

## DISCIPLINARY APPROPRIATION AT THE BEGINNING OF A STATISTICS MAJOR

Florian Berens<sup>1</sup>, Kelly Findley<sup>2</sup>, Nicola Justice<sup>3</sup> and Christopher Kinson<sup>2</sup>

<sup>1</sup>University of Tübingen, Germany

<sup>2</sup>University of Illinois at Urbana-Champaign, United States

<sup>3</sup>Pacific Lutheran University, United States

florian.berens@uni-tuebingen.de

*Demand for skillsets in data analysis and computing has been rising quickly in recent years, but academic programs that prepare students for these positions still struggle with retention issues. In addition, statistics graduates are not representative of all parts of society, with women and people of color, among others, being underrepresented. We therefore look at those who enter a statistics major to understand how they navigate their program and find belonging. For the analysis of incoming undergraduates, we are guided by Levrini et al. (2015), who propose to look at identity development through the lens of disciplinary appropriation. Using the example of three students, we show that the operationalization developed by Levrini et al. (2015) is suitable for examining disciplinary appropriation at the beginning of studies. We present the operationalization and discuss how we made it usable in a domain- and target group-specific way.*

### INTRODUCTION

Many facets of statistical work require creativity, curiosity, and personal judgment to be exercised by the analyst (e.g., Bailyn, 1977; Wild et al., 2018). Statistical analyses are therefore not independent of the statistical analyst, but the analyst plays a substantial role in the statistical process and for the insights gained. The statistical analyst makes decisions in data collection, processing, analysis, and interpretation that influence statistical results and their reception. Therefore, it matters for statistical practice who the statistical analyst is. From this, in turn, it can be deduced that it does matter who becomes a statistical analyst.

However, we know from many STEM fields that the group of people who take up such professions is far from representative of the rest of society (Fry et al., 2021). In almost all societies of the world, women are underrepresented in these areas. The same is true in the Western World for many people of color, who have been historically excluded from needed resources and struggle to find belonging in STEM spaces (Grossman & Porche, 2014; Rainey et al., 2018). Furthermore, students from underrepresented groups who enter STEM degrees have higher attrition rates, and those who do persist often finish with lower grades (National Center for Science and Engineering Statistics, 2023; Whitcomb & Singh, 2021).

However, the problem of lack of representativeness of those who choose STEM professions is only one component to the fundamental problem that fewer people complete a STEM program than are in demand. This problem can also be seen to be composed of too few people choosing STEM programs and higher dropout rates compared to other programs (National Center for Science and Engineering Statistics, 2023; Sithole et al., 2017). STEM disciplines therefore face the challenge of having to attract more young people and to retain them in the field more successfully than before. In this context, the low representation of women and minorities is also an opportunity for the future, as it offers the possibility of achieving higher overall numbers by addressing these groups of people in particular. If one wants to develop such an approach, the questions about the reasons for the lower participation of these groups of people and about the possible successful strategies to win them over are interlinked.

One approach to address this linkage is to look at role models and identity formation. Looking at reasons for low participation in STEM programs, Kricorian and colleagues (2020) show that a lack of role models, especially for women and people of color, is a barrier toward entering a STEM program. To this end, Singer and colleagues (2020) show that young people's identity formation is a major factor that makes young people without role models less likely to enter STEM programs than those with role models within a STEM profession.

Based on these findings, it can be inferred that the role of a STEM discipline in identity formation is an important factor in determining which program of study an individual chooses and how likely he or she is to remain in that field of study. Therefore, it is important to understand how the

process of identity formation plays out in relation to the role of a potential field of study. This paper addresses this challenge by using the example of three case studies of students at the beginning of a statistics major in an attempt to understand how they have already made statistics a part of their identity at the beginning of their studies. Insights into this process could help better guide future students along the path of this identity formation, thus attracting students to statistics and retaining them in statistics.

In this regard, statistics seems to be a suitable discipline for such an investigation because, like other STEM disciplines, it is not very diverse (Lorah & Valdivia, 2021) and is also perceived as homogeneously male and white by young people (Taasoobshirazi et al., 2022).

In statistics education research, it has already been registered in the past two decades that statistics must not be understood as a rigid object, but that its reception is relevant for its visibility as well as its application. For example, Rolka and Bulmer (2005) and Bond et al. (2012) have reported insights into how statistics is perceived by students. More recently, Justice et al. (2020) and Findley and Berens (2020) have developed more advanced conceptualizations in which different belief systems about statistics are conceptualized. What all these studies have in common is that statistics is not understood as a fixed system, but as a domain that is subjectively perceived and understood. At the same time, however, all studies stop at the individual perception of statistics and do not take into account how personal factors and developments shape one's view of statistics.

In this context, Levrini et al. (2015) suggest connecting students' identity development to their personal perception of a discipline. Only then, they argue, can the complex interactions in learners' engagement with a discipline be explored and how learners gradually appropriate the discipline in this process be understood.

#### DISCIPLINARY APPROPRIATION

Following the modern discourse on identity, Levrini et al. (2015) first state that “the self can no longer be perceived as something that is assigned or given in modern societies but rather is something one has to choose and develop through a reflexive construction of one's own personal story” (Levrini et al., 2015, p. 95). This is initially good news for statistics, as it would mean that in an identity that is not fixed, the role of statistics is not fixed either, but evolves in discourse and reflection. Disciplines such as statistics thus have the opportunity to actively influence what role statistics plays in learners' identities by shaping discourse with learners. At the same time, the discursive nature between identity formation and statistics provides an entry point for research in which the process of identity formation becomes observable in interaction with statistics and in reflections on statistics.

In this regard, Levrini et al. (2015) emphasize that this view of identity development involves an important shift in perspective:

*Within science education research, productive science learning usually refers, more or less implicitly, to the ability to participate in scientific discourse. In contrast, our perspective emphasizes that learning is seen as productive if it includes a transformation of scientific discourse that allows science learning to contribute to students' self-identities. (Levrini et al., 2015, p. 96)*

Taking this perspective, the need arises to make visible these contributions to student identity formation in order to determine the extent to which learners have appropriated a discipline such as statistics for themselves. In this context, Levrini et al. (2015) define disciplinary appropriation as follows:

*Appropriation is a complex and reflexive process of transforming scientific discourse (scientific words and utterances) so as to embody it in one's own personal story, respecting disciplinary rules and constraints. The process of transformation involves one populating scientific discourse with one's own intentions, idiosyncratic tastes, and purposes in order to make it sensible not only for oneself but also with respect to one's way of participating in the social context of the class. (Levrini et al., 2015, p. 99)*

One challenge of this definition is its low specificity and its difficult operationalizability. In their research, Levrini et al. (2015) therefore use qualitative material on physics lessons from an extended intervention on thermodynamics in an Italian secondary school class (grade 12) to develop five markers by which disciplinary appropriation can be recognized. These markers function both as a refinement of the definition of disciplinary appropriation and as a tool for operationalizing the definition. The markers can be briefly described as follows:

1. The salience of an idiosyncratic signature idea of the individual student. For example, this makes itself apparent in words or phrases that an individual repeatedly uses to talk about the discipline, while other students do not use them. From these word choices and from the way of talking about the discipline, an individual view of the discipline is established.
2. Disciplinary grounding. The student uses his or her idiosyncratic idea to absorb knowledge provided to him or her and to build his or her own body of knowledge about the discipline.
3. Signs of depth and thickness in thinking about the discipline. In this context, metacognitive reflections on the relevance of the discipline (for the individual) and epistemological reflections are understood as thick thoughts.
4. Carrier of social relations. A sign of disciplinary appropriation here is when students define their own role in the discipline or use the discipline to define a role for themselves as a person in a social structure, for example, wanting to be a consultant for the discipline in a heterogenous team.
5. Repetition of signs of the first four markers across different forms of qualitative data or at multiple points in the same material.

Disciplinary appropriation is then identified by Levrini et al. (2015) in the synopsis of all markers. However, not all markers are to be regarded as equally important and independent. The idiosyncratic idea from the first marker is the basis for the disciplinary grounding in marker two and the formation of social relations in marker four. It thus has a special role. The repetitions from marker five are rather to be understood as general quality assurance, similar to the cross-validation in other methodological approaches. Only the thickness of thinking about the discipline as a marker stands somewhat on its own.

## METHODOLOGY

In this exploratory study, we conducted qualitative case studies to understand the disciplinary appropriation of statistics by first-year students majoring in statistics. In Fall 2022, six students responded to our interview call, which was sent to all incoming first-year statistics students at a large university in the Midwestern United States. We present three of the six who responded to the interview here, one identifying as female, one as male and one as non-binary. We do not attempt to generalize these results to a larger population. We consider the three students presented as what Creswell and Poth (2016) describe as instrumental cases, on which the phenomenon of disciplinary appropriation can be observed.

Between one and two of the authors jointly conducted semi-structured interviews with each student. Similar to the Draw a Scientist Test (Chambers, 1983), we asked, "Who is statistics?" and asked students to draw a picture personifying statistics and to explain their drawings (Malaspina, 2018). We also asked students to share characteristics needed to be a successful statistician, their past experiences with courses or projects in statistics, and to describe their motivation for studying statistics. The semi-structured format allowed for a conversational style where we could follow up on important ideas and encourage students to elaborate further. Interview transcripts and students' drawings formed the primary data sources for the study. Secondary data were the researcher's audio conversations of impressions after the interviews and written notes recorded during data collection and analysis.

The analysis of the data was conducted in five steps. In the first step, four researchers from the field of statistics education conducted independent coding of the interview transcripts, coding both In Vivo and the occurrence of the five markers. As a second step, the researchers met to discuss their codes of the markers. In the third step, the results were aligned with the In Vivo codes. In the fourth step, case characteristics were developed from these on a case-by-case basis to describe and embed the disciplinary appropriation of statistics. In the final fifth step, the utility of the markers to the overall

process was discussed and evaluated. This way, it will be presented here how the five markers presented by Levrini et al. (2015) can be used to investigate disciplinary appropriation of statistics at the beginning of the undergraduate studies.

## RESULTS

As results of our analyses, we first present salient characteristics of the three cases before analyzing the utility of Levrini et al.'s (2015) markers.

### *Sami*

Sami was drawn to statistics through their affinity to mathematics and their positive experiences in Advanced Placement (AP) Statistics. They describe statistics as a discipline that might let them combine their mathematical strengths with real world issues. While Sami did acknowledge the uncertainty that comes with making claims from statistics, their epistemology for statistics centers on using correct procedures. They emphasize the importance of doing one's work carefully, checking that data or calculations were done correctly, and keeping one's work neat and organized. We see Sami as a "Data Inspector" who sees themself managing details to ensure things are done correctly. They self-describe as a "Type A" person who feels satisfaction when working through a math problem and finding the answer. At the same time, they want to feel that they are contributing to a larger mission and that they are making a positive difference. This was reflected in repeated words and phrases like "love", "advocacy," "education," and "options." But when talking more about their experiences with statistics thus far, they use terms like "memorization," "procedures," and "detail-oriented." We found Sami's descriptions of statistics to have shadows of disciplinary authenticity, but without a clear personal signature for how these pieces connected to their passion for advocacy work. They recognize how statistics can be useful in application, but they struggle to articulate how statistical work itself has the power to solve real-world problems.

### *Liam*

Liam is an aspiring baseball analyst, and this aspiration inspires his personal signature idea for statistics—to find hidden patterns in data that inform strategic decisions. His knowledge of statistics and mathematics were encouraged by his high school teachers in AP Statistics and Calculus. He shows an understanding of the discipline that goes beyond introductory - describing the work of statisticians and data scientists through experimental design and research as well as writing reports based on observational studies noting that statisticians and data scientists are "analyzing and interpreting data sets...like studies and experiments...analyzing other ones that have happened and seeing if they could draw their new interpretations from data that was already found." Liam also acknowledges the role of programming and coding within the discipline, stating "I could see working in different coding languages and working on technology to create new tools to help data be analyzed easier." Common repeated words and phrases included "communication," "analytics," and "cooperation." He emphasizes communication and patience. We describe Liam's intended social role as a "Data Interpreter." He takes information and finds meaning in the "hidden numbers." He seems to grasp the metacognitive when reflecting on how the ideas of analytics within baseball extend beyond just that sport as he weakly connects the epistemological dots, "I think that you could apply that knowledge to, maybe not even baseball...and find different ways to use that sort of analysis." He certainly is directed toward sports statistics and analytics, where the truth is in the data, but can vary depending on the interpretation and perspective.

### *Priya*

Priya described rich experiences in a summer internship and watching weekly open international "causal inference seminars" that drew her into a statistics community that she sees as diverse and welcoming. Her statistics person is "open to new foods, new people" and seems to match her keen interest in considering new ideas. Priya sees herself as a novice member of this welcoming community where members can ask questions and consider new ideas. Paired with her relationship-centered perspective, her signature identity is a vision of statistics as a versatile field where she can ask questions and work with a diverse community to answer those questions. Thus, we see Priya's social role in statistics as a "co-investigator." Common repeated words and phrases related to

relationships, work-life balance, and taking time for mentoring/helping others as well as notions of passion, ideas, and interest helped paint her identity as a novice learner in the statistics community. While Priya's experiences exploring alongside other statisticians were rich and grounded in disciplinary notions of collaboration, open-mindedness, curiosity, and the important role of coding, our group wondered whether her enthusiasm for statistics is specific to statistics, or reflects enthusiasm for scientific discovery, more generally. We also did not yet see evidence of thick metacognitive or epistemological dimensions that might support a more established identity as a statistician.

#### *The five markers as operationalization of appropriation of statistics*

Due to the high importance of the idiosyncratic signature idea for disciplinary appropriation in Levrini et al. (2015), it is indeed helpful to start the analysis of the data with this marker. Also, it may explain why our group's discussions tended to frequently gravitate (back) to the signature idea. Following Levrini's analyses, looking for frequently repeated words and phrases proved to be a helpful first approach in our work to identify the idea. Our experience shows that matching the idea contained in these words with, for example, reported experiences and depictions about the "statistics person" reveals a high degree of congruence. Also, the comparison of the initial codes of the four coders showed high agreement. Nevertheless, extensive discussions were necessary to transform shared impressions into communicable expressions. However, once such a version of the idiosyncratic idea was achieved, even this one marker gave a good impression of how the student was potentially appropriating statistics into their identity.

The search for disciplinary groundings in the interviewees' responses also proved clear in our analytical discussions. More difficult to see, however, was how disciplinary grounding relates to the individual's idiosyncratic idea. For example, Sami's perspective of statistics was rooted in more procedural activities of inspecting data and ensuring accuracy, which seemed disconnected from her passion for advocacy and education. We theorize that Sami does not yet have a personal signature idea for statistics that is grounded in the discipline..

For some cases, it was straightforward to identify how students viewed statistics as a carrier of their social role,, while in others, it was quite difficult to tease apart., Priya, for example, hinged her signature idea of statistics as a community rich with research opportunity. This naturally extended to her positioning as a Data Co-investigator. In contrast, Sami's positioning herself as a Data Inspector seemed at odds with her desire to question norms and effect positive change for real-world issues. However, we recognize the limitations of drawing inferences from only one interview, and it likely will take more time and data to see how each student's perspective and sense of belonging in statistics materializes.

The most difficult to identify was the thickness of thinking about statistics. We find very limited metacognition or epistemological reasoning in our data, and what evidence for these markers we did find tended to be implicit. However, this may also be due to the fact that our interviewees are still at the very beginning of their studies and answers gain thickness at later points in time. For these early interviews, however, we have found the marker to be of little use so far in detecting disciplinary appropriation.

We considered the marker of repeated occurrence of findings across different points in the material primarily within our matching of markers to In Vivo codes. Fortunately, this showed that the findings found were mostly consistent across multiple sites. We interpret this as a positive sign for the handling of the markers and for their ability to represent disciplinary appropriation.

## CONCLUSION

The above-mentioned reflections on the use of the markers suggest that the markers are suitable for identifying and describing disciplinary appropriation in students already at the beginning of their studies. They help to elucidate students' evolving disciplinary identity and their early sense of belonging in the field. What role the markers can play in observing identity development will be explored in further analyses of interviews at later time points. We hope to be able to trace the path of identity formation within the study of statistics across their full Bachelor's degree programs and to identify successful pathways into the discipline.

## REFERENCES

- Bailyn, L. (1977). Research as a cognitive process: Implications for data analysis. *Quality and Quantity* 11, 97–117. <https://doi.org/10.1007/BF00151906>
- Bond, M. E., Perkins, S. N., & Ramirez, C. (2012). Students' Perceptions of Statistics: An Exploration of Attitudes, Conceptualizations, and Content Knowledge of Statistics. *Statistics Education Research Journal*, 11(2), 6-25. <https://doi.org/10.52041/serj.v11i2.325>
- Chambers, D. W. (1983). Stereotypic Images of the scientist: The draw-a-scientist test. *Science Education*, 67(2), 255–265. <https://doi.org/10.1002/sce.3730670213>.
- Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Sage.
- Findley, K. P., & Berens, F. (2020). Assessing the disciplinary perspectives of introductory statistics students. *Proceedings of the 23rd Annual Conference on Research in Undergraduate Mathematics Education* (pp. 1090–1095). Boston, MA. <http://sigmaa.maa.org/rume/RUME23.pdf>
- Fry, R., Kennedy, B., & Funk, C. (2021). STEM jobs see uneven progress in increasing gender, racial and ethnic diversity: Higher education pipeline suggests long path ahead for increasing diversity, especially in fields like computing and engineering. Pew Research Center, 1-28.
- Grossman, J. M., & Porche, M. V. (2014). Perceived gender and racial/ethnic barriers to STEM success. *Urban Education*, 49(6), 698-727. <https://doi.org/10.1177/0042085913481364>
- Justice, N., Morris, S., Henry, V., & Fry, E. B. (2020). Paint-by-number or Picasso? A grounded theory phenomenographical study of students' conceptions of statistics. *Statistics Education Research Journal*, 19(2), 76–102. <https://doi.org/10.52041/serj.v19i2.111>
- Kricorian, K., Seu, M., Lopez, D., Ureta, E., & Equils, O. (2020). Factors influencing participation of underrepresented students in STEM fields: matched mentors and mindsets. *International Journal of STEM Education* 7(16). <https://doi.org/10.1186/s40594-020-00219-2>
- Levrini, O., Fantini, P., Tasquier, G., Pecori, B., & Levin, M. (2015). Defining and operationalizing appropriation for science learning. *Journal of the Learning Sciences*, 24(1), 93-136. <https://doi.org/10.1080/10508406.2014.928215>
- Lorah, J. A., & Valdivia, M. (2021). Diversity in Statistics Education at Postsecondary Institutions. *International Journal of Research in Undergraduate Mathematics Education*, 7(1), 21–32. <https://doi.org/10.1007/s40753-020-00120-x>
- National Center for Science and Engineering Statistics (NCSES) (2023). *Diversity and STEM: Women, Minorities, and Persons with Disabilities 2023. Special Report NSF 23-315*. Alexandria, VA: National Science Foundation. <https://nces.nsf.gov/wmpd>.
- Rainey, K., Dancy, M., Mickelson, R., Stearns, E., & Moller, S. (2018). Race and gender differences in how sense of belonging influences decisions to major in STEM. *International Journal of STEM Education*, 5, 1-14. <https://doi.org/10.1186/s40594-018-0115-6>
- Rolka, K., & Bulmer, M. (2005). Picturing student beliefs in statistics. *ZDM Mathematics Education*, 37(5), 412-417. <https://doi.org/10.1007/s11858-005-0030-4>
- Singer, A., Montgomery, G., & Schmoll, S. (2020). How to foster the formation of STEM identity: studying diversity in an authentic learning environment. *International Journal of STEM Education*, 7(57), 1-12. <https://doi.org/10.1186/s40594-020-00254-z>
- Sithole, A., Chiyaka, E. T., McCarthy, P., Mupinga, D. M., Bucklein, B. K., & Kibirige, J. (2017). Student Attraction, Persistence and Retention in STEM Programs: Successes and Continuing Challenges. *Higher Education Studies*, 7(1), 46-59. <http://dx.doi.org/10.5539/hes.v7n1p46>
- Taasobshirazi, G., Wagner, M., Brown A., & Copeland, C. (2022) An Evaluation of College Students' Perceptions of Statisticians. *Journal of Statistics and Data Science Education*, 30(2), 138-153. <https://doi.org/10.1080/26939169.2022.2058655>
- Whitcomb, K. M., & Singh, C. (2021). Underrepresented minority students receive lower grades and have higher rates of attrition across STEM disciplines: A sign of inequity? *International Journal of Science Education*, 43(7), 1054-1089. <https://doi.org/10.1080/09500693.2021.1900623>
- Wild, C.J., Utts, J.M., & Horton, N.J. (2018). What Is Statistics?. In: Ben-Zvi, D., Makar, K., Garfield, J. (eds) *International Handbook of Research in Statistics Education*. Springer International Handbooks of Education. Springer, Cham. [https://doi.org/10.1007/978-3-319-66195-7\\_1](https://doi.org/10.1007/978-3-319-66195-7_1)