MASDERING ATTITUDE RESEARCH IN STATISTICS AND DATA SCIENCE EDUCATION: INSTRUMENTS FOR MEASURING STUDENTS, INSTRUCTORS, AND THE LEARNING ENVIRONMENT

Douglas Whitaker[1], Alana Unfried[2], Leyla Batakci[3],
Marjorie E. Bond[4], April Kerby-Helm[5] and Michael A. Posner[6]
[1]Mount Saint Vincent University, Canada
[2]California State University, Monterey Bay, USA
[3]Elizabethtown College, USA
[4]Pennsylvania State University, USA
[5]Winona State University, USA
[6]Villanova University, USA
douglas.whitaker@msvu.ca

Research about students' affective outcomes (such as attitudes) in statistics courses has proliferated over the past three decades, but questions about the impact of instructors and the learning environment on student attitudes remain open. In data science education, research about students' attitudes is nascent. In many statistics education studies, developing items about individual and course characteristics receives less attention than developing other aspects of the study. Without a reliable way to measure characteristics of individuals and courses we cannot identify barriers to student success in statistics and data science–much less dismantle those barriers. This paper describes the development process that the Motivational Attitudes in Statistics and Data Science Education Research (MASDER) team has used for items measuring individual characteristics to be used across the family of instruments. Further work – including some results from a large data collection in the United States – will be presented at the conference.

INTRODUCTION

The existence of a topic for "Promoting Inclusion in Statistics and Data Science Education" implicitly recognizes that barriers exist and that we, as a field, should be doing better to ensure equitable access for our students. Barriers to educational success exist across disciplines, and the notion that barriers can result in systemic differences in educational outcomes between groups has been widely known for decades ("the achievement gap"; e.g., Ladson-Billings, 2006). While some barriers may present unique challenges for statistics and data science education (such as access to and familiarity with technology), students enrolled in statistics and data science courses face the same barriers as students enrolled in other STEM courses. What distinguishes statistics and data science from other fields, then, is how educators address these barriers in statistics and data science education research and in interventions specific to these fields.

For all the focus on developing psychometrically valid instruments to measure constructs in statistics education, there tends to be a paucity of information provided in research articles about the development of the demographic questionnaires that are completed along with the psychometric instrument. The demographic categories provided are then more likely to be outdated, incomplete, or otherwise specific to a particular context; this would result in the results being more difficult to compare with the results from other studies. These demographic questions should not be an afterthought: instead, they should be intentionally chosen in ways to maximize the quality of the data, comparability of findings, and visibility of underrepresented groups. It is only by appropriately collecting these data that statistics and data science educators will be able to determine where barriers exist and whether their interventions are working to dismantle the barriers.

This paper describes the development of a family of surveys being developed to measure student and instructor attitudes and the environment in which they interact in statistics and data science education. This family of surveys includes both psychometric instruments and questionnaires that provide supplementary information. By making the focus on demographic and other questionnaire-type items explicit and making our final materials publicly available, the research team hopes to support the important work of measuring barriers to student success.

In: EM Jones (Ed.), *Fostering Learning of Statistics and Data Science*
Proceedings of the Satellite conference of the International Association for Statistical Education (IASE),
July 2023, Toronto, Canada. https://iase-web.org

OVERVIEW OF MASDER

The goal of the Motivational Attitudes in Statistics and Data Science Education Research (MASDER) project is to develop a family of six surveys. For statistics and data science, a survey of motivational attitudes (SOMA) will be created for both students (S-SOMAS and S-SOMADS, respectively) and instructors (I-SOMAS and I-SOMADS, respectively). Inventories to document salient aspects of the learning environments, pedagogical choices, and institutional characteristics for statistics and data science courses (E-SOMAS and E-SOMADS, respectively) are also being created. All surveys will be made freely available under an open license, and a website is being developed to facilitate administering the surveys (http://sdsattitudes.com/). The most similar extant survey in terms of population of interest is the Survey of Attitudes Toward Statistics (SATS-36; Schau, 2003) which is widely used in undergraduate statistics education research (Whitaker et al., 2022), and many comparisons that follow will be to the SATS-36 because of its widespread use.

The S-SOMAS/DS and I-SOMAS/DS instruments are primarily psychometric instruments consisting of items measuring distinct psychological constructs, but they also include a questionnaire part to gather demographic data from respondents (referred to as characteristics questions [CQs]). It is these demographic questionnaires that will be explored more in this paper. Additionally, some aspects of the E-SOMAS/DS will be explored. Due to a staggered development process (see Whitaker et al., 2019), the statistics instruments are being developed first to inform the development of the data science instruments. The discussion below will center on the development of the S-SOMAS and I-SOMAS but the S-SOMADS and I-SOMADS will ultimately use the same demographic items. This work is funded by the US National Science Foundation (NSF DUE-2013392), and because of this the work is situated within an explicitly US context. However, we believe that the psychometric aspects of the project will apply outside of this context, and the development process used to develop CQs can be adapted by researchers in other countries to account for local needs.

INDIVIDUAL CHARACTERISTICS QUESTIONS

There are four broad areas in which we want to collect demographic information from respondents: gender/sex, race/ethnicity, language proficiency, and area of study. For each of these areas, the development process for the CQs will be described, and the most recent version of the CQs (as of Summer 2023) will be presented.

Gender/sex

Sex and gender are distinct concepts that have been conflated for many years: a person's *sex* is based on their anatomy and physiology, while a person's *gender* is based on factors such as that person's identity, their behaviors related to sex characteristics and expression, and the sociocultural expectations about sex and gender (National Academies of Sciences, Engineering, and Medicine [NASEM] Committee on Measuring Sex, Gender Identity, and Sexual Orientation et al., 2022). Both sex and gender have historically been treated as binary variables (male/female), though both concepts are in fact multidimensional, and the items used to measure them should reflect this (NASEM, 2022). While sometimes collecting both sex and gender is necessary (such as in clinical settings; NASEM, 2022), the MASDER project is collecting data in educational settings where sociocultural experiences of gender are more salient than an individual's sex traits; therefore, we chose early on in the project to measure only gender rather than sex.

Like other disciplines, statistics education research has conflated sex and gender and used dichotomous variables to measure them. For example, SATS-36 (Schau, 2003) includes a demographic item that reads "Your sex:" with the options "1. Male" and "2. Female" though results from the instrument are often contextualized in terms of gender (e.g., Hilton et al., 2004; Ramirez et al., 2012). In other cases, results are reported with results disaggregated by male/female categories without any information provided about the item(s) used to collect the data. This problem is not merely rhetorical: a recent study by the Pew Research Center (Brown, 2022) estimated that 3% of American adults aged 18-29 identify as nonbinary and a further 2% identify as being a trans man or trans woman. Considering the age of many participants in undergraduate statistics education research studies, evidence that perhaps 5% of respondents will not be presented with a way to answer a question about sex/gender that matches their experience presents – at best – a data credibility problem.

Wherever possible, the MASDER team intends to draw on the best practices in the literature rather than developing idiosyncratic items. Our goal with this characteristic is to identify uniform CQs that can be used on any of the surveys being developed. In the first full S-SOMAS pilot study (with 662 respondents), we used a single item to measure gender that we developed to be an improvement upon the types of items mentioned in the previous paragraph: "What gender do you identify with?" Four response options were provided: "Man", "Woman", "I prefer to self-identify:" [free text box], and "Prefer not to answer". These options were selected by 222, 422, 6, and 12 respondents, respectively. After these data were collected, we found recommendations by Spiel et al. (2019, p. 64) for improving the collection of gender data and adopted their recommendations about response options. Our most recent gender item is therefore "What is your gender?" with the options "Woman", "Man", "Non-binary", "Prefer not to disclose", and "Prefer to self-describe" [free text box]. In the Fall 2022 pre-course data collection with 1967 respondents, these options were selected by 1220, 689, 24, 25, and 9 respondents, respectively. Note that we are seeing more than 1% of respondents who wish to disclose their gender selecting options outside of Woman/Man: an appreciable number of respondents seem to be using the "Non-binary" response option and we believe that its inclusion has improved our data collection. Spiel et al. recommend using a multiple selection item instead of a single selection item, but we opted to use a single selection item instead. This choice was made to simplify the data analysis, and the presence of a free text box allows people with multiple identities to express them.

Race/ethnicity

Categories used for collecting race and ethnicity data differ from country to country, and the MASDER project is focused on data collection within the United States due to our funding source. We hope that by illustrating the process we used to select items to measure race and ethnicity that other researchers in the United States and worldwide will be able to improve their approach to collecting such data in their own studies. On S-SOMAS Pilot 1, we used a single item with categories inspired by the US Government's Office of Management and Budget's (OMB) preferred categories (United States Census Bureau, 2022). While the OMB prefers separate questions about race and ethnicity, we opted for a single question to lower respondent burden and a survey that was already lengthy. The Pilot 1 item was "Please specify your ethnicity. Select all that apply." with the following response options: 1) "Caucasian/White", 2) "Black or African American", 3) "Latino or Hispanic", 4) "Asian", 5) "Indigenous American or American Indian or Alaska Native", 6) "Native Hawaiian or Pacific Islander", 7) "Other Ethnicity:" [free text box], and 8) "Prefer not to answer". After this pilot administration, we met with other statistics education researchers who had conducted large studies (e.g., Chance et al., 2022) to discuss their approach to various considerations; they gave us permission to use their item. We chose to adopt their item for Pilots 2, 3, and 4:

What is your race or origin? Select all that apply.

- *White*: German, Irish, Lebanese, Egyptian, etc. (1)
- *Black or African-American*: African American, Haitian, Nigerian, etc. (2)
- *Hispanic, Latino or Spanish origin*: Mexican, Mexican American, Puerto Rican, Cuban, Argentinean, Dominican, Salvadoran, Spaniard, etc. (3)
- *American Indian or Alaskan Native*: Navajo, Mayan, Tlingit, etc. (4)
- *Asian*: Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, Hmong, Laotian, Thai, Pakistani, Cambodian, etc. (5)
- *Native Hawaiian or Pacific Islander*: Native Hawaiian, Guamanian, Somoan, Fijian, etc. (6)
- *Some other race or origin*: Provide race(s) or origin(s) below. [free text box] (7)
- Prefer not to answer (8) (B. Chance & N. Tintle, personal communication, September 21, 2021).

On S-SOMAS Pilot 2 we also asked a separate question "Are you of Hispanic, Latino, or Spanish origin?" with the options "Yes, I am of Hispanic, Latino, or Spanish origin.", "No, I am not of Hispanic, Latino, or Spanish Origin.", and "I prefer not to answer." because of the OMB's preferred two-question approach to race and ethnicity. However, this seemed to confuse some survey respondents and we dropped the item in subsequent data collection rounds.

A considerable amount of time was spent by the MASDER team in determining how race/ethnicity should be measured. While the US government has an official two-item format, there is ongoing work to improve the official collection of this data to account for societal changes (United States Census Bureau, 2022), and so we did not feel that we should restrict ourselves to an official pair of items if we have an opportunity to collect better data. We did feel that it was important for the options presented to reflect officially recognized categories, which all of the items we piloted did. We found that Chance and Tintle's item with the examples helped to clarify the categories for respondents; we believe this to be especially important for international students who may be unfamiliar with US conceptions of race/ethnicity. We also considered official items from the US Census, but we found that these items tended to be much more detailed than the data we were looking for (e.g., providing separate options for "Japanese", "Chinese", "Korean", and "Vietnamese") – analyzing select-all-that-apply data is already challenging, and too much specificity does not further our research goals. We also opted to drop the term "Caucasian" because of its origins as a term "in a beauty-based hierarchy with implied superiority" (Rambachan, 2018, p. 907). Based on data from the free text options, we considered adding a "Middle Eastern countries" category, but ultimately decided not to because we wanted to keep parity with the official OMB categories, we were already capturing this data with the free text options, and we believed that we would likely end up aggregating this category anyway. Future researchers may wish to consider the merit of adding additional categories.

Language proficiency

English language proficiency is a common barrier to educational success at the university level, particularly for international students (Wu et al., 2015). Though not widely studied, there is a growing literature on issues surrounding language proficiency in statistics education (e.g., Lesser & Winsor, 2009; Sharma, 2019). In some prior studies of students' attitudes using the SATS-36, language proficiency has not received explicit attention; instead, the SATS-36 includes an item which could be used as a proxy for language proficiency: "Your citizenship:" with the options "US citizen", "Foreign student", and "Other". The MASDER team wished to avoid a question that could be interpreted by respondents as being about their legal status (which may increase nonresponse bias, among other issues). Instead, we endeavored to find an appropriate question that focused on language proficiency directly.

In S-SOMAS Pilot studies 1, 2, and 3, we included the item "Which language(s) do you speak fluently? Select all that apply." with the response options "English", "Spanish", "French", "Chinese", "Other language(s):" [free text box], and "Prefer not to answer". These response options were chosen based on the team's perception of what common responses might be. The research team periodically analyzed the free text responses to consider other languages that might be included in the list based on common responses; the following languages were entered by at least 20 students during a pilot study: Arabic, German, Hindi, Japanese, Korean, Polish, and Vietnamese. The research team was hesitant to add additional languages though because we anticipated aggregating all of the non-English responses anyway because English-language proficiency was our primary focus.

For S-SOMAS Pilot 4 and I-SOMAS Pilot 1, we focused on rewriting this question to focus on English-language proficiency. However, we ultimately decided that determining English-language proficiency was *not* the primary research need: instead, we need to determine students' proficiency in the language that their statistics course is taught in. This began as a hypothetical distinction, but while investigating university websites for information about their statistics courses as part of developing a plan for randomly sampling universities we discovered that there are introductory statistics courses that are routinely taught in Spanish in the United States. With this in mind, we reconceptualized the student item as "This course is taught in a language in which" with the response options "I am a native.", "I am fluent, but not a native.", "I am proficient.", "I am conversant at an intermediate level.", "I have only basic or little knowledge.", and "I have no knowledge." For I-SOMAS Pilot 1, we framed the question as "The language in which I teach is…" with the response options "My native language", "A language in which I am fluent", "A language in which I am proficient", and "Other" [free text box]. Because this item is intended for instructors, we did not believe that including options such as "I have only basic or little knowledge" was appropriate. These options were developed by the research team; we considered using terms from standardized language tests such as the TOEFL but there did not seem to be a single set of labels that felt appropriate for this item. Moreover, we were

concerned that basing the response options too closely on a standardized test might confuse students who are native speakers of English. These response options might be adjusted by future researchers by using terms that have some external meaning. Moreover, if researchers are conducting a study where it is known that all courses will be taught in a single language, such a question could be reframed to focus on knowledge of that language. (The research team is aware that developing an English-language survey that accounts for use in non-English language courses is not ideal, and perhaps these questions will be revisited if there is interest in translated versions of the instruments.)

Area of study

An astounding number of areas of study are available at the undergraduate level across all universities. One approach to collecting data about this is to provide students with a list of majors or areas of study that have been curated by the researchers. For example, the SATS-36 includes a list of 12 areas of study: "Arts/Humanities, Biology, Business, Chemistry, Economics, Education, Engineering, Mathematics, Medicine/Pre-Medicine, Psychology, Sociology/Social Work, Statistics, Other" (Schau, 2003, p. 4). Other statistics education researchers have asked students to select broad categories such as "Social Sciences", "Natural and Applied Sciences", "Arts and Humanities", "Undeclared/Undecided", and "Other (please specify)" [free text box]; examples of specific majors were provided for the first three broad categories (B. Chance & N. Tintle, personal communication, September 21, 2021). The MASDER team wanted to identify a list to present students that was sufficiently specific as to allow nearly all students to find a category that fit them well and to allow distinguishing between majors such as Statistics, Mathematics, and Data Science.

In S-SOMAS Pilot 1, we opted to include only a single free response question: "What is your major(s)?" [free text box] with the intention of using the responses to help develop this list empirically. The US National Center for Education Statistics maintains a Classification of Instructional Programs (CIP; 2020) where each program of study is assigned a code; these codes are hierarchical and at the third level are quite granular. We had originally considered having students use a drill-down item to choose their specific program within this hierarchy, but this was determined to be too complicated. Ultimately, the results from the free response question, the CIP code hierarchy levels, and specific research goals (such as distinguishing between Statistics and Data Science) were used to devise a list of 40 areas of study that are grouped by area. This list is used three times to allow students to select a first major, second major, and a minor. For the first major, the item is "Pick the field that best describes your major or intended major." with these response options: Undecided; Computer Science; Data Analytics/Business Analytics; Data Science; Mathematics; Statistics; Accounting; Business Administration, Management; Economics; Human Resources; Management Information Systems; Marketing, Advertising; Agricultural, Animal, Plant, and Veterinary Science; Biology; Chemistry; Engineering; Environmental Science and Natural Resource Studies; Geological and Earth Sciences; Health Professions and Related Programs; Kinesiology and Fitness Studies; Physics; Physiology; Anthropology; Homeland Security, Criminal Justice, and Related Fields; Military Science and Leadership; Political Science; Psychology; Public Administration; Social Work and Human Services; Sociology; Architecture; Area, Ethnic, Cultural, Gender, and Group Studies; Communication, Journalism, and Related Programs; Education; English Language and Literature; Foreign Languages, Literatures, and Linguistics; History; Liberal Arts and General Studies; Philosophy and Religious Studies; Visual and Performing Arts; Other (Please Specify) [free text box]. The presentation of this list is still not optimal, but we believe that these categories provide sufficient granularity to capture many students' majors; we can then aggregate categories with few responses based on similarity of fields (e.g., by using the CIP hierarchy as a guide). Not all students select one of these options even when we feel that there is a clear match: for example, 45 respondents indicated that their field of study is "nursing" using the free text box despite the "Health Professions and Related Programs" option appearing in the list. We will continue to tweak the presentation of this list (and the list itself) while understanding that there will be data cleaning tasks to perform manually.

DISCUSSION

Asking students questions about their gender, race or origin, language proficiency, and major may seem straightforward, but these are all multidimensional constructs that become increasingly difficult to measure in standardized ways that can be used outside of a local context (e.g., nationally).

Without such standardized items, the results of studies can be difficult to compare, and identifying barriers to education and inclusivity becomes harder – much less documenting progress on dismantling these barriers. The MASDER team has carefully developed items for measuring these characteristics that we hope will inform other research projects and result in more inclusive learning environments for all statistics and data science students. Even if the specific items presented here are not appropriate for a researcher's study, we hope that the process we followed and considerations we accounted for can be helpful for crafting new or modified questions.

Beyond these questions focused on individuals, the E-SOMAS/DS questionnaires are being similarly developed to document characteristics of statistics and data science courses in a standardized way. By measuring course characteristics in a standardized way, we hope that other barriers to education can be identified and dismantled – and we anticipate that these questionnaires might even have utility outside of statistics and data science education.

REFERENCES

Brown, A. (2022, June 7). *About 5% of young adults in the U.S. say their gender is different from their sex assigned at birth*. Pew Research Center. https://pewrsr.ch/3Qi2Ejd

Chance, B., Tintle, N., Reynolds, S., Patel, A., Chan, K., & Leader, S. (2022). Student performance in curricula centered on simulation-based inference. *Statistics Education Research Journal*, *21*(3), 4. https://doi.org/10.52041/serj.v21i3.6

Committee on Measuring Sex, Gender Identity, and Sexual Orientation, Committee on National Statistics, Division of Behavioral and Social Sciences and Education, & National Academies of Sciences, Engineering, and Medicine. (2022). *Measuring Sex, Gender Identity, and Sexual Orientation* (N. Bates, M. Chin, & T. Becker, Eds.; p. 26424). National Academies Press.

Hilton, S. C., Schau, C., & Olsen, J. A. (2004). Survey of Attitudes Toward Statistics: Factor Structure Invariance by Gender and by Administration Time. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*(1), 92–109. https://doi.org/10.1207/S15328007SEM1101_7

Ladson-Billings, G. (2006). From the Achievement Gap to the Education Debt: Understanding Achievement in U.S. Schools. *Educational Researcher*, *35*(7), 3–12.

Lesser, L. M., & Winsor, M. S. (2009). English language learners in introductory statistics: Lessons learned from an exploratory case study of two pre-service teachers. *Statistics Education Research Journal*, *8*(2), 5–32.

National Center for Education Statistics. (2020). *The Classification of Instructional Programs*. Institute of Education Sciences. https://nces.ed.gov/ipeds/cipcode/default.aspx?y=56

Rambachan, A. (2018). Overcoming the Racial Hierarchy: The History and Medical Consequences of "Caucasian." *Journal of Racial and Ethnic Health Disparities*, *5*(5), 907–912.

Ramirez, C., Schau, C., & Emmioğlu, E. (2012). The Importance of Attitudes in Statistics Education. *Statistics Education Research Journal*, *11*(2), 57–71.

Schau, C. (2003). *Survey of Attitudes Toward Statistics (SATS-36)*. http://evaluationandstatistics.com/

Sharma, S. (2019). Language Challenges and Strategies for English Language Learners in Statistics Education: An Overview of Research in This Field. *Education Quarterly Reviews*, *2*(3), 651–665. https://doi.org/10.31014/aior.1993.02.03.96

Spiel, K., Leuven, K., & Haimson, O. L. (2019). How to Do Better with Gender on Surveys: A Guide for HCI Researchers. *Interactions*, *26*(5), 62–65. https://doi.org/10.1145/3338283

United States Census Bureau. (2022, June 9). *Research to Improve Data on Race and Ethnicity*. Census.Gov. https://www.census.gov/about/our-research/race-ethnicity.html

Whitaker, D., Unfried, A., & Bond, M. (2019). Design and validation arguments for the Student Survey of Motivational Attitudes toward Statistics (S-SOMAS) instrument. In J. D. Bostic, E. E. Krupa, & J. C. Shih (Eds.), *Assessment in Mathematics Education Contexts: Theoretical Frameworks and New Directions* (1st ed., pp. 120–146). Routledge.

Whitaker, D., Unfried, A., & Bond, M. E. (2022). Challenges associated with measuring attitudes using the SATS family of instruments. *Statistics Education Research Journal*, *21*(1), Article 4. https://doi.org/10.52041/serj.v21i1.88

Wu, H., Garza, E., & Guzman, N. (2015). International Student's Challenge and Adjustment to College. *Education Research International*, *2015*, 1–9. https://doi.org/10.1155/2015/202753