there are additional cultural barriers to entry such as lack of early exposure, limited understanding of CS careers, and minimal professional and peer representation [Vachovsky et al., 2016]. A majority of the efforts to increase diversity focuses on K through college; however, attrition indicates retention problems further along the pipeline, which not only impedes representation, but overall advancement of a field as well [Hunt et al., 2020, Hofstra et al., 2020].

1.2 How Federal TRIO Programs decode the cultural barriers in science

Much of the work in overcoming the social exclusivity of science is focused on early education through college, with less work focused on graduate education and early career development or the support that leads to a successfully connected researcher in a field like AI. Some of the longest standing U.S. federally funded educational opportunity programs are classified as TRIO, which only programmatically supports undergraduate college students towards pursuing Masters and PhDs.

> Due to structural inequity, White students, those whose parents attained a university degree, and those from upper/middle class households are more likely to attain a doctoral degree. One federal program, the Ronald E. McNair Post-baccalaureate Achievement Program, provides undergraduates with academic and financial support to help marginalized students enroll and succeed in graduate school. However, little research has examined how this program has helped students attain the ultimate goal of a PhD. [Renbarger et al., 2021]

In addition to research funding, programs like McNair provide both skill and cultural gap training [Gittens, 2013]. Essential skills for entering post-baccalaureate studies may include being able to perform well in standardized testing like the GREs, understanding the foundations of research, and public speaking. However, McNair programs also provide opportunities to overcome cultural barriers through programs that helped students communicate and understand their research advisors, engage in dinner-time debates and discussions around current events, and go on graduate school visits, after helping these scholars purchase their first business suit [Fifolt et al., 2014]. Understandably, these culture gaps persist beyond PhD programs, and as researcher careers advance, so does the elitism [Abbink and Harris, 2019]. While a skills-wise proficiency of a field should advance, cultural elitism proves to be a nontrivial and less-acknowledged hindrance in broadening research collaboration.

1.3 What we can learn about cultural gaps from the 5 paragraph essay

A better researched example of cultural gaps can be found through studies on learning English as a second language [Giridharan and Robson, 2011]. Educators are often torn by the cost-benefit of teaching to a standard 5-paragraph position structure [Smith, 2006]. On the one hand, it can make writing formulaic and dry [Brannon et al., 2008]; however, such formulas, whether explicitly outlined or internalized from speaking English as a first language, give traction to those learning the English language later in life [Kos and Sims, 2014]. While "teaching to the test" may be frowned upon, having a standard framework not only provides a proxy for evaluating the equitability of an ecosystem, but also helps individuals identify cultural gaps on all sides of the divides.

We see standardization in tech through efforts poured into training students to have the skillsets to pass technical interviews [Griffin et al., 2022]; however, in tandem, research has also shown that "economic connectedness" generates the motivation that automatically overcomes skill gaps [Rosalsky and Woods, 2022]. We've observed that while skills enable early-career scientists and engineers into STEM research, overcoming the culture gap is required for retention, belonging, and an equitably collaborative ecosystem. Furthermore, identifying standards (or structures) for cultural understanding brings forth accessibility to what was only implicitly understood.

2 Discussion

Once students arrive at their graduate programs, they are often expected to have "made it" [Duncan and Murnane, 2011, Bailey and Dynarski, 2011]. Standardization of educational degrees poses a convenience when measuring impact through student retention versus the impact of later careers. For a program like McNair, we can measure a program's success by how many of them go on to earn PhDs [Renbarger et al., 2021]. Post-education, the success measures are less pronounced, which widens cultural divisions. The following subsections describe three qualities and allow for measurable impact towards increased participation in top research communities.

2.1 Better measures of success

The problems outlined in the previous sections have complexities and nuances that can be challenging to solve for. To better understand the problems we are solving, we pose the following structural questions:

- What does it mean to be a successful ML researcher?
- What are the practices required to be a successful scientist in ML?
- What types of collaborators exist that benefit ML research?
- What are minimum requirements for various types of research collaborations in ML?

As we saw with the 5-paragraph essay, rules and laws are approximations of how we want humans to cooperate. Some people are able to internalize these frameworks from a young age and benefit from that their entire career [Bailey and Dynarski, 2011, Morgan et al., 2022]. Less-affluent backgrounds (more so) benefit from a continued outline of standards at every career stage, making explicit what are implicit barriers of entry to collaboration [Baykut et al., 2022, Koutsouris et al., 2021].

2.2 Support during PhD programs and beyond

While it is helpful to make explicit the collaboration barriers, what does that support look like? There are examples from broader programs, like the National Center for Faculty Development and Diversity and school specific programs like Stanford's First-Generation and/or Low-Income Graduate Student Office. For funding, there are various grants like the Gates Millennium Scholarship Program, NSF GAANN, and GEM fellowships, which often come with additional career support [DesJardins and McCall, 2014].

Specific to ML, there are organizations and communities designed to support those not typically represented in ML research circles, like Women in ML and Black in AI. With increase of AI-specific residencies at some of the leading research groups (mostly in industry), programs are also emerging that enlist AI researchers from less-traditional backgrounds [Hooker, 2022]. In a similar vein to new movements in open research collaborations [Willen, 2018], researchers are affiliating more with organizations like ML Collective.

2.3 Addressing the bi-directional culture gap between the in-group and out-groups

Better representation is yet another quality that would broaden research collaborations. The burden of responsibility, however, is not solely for the underrepresented to bridge. In fact, it may potentially be more effective for the in-group to fill their gaps of the out-groups (and not the other way around). However, from what we've surveyed, programs typically expect the out-groups to overcome the gaps programatically through:

1. Inspiring scientists with global marketing campaigns and summer activities
2. Discovering scientists with internships and fellowships
3. Cultivating scientists with scholar and support programs
4. Retaining scientists with advocacy and allyship communities

Commonly seen in early-education programs, being connected to successful ML role models of underrepresented backgrounds bridges the cultural gaps that increasingly discourage the advancement of students into early career and beyond. In absence of role models, underrepresented students have a harder time imagining themselves working in careers like Machine Learning, creating a cyclical challenge for programs trying to discover talent in the first place, much less cultivate it [Cruz, 2015].

To improve the lifecycle of talent listed above, we can (1) retain more representation (of underrepresented researchers) which enable self-conceptual career pathways for prospective students, and (2) continue to identify and fill the cultural/skill gaps, building bridges from both sides of the disparity; in hopes that this, in turn, may create a curb cut effect [Blackwell, 2017] for collaboration efficacy, providing benefits from inspiration to retention, reducing attrition, and potentially benefiting those along all sides of the divides.

3 Future Work

Applying what we learn from programs like McNair, there are four areas where we can enable more equitable research interactions through post-education support.

(1) **Demystify Academic and Research pathways with standards.** Programs like McNair rely on metrics to understand trajectories and progress [Ishiyama and Hopkins, 2003]. In order identify the barriers that impeded research collaboration, explicit frameworks for ML research efficacy can broaden pathways to success. Specifically, finding standards for every stage of an ML researcher's career, as well as for every type of collaboration, can enable a more open research ecosystem. For example, having clearer ML research standards can inform collaboration standards, which will inform standards for technical and soft skills, and ultimately, cultural standards [Posselt and Black, 2012]. Furthermore, studies show that without clear rubrics, bias and stereotypes create unnecessary barriers to entry [Bertrand and Mullainathan, 2004, Latu et al., 2015].

(2) **More social studies on cultural gaps.** It follows that standards and metrics are, at best, projections of a preferred research environment. At worst, bias and discrimination can be baked into the structures (and models) we establish. Deeper understanding around cultural discrepancies enable programs like McNair to fill the gaps through activities as direct as upper-class etiquette training [Grimmett et al., 1998, Posselt and Black, 2012]. Finding solutions like difference education [Chang, 2018], help bring qualitative understanding to culturally complex dynamics like research . Particularly, studies to uncover the rules of the in-group help identify bias and create more equitable inroads for collaboration.

(3) **Reversing the "culture-fit" criteria.** Rather than imposing culture through assimilation, finding cultural similarities as we work towards better representation is a means of broadening research collaboration. Clearer measures like racial underrepresentation in tech have incentivised companies to meet students of color where they are at, rather than imposing their own culture-fit as a gatekeeper [Ryce, 2022, Phelan, 2022]. While programs like McNair aim to help students catch-up, a more shared ownership of cultural gaps can redirect efforts towards propelling the out-groups beyond the status quo, broadening the pool of collaborators [Hofstra et al., 2020, Yang et al., 2022].

(4) **Interdisciplinary exposure through domain-impact foci.** A by-product of the McNair program design is the interdisciplinary nature of the scholarly cohorts being uplifted. Outside of these multi-disciplinary peer groups, scientists often move towards more focused, and as a result, more homogeneous collaborators. Two approaches to broadening expertise is through explicitly interdisciplinary spaces like Media Lab, or through domain-specific impact where a specific problem, like Addiction Medicine (asam.org/), or shared passions, like ML for Creativity and Design, draw together scholarly communities. Connecting research to broader outcomes expands research communities through broadening the aspirations for ML advancement [Vachovsky et al., 2016].

4 Conclusion

Cultural and skill gaps can create a perceived permanent exclusion from participating in leading scientific communities and, while education programs increase representation for early-career scientists, these trends suggest that additional post-educational support and standardization would improve the retention, belonging, and advancement ecosystem for researcher collaboration in AI. We propose four approaches towards filling culture gaps to enable better skill acquisition and a more equitable collaboration ecosystem for ML, making explicit what's implicitly understood from within.

References

K. Abbink and D. Harris. In-group favouritism and out-group discrimination in naturally occurring groups. *PloS one*, 14(9):e0221616, 2019.

M. J. Bailey and S. M. Dynarski. Gains and gaps: Changing inequality in us college entry and completion. Technical report, National Bureau of Economic Research, 2011.

S. Baykut, C. Erbil, M. Ozbilgin, R. Kamasak, and S. H. Bağlama. The impact of the hidden curriculum on international students in the context of a country with a toxic triangle of diversity. *The Curriculum Journal*, 33(2):156–177, 2022.

A. Bell, R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen. Who becomes an inventor in america? the importance of exposure to innovation. *The Quarterly Journal of Economics*, 134(2):647–713, 2019.

M. Bertrand and S. Mullainathan. Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American economic review*, 94(4):991–1013, 2004.

A. G. Blackwell. The curb-cut effect (ssir), 2017. URL `https://ssir.org/articles/entry/the_curb_cut_effect`.

L. Brannon, J. P. Courtney, C. P. Urbanski, S. V. Woodward, J. M. Reynolds, A. E. Iannone, K. D. Haag, K. Mach, L. A. Manship, and M. Kendrick. Ej extra: The five-paragraph essay and the deficit model of education. *The English Journal*, 98(2):16–21, 2008.

A. Chang. The subtle ways colleges discriminate against poor students, explained with a cartoon. *September*, 12:2018, 2018.

I. Cruz. Reimagining the ronald e. mcnair scholars program through the lens of intellectual entrepreneurship. *Planning for Higher Education*, 43(2):33–39, 2015.

S. L. DesJardins and B. P. McCall. The impact of the gates millennium scholars program on college and post-college related choices of high ability, low-income minority students. *Economics of Education Review*, 38:124–138, 2014.

G. J. Duncan and R. J. Murnane. *Whither opportunity?: Rising inequality, schools, and children's life chances*. Russell Sage Foundation, 2011.

M. Fifolt, J. Engler, and G. Abbott. Bridging stem professions for mcnair scholars through faculty mentoring and academic preparation. *College and University*, 89(3):24, 2014.

A. Freire, L. Porcaro, and E. Gómez. Measuring diversity of artificial intelligence conferences. In *Artificial Intelligence Diversity, Belonging, Equity, and Inclusion*, pages 39–50. PMLR, 2021.

B. Giridharan and A. Robson. Identifying gaps in academic writing of esl students. In *Enhancing Learning: Teaching and learning conference 2011 proceedings*. Enhancing Learning: Teaching and Learning Conference 2011, Curtin University ..., 2011.

C. B. Gittens. The mcnair scholars program as an agent of socialization in the doctoral experience. *ProQuest LLC*, 2013.

J. Griffin, L. Burge, and S. Goldman. Innovative courses that broaden awareness of cs careers and prepare students for technical interviews. 2022.

M. A. Grimmett, J. R. Bliss, D. M. Davis, and L. Ray. Assessing federal trio mcnair program participants' expectations and satisfaction with project services: A preliminary study. *Journal of Negro Education*, pages 404–415, 1998.

B. Hofstra, V. V. Kulkarni, S. Munoz-Najar Galvez, B. He, D. Jurafsky, and D. A. McFarland. The diversity–innovation paradox in science. *Proceedings of the National Academy of Sciences*, 117 (17):9284–9291, 2020.

S. Hooker. Introducing the cohere for ai scholars program: Your research journey starts here, Sep 2022. URL https://txt.cohere.ai/.

V. Hunt, S. Prince, S. Dixon-Fyle, and K. Dolan. Diversity wins. Technical report, McKinsey, 2020.

J. T. Ishiyama and V. M. Hopkins. Assessing the impact of a graduate school preparation program on first-generation, low-income college students at a public liberal arts university. *Journal of College Student Retention: Research, Theory & Practice*, 4(4):393–405, 2003.

C. Joffe-Walt and I. Glass. This american life 550: Three miles, Mar 2015. URL https://www.thisamericanlife.org/550/three-miles.

B. A. Kos and E. Sims. Infographics: The new 5-paragraph essay. *Rocky Mountain Celebration of Women in Computing*, 23, 2014.

G. Koutsouris, A. Mountford-Zimdars, and K. Dingwall. The 'ideal'higher education student: understanding the hidden curriculum to enable institutional change. *Research in Post-Compulsory Education*, 26(2):131–147, 2021.

I. M. Latu, M. S. Mast, and T. L. Stewart. Gender biases in (inter) action: The role of interviewers' and applicants' implicit and explicit stereotypes in predicting women's job interview outcomes. *Psychology of Women Quarterly*, 39(4):539–552, 2015.

A. C. Morgan, N. LaBerge, D. B. Larremore, M. Galesic, J. E. Brand, and A. Clauset. Socioeconomic roots of academic faculty. *Nature Human Behaviour*, pages 1–9, 2022.

M. Muro, A. Berube, and J. Whiton. Black and hispanic underrepresentation in tech: It's time to change the equation. *The Brookings Institution*, 2018.

R. T. Palmer, D. C. Maramba, and T. E. Dancy II. A qualitative investigation of factors promoting the retention and persistence of students of color in stem. *Journal of Negro Education*, 80(4):491–504, 2011.

D. Phelan. Apple's tim cook launches new initiative: 'it's about values', Sep 2022. URL https://www.forbes.com/sites/davidphelan/2022/09/26/apples-tim-cook-launches-new-initiative-its-about-values/?sh=6070f1622a50.

J. R. Posselt and K. R. Black. Developing the research identities and aspirations of first-generation college students: Evidence from the mcnair scholars program. *International Journal for Researcher Development*, 2012.

R. Renbarger, T. Talbert, and T. Saxon. Doctoral degree attainment from ronald e. mcnair scholars program alumni: An explanatory embedded case study. *Educational Policy*, page 08959048211042569, 2021.

G. Rosalsky and D. Woods. The secret to upward mobility: Friends, Aug 2022. URL https://www.npr.org/2022/08/08/1116398427/the-secret-to-upward-mobility-friends.

R. D. Roy. Science still bears the fingerprints of colonialism. *Smithsonian Magazine*, 365:366, 2018.

W. Ryce. Monterey jazz festival and csumb showcase top jazz students in concert, Sep 2022. URL https://csumb.edu/news/news-listing/monterey-jazz-festival-and-csumb-showcase-top-jazz-students-in-concert/.

S. Seth. Putting knowledge in its place: science, colonialism, and the postcolonial. *Postcolonial studies*, 12(4):373–388, 2009.

K. Smith. In defense of the five-paragraph essay. *English Journal*, 95(4):16, 2006.

N. M. Stephens, M. G. Hamedani, and M. Destin. Closing the social-class achievement gap: A difference-education intervention improves first-generation students' academic performance and all students' college transition. *Psychological science*, 25(4):943–953, 2014.

M. E. Vachovsky, G. Wu, S. Chaturapruek, O. Russakovsky, R. Sommer, and L. Fei-Fei. Toward more gender diversity in cs through an artificial intelligence summer program for high school girls. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, pages 303–308, 2016.

R. Willen. 'new academia' - a safe harbour for researchers who love science, Jul 2018. URL `https://igdore.medium.com/new-academia-a-safe-harbour-for-researchers-who-love-science-c4baa87c1ebe`.

Y. Yang, T. Y. Tian, T. K. Woodruff, B. F. Jones, and B. Uzzi. Gender-diverse teams produce more novel and higher-impact scientific ideas. *Proceedings of the National Academy of Sciences*, 119 (36):e2200841119, 2022.

A Appendix

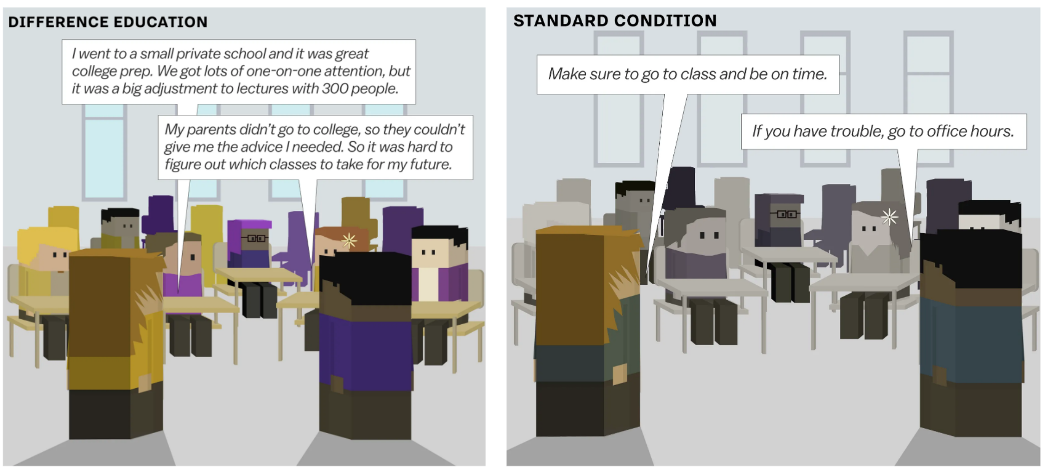

Figure 1: Vox article cartoon citing a study on difference education [Chang, 2018]

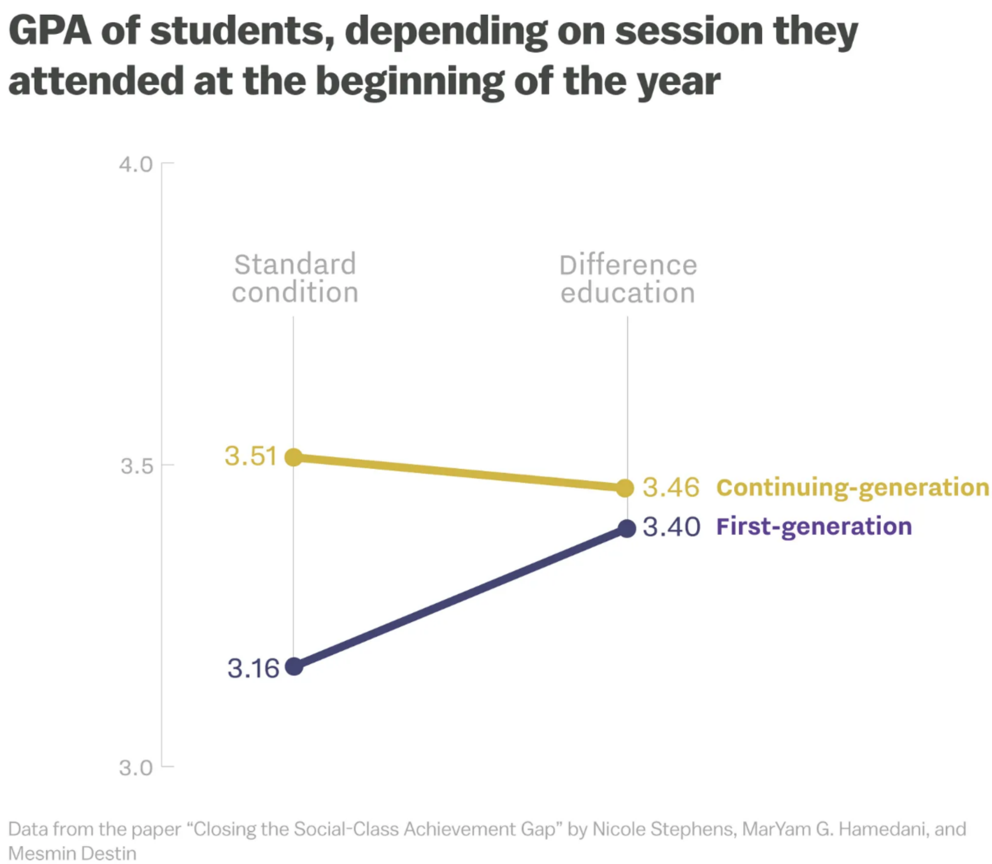

Figure 2: Difference education study results [Chang, 2018]

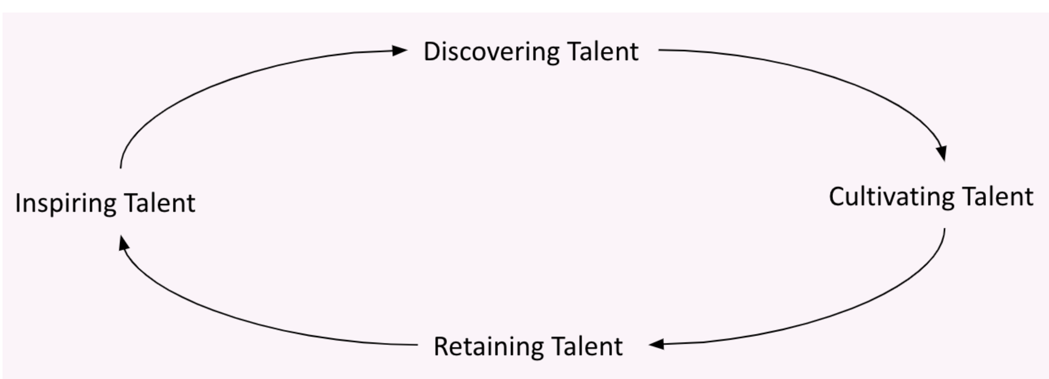

Figure 3: Visualization of a life-cycle for talent from Section 2.3