Shaping the STEM Teacher Workforce: What University Faculty Value about Teacher Applicants

Dan Goldhaber
Roddy Theobald
Amy Roth McDuffie
David Slavit
Jennifer Dechaine-Berkas
John M. Krieg
Emma Dewil

March 2024

WORKING PAPER No. 295-0324

CALDER
National Center for Analysis of Longitudinal Data in Education Research

AIR®

Shaping the STEM Teacher Workforce: What University Faculty Value about Teacher Applicants

Dan Goldhaber
American Institutes for Research / CALDER
University of Washington

Roddy Theobald
American Institutes for Research / CALDER

Amy Roth McDuffie
Washington State University

David Slavit
Washington State University

Jennifer Dechaine-Berkas
Central Washington University

John M. Krieg
Western Washington University

Emma Dewil
American Institutes for Research / CALDER

Contents

Acknowledgments

This research was supported by the National Science Foundation, DUE # 1950030. The findings do not necessarily represent the views of NSF or the organizations with which the authors are affiliated. We thank project advisory board members Tariq Akmal, Vivien Chen, Edward Geary, Mary Jo Larsen, Bruce Larson, Cap Peck, Allyson Rogan-Klyve, Shannon Seidel, Patrick Sexton, Mark Windschitl, and Jim Wyckoff for comments that improved the paper, and Jason Giersch for support with the design of the conjoint analysis.

CALDER • American Institutes for Research
1400 Crystal Drive 10th Floor, Arlington, VA 22202
202-403-5796 • www.caldercenter.org

Shaping the STEM Teacher Workforce: What University Faculty Value about Teacher Applicants

Dan Goldhaber, Roddy Theobald, Amy Roth McDuffie, David Slavit, Jennifer Dechaine-Berkas, John M. Krieg, & Emma Dewil

CALDER Working Paper No. 295-0324
March 2024

Abstract

Who ends up in the teacher workforce is greatly influenced by who is admitted into teacher education programs (TEPs). To better understand how the preferences of teacher education faculty might shape admissions of STEM teacher candidates, we surveyed faculty who teach content or methods courses to STEM teacher candidates across five universities.
Faculty reported that they most value information collected from individual interviews with applicants and data on the number of STEM courses taken in college and their performance in these courses, and least value data on university admissions tests, high school GPA, and teacher licensure test scores. When we investigate faculty members' revealed preferences through a conjoint analysis, we find that faculty most value applicants who have worked with students from diverse backgrounds and applicants from a marginalized racial or ethnic community, and least value whether they received high grades in math and/or science courses. Finally, we find significant variation in these perceptions across respondents in different faculty roles, who teach different courses, and from different institutions: for example, Arts and Sciences faculty tend to value TEP applicants' performance in college STEM courses relatively more than STEM education faculty, while STEM education faculty tend to value applicants' race and ethnicity relatively more than Arts and Sciences faculty.

1. Introduction

More than 85% of the nation's science, technology, engineering, and math (STEM) teachers graduated from traditional teacher education programs (TEPs) (U.S. Department of Education, 2016). Hence it is no exaggeration to describe admissions to TEPs as the primary gateway to the teaching profession. Yet there is remarkably little known about TEP admission processes and the extent to which different applicant attributes are valued by TEP faculty.

The lack of evidence on what faculty value is problematic. There is now a vast amount of evidence showing that in-service teachers have large and varied impacts on both students' test-based (e.g., Aaronson et al., 2007; Chetty et al., 2014; Rivkin et al., 2005) and non-test (e.g., Backes et al., 2023; Kraft, 2019; Jackson, 2018) outcomes. These teacher impacts are not strongly associated with the credentials used to determine employment or compensation (Gordon et al., 2006), but TEP faculty have access to much more nuanced information about teacher applicants when they apply to preparation programs which may be used to shape the future teacher labor force.

In this paper we describe analyses of an original survey of university faculty involved with the preparation of teacher candidates in mathematics and science. This survey was administered to all faculty who participated in STEM teacher education at the five participating institutions for higher education (IHEs), *whether or not they participated directly in the admissions process.*[1] The survey included background questions about faculty roles in their departments and the courses they teach, and also included questions designed to better understand faculty preferences about how useful and important they think different applicant

[1] Note that we include programs in elementary education because they are preparing elementary teachers for math and science instruction and that we use the term "STEM" throughout to refer to the programs and faculty involved in mathematics and science preparation.

attributes and dispositions are for assessing potential future STEM teachers. Finally, to assess faculty revealed preferences about specific applicant attributes, the survey also presented respondents with pairs of hypothetical applicants with different attributes (high school/college GPA, prior teaching experiences, etc.) and asked recipients to pick between applicants with different profiles.

We used these survey data to answer the following overarching questions:

1. What are faculty members' *stated preferences* for various applicant attributes and dispositions for admissions considerations and assessment of an applicant's future potential as a teacher?
2. What are faculty members' *revealed preferences* (based on the conjoint experiment) for specific applicant attributes?
3. To what extent is there variation in stated or revealed preferences according to faculty roles, courses taught, or by institution?

For RQ1, faculty reported that they most value information collected from individual interviews with applicants and data on the number of STEM courses taken in college and their performance in these courses. Faculty report value data on university admissions tests, high school GPA, and teacher licensure test scores. More than half of faculty also agreed that some specific potential areas of concern represent "red flags" that they would want to investigate further before making admissions offers: deficit-based attitudes about specific groups of students, negative attitudes towards culturally responsive pedagogy, negative attitudes towards inquiry-based learning, and a lack of professionalism. Finally, faculty tended to value applicants who view math and/or science as a way to understand the world, those who persevere to solve math and/or science problems, and those who view math problem-solving and/or scientific inquiry more positively.

For RQ2, we investigated faculty members' revealed preferences through a conjoint analysis and find that faculty most highly value applicants who have worked with students from diverse backgrounds and applicants from a marginalized racial or ethnic community. Interestingly, faculty tended to slightly prefer applicants who favored teacher-centered approaches over inquiry-based approaches, on average, and whether applicants received high grades in math and/or science courses was not significantly predictive of whether faculty preferred one applicant over another.

Finally, for RQ3, we find significant variation in these perceptions across respondents in different faculty roles, who teach different courses, and from different institutions. For example, Arts and Sciences (A&S) faculty tend to value TEP applicants' performance in college STEM courses relatively more than STEM education faculty, while STEM education faculty tend to value applicants' race and ethnicity relatively more than A&S faculty. Faculty who primary taught education courses tended to value teacher-centered approaches, relative to inquiry-based approaches, more than faculty who taught content courses for teachers or other courses. Finally, faculty preferences for applicants from marginalized racial or ethnic communities varied significantly across the specific TEPs participating in this study.

2. Background

There is overwhelming evidence that teacher quality influences student achievement, both in terms of test (Aaronson et al., 2007; Goldhaber et al., 1999; Kane et al., 2013; Nye et al., 2004; Rivkin et al., 2005) and non-test (Backes and Hansen, 2018; Jackson, 2018; Kraft, 2019; Liu and Loeb, 2021) outcomes, and that the influence of teachers has implications for longer run student success in college and the workforce (Backes et al., 2023; Chetty et al., 2014). It is not surprising, therefore, that policy makers focus on teachers and teacher education as a lever for improving student achievement, and, in particular, STEM outcomes (e.g., National Academies of Sciences, Engineering, and Medicine, 2020; President's Council of Advisors on Science and Technology, 2010).

While there is little evidence about what the faculty involved in the selection of teacher candidates into programs value about applicants to TEPs, scholars have identified dispositions of effective teachers that could be observed prior to teacher preparation (e.g., Helm, 2006), as well as the role that identifying these dispositions *should* play in teacher education admissions processes (e.g., Goodlad, 1990; Jacobowitz et al., 2000). That said, scholars have recently challenged the theory underpinning this process (e.g., Borko et al., 2007; Childs et al., 2011; Deluca, 2012); i.e., that there is insufficient time in teacher preparation to develop these dispositions (Jacobowitz, 1994), so admissions should disproportionately focus on identifying candidates with these specific dispositions. Conversely, Childs and Ferguson (2015) identify the specific types of attributes that TEP admissions processes are meant to "filter out," including those with poor skills or potentially problematic attitudes, as well as goals that TEP admissions can be used to promote (e.g., workforce diversity).

The lack of information about faculty values in TEP admissions is surprising given that TEPs have faced significant criticism for lacking rigor in their teacher candidate selection processes. For instance, in a widely cited review of TEP practices, Levine (2006) concludes that "The [academic] standards for admitting students to the nation's teacher education programs are too low" (p. 60). More than a decade later, this criticism still surfaces: "Too few graduate and alternative route teacher education programs screen elementary and secondary applicants rigorously for their academic caliber" (Rickenbrode et al., 2018, p. 30).[2] Moreover, both academics and pundits point to the fact that U.S. teachers—as opposed to teachers in countries with better student performance—tend not to be drawn from the top of the academic performance distribution and then link this fact to the poor performance of U.S. students relative to their peers in other industrialized countries (Felton, 2016; Greenberg, Walsh, & McKee, 2014, 2015; New York Times Editorial Board, 2013).

The above critiques notwithstanding, there is limited evidence supporting the commonly held belief that TEP admissions processes are not rigorous enough and that the recruitment and selection of more academically proficient teacher candidates would ultimately result in a higher quality teacher workforce (Cochran-Smith et al., 2017; Goldhaber, 2019). But it is also not possible to know whether the selectivity critique leveled at TEPs is ultimately a reflection of faculty values or related to institutional demands. Levine (2006), for instance, describes situations where TEPs feel pressure to have virtually open admissions processes because TEPs sometimes serve as "cash cows" for universities.[3]

Another common critique of teacher education is that there should be stronger connections and better integration between faculty in colleges of education and faculty in arts and sciences departments (AMTE, 2017).[4] While a lack of connections is often a critique leveled at universities with TEPs, it is difficult to know whether this reflects a lack of alignment of

---

[2] Criticisms of TEP selectivity in terms of admissions is not new (e.g., Barnard & Thornburg, 1980; Laman and Reeves, 1983; Watts, 1980).
[3] It is not clear that this idea of TEPs as cash cows is systemically borne out; see Hemelt et al., 2018.
[4] This critique is perhaps best exemplified by the Teachers for a New Era initiative (TNE, 2001), whose second design principle is that "Faculty in the disciplines of the arts and sciences should be fully engaged in the education of prospective teachers, especially in the areas of subject matter understanding and general and liberal education." This critique also is consistent with both a 1996 advisory committee to the National Science Foundation (NSF, 1996) and a National Research Council report (NRC, 2000) on STEM teacher education, both of which emphasized the need for better coordination of faculty across departments so that teacher candidates themselves can integrate pedagogical and content knowledge.

values across different parts of the faculty or institutional constraints making coordination across departments/schools challenging. We were unable to find even a single study that speaks to the extent to which faculty across departments/schools share common values about admission of teacher candidates.

This quantitative study builds most closely on two prior qualitative studies done in partnership with the same IHEs participating in this study (described in more detail below). The first of these studies (Roth McDuffie et al., 2024) investigated teacher education program admissions processes by interviewing faculty and analyzing program documents with particular attention to how faculty attend to diversity, equity, inclusion, and social and racial justice (DEIJ) issues. Specific foci included applicant recruitment and selection, components of applications (e.g., forms, essays, interviews), and how applicants' DEIJ-related information and orientations factor into admissions. The authors found that all TEPs collected information related to DEIJ (e.g., applicants' ethnoracial backgrounds, citizenship). Interviewed faculty expressed an interest in increasing the diversity of applicants and admitted students, and faculty expressed preferences for applicants who evidenced positive DEIJ orientations.

The second of these studies (Slavit et al., 2023) focused on math and science content background and how university faculty perceive and value this information in admissions decisions. The authors found that broad measures of mathematics and science content background (e.g., achievement test scores, past mathematics and science courses taken) were used more frequently than information on applicants' specific mathematics and science content knowledge and dispositions. In many cases, application components (such as interviews and personal essay statements) were perceived by faculty as helpful for assessing applicants' content knowledge and dispositions; however, these application components were not constructed or employed in a way that allowed them to attain this information.

This study builds on this prior work in several ways. First, the qualitative evidence described above was used to develop surveys of faculty from these TEPs, which were then sent to *all* faculty at these TEPs who are involved in STEM teacher education (as defined below) resulting in evidence from a wider range of respondents covering a narrower scope of topics. We explicitly asked respondents to rank attributes of prospective applicants allowing for quantitative analysis that identifies faculty preferences for applicant characteristics. Finally, we revealed these preferences by implementing a conjoint analysis of faculty preferences for specific applicant attributes.

3. Data

The primary data we use in the analysis were collected as part of a survey administered to faculty in 31 mathematics and science teacher education programs housed within five institutes of higher education in a northwestern state. Collectively these institutions provide about 50% of new teachers prepared from institutions within the state, but they also differ from each other in key ways. For instance, of the 31 TEPs across the five participating IHEs, 6 are undergraduate elementary programs, 3 are undergraduate secondary programs, 6 are masters elementary programs, 8 are masters secondary programs, 6 are alternative certification elementary programs, and 2 are alternative certification secondary programs. In terms of selectivity, TEPs from one participating IHE admit only about 50% of applicants while another (undergraduate program) admits any applicant already enrolled in the university.[5]

[5] For more details on the materials that applicants to preparation programs are required to submit as part of their application, see Roth McDuffie et al. (2024).

The survey includes four primary categories of items: background information on respondents, TEP admissions processes, items designed to assess the applicant attributes faculty value, and the role faculty perceive they play in admissions processes. The specific questions were developed by the authors based on results of the prior qualitative work described in Section 2 (see Roth McDuffie et al., 2024 and Slavit et al., 2023 for more details).

The descriptive analysis of RQ1 focuses on three questions in the faculty survey. The first question asks, “Regardless of current admissions requirements or how the admissions process currently works at your institution, how useful do you think the following types of information are or would be for assessing the potential of applicants to eventually become successful STEM teachers?”, with Likert scale responses ranging from “0 - Not Important” to “3 - Very Important” and the option to identify up to three sources of information as “One of the three most important.”

The second question provides the following prompt to respondents:

“We have found that faculty describe “red flags” and “yellow flags” related to issues of concern when talking to applicants to their teacher education programs. “Red flags” represent a concern that would lead you to not want the candidate in the program. “Yellow flags” represent a concern that you might want to find out more information, make sure to monitor the candidate within the program, and/or be cautious about moving forward with the candidate. With these descriptions in mind, please consider each of the below potential areas of concern, and check one for the following: red flag, yellow flag, no concern.”

Respondents are then able to characterize each potential area of concern as a “red flag,” “yellow flag,” or “not a concern.”

The final question asks respondents, “In your opinion, how important is it for applicants to have the following experiences or dispositions prior to entering your teacher education program?” with Likert scale responses ranging from “0 - Not Important” to “3 - Very Important” and the option to identify up to three sources of information (e.g., “views math and/or science as a way of understanding the world”) as “one of the three most important.” In heterogeneity analyses aligned with RQ3 and the balance tests below, we also use faculty responses to questions about: faculty members’ predominant role at their universities, which we group into “A&S STEM faculty” or “STEM education faculty”;[6] the primary type of course faculty reported teaching;[7] and how much influence faculty said they had over different aspects of the admissions process.[8]

To better understand how applicants presented with *combinations* of attributes would be valued as part of RQ2, we also included a conjoint experiment (Giersch and Dong, 2018) portion of the survey instrument. This part of the survey presented faculty with a series of hypothetical applicants to their program, with different combinations of attributes, and asked them to choose which applicant to admit based on the following prompt:

> “This part of the survey is about your thoughts on attributes of STEM applicants. In each of the following questions, you will be asked to compare **hypothetical applicants** to a

[6] For this we categorized based on the following question: “Please choose one of the following that best describes your predominant role at your university” to identify content and education faculty. Dual appointed faculty in education and another department are labeled as content and faculty members who indicated they had a role other than the options provided were categorized based on their response to a free text question asking about their role.
[7] We categorized this based on the following question: “With respect to your role at the university, please rank the following responsibilities by moving the items to the right side so the courses you teach the most often are at the top, and least often at the bottom. If you do not teach a certain type of course, do not rank it.” The responsibilities to be ranked are “Teach education/methods courses,” “Teach content-specific courses to prepare teachers,” and “Teach content-specific courses that are not necessarily designed for teacher education.” Faculty are categorized according to the type of course they rank first.
[8] For this we categorized based on the following question: “How would you rate your level of *influence* in different parts of the admissions process?” with a 4-point Likert scale response option, “none,” “a little,” “some,” “a lot,” and “I don’t know.” The question asked about “Influence over TEP admission requirements,” “Influence over applicant selection criteria,” and “Influence over specific admissions decisions.”

teacher education program at your university. Please select the applicant **you would prefer to admit to the program**, regardless of how the admissions process currently operates. Please consider all four applicant attributes, **regardless of whether this specific applicant information is currently collected as part of your program's admissions process**." (emphasis from survey)

A power analysis conducted prior to the survey suggested that we had sufficient power with five comparisons per respondent (and an assumed 50% response rate) to distinguish between faculty preferences between four binary variables, so we selected the following pairs of potential attributes:

1) Prior experience ("Has/has not worked with students from diverse backgrounds")
2) Grades ("Did/did not receive high grades in her past math and/or science courses")
3) Disposition ("Prefers inquiry-based/teacher-centered approach")
4) Racial/ethnic background ("Is white/Is from a marginalized racial or ethnic community")

Figure 1 provides an example of how respondents were presented with each of five hypothetical pairs of applicants. Pairs of applicants were displayed with these attributes in a randomized order top to bottom in case attribute order impacted responses. Due to a coding error, one of the 16 combinations of attributes was never presented to any respondent, though we view this as impacting our power rather than our estimates because other combinations of attributes were equally likely to be presented to respondents.

The sampling frame for the survey was collected directly from "admissions team" members at participating institutions, in most cases an Associate Dean but always someone with in-depth knowledge of faculty at the institution. We asked these admissions team members in

Spring 2022 to provide complete lists of faculty from their IHEs who belonged to three categories: 1) Faculty who focus on mathematics or science content courses in their discipline, and at least part of their work includes mathematics or science content courses for teacher candidates ("STEM content faculty"); 2) Faculty who focus on mathematics and/or science teacher education ("STEM education faculty"); and 3) Faculty who focus on teacher education and teach teacher education courses that have elementary or secondary mathematics and science teacher candidates in them ("other faculty"). In cases where faculty held multiple roles, we asked admissions team members to make a judgment call about the category that most aligned with their job responsibilities. Through this process we identified 113 faculty across the five participating IHEs who were eligible to take the survey.[9]

We sent a first round of survey invitations to faculty before the end of the 2021-22 school year in Spring 2022 and again at the beginning of the 2022-23 school year in Fall 2022. By the end of Fall 2022, 70 of the 113 eligible faculty had responded to the complete survey for a response rate of 62%.[10] Panel A of Table 1 shows how these response rates vary across participating IHEs and the three faculty groups identified by *admissions team members* (thus observable for all faculty regardless of whether they responded to the survey). Institution-level response rates ranged from 44% to 81%, though a Chi-square test did not find significant differences in response rates across institutions (df=4, p=0.13). Likewise, while response rates within the three faculty categories identified above ranged from 54% to 81%, a Chi-square test again failed to reject the null hypothesis that the probability of responding was the same across institutions (df=2, p=0.26). To explore these differences in response rates further, Panel B of

[9] An additional nine faculty were provided by admissions team members who were later determined to be retired, on leave, or otherwise no longer teaching. These faculty are not considered part of the sampling frame.
[10] Five faculty members had partial survey responses, which we consider as non-responses for the purposes of this analysis.

Table 1 formally tests the proportion of respondents and non-respondents from each IHE and faculty group and shows that response rates were significantly higher for one of the five participating IHEs and for the "education" faculty group (i.e., group 2 described above).

Given that many of our analyses disaggregate by faculty self-reported categories of "A&S STEM" and "STEM education" faculty, we also explore the extent to which these self-categorizations align with admissions team members' descriptions of faculty roles in Table 2. For the most part they align well: most faculty identified as "STEM content faculty" by admissions team members self-identify as A&S STEM faculty in the survey, while most faculty identified as "other faculty" identify as STEM education faculty. Interestingly, faculty identified by admissions team members as "STEM education faculty" are not significantly more or less likely to self-identify as A&S STEM or STEM education faculty, which perhaps illustrates the dual role that many of these faculty play. From this point on in the analysis, we focus on faculty members' self-identified categories.

4. Analytic Approach

Our analysis of the Likert scale survey questions above is purely descriptive, as we calculate the distribution of responses for each component of these survey items and test whether there are significant differences between the responses of STEM education and STEM content faculty on each component. We use both Chi-Square tests of the overall distribution of responses between STEM education and STEM content faculty as well as two-sided t-tests of the proportion of responses within given categories; our primary results test the proportion of responses in the top two categories of the Likert scale.

Our analysis of the conjoint analysis data follows Giersch and Dong (2018), who use a similar experimental design to study the preferences of hiring principals for various applicant teacher attributes. The outcome in this analysis is a series of binary indicators Y_{ijk} indicating whether respondent *i* when presented with a pair of profiles *j* (*j* = 1, …, 5 since each respondent was presented with five pairs of profiles) selected profile *k* (where *k* = 1 or 2 and $Y_{ij1} + Y_{ij2} = 1$ since each respondent had to select one of the pairs of profiles). We model the choice of Y_{ij1} or Y_{ij2} as a conditional logit model, estimated as a function of the attributes in X_{ijk1} through X_{ijk4}. These attributes are also binary indicators indicating whether hypothetical applicant in that profile: Has/has not worked with students from diverse backgrounds (X_{ijk1}); Did/did not receive high grades in her past math and/or science courses (X_{ijk2}); Prefers inquiry-based/teacher-centered approach (X_{ijk3}); and Is white/is from a marginalized racial or ethnic community (X_{ijk4}). The formal model is as follows:

$$\Pr\left(Y_{ijk} = 1 | \boldsymbol{X_{ijk}} \right) = \frac{\exp\left(\beta_{ij}+\beta_1 X_{ijk1}+\beta_2 X_{ijk2}+\beta_3 X_{ijk3}+\beta_4 X_{ijk4}\right)}{1+\exp\left(\beta_{ij}+\beta_1 X_{ijk1}+\beta_2 X_{ijk2}+\beta_3 X_{ijk3}+\beta_4 X_{ijk4}\right)} \qquad (1)$$

We transform the coefficients β_n to marginal effects that can be interpreted as the marginal increase in the probability that the average respondent chooses a given profile associated with the specific attribute in the profile that they saw, all else equal. The "all else equal" portion of this regression is critical, as the experimental design is intended to tease apart the *relative* importance of different applicant attributes in predicting which hypothetical profiles are chosen more often than others.

We first estimate the model in equation 1 across all responding faculty and then with interactions between each of the *X* variables and indicators for different faculty indicators (e.g., A&S STEM vs. STEM education faculty), testing the coefficients between these groups to see if one group differentially prefers some attributes more than others. In extensions we also test differences by faculty teaching responsibilities, faculty institution, and faculty role in admissions.

5. Results

5.1 Descriptive Survey Analysis

We present faculty responses to the three focal questions identified in Section 3 in a series of figures, beginning with Figure 2 which pools faculty responses to the question "**Regardless of current admissions requirements or how the admissions process currently works at your institution**, how useful do you think the following types of information are or would be for assessing the potential of applicants to eventually become successful STEM teachers?" (emphasis from survey). The bars in Figure 2 are ordered from left to right in terms of proportion of respondents who selected the source of information as "Very important" or "One of the three most important."

By this measure of faculty preferences, the four most important applicant attributes according to faculty respondents are "information from individual interview," "number of college STEM courses," "application essay or open-ended response," and "performance in college STEM courses." In all four cases, more than half of respondents identified this source of information as "Very important" or "One of the three most important." By this same measure, the four least important applicant attributes are "teacher certification tests," "performance in high school STEM courses," "high school GPA," and "university admissions tests"; in each case, less than 15% of respondents indicated that the source of information was at least "Very Important."

In Figure 3, we turn to the second focal question from the survey about "potential areas of concern" that faculty could label as "red flags," "yellow flags," or "not a concern" (see Section 3 for a full description of the prompt for this question). Two potential areas of concern were easily the most commonly identified as a red flag: "has deficit-based perspective about specific groups of students, families, and/or students' potential to learn" and "has negative attitude toward culturally-responsive pedagogy"; in each case, about 75% of respondents labeled these areas as a red flag. On the opposite end of the spectrum, 10% or fewer respondents labeled the following as red flags: “not a US Citizen” (0%), “low university admissions test scores” (3%), “low teacher certification test scores” (4%), “poor interview skills” (6%), “not fluent in English” (7%), and “low GPA” (10%).

Finally, we display the distribution of faculty perspective on the importance of specific applicant dispositions in Figure 4**.** About two-thirds of respondents report that the dispositions “views math and/or science as a way to understand the world,” “perseveres to solve math and/or science problems,” and “views math problem solving and/or scientific inquiry as important” to be either “Very important” or “One of the three most important” dispositions. On the other end of the spectrum, only 14% of respondents place “views math and/or science primarily as a way to systematically arrive at correct answers” in one of these top two categories.

5.2 Conjoint Analysis

In column 1 of Table 3, we report estimated average marginal effects from the conditional logit regression models in equation 1. The model in the first column is estimated across all faculty respondents, and we present the estimated marginal effects and confidence intervals from these models in Figure 5 to permit visual inspection of the relative importance of each item. These marginal effects can be interpreted as follows: a profile that states that the applicant has worked with students from diverse backgrounds was 23.7 percentage points more likely to be selected by respondents than a profile that states that the applicant has not worked with students from diverse backgrounds, all else equal. The next highest marginal effect was for candidate race and ethnicity: a profile that states that the applicant is from a marginalized racial or ethnic minority was 14.1 percentage points more likely to be selected than a profile that states that the applicant is White, all else equal. The only other statistically significant coefficient is for applicant dispositions: a profile that states that the applicant prefers a teacher-centric approach was 9.3 percentage points more likely to be selected than a profile that states that the applicant prefers an inquiry-based approach, all else equal. Perhaps surprisingly, there was no significant difference in the probability that a profile stating that the applicant received high grades in past math/science courses was selected relative to a profile stating that the applicant did not receive high grades in prior math/science courses, all else equal.

5.3 Testing for Heterogeneity of Valued Attributes

It has long been argued that one of the factors limiting improvement in teacher preparation is divergent views about what preparation should entail across education and arts and sciences departments and too little collaboration on improvement activities (TNE, 2001). Hence we begin the heterogeneity analysis by splitting the sample into A&S STEM and STEM education faculty to access for divergent views about admissions.

In Figure 6 we compare responses between the 49 STEM education faculty (top panel) and 21 A&S STEM faculty (bottom panel) who responded to the survey.[11] Each panel is ordered from left to right according to proportion of *STEM Education Faculty* respondents who selected the source of information as "Very important" or "One of the three most important," so differences from this ordering in the bottom panel provide a visual inspection of places where A&S STEM faculty provided different rankings than STEM education faculty. We focus on the only two responses for which there was a statistically significant difference between the proportion of respondents in these top two categories: A&S STEM faculty were more likely to place importance on performance in college STEM courses (71% in top two categories) than

[11] For this we categorized based on the following question: "Please choose one of the following that best describes your predominant role at your university" to identify content and education faculty. Dual-appointed faculty in education and another department are labeled as content and faculty members who indicated they had a role other than the options provided were categorized based on their response to a free text question asking about their role.

STEM education faculty (47% in top two categories), while STEM education faculty were more likely to place importance on an applicant's race/ethnicity (43% in top two categories) than A&S STEM faculty (14% in top two categories). When we estimate ordered logistic regressions predicting these responses as a function of both STEM faculty status and institution, only the difference in responses about "performance in college STEM courses" is still statistically significant when making comparisons between faculty from the same institution.

We then compare the distribution of responses about red flags and yellow flags between STEM education and A&S STEM faculty in Figure 7, again ordering both panels in terms of the proportion of STEM education faculty who identify the potential area of concern as a red flag. Two areas were rated significantly differently between the two groups: A&S STEM faculty were significantly more likely to label "has negative attitude towards inquiry-based learning" (71%) and "poor grades in math and science content courses" (33%) as red flags than STEM education faculty (47% and 8%, respectively). Both differences are still statistically significant in ordered logistic regression models with institution fixed effects, meaning that A&S STEM faculty are more likely to identify these as red flags as STEM education faculty from the same institution.

Finally, when we compare the distribution of perspectives on valued experiences or dispositions between STEM education faculty and A&S STEM faculty in Figure 8, we do not find significant differences between the responses for these groups. With that said, when we control for respondents' institutions in ordered logistic regression models, we find that A&S STEM faculty are more likely to place higher values on the attributes "Views math problem solving and/or scientific inquiry as important," "Has deep understandings of mathematics and/or science concepts, principles, and practices," and "Views math and/or science as a way to understand the world" than STEM education faculty from the same institution.

We now turn to the conjoint analysis, interacting different applicant attributes with observable faculty characteristics to test whether these applicant attributes are valued by some faculty more than others. Columns 2 and 3 of Table 3 include interactions for A&S STEM faculty and STEM education faculty; we report the marginal effects for each faculty group from these models and plot these estimates in Figure 9 to permit comparisons between the groups. Only one coefficient is significantly different between the two groups: the expected increase in the probability of selecting an applicant that has worked with students from diverse backgrounds, relative to an applicant that has not worked with students from diverse backgrounds, was significantly greater for A&S STEM faculty (35.8 percentage point increase) relative to STEM education faculty (19.3 percentage point increase).

The remaining columns of Table 3 disaggregate the sample by additional self-identified faculty attributes described in Section 3. Columns 4-6 of Table 3 present marginal effects separately for faculty who primarily teach education courses, faculty who primarily teach content courses for teachers, and faculty who primarily teach courses not designed for teachers. There are significant differences across these faculty types for each of the first two attributes: faculty who primarily teach content courses for teachers more highly value applicants who have worked with students from diverse backgrounds than the other groups, while faculty who teach primarily education courses value teacher-centered approaches much more than faculty from the other groups. Finally, columns 7 through 11 present results separately by institution, illustrating that preferences for applicants who have worked with students from diverse backgrounds are quite consistent across institutions while preferences for the other attributes vary quite considerably across institutions.[12]

[12] While we do not identify individual institutions as we promised anonymity to participating institutions under the terms of our IRB approval, estimates from conjoint models estimated separately by institution indicate that

Finally, faculty have quite varied perspectives on their influence on three elements of the admissions process (admissions requirements, admissions criteria, and admissions decisions).[13] Indeed, in the case of requirements and criteria, a majority of faculty respondents report they have no influence over admissions. When we explore heterogeneity for these faculty (column 7) relative to faculty who report having any influence over admissions (column 8), we find a statistically significant difference between the groups in terms of their preference for applicants' prior grades: faculty who report having no influence over admissions are actually *less likely* to prefer an applicant with higher grades, while faculty with some influence over admissions are no more likely to prefer applicants with stronger grades than weaker grades.

6. Discussion and Conclusions

We draw three broad conclusions from this analysis, each aligned with one of the research questions we posed. The first conclusion is that, consistent with prior qualitative work (Slavit et al., 2023), faculty tend to value information collected directly from applicants (e.g., from interviews, which were not conducted by all TEPs) and information more proximate to their enrollment (e.g., recent course-taking), while they place little value on testing data commonly used for university admissions and teacher licensure (e.g., SAT, GPA, licensure test scores). As discussed in Slavit et al. (2023), a key implication from faculty preferences expressed in interviews is that TEPs need to ensure that the information collected from admissions interviews actually allows them to assess whether applicants demonstrate the skills and dispositions that will make them effective STEM teachers from the faculty's perspective. On the

preferences for applicants who have worked with students from diverse backgrounds are quite consistent across institutions while preferences for the other attributes (e.g., candidate race/ethnicity) vary considerably across institutions.

[13] We categorized faculty based on the following question: "How would you rate your level of *influence* in different parts of the admissions process?" with a 4-point Likert scale response option, "none," "a little," "some," "a lot," and "I don't know." The question asked about "Influence over TEP admission requirements," "Influence over applicant selection criteria," and "Influence over specific admissions decisions."

other side, that faculty say that they want to reduce the importance of testing data as a source of data is notable because universities and states are increasingly relaxing testing requirements for admissions and licensure, respectively. Although diminishing the role of testing appears to reflect faculty values in the case of TEP admissions, more research is necessary about the sources of data (e.g., essays, interviews) that could replace testing data in these processes.

The second broad conclusion is that the conjoint analysis showed that, again consistent with prior qualitative work (Roth-McDuffie et al., 2024), faculty value diversity both in terms of applicants' prior experiences (e.g., working with diverse populations, interacting with marginalized communities) *and* in terms of the applicants' own ethnoracial and cultural background. The key contribution of the conjoint analysis is disentangling these preferences from other preferences (e.g., for course grades and/or dispositions) that faculty also state that they value. The conjoint analysis suggests that, on average and all else equal, faculty *most* value applicants who have worked with students from diverse backgrounds, and *next* value applicants from a marginalized racial or ethnic community. After controlling for these preferences, we found that high grades in mathematics and/or science courses is *not* predictive of which applicants faculty prefer. This was true even though faculty stated a preference for grades during interviews (Slavit et al., 2023) and in their responses to our survey questions.

Finally, the significant variation in these preferences across A&S faculty and STEM education faculty speaks directly to their differences in perceptions about teacher preparation, something that has been documented in prior work (TNE, 2001). Specifically, we document quantitatively that A&S faculty place a greater value on prior measures of academic achievement, while education faculty place a greater value on applicants' ethnoracial background. These preferences can co-exist, but these diverging perceptions further motivate recent calls to develop and facilitate communication between these groups of faculty (AMTE, 2017).

References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics, 25,* 95–135. https://doi.org/10.1086/508733

Association of Mathematics Teacher Educators (AMTE). (2017). *Standards for preparing teachers of mathematics*. https://amte.net/standards.

Backes, B., Cowan, J., Goldhaber, D., & Theobald, R. (2023). *How to measure a teacher: The influence of test and nontest value-added on long-run student outcomes*. CALDER Working Paper No. 270-0423-2. https://caldercenter.org/publications/how-measure-teacher-influence-test-and-nontest-value-added-long-run-student-outcomes

Backes, B., & Hansen, M. (2018). *Reaching further and learning more? Evaluating Public Impact's Opportunity Culture initiative*. CALDER Working Paper No. 181. https://caldercenter.org/publications/reaching-further-and-learning-more-evaluating-public-impacts-opportunity-culture

Barnard, H. V., & Thornburg, K. R. (1980). Selective admissions in teacher education: A critical reappraisal. *Teacher Education and Special Education, 3*(2), 36–42. https://doi.org/10.1177/088840648000300205

Borko, H., Liston, D., & Whitcomb, J. A. (2007). Apples and fishes: The debate over dispositions in teacher education. *Journal of Teacher Education*, *58*(5), 359–364. https://doi.org/10.1177/0022487107309977

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review, 104*(9)*,* 2633–2679. http://www.doi.10.3386/w19424

Childs, R., Broad, K., Gallagher-Mackay, K., Sher, Y., Escayg, K.-A., & McGrath, C. (2011). Pursuing equity in and through teacher education program admissions. *Education Policy Analysis Archives*, *19*, 24. https://doi.org/10.14507/epaa.v19n24.2011

Childs, R., & Ferguson, A. K. (2015). Changes in, to, and through the initial teacher education program admission process. In Thomas, L. & Hirschcorn, M. (Eds.), *Change and progress in Canadian teacher education: Research on recent innovations in teacher preparation in Canada* (pp. 420–440). Canadian Association for Teacher Education.

Cochran-Smith, M., Baker, M., Burton, S., Chang, W.-C., Cummings Carney, M., Fernández, M. B., Stringer Keefe, E., Miller, A. F., & Sánchez, J. G. (2017). The accountability era in US teacher education: Looking back, looking forward. *European Journal of Teacher Education*, *40*(5), 572–588. https://doi.org/10.1080/02619768.2017.1385061

Deluca, C. (2012). Promoting inclusivity through and within teacher education programmes. *Journal of Education for Teaching*, *38*(5), 551–569. https://doi.org/10.1080/02607476.2013.739792

Felton, E. (2016, July 20). *We'd be better at math if the U.S. borrowed these four ideas for training teachers from Finland, Japan and China*. The Hechinger Report. https://hechingerreport.org/wed-be-better-at-math-if-the-u-s-borrowed-these-four-ideas-for-training-teachers-from-finland-japan-and-china/

Giersch, J., & Dong, C. (2018). Principals' preferences when hiring teachers: A conjoint experiment. *Journal of Educational Administration*, *56*(4), 429–444. https://doi.org/10.1108/JEA-06-2017-0074

Goldhaber, D. (2019). Evidence-based teacher preparation: Policy context and what we know. *Journal of Teacher Education, 70*(2), 90–101. https://doi.org/10.1177/0022487118800712

Goldhaber, D., Brewer, D., & Anderson, D. (1999). A three-way error components analysis of educational productivity. *Education Economics, 7*(3), 199–208. https://doi.org/10.1080/09645299900000018

Goodlad, J. I. (1990). *Teachers for our nation's schools*. Jossey-Bass.

Greenberg, J., Walsh, K., & Mckee, A. (2014). *2014 teacher prep review: A review of the nation's teacher education programs.* National Council on Teacher Quality. https://eric.ed.gov/?id=ED545343

Helm, C. M. (2006). The assessment of teacher dispositions. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *79*(6), 237–239. https://doi.org/10.3200/TCHS.79.6.237-239

Hemelt, W. S., Stange, M. K., Fernando, F., Simon, A., & John, E. S. (2018). *Why is math cheaper than English? Understanding cost differences in higher education* (NBER Working Paper No. 25314). National Bureau of Economic Research. https://www.nber.org/papers/w25314

Jacobowitz, T. (1994). Admission to teacher education programs: Goodlad's Sixth Postulate. *Journal of Teacher Education, 45*(1), 46–52. https://doi.org/10.1177/0022487194045001007

Jacobowitz, T., Delorenzo, L., & Adirim, M. (2002). Incorporating concerns for diversity and social justice in admissions decisions. *Academia.* https://www.academia.edu/28372341/Incorporating_Concerns_for_Diversity_and_Social_Justice_in_Admissions_Decisions

Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non–test score outcomes. *Journal of Political Economy*, *126*(5), 2072–2107. https://doi.org/10.1086/699018

Kane, T. J., McCaffrey, D. F., Miller, T., & Staiger, D. O. (2013). *Have we identified effective teachers? Validating measures of effective teaching using random assignment.* Measures of Effective Teaching (MET) Project. Bill & Melinda Gates Foundation. http://eric.ed.gov/?id=ED540959

Klassen, R. M., & Kim, L. (2019). Selecting teachers and prospective teachers: A meta-analysis. *Educational Research Review*, *26*, 32–51. https://doi.org/10.1016/j.edurev.2018.12.003

Kraft, M. A. (2019). Teacher effects on complex cognitive skills and social-emotional competencies. *Journal of Human Resources, 54*(1), 1–36. http://dx.doi.org/10.3368/jhr.54.1.0916.8265R3

Laman, A. E., & Reeves, D. E. (1983). Admission to teacher education programs: The status and trends. *Journal of Teacher Education, 34*(1), 2–4. https://doi.org/10.1177/002248718303400101

Levine, A. (2006). *Educating school teachers*. Education Schools Project. www.edschools.org/teacher_report.htm

Liu, J., & Loeb, S. (2021). Engaging teachers: Measuring the impact of teachers on student attendance in secondary school. *Journal of Human Resources*, *56*(2), 343–379. https://doi.org/10.3368/jhr.56.2.1216-8430R3

National Academies of Sciences, Engineering, and Medicine. 2020. *Changing Expectations for the K-12 Teacher Workforce: Policies, Preservice Education, Professional Development, and the Workplace*. The National Academies Press. https://doi.org/10.17226/25603.

National Council on Teacher Quality. (2016). *Within our grasp: Achieving higher admissions standards in teacher prep* (2016 State Teacher Quality Yearbook Report Series). https://www.nctq.org/dmsView/Admissions_Yearbook_Report

National Research Council (NRC). (2000). *Educating teachers of science, mathematics, and technology: New practices for the new millennium*. The National Academies Press. https://doi.org/10.17226/9832

National Science Foundation (NSF). (1996). *National science education standards*. National Academy Press. http://books.nap.edu/catalog/4962.html

New York Times Editorial Board. (2013, December 18). Why other countries teach better. *The New York Times*. http://www.nytimes.com/2013/12/18/opinion/why-students-do-better-overseas.html

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*(3), 237–257. https://doi.org/10.3102/01623737026003237

President's Council of Advisors on Science and Technology. (2010). *Prepare and inspire: K–12 science, technology, engineering, and math (STEM) education for America's future.* https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/pcast-stem-ed-final.pdf

Rickenbrode, R., Graham, D., Pomerance, L., & Walsh, K. (2018). *2018 teacher prep review*. National Council on Teacher Quality. https://eric.ed.gov/?id=ED590748

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458. doi:10.1111/j.1468-0262.2005.00584.x

Roth McDuffie, A., Slavit, D., Goldhaber, D., Theobald, R., & Griggs, N. (2024). Attention to equity in teacher education admissions processes. *Journal of Teacher Education*, *0*(0). https://doi.org/10.1177/00224871241230306

Slavit, D., Roth McDuffie, A., Griggs, N., Goldhaber, D., & Theobald, R. (2023). Faculty perspectives and values toward mathematics and science content information used in teacher preparation admissions processes. *School Science and Mathematics*, 1–18. https://doi.org/10.1111/ssm.12627

Teachers for a New Era (TNE). (2001). *Teachers for a new era: A national initiative to improve the quality of teaching.* Carnegie Corporation. https://www.carnegie.org/publications/teachers-for-a-new-era-a-national-initiative-to-improve-the-quality-of-teaching/

Watts, D. (1980). Teacher education: For teachers only. *Journal of Teacher Education, 31*(5), 11–13. https://doi.org/10.1177/002248718003100504

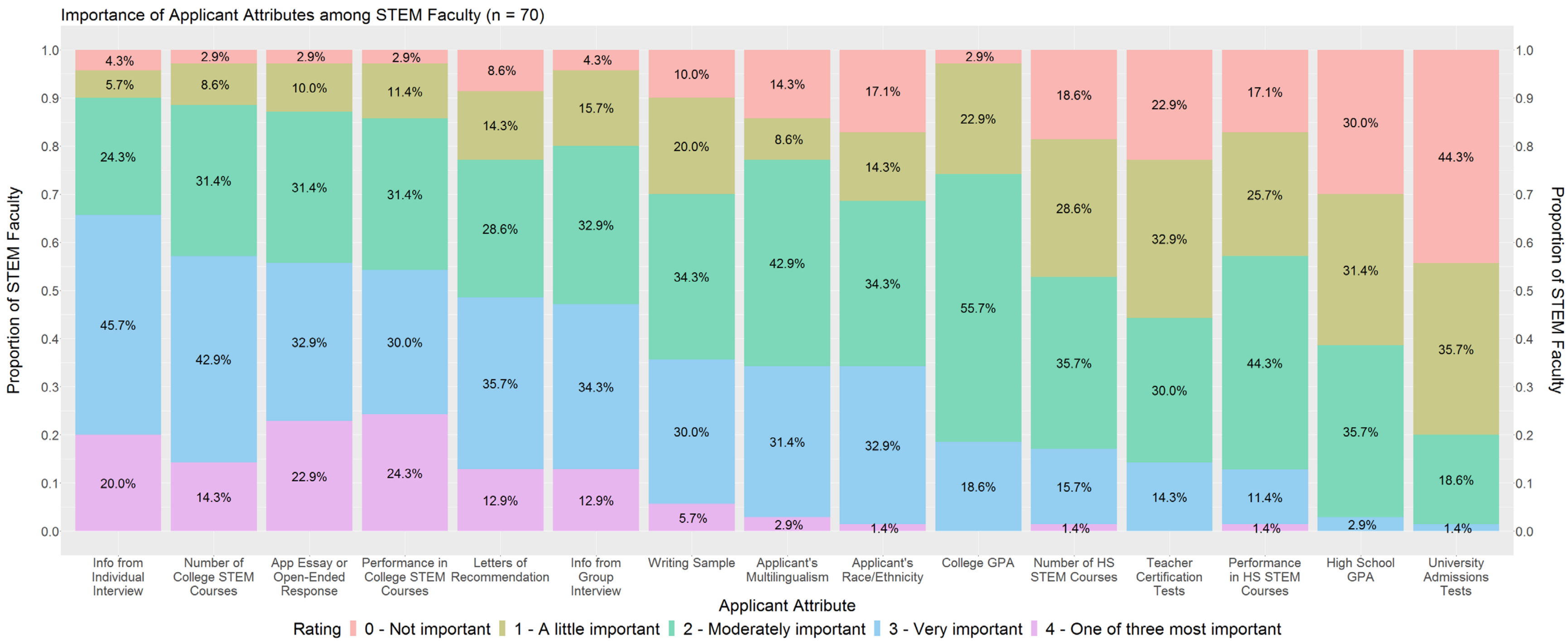
Figures and Tables

Figure 1. Example Conjoint Analysis Comparison

4. Which applicant would you prefer to admit?

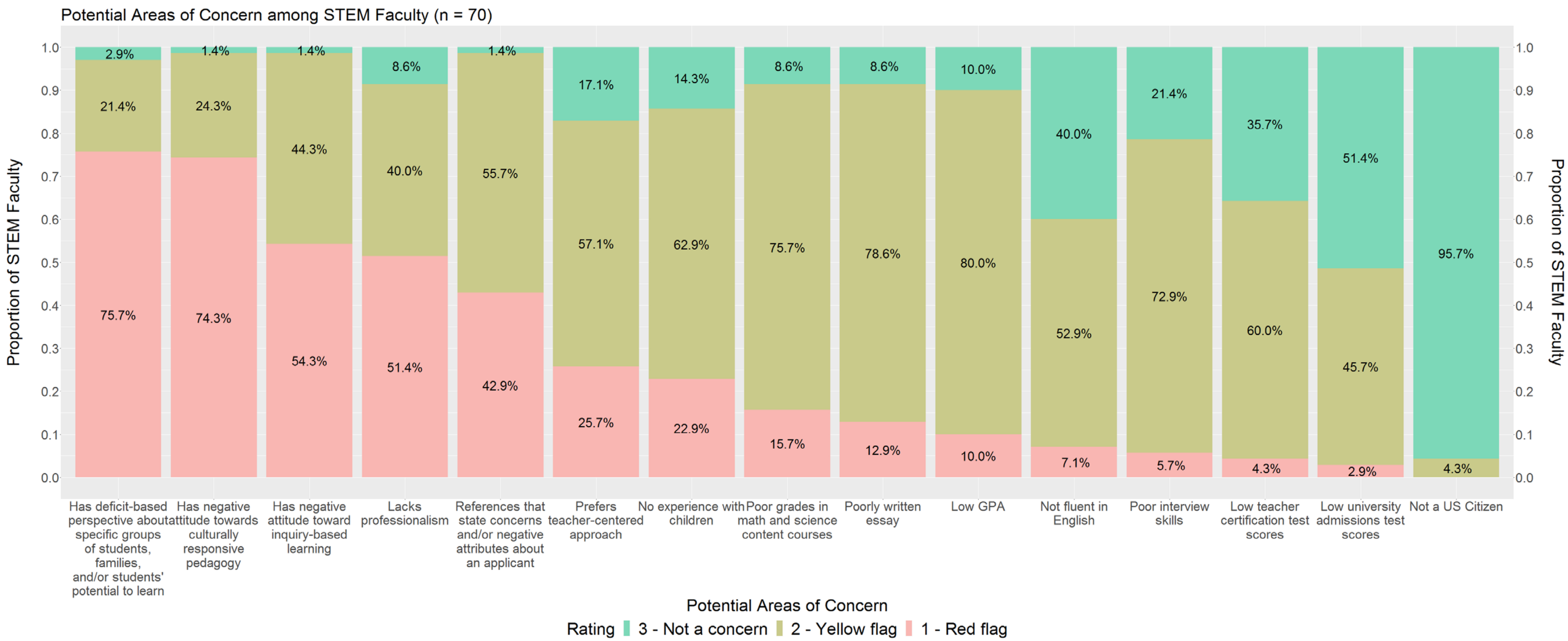
	Applicant A	**Applicant B**
Disposition	Prefers inquiry-based approach.	Prefers teacher-centered approach.
Grades	Received high grades in her past math and/or science courses.	Received high grades in her past math and/or science courses.
Racial/ethnic background	Is white.	Is from a marginalized racial or ethnic community.
Prior experience	Has not worked with students from diverse backgrounds.	Has not worked with students from diverse backgrounds.
Preference:		

Figure 2. Importance of Applicant Attribute Among All STEM Faculty Respondents



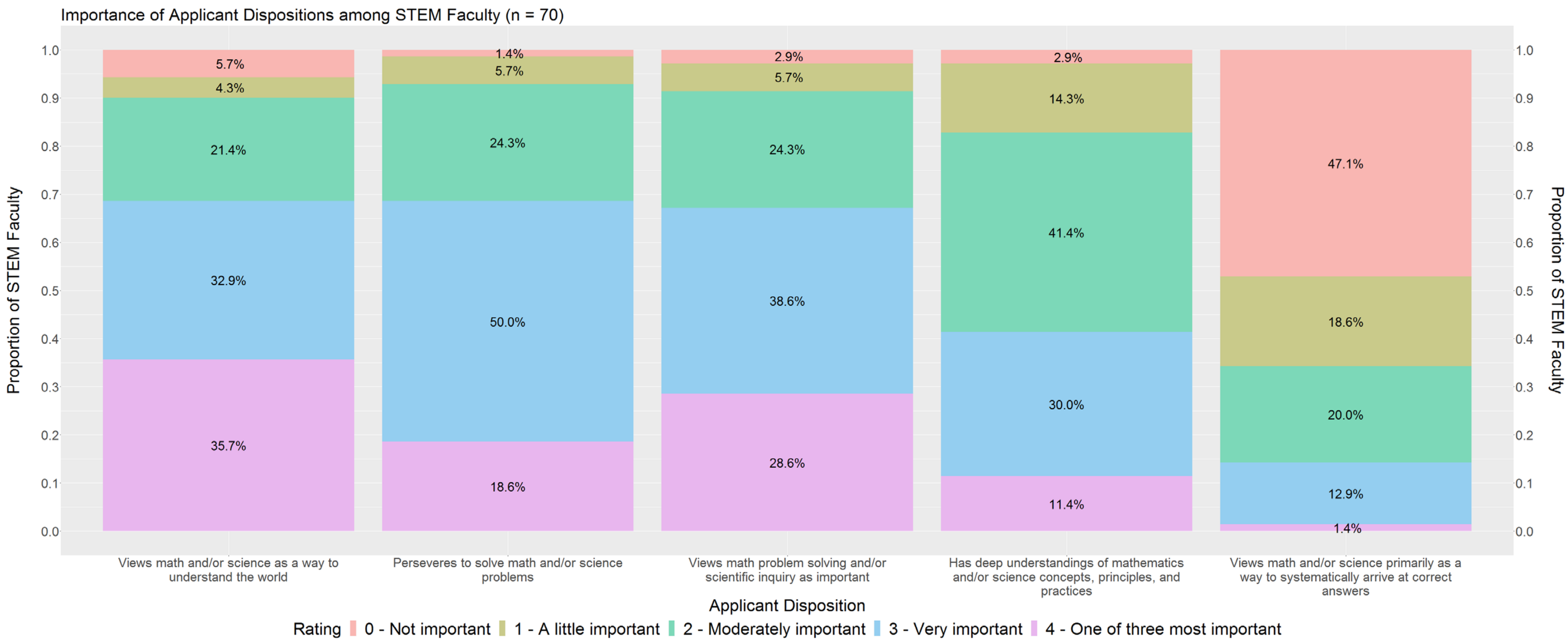
Note. Bars ordered from left to right in terms of proportion of respondents who selected the source of information as "Very important" or "One of the three most important"

Figure 3. Potential Areas of Concern Among All STEM Faculty



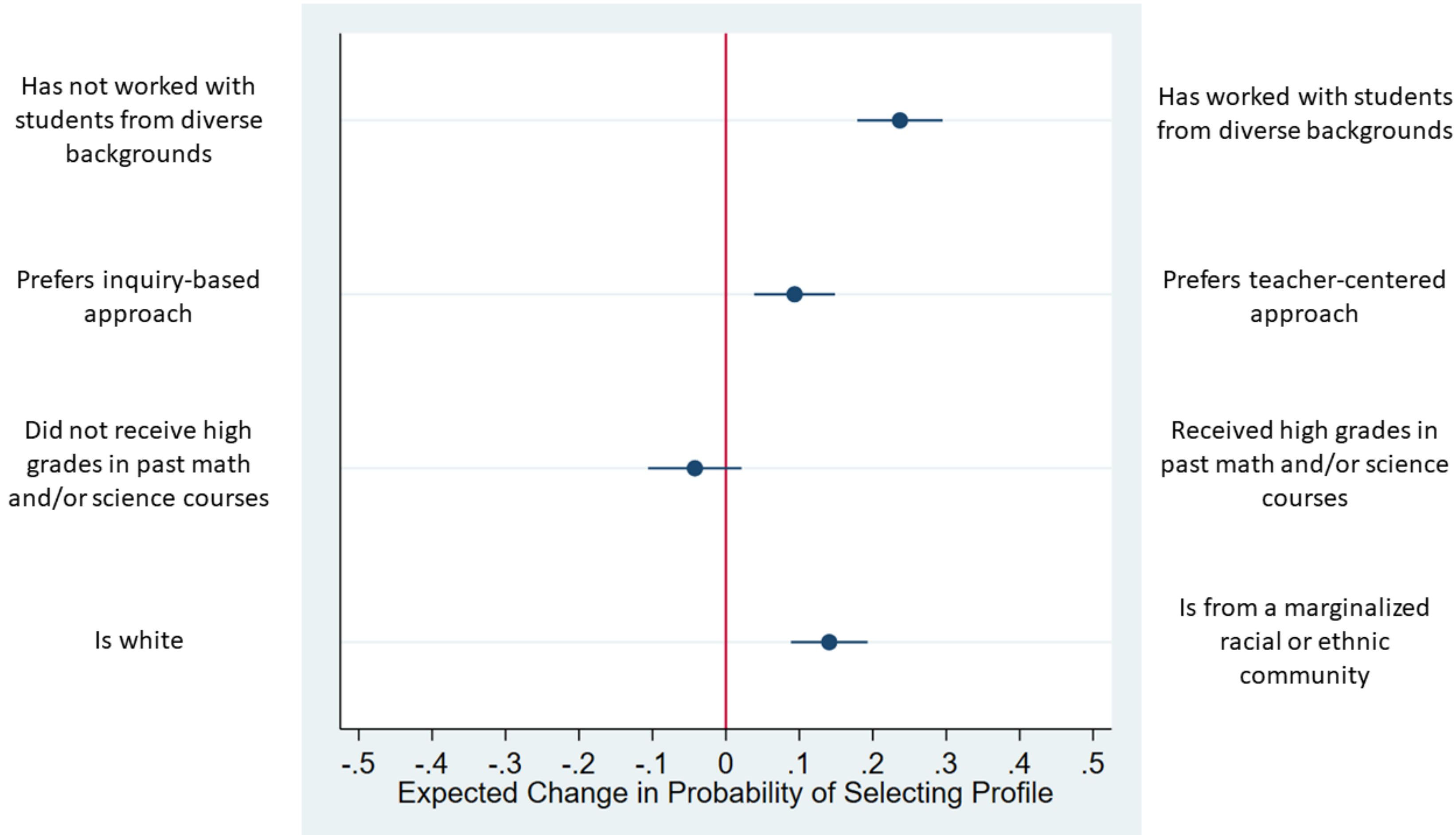
Note. Bars ordered from left to right in terms of proportion of respondents who selected potential area of concern as a "red flag"

Figure 4. Valued Experiences or Dispositions Among All STEM faculty



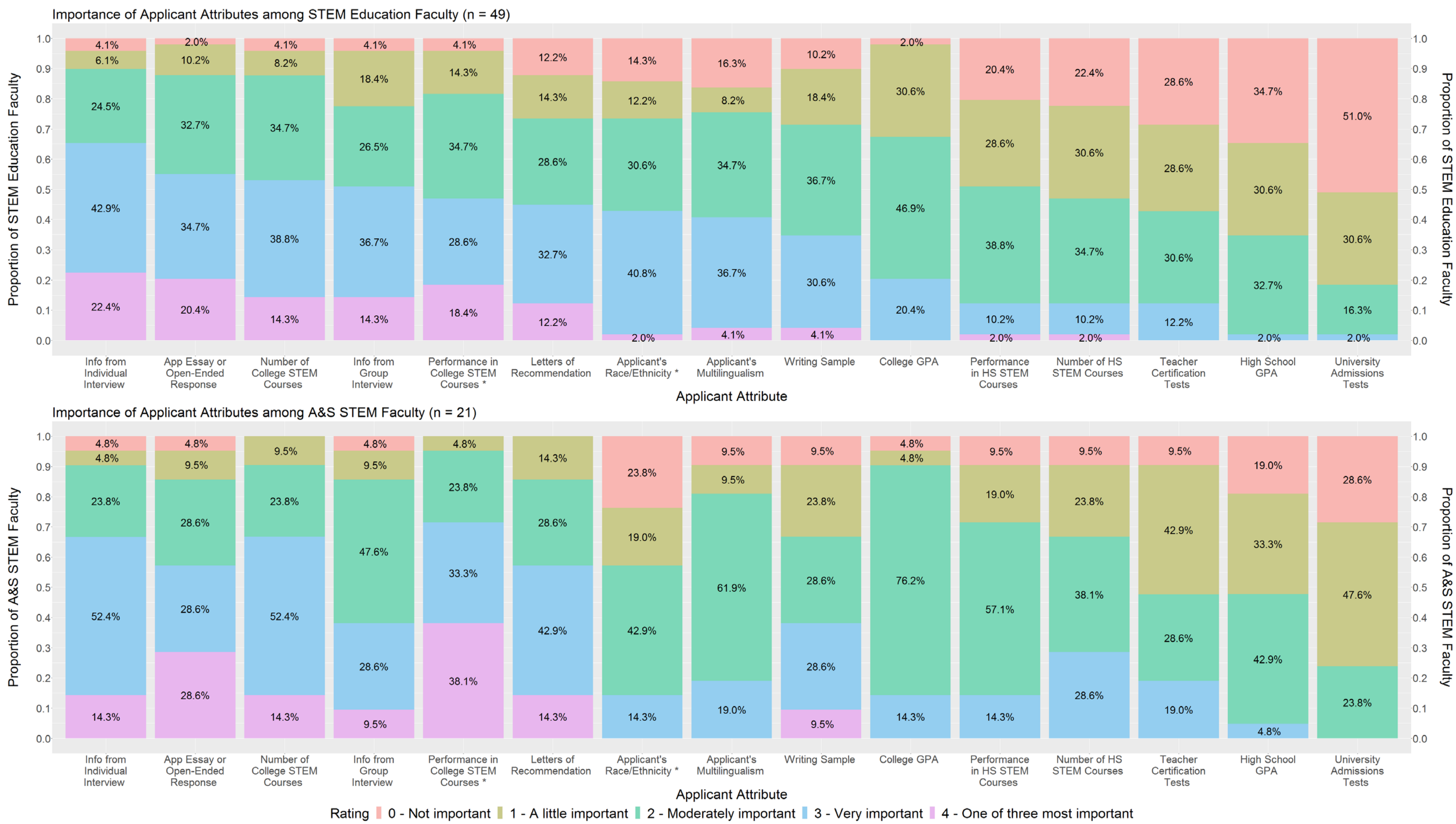
Note. Bars ordered from left to right in terms of proportion of respondents who selected the experience or disposition as "Very important" or "One of the three most important"

Figure 5. Conjoint Analysis Results, All STEM Faculty



Note. Marginal effects from conditional logistic regressions reported in column 1 of Table 3.

Figure 6. Importance of Applicant Attributes Among STEM Education and A&S STEM Faculty



Note. *Bars ordered from left to right in terms of proportion of STEM Education Faculty respondents who selected the source of information as "Very important" or "One of the three most important." P-values from two-sided t-test of the difference in the proportions who select source of information as "Very important" or "One of the three most important" between the STEM Education and A&S STEM groups: *p<.05.*

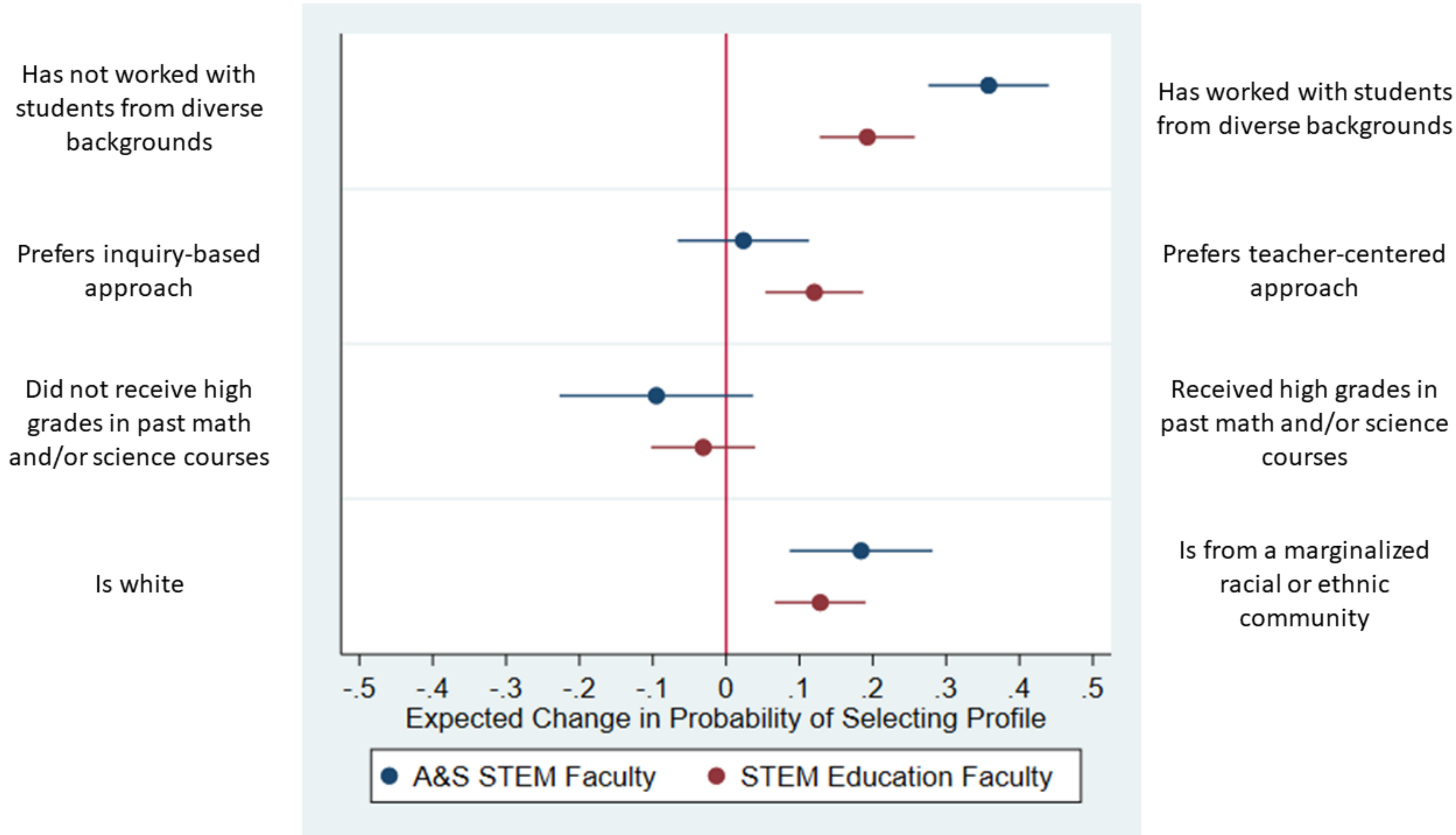
Figure 7. Potential Areas of Concern Among STEM Education and A&S STEM Faculty

*Note. Bars ordered from left to right in terms of proportion of STEM Education Faculty respondents who selected the potential area of concern as a "red flag." P-values from two-sided t-test of the difference between the proportion who select the potential area of concern as a "red flag" between the STEM Education and A&S STEM groups: *p<.05.*

Figure 8. Valued Experiences or Dispositions Among STEM Education and A&S STEM Faculty

Note. *Bars ordered from left to right in terms of proportion of STEM Education Faculty respondents who selected the experience or disposition as "Very important" or "One of the three most important." P-values from two-sided t-test of the difference in the proportions who select experience or disposition as "Very important" or "One of the three most important" between the STEM Education and A&S STEM groups: *p<.05.*

Figure 9. Conjoint Analysis Results, STEM Education and A&S STEM Faculty



Note. Marginal effects from conditional logistic regressions reported in columns 2 and 3 of Table 3.

Table 1. Survey Summary Statistics, Full Sample

Panel A. Response Rates			
	Respondents (N)	Non-Respondents (N)	Response Rate
Participating Institutions			
Institution 1	16	10	0.62
Institution 2	7	9	0.44
Institution 3	10	9	0.53
Institution 4	16	10	0.62
Institution 5	21	5	0.81
Faculty Groups			
Group 1 (STEM faculty)	20	17	0.54
Group 2 (Education faculty)	22	5	0.81
Group 3 (Other faculty)	28	21	0.57
All faculty	70	43	0.62
Panel B. Non-Random Response Tests			
	Respondents (prop)	Non-Respondents (prop)	t-statistic
Participating Institutions			
Institution 1	0.23	0.23	-0.05
Institution 2	0.10	0.21	-1.51
Institution 3	0.14	0.21	-0.88
Institution 4	0.23	0.23	-0.05
Institution 5	0.30	0.12	2.48*
Faculty Groups			
Group 1 (STEM content faculty)	0.29	0.40	-1.18
Group 2 (STEM education faculty)	0.31	0.12	2.65*
Group 3 (Other faculty)	0.40	0.49	-0.91

Note. p-values in Panel B based on two-sided t test, *p<.05; **p<.01; ***p<.001

Table 2. Agreement Between Admissions Team and Self-Reported Categories

	A&S STEM faculty (self-identified)	STEM education faculty (self-identified)	t-statistic
Faculty Groups			
Group 1 (STEM content faculty)	0.62	0.14	3.98***
Group 2 (STEM education faculty)	0.24	0.35	-0.93
Group 3 (Other faculty)	0.14	0.51	-3.45***
Observations	21	49	

Note. p-values based on two-sided t test, *p<.05; **p<.01; ***p<.001.

Table 3. Conjoint Analysis Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Categories	All	By faculty role		By most common course			By admissions influence	
Faculty Group	All	A&S STEM	STEM education	Education	Content for teachers	Not for teachers	No influence	Any influence
Worked vs. has not worked with students from diverse backgrounds	0.237***	0.358***	0.193***	0.175***	0.376***	0.247**	0.317***	0.164***
	(0.0296)	(0.0419)	(0.0331)	(0.0315)	(0.0656)	(0.0755)	(0.0385)	(0.0339)
Teacher-centered approach vs. inquiry-based approach	0.0935***	0.0235	0.120***	0.178***	0.00773	0.00486	0.0764	0.113***
	(0.0281)	(0.0458)	(0.0340)	(0.0389)	(0.0498)	(0.0670)	(0.0437)	(0.0334)
Received vs. did not receive high grades in past math and/or science courses	-0.0423	-0.0951	-0.0312	-0.0453	-0.102	0.00150	-0.113*	0.0224
	(0.0324)	(0.0673)	(0.0362)	(0.0412)	(0.0625)	(0.0889)	(0.0455)	(0.0372)
Is from a marginalized racial or ethnic community vs. Is white	0.141***	0.184***	0.128***	0.136***	0.153**	0.122	0.110**	0.161***
	(0.0266)	(0.0497)	(0.0317)	(0.0377)	(0.0478)	(0.0804)	(0.0399)	(0.0339)
N	700	210	490	380	210	110	380	320

Note. Average marginal effects calculated from conditional logistic regressions predicting the profile selected as a function of the four sets of potential comparisons. P-values based on two-sided t test: *p<.05, **p<.01, ***p<.001.