

Measuring educational disadvantage in economics education: A research practitioners' toolkit for international research

Jana Sadeh^{a,*}, Malte Ring^b, Douglas McKee^c

^a*University of Southampton, Department of Economics, United Kingdom*

^b*University of Tübingen, Department of Economics Education, Germany*

^c*Cornell University, Department of Economics, USA*

Abstract

Sociodemographic inequalities shape economics students' performance and experience, yet their measurement remains fragmented and context-dependent. Differences in how background is conceptualized and operationalized across countries and institutional contexts limit comparability and hinder the identification of systematic inequalities and subgroup effects. This paper develops a harmonized toolkit for measuring sociodemographic characteristics in international higher economics education research. Drawing on a structured review, we identify core dimensions of inequality in higher education and outline key measurement challenges, including trade-offs between different indicator types and practical constraints in survey design. Based on these considerations, we propose a compact set of ten items covering gender, language, caregiver education, migration background, educational experience, minority status, responsibilities outside of studies, financial concerns and health. The toolkit is evaluated through an expert validation study combining quantitative ratings and international instructor feedback, providing a flexible and standardised baseline for comparative research on inequality in economics education.

Keywords: Scholarship research, survey design, diversity

*The authors would like to thank all the economics educators who participated in the expert validation of this toolkit, especially those who took a substantial amount of time to participate in our focus groups. Their input has been invaluable. The authors would also like to thank Laura Ahlstrom for her comments on an earlier version of the manuscript.

*Corresponding author.

Email address: jana.sadeh@soton.ac.uk (Jana Sadeh)

1. Introduction

Ask any economics educator whether background shapes student success and the answer is an unequivocal yes, but student background is complex, and the better we measure it the better we can address sources of potential educational disadvantage. A substantial body of research shows that gender, race and ethnicity, parental education, and socioeconomic background all predict participation, field choice, and attainment in higher economics education (Arnold, 2020; Happ et al., 2021; Paredes Fuentes et al., 2023; Simson, 2021; Vincent and Hlatshwayo, 2018). These characteristics matter because they proxy deeper structural disadvantages that shape not only academic performance but also students’ attitudes, expectations, and mindsets (Bahník and Vranka, 2017). The practical stakes are high: knowing which groups fall behind is a precondition for directing institutional support effectively, and because interventions such as active learning or small-group tutoring differ in their effectiveness across student groups, comparable measures of background are needed to judge which teaching approaches actually narrow achievement gaps. Measuring sociodemographic background well is therefore not a peripheral methodological concern—it is foundational to almost everything economic educators do.

Yet the measurement of sociodemographic background in economics education research is highly fragmented. Existing instruments are typically anchored in a single national context, with categories, definitions, and question formats that do not travel well. Race and ethnicity categories designed for the United States, and derived constructs such as “underrepresented minority,” are uninformative or not applicable in much of the world. In some countries the salient social divisions are caste or religion rather than race, while in others historical sensitivities make ethnic self-categorisation itself problematic (e.g., Morning, 2008). Common proxies for educational disadvantage, such as first-generation university attendance, are too coarse in some systems and too demanding in others, and seemingly simple items such as parental education require response scales that span contexts where many parents have not completed formal schooling and others where graduate degrees are common (Spiegler and Bednarek, 2013). Meanwhile, several other dimensions of disadvantage—health, language, migration background, household resources—are often not captured at all, even in data-rich settings.

This fragmentation carries real costs. Most evidence in economics education comes from single institutions, single courses, and almost always single countries, so the external validity of any one study is confined to the setting that produced it. When the same construct is operationalised differently across studies, data cannot be pooled into the larger, more representative international samples that would permit stronger inference, and it becomes impossible to distinguish globally relevant educational dimensions from context-specific ones. The consequences extend beyond research: educators seeking institutional investment in evidence-based teaching reforms are routinely asked whether the underlying evidence can be trusted, and larger, more externally valid studies are central to answering that question credibly.

Responding to this fragmentation is particularly challenging as it requires a delicate tradeoff between breadth and depth. Detailed, country-specific, survey questions maximise local relevance at the expense of comparability, while generic questions risk missing the inequalities that matter most in a given context. More detailed questions quickly become context-dependent, undermining the very comparability they were meant to serve. Attempting to capture greater depth and breadth in one go is similarly self-defeating as the result is a long and onerous survey which depresses response rates through survey fatigue, particularly when these background items are added to a set of study-specific survey questions. What is needed instead is a concise, harmonised core: a small set of items capturing the major dimensions of potential educational disadvantage, designed so that standard national measures—such as United States race categories or first-generation status—can

be recovered from the harmonised responses, allowing the same items to serve cross-national pooling and single-country research alike.

Our study contributes an empirically informed and practically usable instrument that supports consistent data collection in international economics education research and could be adapted for use across discipline-based education research more broadly. The toolkit is deliberately foundational rather than exhaustive: researchers requiring depth in a particular dimension can extend it, and although it is not itself a composite index, the comparable measures it uses enable the construction of one.

To develop the toolkit, we have drawn on the international literature on higher economics education and broader education research, on existing survey instruments, and on the expert knowledge of educators, constructing an initial set of ten items measuring potential educational disadvantage. These span gender, language, migration background, parental and caregiver education, minority status (including social and religious minorities), health, and financial resources. We have evaluated this initial toolkit through international expert validation, combining quantitative ratings with qualitative feedback from economics educators working across diverse educational systems. In this paper we document both the design choices and the dimensions deliberately excluded, such as expanded gender categories, the separation of physical and mental health, and technology access.

2. Capturing Educational Disadvantage in an International Context

2.1. Why do sociodemographics matter in surveys?

In economics education, sociodemographic background shapes multiple dimensions of students' trajectories. Differences emerge in participation and continuation (Mare, 1980), field and institutional choice (Emerson et al., 2012), and academic performance (Bottan et al., 2022). Where systematic disparities are not explained by ability, this indicates a misallocation of talent with long-run consequences for productivity, growth, and equity. When access to and success in high-return fields such as economics depends on background rather than potential, human capital investment is unlikely to be efficient. Accurate measurement of background is therefore essential for both research and practice. Including background-related items in student surveys allows researchers to distinguish constraints and barriers from genuine preferences, identify differences between substantively meaningful student groups, and analyze heterogeneous effects of pedagogical interventions. Without such information, observed differences in field choice or academic performance risk being misattributed to motivation or ability rather than different starting points or differential access to resources, information, or institutional support. Yet which background characteristics are relevant, and how they should be measured, varies substantially across institutional and cultural contexts.

2.2. Sociodemographic characteristics in economics education

The uneven distribution of economics education scholarship studies across regions illustrates the asymmetries in the development of the field, shaping how sociodemographic background is documented and analyzed. Additionally, the international visibility of such research might differ between contexts. In some systems, relevant studies may be published primarily in national languages or embedded in local policy debates, and might therefore be underrepresented in the international literature. From the research that is available, the following dimensions emerge as key themes for educational disadvantage.

Gender is a global category when investigating economics learning outcomes. Research frequently examines gender differences in subject choice, continuation, learning experiences, and academic performance (Emerson et al., 2012). The mechanisms underlying these disparities may differ across contexts and include subject stereotypes, prior curricular choices, or differences in confidence

or sense of belonging (Bayer et al., 2020; Allgood and McGoldrick, 2023). Nevertheless, a substantial international literature documents persistent gender-related inequalities in participation and outcomes within economics (Haag and Brahm, 2025).

Ethnic, migration and language-related inequalities are highly context dependent. In North America, race and ethnicity are central categories and are usually measured using federally standardized classifications (Bayer and Wilcox, 2019) that allow for within-country comparisons but might not translate directly to other contexts. In Europe, nationality and migration background (of students and their families) are often investigated, reflecting immigration histories, integration debates, and inequalities linked to prior schooling (Happ et al., 2021). This is connected to home language, which can play an important role, particularly in countries with multiple official or spoken languages that differ from the language of instruction and/or assessment (Xulu-Gama and Hadebe, 2022). In other settings, indigenous status, caste, or religious minority status may play a similar role (Tierney et al., 2019).

Across regions, **socioeconomic background** is among the consistently relevant dimensions of inequality in education. Students from less financially advantaged families may face constraints, weaker access to information, fewer educational resources and lower familiarity with higher education pathways. Research commonly draws on indicators such as parental education, occupation, income or eligibility for financial-aid to approximate these underlying resource differences (Stansbury and Schultz, 2023). Regarding parental education, studies in higher education in the global north often focus on first generation students, whereas studies in the global south might include completion of primary or secondary schooling, reflecting lower average educational attainment in older generations.

This is closely connected to **regional differences and international mobility**. Within larger countries, students from rural or peripheral areas may face lower school quality, fewer nearby universities, higher transport or relocation costs, and weaker access to information networks (Simson, 2021). International students are also an important group of interest, although highly heterogeneous themselves: they may face challenges related to language, educational system familiarity or social integration (Dolan, 2012) but in some contexts (particularly high-fee systems) they may also represent comparatively advantaged groups from their countries of origin.

Additional constraints not captured by traditional sociodemographic categories can also shape students' ability to participate and perform. Students who combine study with paid work (Callender, 2008) or care responsibilities (Armstrong-Carter et al., 2022) face competing time demands that reduce flexibility and academic engagement. Finally, **health conditions** can limit attendance, concentration, or continuity of study (El Ansari and Stock, 2010). While these dimensions are less frequently operationalized as background variables in economic education, they constitute real structural barriers to equal participation in higher economics education.

Taken together, the literature reveals some consensus on the main dimensions of background characteristics in higher economics education, including gender, socioeconomic background, ethnic classification or migration history, prior educational opportunity, responsibilities outside of studies, and health-related barriers. At the same time, these dimensions are rarely integrated in a consistent manner within empirical studies due to differences in conceptual frameworks, as well as variation in data availability, institutional contexts, and measurement approaches across countries and studies. As a result, evidence on student disadvantage is often fragmented across separate strands of the literature, each focusing on specific aspects of background. This fragmentation complicates cross-study and cross-country comparison and raises challenges for developing harmonized instruments that can capture multiple dimensions of inequality in a consistent way. In an educational scholarship landscape where international research is becoming more common, this is an important gap.

2.3. Measurement challenges and practical constraints

Measuring sociodemographic background in higher education involves a set of fundamental trade-offs that shape how inequality can be captured across contexts. These trade-offs concern technical measurement decisions and questions about what it means to operationalise “background” in internationally comparable research. We focus on two interrelated challenges: objective versus subjective indicators and practical constraints in higher education survey settings. Together, these considerations form the basis for the design of a harmonised toolkit.

Methodologically, the measurement of socioeconomic variables raises a fundamental distinction between objective indicators and subjective measures (Avvisati, 2020). Objective indicators aim to capture observable or externally verifiable characteristics such as ethnic categories, household income, financial-aid eligibility or parental education. Their main advantage lies in their transparency: categories are typically clear, interpretable and closely connected to institutional realities within a given country. At the same time, the meaning of objective indicators depends on national institutions, administrative classifications and welfare systems and therefore complicate cross-country comparability. Moreover, objective measures do not necessarily capture how individuals experience their own position. Students in different countries with similar formal resources may perceive their opportunities, constraints, or relative disadvantage quite differently depending on local norms and peer environments. Subjective measures, by contrast, ask respondents to evaluate their own circumstances, for example, whether their family was financially comfortable or whether they see themselves as belonging to a lower or higher social status group (Goodman et al., 2001). A key advantage of subjective measures is that they can be formulated more consistently across countries and capture perceived disadvantages and relative deprivation. These measures, however, depend on respondent’s ability and willingness to assess their own status accurately. Perceptions are shaped by reference groups, expectations, social desirability and cultural norms. As a result, students with similar objective circumstances may report different levels of disadvantage, while individuals facing substantial structural barriers may normalize their situation and underreport hardship. Additionally, subjective categories can be difficult to translate into targeted policy responses.

A second set of challenges arises from the practical context of surveying students in economics education. Survey length constraints require a limited number of items, creating a trade-off between coverage and feasibility. Instruments that are too extensive may reduce response rates and data quality, and are also less likely to be adopted by other researchers or included within existing survey instruments. This is particularly relevant as sociodemographic background questions are often added as supplementary modules to broader student surveys focusing on other research questions. In such settings, researchers may be reluctant to include items that are not tailored to their specific institutional context. A further practical challenge concerns the inclusion of survey items originating from secondary (middle/high school) education research in studies in higher education settings. Secondary students typically still live in their parental household. In that context, questions about “home” or “household” circumstances provide a relatively direct proxy for family resources. In higher education, however, students are more heterogeneous in their living arrangements, and identical questions may no longer clearly refer to parental background, current living situation, or financial dependency. This creates ambiguity not only in interpretation but also in response accuracy. As a result, researchers face a trade-off between conceptually relevant but potentially unknown information (e.g., parental income) and more accessible but less precise proxies (e.g., perceived financial constraints). Finally, higher education samples are themselves selected populations: students already represent a filtered subgroup due to prior educational sorting processes. As a result, observed group differences within higher education do not necessarily reflect population-level inequalities, complicating interpretation of background effects.

These methodological and practical trade-offs are reflected in the design of existing higher education survey instruments. Large-scale surveys of student well-being, such as EUROSTUDENT (Hauschildt et al., 2024), have long faced the challenge of collecting socio-demographic information that is sufficiently detailed to capture student diversity while remaining feasible to administer across institutions and countries. As a result, they provide useful examples of how researchers have balanced objective and subjective indicators, questionnaire length, and cross-context comparability in practice. Although these instruments are primarily designed for monitoring student populations rather than evaluating pedagogical interventions, they provide well-tested question wording and operationalisations that informed the development of the present toolkit. At the same time, these surveys typically contain substantially larger socio-demographic modules than are practical for studies in which background characteristics are collected primarily as control variables, reinforcing the need for a concise harmonised question set.

Taken together, these methodological and practical challenges indicate that no single measurement strategy can fully capture socio-demographic background in internationally comparable higher education research. They therefore motivate the development of a concise harmonised toolkit, whose design principles are outlined below.

2.4. Design criteria for the toolkit

In response to these challenges, the toolkit was developed according to a set of pragmatic design principles intended to balance conceptual relevance, cross-context comparability, and practical feasibility.

- First, the toolkit should provide broad coverage of key sources of disadvantage observed across educational systems, while acknowledging that no fixed instrument can fully represent all locally salient groups or context-specific forms of inequality. It is therefore intended as a common baseline that can be supplemented by country- or project-specific items where necessary.
- Second, the toolkit should remain brief. Because sociodemographic measures are often included as secondary components of larger student surveys, excessive length reduces implementation and response quality. The instrument was therefore designed as a compact module of approximately ten items.
- Third, the toolkit should demonstrate cross-context validity. Items should be understandable, culturally interpretable, and substantively relevant across different countries and higher education systems. This requires avoiding overly context-bound terminology while retaining enough specificity to measure meaningful differences.
- Fourth, the toolkit should support standardized quantitative analysis. Wherever possible, response formats were designed to allow direct coding and comparison across studies, minimizing the need for extensive post-hoc recoding of open-text responses.
- Lastly, the toolkit is designed as a set of complementary items, where individual questions capture distinct dimensions of background, while their combined use provides a more holistic representation of sociodemographic disadvantage than any single indicator could achieve.

Taken together, these criteria reflect a deliberate trade-off between exhaustiveness and usability. In practice, highly detailed country-specific measures often improve local precision but reduce comparability and implementation, whereas overly generic instruments risk losing substantive relevance. The toolkit therefore aims to capture the core dimensions of sociodemographic background in a form that researchers are likely to adopt in international economics education research.

3. Expert Validation Methodology

Building on the literature review and the design criteria outlined above, we developed an initial set of harmonised sociodemographic items intended to capture core dimensions of inequality in higher economics education (Appendix A). Rather than treating the toolkit as fixed from the outset, we used the initial version as a point of departure and repeatedly revised it in response to contextual feedback and empirical validation. To evaluate the clarity, relevance, and applicability of the proposed items across different higher education settings, we conducted a multi-stage validation process combining quantitative survey feedback with qualitative focus group discussions involving economics educators from different geographical regions.

3.1. Instructor Survey

The first draft of the toolkit was hosted on the Qualtrics survey platform and disseminated in a regionally targetted manner to economics educators around the world. They were asked to evaluate each of the 10 questions on three dimensions, the justification for inclusion of the question (is that particular dimension of interest to economics scholarship?), the relevance of the question to their educational context (would it provide insight into their student cohort?) and the clarity of the question wording (is the question easy to understand?). Respondents rated each dimension on a 5-point Likert scale from “Not well at all” to “Extremely well.” In addition, we asked open ended questions after presenting each of the 10 survey questions to allow respondents to suggest amendments specific to that question. We also asked respondents to evaluate the set as a whole, whether it was relevant to their institutional context and to rank the questions from most to least relevant in their particular institutional context. Finally we collected some information about the respondents themselves, the country they were from, the institution they worked in and the number of years they had been teaching economics in higher education. The full survey text can be found in Appendix B. Following this stage we used both the quantitative responses as well as the qualitative comments to redesign the 10-question instrument to address the concerns and feedback raised. This revised questionnaire was then evaluated by three additional focus groups.

3.2. Focus groups

We conducted three focus groups, all administered online. Respondents were recruited using maximum variation sampling (Patton, 2014), which allowed us to capture a range of educational perspectives across a wide geographic spread. In each session, we presented participants with the revised 10-question instrument and asked a set of structured questions designed to establish whether the revisions were substantial enough to make the instrument usable in their own institutional settings. By the third session we had reached saturation: participants requested very few further changes, and responses converged on most items in the toolkit. The full protocol appears in Appendix D.

4. Data and Summary Statistics

4.1. Survey Respondents Profile

We received 38 usable responses to the initial survey, with a broad range of countries participating. A list of regions and the number of responses received from each can be seen in Table 1

Region	Count
Africa	4
Asia	6
Australasia	3
Europe	13
Central & South America	2
North America	10
Total	38

Table 1: Counts of survey responses by region

The educators responding to the survey were mostly very experienced, with 29 of the 38 having over 10 years of experience in higher education as can be seen in Figure 1.

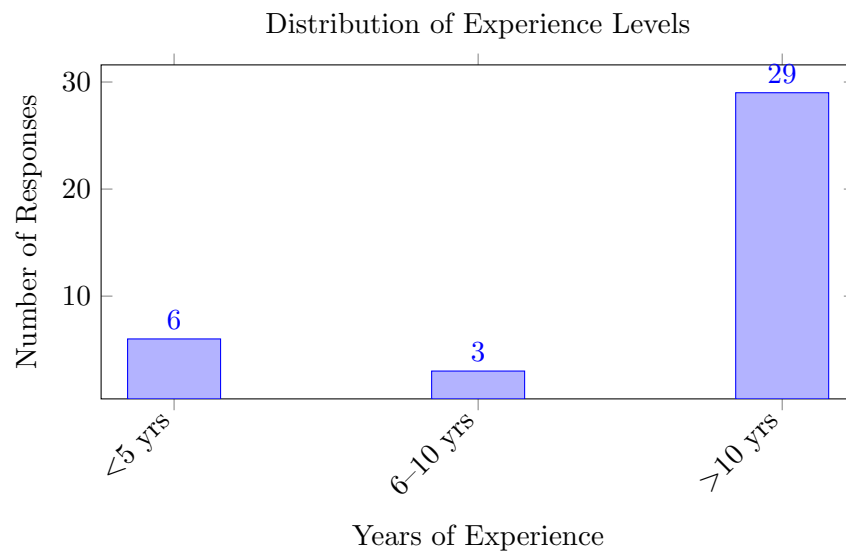


Figure 1: Survey Responses by Experience Level

4.2. Results of First Draft Question Evaluation

The stacked bar charts below show how respondents rated each of the ten toolkit questions along three dimensions: justification (is that particular dimension of interest to economics scholarship?), relevance (would it provide insight into their student cohort?) and clarity (is the question easy to understand?).

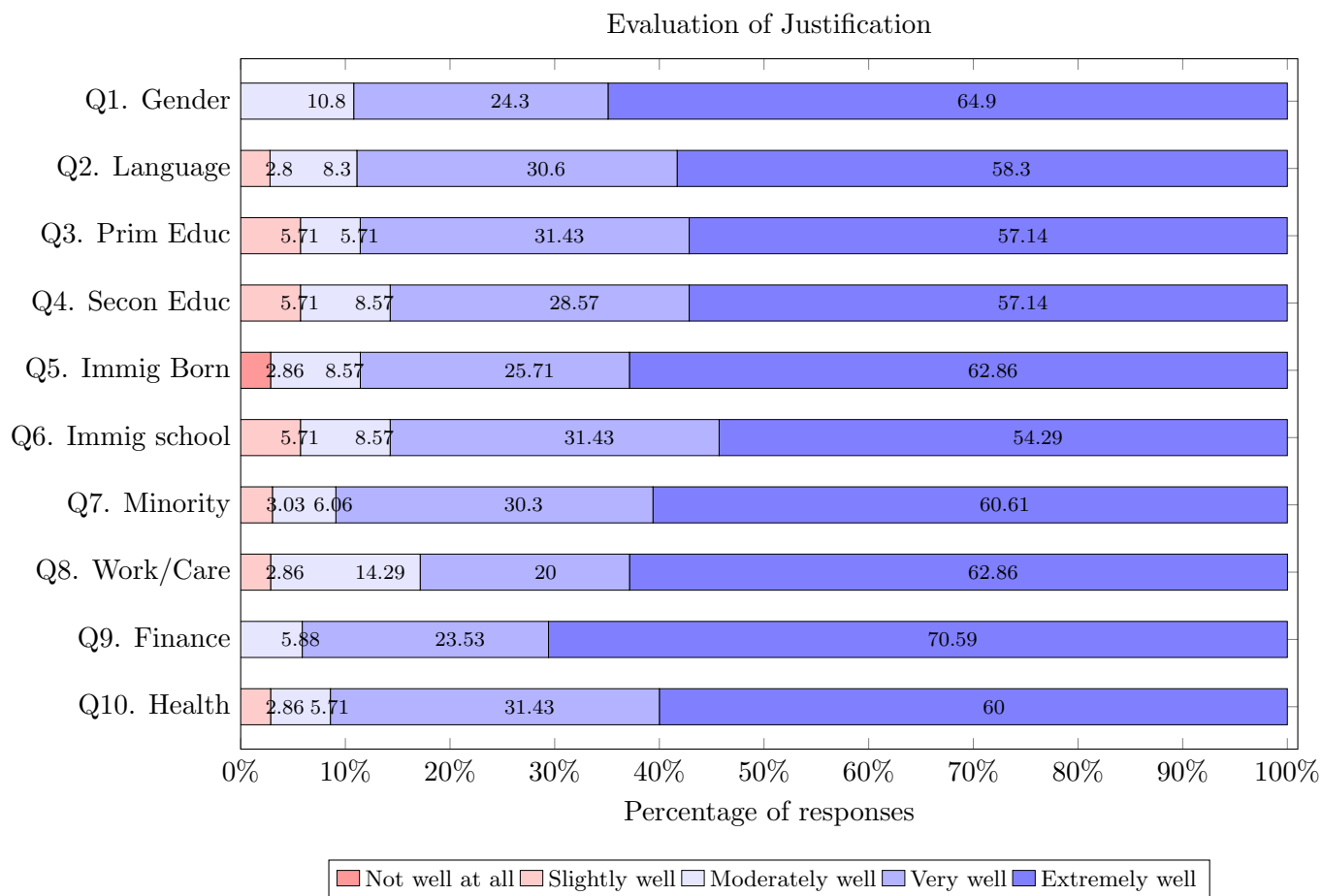


Figure 2: Stacked percentage distribution of Justification ratings across Q1–Q10

From the responses we can see that there is strong support for the inclusion of each of the 10 questions in the toolkit. The question on Finance has the highest level of approval with 94.1% selecting *very well* or *extremely well*. Work/care responsibilities questions scored the lowest, with 82% selecting *very well* or *extremely well*.

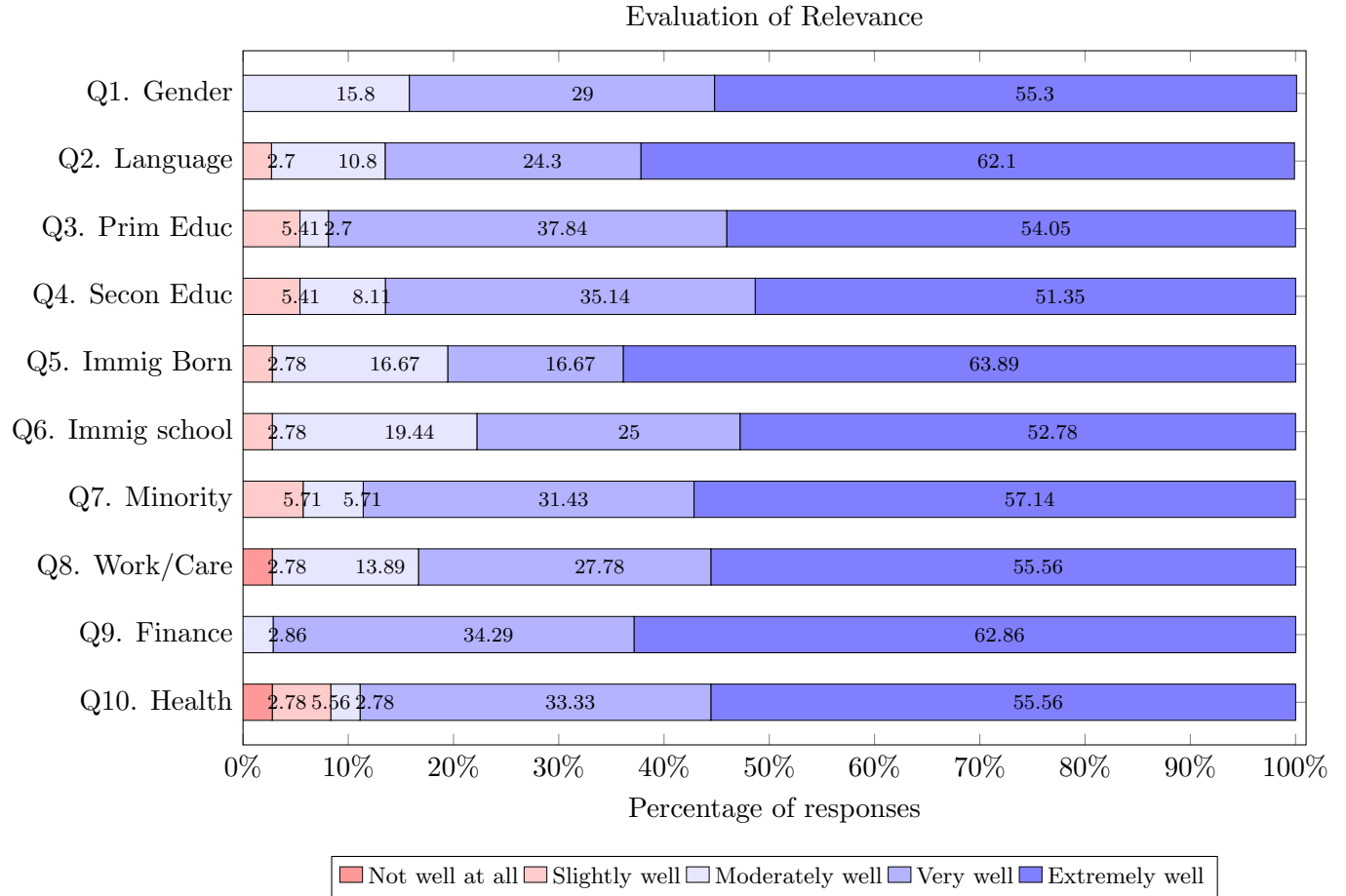


Figure 3: Stacked percentage distribution of Relevance ratings across Q1–Q10

When evaluating the relevance of the questions to their particular student cohort, finance again scores highest, with 97.2% selecting *very well* or *extremely well* while immigration (school) scores lowest with 77.8% selecting *very well* or *extremely well*. Based on respondents' satisfaction with justification and relevance of the questions, we retained all ten, and turned to refining the wording of individual items.

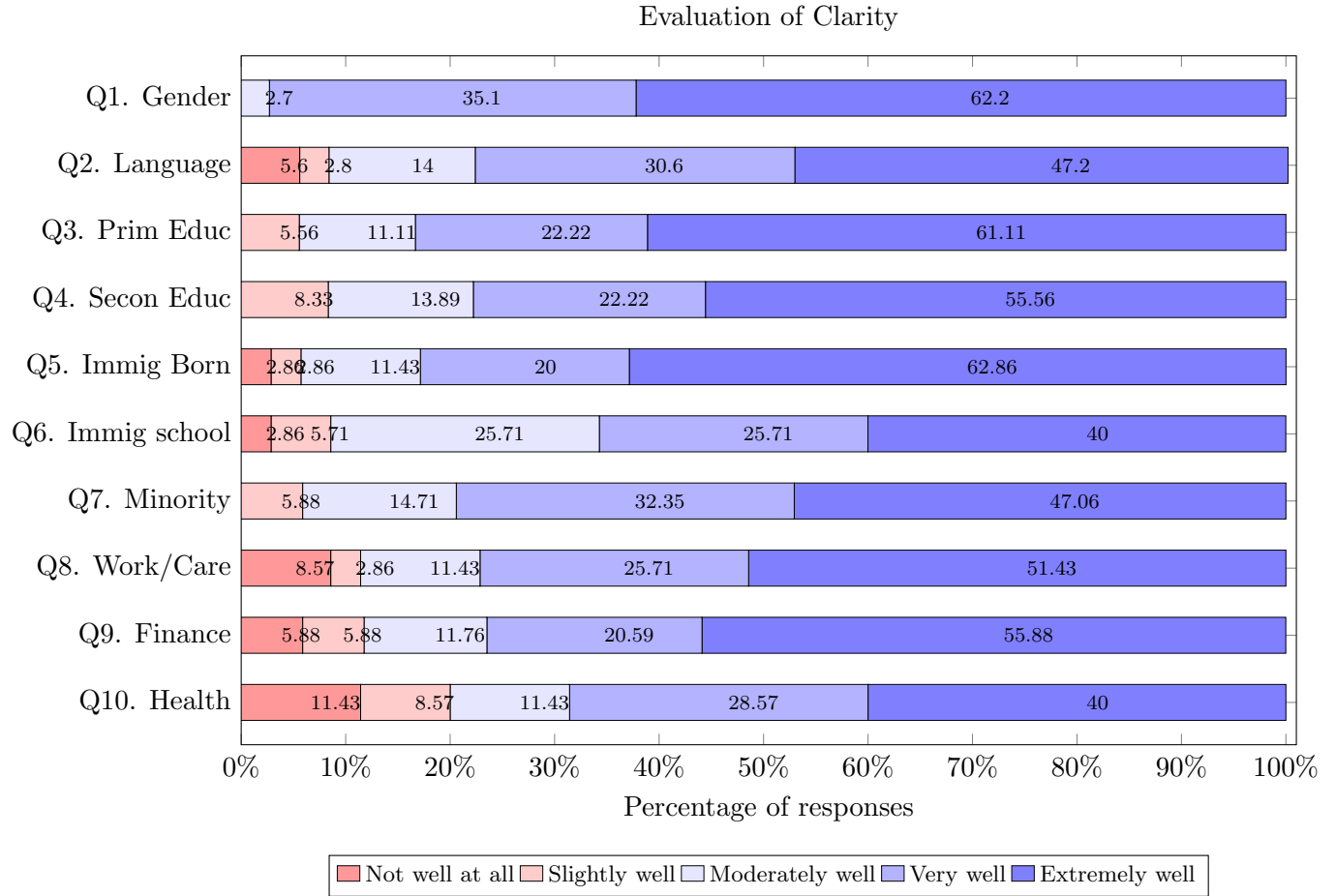


Figure 4: Stacked percentage distribution of Clarity ratings across Q1–Q10

Clarity was the dimension where ratings varied most. The gender question was rated clear by nearly all respondents (97.3% selecting *very well* or *extremely well*), while the immigration (school) question reached only 65.7%. Across the ten items the average was 78.6%, with only the immigration (school) and health questions falling below 70%. The pattern suggested that what remained was a wording problem rather than a selection problem, and the subsequent focus groups took up the wording of these items directly.

4.3. Focus Group Output

The focus groups were carried out sequentially, so that each round could evaluate the toolkit and refine the questions further for comprehension across national contexts. Participants were educators in economics departments around the world, several of whom drew on perspectives from their home countries as well as their current institutions. Between them, the groups involved 13 educators who spanned Canada, China, Colombia, Egypt, Germany, India, Malaysia, New Zealand, Pakistan, Romania, South Africa, Spain, and the UK.

In the first round, three questions were approved by the focus group without any request for amendment. Question 1 (Gender), Question 8 (Working and Caring Responsibilities), and Question 5 (born domestic) were universally accepted. Two questions were approved following targeted and relatively contained revisions. Question 9 (Financial Barriers), originally listing specific expense types (tuition, fees, books, living expenses), was simplified to ask directly how often the student experiences financial concerns while studying, removing complications arising from cross-national variation in tuition fee structures. Question 10 (Health), which had asked how often health reasons prevented the student from performing comparably to peers, was revised to replace the comparative framing with a direct frequency scale — very frequently through never — on the grounds that students cannot meaningfully benchmark their own health difficulties against others.

Five questions generated more substantial debate and reformulation. Question 2 (Language) underwent the most significant conceptual reorientation: the original wording asked whether the language of instruction matched the student’s home language, but participants argued that the question should instead ask whether the student experiences language as a barrier, using a scale from “it is a problem” to “not a problem at all.” A secondary concern about whether instruction and assessment languages should be treated separately was resolved within this subjective-difficulty framing. Questions 3 and 4 (Caregiver Education) required multiple concurrent changes: categories needed to clearly separate vocational post-secondary education from university education (given that the instrument is specifically designed to identify first-generation university students); the term “degree” in Q3 needed replacing with “level of education” for consistency with Q4 and to avoid implying tertiary qualifications only; and instructions for respondents with equally responsible caregivers of the same gender needed to be made explicit across both questions. For Question 6 (Prior Education Domestic), participants noted that the reference to previous “school” was potentially ambiguous. Following a discussion on several alternative formulations the question was revised to ask about the amount of time students had spent in the educational system of the country in which they are currently studying. Question 7 (Minority Status) proved to be the most contentious, with extended debate centering on whether self-identification as an ethnic, racial, or cultural minority is a valid proxy for educational disadvantage. Participants raised cases in which minority groups hold dominant economic or social power, and while a number of proposed alternatives were discussed no consensus on a possible improved question wording was reached. The conclusion was to rewrite the accompanying justification text to acknowledge that the relationship between minority status and educational outcomes may be multi-directional.

In the second focus group, six questions were approved following minor discussion. Question 1 (Gender), Question 2 (Language of Instruction), Question 5 (Born Domestic), Question 6 (Prior Education Domestic), Question 10 (Health) and Question 7 (Minority Status).

Four questions required minor tweaks. Question 9 (Financial Concerns), originally asking how often the student experiences financial concerns while studying, was revised following unanimous agreement to reframe it as “how often does your financial situation interfere with your studies?”, mirroring the threshold-based framing of the health question and avoiding capture of trivial financial worries. Questions 3 and 4 (Caregiver Education) attracted multiple concurrent comments. Participants proposed a preamble question identifying who the primary caregiver is before Q3 is

answered, preserving both the gender-neutral framing valued across diverse national contexts and the analytical distinction between caregiver types valued in developing-country research. Additionally, participants identified the need to add an explicit “no secondary caregiver” response option to Q4, to clarify across both questions that categories refer to completed. Question 8 (Working and Caring Responsibilities) prompted discussion about whether a fixed category list — paid work, caring responsibilities, volunteering, and sports commitments — risked overcomplicating the question while still failing to capture locally prevalent activities such as unpaid work in family businesses. There was strong consensus to simplify the base question to a yes/no response asking whether the student is engaged in activities outside their studies that take ten or more hours per week, with an open-text or researcher-added sub-question available for studies in which the type of activity is analytically relevant.

In addition to the discussions above, it was recommended to move away from a 5 point Likert scale to a movable toggle with always, and never at the extremities as frequency labels such as “rarely,” “sometimes,” and “often” carry different connotations across languages and cultural contexts, generating measurement non-equivalence that undermines cross-national comparison. This change affected 3 questions, Question 2 (Language of Instruction), Question 9 (Financial Concerns) and Question 10 (Health).

In the third round, five questions were approved following minor discussion. Question 1 (Gender), Question 2 (Language of Instruction), Questions 5 (Born Domestic), Question 6 (Prior Education Domestic), and Question 10 (Health). Three questions prompted recommendations for minor revision. For Questions 3 and 4 (Caregiver Education), participants recommended replacing “caregiver” with a term such as “parent/guardian” as more universally intelligible, and proposed adding a category below “no formal education” to capture illiteracy prevalent in some contexts. For Question 8 (Working and Caring Responsibilities), participants recommended clarifying that caring responsibilities extend to elderly relatives and other family members beyond children. Compared to the previous group the participants highlighted the need to separate different responsibilities rather than have a yes-no-base question. For Question 9 (Financial Concerns), participants raised two points worth addressing in the wording: whether framing the question around the individual student alone might cause disengagement in contexts where families rather than individuals manage study costs, and whether the question would benefit from illustrative examples such as technology and study materials. One question remained contested. Question 7 (Minority Status) generated the most substantive debate, with participants raising concerns about political sensitivity and IRB risk in contexts where racial and ethnic categories are officially discouraged, the absence of religious and cultural minority identity as meaningful self-descriptors in some settings, and the risk of near-universal non-response where the word minority carries politically charged or legally complex connotations. A proposal to disaggregate the item into separate open-text fields for ethnicity, religion, race, and cultural background was discussed, as students would be able to click not relevant, if a category would not make sense in their context.

All focus group participants expressed appreciation for the nature of this research, providing evidence of a gap in the literature for sociodemographic questions add-on for research that can be flexible enough to span multiple national contexts.

“We typically are interested in underrepresented groups, so we want to see if there is a difference between these underrepresented groups and the majority. But of course, internationally, it becomes difficult to compare because what’s an underrepresented group changes a lot from country to country. And so it’s just the ability to compare, which becomes really, really crucial.”

“I think this is a fascinating study and actually trying to get on top ... trying to find comparable patterns across countries. I think that would be really fascinating.”

“A study like this would be so important.”

“... I think I could run this here in Spain and run this in Colombia and I would get nice results, which I think is the goal.”

4.4. From Expert Feedback to Final Toolkit

The three focus groups produced a convergent set of recommendations alongside a smaller number of unresolved tensions. Translating this feedback into a final instrument required three further steps beyond straightforward implementation of agreed revisions. Rather than treating focus-group recommendations as binding, the development process viewed them as one source of evidence to be considered alongside the predefined design principles and methodological literature.

The first step was to evaluate all recommendations against the toolkit’s design principles, particularly the goals of international comparability, brevity, and broad applicability. Most wording suggestions that improved clarity without compromising these principles were incorporated directly. Where recommendations conflicted with the underlying design philosophy, explicit trade-offs had to be made. The most significant instance of this was the recurring suggestion that different versions of the toolkit might better serve contexts where several items generate low variance or carry culturally misaligned connotations. Two items were at issue. For the minority question, multiple respondents noted that the category is highly context dependent and that most previous research has used country-specific examples (Morning, 2008). Similarly, several respondents argued that the educational levels in the caregiver education question might not translate well across systems; rather than providing universal categories only, they recommended using ISCED levels alongside country-specific examples. While these concerns are substantively valid, we retained a single universal instrument in service of the comparability goal that motivates the toolkit: country-specific variants would yield data that could not be pooled or compared across contexts.

The second step was to return to the methodological literature on specific design questions raised by the focus groups. The most prominent of these concerned the response format for the frequency items. The second focus group recommended replacing the labeled Likert scale with a movable slider anchored at never and always, on the grounds that intermediate frequency labels carry different connotations across languages. This recommendation was partially endorsed: the decision to anchor only the endpoints and leave intermediate points unlabeled was retained. However, a slider format was not adopted as it is not a standard format and might also be poorly suited to paper-based administration, which remains common in several of the contexts the toolkit might serve. A five-point scale with endpoint labels only was therefore adopted as the standard format.

The third step was a systematic review of the full item set for internal consistency and natural language. Across three rounds of focus group discussions, individual items had been revised in response to specific concerns, sometimes producing inconsistencies in framing, terminology, or response format across questions. This step aimed to harmonise the wording so that items read as a coherent module rather than an accumulated set of patches. In particular, the three frequency items were aligned around a common stem and explanation structure and the terminology was standardised across the toolkit. To improve readability and reduce cognitive burden, questions that required additional clarification were restructured to combine a concise question with a short explanatory description.

The resulting instrument is presented in the following section.

5. Final Toolkit

Table 2: Toolkit item Q1

Q1 Gender
What is your gender? <i>Please select the option that best describes you.</i> ☐ Woman ☐ Man ☐ Prefer to self-describe: _____ ☐ Prefer not to answer
Justification Gender is a key dimension of educational inequality for higher education in general as well as in economics specifically. Including options beyond the binary follows inclusive survey practices and supports equity analyses across diverse populations.

Table 3: Toolkit item Q2

Q2 Language
How often does the language of instruction and/or assessment interfere with your studies? <i>This includes situations where the language used in lectures, seminars, or exams is not your strongest language, making it harder to follow, participate, or express yourself fully.</i> Never ☐ ☐ ☐ ☐ ☐ Always ☐ Prefer not to answer
Justification Language-related barriers can influence access to course materials, participation in class, communication with peers and instructors, and overall academic success. Compared to questions about language spoken at home, this item focuses more directly on experienced academic difficulties and is therefore more applicable across diverse educational settings.

Table 4: Toolkit items Q3/Q4

Q3/Q4 Caregiver Education		
Who are the two people most responsible for raising you?		
	Caregiver 1	Caregiver 2
Mother	☐	☐
Father	☐	☐
Other relative	☐	☐
Other non-related person	☐	☐
Not applicable	☐	☐
Prefer not to answer	☐	☐
What is the highest level of education they have completed?		
Finished education	Caregiver 1	Caregiver 2
No formal education	☐	☐
Primary school	☐	☐
Lower secondary school	☐	☐
Upper secondary school	☐	☐
Postsecondary vocational education	☐	☐
Postsecondary academic education	☐	☐
Postgraduate academic education	☐	☐
I do not know	☐	☐
Not applicable	☐	☐
Prefer not to answer	☐	☐
Justification Caregiver education is a widely used indicator of socioeconomic background and educational capital in both secondary and higher education research. Identifying caregivers separately allows the item to accommodate diverse family structures while maintaining a consistent measure of educational background across cultural contexts.		

Table 5: Toolkit item Q5

Q5 Born domestic			
Were you or your caregivers born in the country where you are currently studying?			
	Me	Caregiver 1	Caregiver 2
Yes	☐	☐	☐
No	☐	☐	☐
I do not know	☐	☐	☐
Not applicable		☐	☐
Prefer not to answer	☐	☐	☐
Justification Migration background may influence students' familiarity with the local education system, language environment, and institutional expectations, potentially shaping academic adjustment and participation in higher education.			

Table 6: Toolkit item Q6

Q6 Prior Education domestic
How many years were you formally educated in the country where you are currently studying? ☐ 0–1 year ☐ 1–3 years ☐ 3–5 years ☐ More than 5 years ☐ Prefer not to answer
Justification Prior educational experience within the current national education system may shape familiarity with academic expectations, institutional structures, and assessment practices. This item complements migration background by focusing on educational pathway rather than country of birth alone.

Table 7: Toolkit item Q7

Q7 Minority				
Do you consider yourself to be part of an ethnic/racial, religious, cultural, or geographical minority in the country where you are currently studying?				
	Yes (please specify)	No	Not applicable	Prefer not to answer
Race/Ethnicity	☐ _____	☐	☐	☐
Religious	☐ _____	☐	☐	☐
Cultural	☐ _____	☐	☐	☐
Geographical (e.g., rural)	☐ _____	☐	☐	☐
Other	☐ _____	☐	☐	☐

Justification
Self-identification as part of an ethnic, racial, religious, cultural or geographical minority provides a context-sensitive measure of group belonging that avoids imposing externally defined categories that may not be comparable across countries. The question does not assume disadvantage but allows researchers to examine whether and under which conditions minority status relates to educational experiences and outcomes.

Table 8: Toolkit item Q8

Q8 Responsibilities outside of studies			
Do you have significant responsibilities outside of your studies?			
<i>By significant responsibilities, we mean obligations that together require more than 10 hours per week and would have serious personal, financial, or social consequences if they were not fulfilled.</i>			
	Yes	No	Prefer not to answer
Necessary paid or unpaid work	☐	☐	☐
Caring for children or family members	☐	☐	☐
Formal sport commitments	☐	☐	☐
Other (please specify)	☐ _____	☐	☐

Justification
This item captures substantial responsibilities outside of study that may compete with students' time, energy, and available resources. By assessing different types of responsibilities separately while applying a common threshold of more than 10 hours per week, the item distinguishes significant obligations from routine extracurricular activities.

Table 9: Toolkit item Q9

Q9 Financial concerns
How often do financial concerns interfere with your studies? <i>By financial concerns we mean concerns about affording study-related costs, for example tuition, fees, books, technology, accommodation, or other living expenses.</i> Never ☐ ☐ ☐ ☐ ☐ Always ☐ Prefer not to answer
Justification Financial concerns may affect students’ access to learning resources, housing stability, academic participation, and persistence in higher education. The item captures perceived financial strain rather than exact income or wealth, making it applicable across diverse contexts.

Table 10: Toolkit item Q10

Q10 Health
How often does a serious health condition interfere with your studies? <i>By serious health condition we mean a long-standing or recurring condition, for example chronic illness, mental health difficulties, mobility or sensory impairment, or learning disabilities.</i> Never ☐ ☐ ☐ ☐ ☐ Always ☐ Prefer not to answer
Justification Serious health conditions can significantly affect students’ ability to attend classes, study consistently, and complete assessments. By focusing on functional academic impact rather than diagnosis, the item enables comparable measurement across contexts with different health systems and disclosure norms.

6. Conclusion

This study has proposed a harmonised toolkit for measuring the characteristics and background factors that may place a student at an educational disadvantage in higher economics education. The toolkit responds to the fragmentation of indicators across studies, institutions, and countries documented above. We do not claim to have resolved the measurement problem, and we doubt any single instrument could. Rather, we offer the toolkit as a pragmatic starting point: a concise, usable set of items sufficient for deployment in any economics education research project that spans multiple countries, and designed to be tested and revised as evidence from that deployment accumulates. It draws together several core dimensions of potential educational disadvantage — gender, language-related barriers, caregiver education, migration and educational background, minority status, work and care responsibilities, financial concerns, and health-related barriers — in a format short enough to embed in studies whose primary focus lies elsewhere. As such, it should permit more consistent operationalisation of background characteristics in international pedagogic research and support more systematic analysis of heterogeneity in educational outcomes.

The initial validation with economics education experts across a diverse range of educational systems is cautiously encouraging: the proposed items are broadly considered relevant and usable across institutional contexts. It also illustrates the design tension that any such instrument must manage. Almost every suggested addition—finer-grained gender categories, the separation of physical and mental health, measures of technology access—is individually well motivated, yet each lengthens the instrument, risks fatiguing or alienating respondents, and introduces detail that is more context-dependent and therefore less comparable. The compromises we have struck, including the cap on the number of items, are informed by the literature and by the accumulated experience of the educators who fed into the process, but they are judgement calls rather than settled science, and we expect some of them to be revised following additional validation. Researchers whose research questions demand depth in a particular dimension can extend the core set with context-specific items, so that standardisation and local/project relevance can coexist. For the same reason, the toolkit is not a composite index of disadvantage, although the comparable measures it provides may eventually support the construction of one.

The most significant limitation of this work is that the validation reported here reflects the perspective of educators and researchers, not of the students who will answer the questions. An essential next step, once the toolkit is deployed, is student-level validation focusing on item comprehension, response processes, and cross-context comparability with more objective measures of these dimensions. Whether established national constructs—such as first-generation status or nationally meaningful ethnicity classifications—can in fact be recovered from the harmonised items without material loss of information is an empirical question that field deployment will answer, as is the behaviour of the items once translated beyond English. Linking the toolkit to administrative or performance data will then allow an assessment of construct validity and of how the different background dimensions relate to achievement, participation, and progression in economics education.

This toolkit should serve as a foundation for more comparable empirical research on inequality in higher economics education. In doing so, it can help distinguish robust patterns of disadvantage from context-specific ones, and invite educators and researchers in settings currently underrepresented in economics education research to take up an instrument designed with their contexts in mind. By lowering the barriers to cross-national work, we hope it will make such research more common, and more genuinely comparable, than it is today.

Ethical approval

Research ethics approval for this study was received from the University of Southampton's Research Ethics Committee (108850).

Generative AI Declaration

During the preparation of this work the authors used Anthropic Claude in order to improve readability and language, for condensing anonymised focus group transcripts and for minor latex coding support. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of Competing Interest

The authors have no competing interests to declare.

References

- Allgood, S., McGoldrick, K., 2023. Understanding sex differences when majoring in economics: What little we know, reasons for knowledge gaps, and a research agenda of unanswered questions. *Advances in Economics Education* 2, 4–25.
- Armstrong-Carter, E., Panter, A., Hutson, B., Olson, E.A., 2022. A university-wide survey of caregiving students in the us: Individual differences and associations with emotional and academic adjustment. *Humanities and Social Sciences Communications* 9, 1–10.
- Arnold, I.J., 2020. Gender and major choice within economics: Evidence from europe. *International Review of Economics Education* 35, 100191.
- Avvisati, F., 2020. The measure of socio-economic status in pisa: A review and some suggested improvements. *Large-Scale Assessments in Education* 8, 8.
- Bahník, Š., Vranka, M.A., 2017. Growth mindset is not associated with scholastic aptitude in a large sample of university applicants. *Personality and Individual Differences* 117, 139–143.
- Bayer, A., Bhanot, S.P., Bronchetti, E.T., O’Connell, S.A., 2020. Diagnosing the learning environment for diverse students in introductory economics: An analysis of relevance, belonging, and growth mindsets. *AEA Papers and Proceedings* 110, 294–98. URL: <https://www.aeaweb.org/articles?id=10.1257/pandp.20201051>, doi:10.1257/pandp.20201051.
- Bayer, A., Wilcox, D.W., 2019. The unequal distribution of economic education: A report on the race, ethnicity, and gender of economics majors at us colleges and universities. *The Journal of Economic Education* 50, 299–320.
- Bottan, D., McKee, D., Orlov, G., McDougall, A., 2022. Racial and gender achievement gaps in an economics classroom. *International Review of Economics Education* 40, 100239.
- Callender, C., 2008. The impact of term-time employment on higher education students’ academic attainment and achievement. *Journal of education policy* 23, 359–377.
- Dolan, M., 2012. Supporting international students of economics in UK higher education. URL: <https://doi.org/10.53593/n2264a>, doi:10.53593/n2264a. edited by Dimitra Petropoulou. Published in the Handbook for Economics Lecturers, The Economics Network.
- El Ansari, W., Stock, C., 2010. Is the health and wellbeing of university students associated with their academic performance? cross sectional findings from the united kingdom. *International journal of environmental research and public health* 7, 509–527.
- Emerson, T.L.N., McGoldrick, K., Mumford, K.J., 2012. Women and the choice to study economics. *The Journal of Economic Education* 43, 349–362. doi:10.1080/00220485.2012.714306.
- Goodman, E., Adler, N.E., Kawachi, I., Frazier, A.L., Huang, B., Colditz, G.A., 2001. Adolescents’ perceptions of social status: development and evaluation of a new indicator. *Pediatrics* 108, e31–e31.
- Haag, L., Brahm, T., 2025. The gender gap in economic and financial literacy: A review and research agenda. *International Journal of Consumer Studies* 49, e70031.

- Happ, R., Nagel, M.T., Zlatkin-Troitschanskaia, O., Schmidt, S., 2021. How migration background affects master degree students' knowledge of business and economics. *Studies in Higher Education* 46, 457–472.
- Hauschildt, K., Gwosc, C., Schirmer, H., Mandl, S., Menz, C., 2024. Social and economic conditions of student life in Europe: Eurostudent 8 synopsis of indicators 2021-2024. wbv Media GmbH & Company KG.
- Mare, R.D., 1980. Social background and school continuation decisions. *Journal of the American Statistical Association* 75, 295–305. doi:10.1080/01621459.1980.10477466.
- Morning, A., 2008. Ethnic classification in global perspective: A cross-national survey of the 2000 census round. *Population research and policy review* 27, 239–272.
- Paredes Fuentes, S., Burnett, T., Cagliesi, G., Chaudhury, P., Hawkes, D., 2023. Who studies economics? an analysis of diversity in the uk economics pipeline .
- Patton, M.Q., 2014. *Qualitative research & evaluation methods: Integrating theory and practice*. Sage publications.
- Simson, R., 2021. Regional inequality in university attainment in seven african countries since 1960. *International Journal of Educational Development* 84, 102427.
- Spiegler, T., Bednarek, A., 2013. First-generation students: What we ask, what we know and what it means: An international review of the state of research. *International studies in sociology of education* 23, 318–337.
- Stansbury, A., Schultz, R., 2023. The economics profession's socioeconomic diversity problem. *Journal of Economic Perspectives* 37, 207–230.
- Tierney, W.G., Sabharwal, N.S., Malish, C., 2019. Inequitable structures: Class and caste in indian higher education. *Qualitative Inquiry* 25, 471–481.
- Vincent, L., Hlatshwayo, M., 2018. Ties that bind: The ambiguous role played by social capital in black working class first-generation south african students' negotiation of university life. *South African Journal of Higher Education* 32, 118–138.
- Xulu-Gama, N., Hadebe, S., 2022. Language of instruction: A critical aspect of epistemological access to higher education in south africa. *South African Journal of Higher Education* 36, 291–307.

Appendix A. 10-Question Student Survey Instrument - version1.0

#	Question	Justification
1	Gender What is your gender identity? (Please select the option that best describes you.) – Woman – Man – Non-binary / Other – Prefer not to say	Gender identity remains a key dimension of educational inequality for higher education in general as well as in economics specifically. Including options beyond the binary follows inclusive survey practices and supports equity analyses across diverse populations.
2	Family language Is the primary language spoken in your family home(s) the same as the language of instruction in your study program? – Yes – No – Prefer not to answer	Language spoken at home can influence access to course materials, participation, and overall academic success, especially when it differs from the language of instruction. Regarding test-performance in economics, this might be more relevant in countries with different languages (of instruction).
3	Maternal education Did your mother (or primary caregiver) complete a university-level qualification? – Yes – No – I don't know – Prefer not to answer	Maternal education is a core component of intergenerational socioeconomic transmission and correlates with children's educational attainment and aspirations. Including this measure aligns with findings from both adolescent and higher education research showing its predictive power over academic success.
4	Paternal education Did your father (or secondary caregiver) complete a university-level qualification? – Yes – No – I don't know – Prefer not to answer	Paternal education similarly contributes to students' educational capital and helps build a comprehensive socioeconomic background measure. Empirical studies underscore that both parents' education levels independently influence educational outcomes.
5	Immigration (First Generation) Were you born in the country where you are currently studying? – Yes – No – Prefer not to answer	Being foreign-born can involve navigating unfamiliar education systems and cultural expectations, which often impacts academic adjustment and performance in higher education. Identifying first-generation immigrant students helps highlight structural barriers related to adaptation and support needs.

Table A.11: Survey Questions and Justifications (1–5)

#	Question	Justification
6	International students Is your previous school in the country where you are currently studying? – Yes – No – Prefer not to answer	Completing secondary education abroad may affect academic preparedness and orientation to higher education norms in the host country. This distinction is relevant given evidence showing varying academic trajectories based on schooling background, even among non-migrant students.
7	Minority Do you consider your family to be part of an ethnic, racial, or cultural minority in the country where you are currently studying? – Yes – No – Not sure – Prefer not to answer	Self-identification as part of an ethnic, racial, or cultural minority can indicate exposure to bias or marginalization, which may influence educational opportunities and experiences. Using a self-identification approach offers more validity across diverse international contexts than administratively defined categories.
8	Working Are you working part-time or full-time to support your studies? – Yes – No – Prefer not to answer	Working alongside studies can reflect both financial necessity and time constraints that may negatively impact academic performance. This item helps capture students who may face opportunity costs and competing demands on their study time. It does not distinguish between formal and informal work but focuses on its presence as a potential barrier.
9	Financial barriers How often do you have concerns about affording tuition, books, or living expenses while studying? – Always – Often – Sometimes – Rarely – Never – Prefer not to answer	Financial concerns are a major source of stress and can directly limit students' access to learning resources, participation in academic activities, or even program completion. This question focuses on perceived financial strain rather than exact income or wealth, making it more suitable for international comparison. It complements the working status question by capturing subjective financial pressure.
10	Health Compared to other students, how often are you unable to perform as a student (attend classes, study, assessments...) due to illness or health-related reasons? – Clearly above average – Slightly above average – About average – Slightly below average – Clearly below average – Prefer not to answer	Chronic illness, mental health challenges, or frequent sickness-related absences can significantly impact academic participation and performance. This item uses a comparative framing to allow students to self-assess their health-related absences relative to peers, which may increase cross-national validity.

Table A.12: Survey Questions and Justifications (6–10)

Appendix B. Qualtrics Instructor Survey

Study Information

International URM Harmonisation

Information Goal: The goal of this project is to develop and test a set of internationally applicable sociodemographic questions that help identify student groups who may face disadvantages in higher educational systems in future EENE studies. Rather than relying solely on indicators of underrepresentation, the focus is on capturing relevant aspects of household background, language, migration, and access to resources. The aim is to improve the ability of future EENE studies to analyze educational inequalities across diverse national contexts, using questions that are both sensitive to local realities and comparable across countries.

Question development and use in EENE: The questions were developed drawing on literature on educational disadvantage, including international assessments such as PISA and studies on underrepresented groups in higher education. They are intended for universal use across future EENE studies to ensure comparability and capture key dimensions of disadvantage in a broad, efficient, and internationally relevant way. In designing them, we aimed to balance robustness with feasibility: while some items rely on self-assessment rather than precise economic indicators, they are more accessible to students, reduce missing data, and support cross-country comparability, even if further validation is needed.

Applicability to international education research: While the immediate goal is to develop a consistent question set for future EENE studies, the overall approach is equally relevant for other international education research projects seeking to capture and compare structural inequalities in access and participation. The questions can be adapted or integrated into broader survey designs in economics and related disciplines to enhance comparative analyses of equity and inclusion in higher education.

Your contribution: We would like to evaluate whether the questions successfully capture what we intend to measure for different countries. We are running this survey with faculty teaching in HE across the world to assess this and to further refine our question set.

Consent

Q. If you want to participate, please read the information sheet found here and tick all the statements below to confirm your consent.

Study Title: International URM Harmonisation

Ethics/ERGO number: 108850

Version and date: V2 17th Oct 2025

- I confirm that I read the Participant Information above (version 2, dated 17th October 2025) explaining the study and I understand what is expected of me.
- I was given the opportunity to consider the information, ask questions about the study, and all my questions have been answered to my satisfaction.
- I agree to take part in this study and understand that data collected during this research project will be used for the purpose of this study.
- I understand that my participation is voluntary and that I am free to withdraw from this study at any time without giving a reason. If the survey is submitted anonymously, then I understand that I will not have the right to withdraw my data once I submit my responses.

Survey Questions

Q. The following are the 10 questions we have designed to identify underrepresented or educationally disadvantaged groups within the higher education student body. For each question, we ask you to evaluate its **clarity** (i.e., whether students in your context are likely to understand it as intended), its **relevance** (i.e., whether it captures meaningful aspects of disadvantage in your national or institutional setting), and its **justification** (i.e., whether the provided rationale for including the question aligns with what it actually measures).

For each dimension, responses are given on a five-point scale: *Not well at all* (1), *Slightly well* (2), *Moderately well* (3), *Very well* (4), *Extremely well* (5).

1. **Gender**

Justification: Gender identity remains a key dimension of educational inequality for higher education in general as well as in economics specifically. Including options beyond the binary follows inclusive survey practices and supports equity analyses across diverse populations.

Evaluation dimensions: Gender clarity; Gender relevance; Gender justification.

Q. Suggestions for how we could reword, adapt, or improve this question:

2. **Language spoken at home**

Justification: Language spoken at home can influence access to course materials, participation, and overall academic success, especially when it differs from the language of instruction. Regarding test performance in economics, this might be more relevant in countries with different languages of instruction.

Evaluation dimensions: Language clarity; Language relevance; Language justification.

Q. Suggestions for how we could reword, adapt, or improve this question:

3. **Maternal education**

Justification: Maternal education is a core component of intergenerational socioeconomic transmission and correlates with children's educational attainment and aspirations. Including this measure aligns with findings from both adolescent and higher education research showing its predictive power over academic success.

Evaluation dimensions: Maternal education clarity; Maternal education relevance; Maternal education justification.

Q. Suggestions for how we could reword, adapt, or improve this question:

4. **Paternal education**

Justification: Paternal education similarly contributes to students' educational capital and helps build a comprehensive socioeconomic background measure. Empirical studies underscore that both parents' education levels independently influence educational outcomes.

Evaluation dimensions: Paternal education clarity; Paternal education relevance; Paternal education justification.

Q. Suggestions for how we could reword, adapt, or improve this question:

5. **Immigration (Born In)**

Justification: Being foreign-born can involve navigating unfamiliar education systems and cultural expectations, which often impacts academic adjustment and performance in higher education. Identifying first-generation immigrant students helps highlight structural barriers related to adaptation and support needs.

Evaluation dimensions: Immigration clarity; Immigration relevance; Immigration justification.

Q. Suggestions for how we could reword, adapt, or improve this question:

6. **International (Previous Schooling)**

Justification: Completing secondary education abroad may affect academic preparedness and orientation to higher education norms in the host country. This distinction is relevant given evidence showing varying academic trajectories based on schooling background, even among non-migrant students.

Evaluation dimensions: International clarity; International relevance; International justification.

Q. Suggestions for how we could reword, adapt, or improve this question:

7. **Minority (Self-identification)**

Justification: Self-identification as part of an ethnic, racial, or cultural minority can indicate exposure to bias or marginalisation, which may influence educational opportunities and experiences.

Using a self-identification approach offers more validity across diverse international contexts than administratively defined categories.

Evaluation dimensions: Minority clarity; Minority relevance; Minority justification.

Q. Suggestions for how we could reword, adapt, or improve this question:

8. Working / Caring Responsibilities

Justification: This question captures competing responsibilities that reduce time available for studying, including both paid work and care for family members. While the financial aspect of paid work is partially addressed by the financial barriers question, care responsibilities represent an additional, non-financial constraint on students' academic engagement. Including both in a single item ensures we measure time-related pressures relevant to academic performance without overloading the survey.

Evaluation dimensions: Responsibilities clarity; Responsibilities relevance; Responsibilities justification.

Q. Suggestions for how we could reword, adapt, or improve this question:

9. Financial barriers

Justification: Financial concerns are a major source of stress and can directly limit students' access to learning resources, participation in academic activities, or even programme completion. This question focuses on perceived financial strain rather than exact income or wealth, making it more suitable for international comparison. It complements the working status question by capturing subjective financial pressure.

Evaluation dimensions: Financial clarity; Financial relevance; Financial justification.

Q. Suggestions for how we could reword, adapt, or improve this question:

10. Health

Justification: Chronic illness, mental health challenges, or frequent sickness-related absences can significantly impact academic participation and performance. This item uses a comparative framing to allow students to self-assess their health-related absences relative to peers, which may increase cross-national validity.

Evaluation dimensions: Health clarity; Health relevance; Health justification.

Q. Suggestions for how we could reword, adapt, or improve this question:

Q. Do you have any overarching suggestions on how we could reword, adapt, or improve these questions for international use? _____

Q. Any other feedback? _____

Overall Assessment of the Question Set

Q. Looking at the ten questions as a set — we expect that together they will allow us to identify (and control for) the subset of students that are at an educational disadvantage.

Q. Thinking of your particular group of underrepresented / minority / disadvantaged students, would these 10 questions manage to distinguish them?

- Yes
- No
- I don't know

Q. Please rank which questions are most relevant to your particular group of underrepresented / minority / disadvantaged students, putting the most relevant questions first and the least relevant questions last.

1. Gender

2. Family language
3. Maternal education
4. Paternal education
5. Immigration
6. International
7. Minority
8. (Care) Work
9. Financial
10. Health

About You

Q. Lastly, we would like to know a little bit about you as an economics educator. This information will only be used to help us interpret the survey results (e.g., by comparing perspectives across different countries and levels of experience).

Q. What country are you from? _____

Q. What is the name of the institution you work in? _____

Q. How long have you been teaching economics in higher education?

- 0–5 years
- 6–10 years
- More than 10 years

Q. To further validate the questions in different countries, we are looking for interview partners. Please type your email here if you would be interested in discussing the questions in your context:

Appendix C. Synthesis of Qualitative Comments in Instructor Survey

Appendix C.1. Gender

Respondents recommended increasing response granularity and replacing “other” with a self-describe option. Several noted that the term “identity” may be unnecessary and suggested simplifying the item to “What is your gender?”. A broader concern relates to how much granularity is meaningful to respondents and researchers, and whether highly differentiated categories could provoke discomfort in some cultural contexts.

Appendix C.2. Language

Feedback emphasised that many students use multiple languages or dialects both at home and within their university setting, where instruction may shift across languages. Some questioned referencing “family home(s)” given that students may not currently live with family; however, this does not undermine the construct since language fluency typically develops in childhood. The key challenge is ensuring the question distinguishes between language(s) of instruction and students’ actual fluency, thereby capturing potential linguistic barriers to study.

Appendix C.3. Education of Primary Caregiver

Two main issues emerged. First, in Western contexts, linking “mother” with “primary caregiver” was viewed as inappropriate; in other cultural contexts, however, “primary/secondary caregiver” terminology may itself be unclear. Second, respondents argued that the educational categories provided do not align well with global educational structures, and in countries where few parents attended university, parental education may not discriminate meaningfully.

More importantly, many respondents felt the purpose of this question was insufficiently explained: the aim is to identify first-generation students and the informational capital available to them, not to measure parental qualifications per se. Alternatives, including regional customisation (e.g., versions with or without “mother”), were proposed as acceptable minimal adaptations.

Appendix C.4. Education of Secondary Caregiver

Similar concerns arose around gender assumptions and cultural variation in caregiver roles. Respondents also highlighted that some students lack a secondary caregiver or have multiple caregivers. Adding “caregiver(s)” and a “not applicable” option was suggested as a potential improvement.

Appendix C.5. Immigration, born in

Respondents argued that the current wording fails to distinguish students born abroad who relocated at different ages, leading to different socialisation experiences. A proposed alternative categorisation differentiates between age-of-arrival bands (e.g., ≤5, 5–10, 10–18) and “born in country of study,” which may better capture meaningful variation.

Appendix C.6. Immigration, previous school

Participants noted ambiguity in the term “previous school,” which could refer to high school or prior tertiary study. Clarifying that any prior schooling in the host country is relevant—because it reflects familiarity with the local educational system—may resolve this. Combined with the birthplace/age-of-arrival question, respondents generally felt this pair of items could capture meaningful differences in educational experience.

Appendix C.7. Minority Self evaluation

Concerns focused on ambiguity for students from mixed backgrounds and the lack of clarity around what constitutes a minority. Participants recommended offering examples and acknowledging that minority status may be regional rather than national. Shifting the focus to self-identification rather than family background may improve clarity and cross-cultural applicability; limited, locally modifiable wording may be required.

Appendix C.8. Working/caring responsibilities

Respondents emphasised that unpaid labour (e.g., family business work) and substantial volunteering may impose time burdens similar to paid work. Some suggested differentiating motivations for working, though others noted this distinction is unnecessary for identifying students balancing work and study. A simplified yes/no question capturing all forms of work and caring responsibilities was proposed, with clearer justification for the construct.

Appendix C.9. Financial

Participants noted that the term “tuition” may not be universally applicable and that contemporary student expenses increasingly relate to software and technology rather than books. Several suggested adopting a temporal frame (e.g., “have you ever...”, “in the past semester...”) to reflect fluctuations in financial stress. Alternatives to “do you have concerns” such as “do you worry” or “are you anxious” were proposed, and respondents debated whether frequency categories or broad descriptors (e.g., never/sometimes/often/always) better capture variation.

Appendix C.10. Health

Respondents questioned comparison to an “average student,” noting students may lack reliable information about peers’ experiences. They also found the existing response scale insufficiently progressive. Clarity could be improved by explicitly referencing both mental and physical health and by reframing the question to focus on functional impact—e.g., frequency of missed classes or deadlines due to health issues—while avoiding overcounting minor acute illnesses.

Appendix C.11. General feedback on more than one question

Some respondents advocated for optional open-text responses to provide richer context, though many noted that such data are rarely analysed and significantly increase respondent burden, reducing completion rates. Broader reflections emphasised that while individual questions may have limitations in particular contexts, the value of the survey lies in combining indicators to produce a multidimensional picture of disadvantage.

Appendix D. Questions used for Focus Group

Opening / Warm-up

Warm-up: Thinking about this survey toolkit generally, what will make it a useful instrument for international HE research?

Looking at the revised questions individually.

1. Gender

Goal: Balance clarity, respect, comparability, and cultural sensitivity.

- Wording: Is the revised wording clear and respectful across contexts? How can it be improved?
- Probes: Does removing the word “identity” improve clarity? Is “What is your gender?” appropriate?
- Granularity: How much granularity is helpful (e.g., self-describe vs. “Other”)? Is there a risk of over-granularity?

2. Language

Goal: Capture multilingual contexts and language of instruction fluency.

- Wording: Is the revised wording supportive of settings with multiple languages or dialects spoken at home?
- Probes: Best format—select all that apply? Write-in?
- Instructional language(s): How can we clarify whether the language(s) of instruction are languages the student is fluent in?
- Residence wording: Is using “family home(s)” confusing or unnecessary?

3. Education of Primary Caregiver

Goal: Measure first-generation status and informational capital, not just credentials.

- Wording: Is the revised wording consistent with the aim to identify first-generation students? If not, what is the clearest way to ask this?
- Cultural fit: Do the standard education levels used map poorly in some countries?
- Probes: What alternative categories or examples would be more accurate?
- Terminology: Does the use “primary caregiver” instead of “mother,” work cross-culturally?

4. Education of Secondary Caregiver

Goal: Acknowledge absence/multiplicity of caregivers and reduce ambiguity.

- Wording: Is the revised wording flexible enough to account for no secondary caregiver or multiple caregivers?
- Probes: Is “caregiver(s)” helpful? Should we include “not applicable / no secondary caregiver”?
- Data utility: Does asking both primary and secondary caregiver education add value, or is first-generation status sufficient?

5. Immigration – Born In

Goal: Distinguish between students who are familiar with their current educational context. This question needs to be paired with the next one to be meaningful.

- Wording: Is the wording clearly identifying students who were born domestically? Would categories like “born abroad and lived there until < 5 / 5–10 / 10–18” vs. “born in country of study” better capture meaningful differences?
- Do we get information that is relevant even if the child moved shortly after birth?
- Interpretability: Would respondents understand and answer these categories reliably? Are gains in granularity at the cost of clarity? Is the granularity needed?

6. Immigration – Previous Schooling

Goal: Distinguish between students who are familiar with their current educational context. This question needs to be paired with the previous one to be meaningful.

- Definition clarity: Do we need to define “previous school” to avoid confusion (e.g., high school vs. prior degree)? Does the distinction matter to familiarity?
- Outcome relevance: Does measuring any experience in the local system capture what matters (familiarity with the system)?
- Probes: Should we add a duration (e.g., at least one year)?

7. Minority – Self Evaluation

Goal: Handle mixed backgrounds, regional vs. national minorities, and clarity.

- Level of analysis: Should the question focus on self-identification rather than clear imposed groups?
- Probes: Does this reduce ambiguity for mixed-heritage students?
- Granularity: How do we capture granularity without losing focus?

8. Working / Caring Responsibilities

Goal: Capture all forms of time burden without unnecessary complexity.

- Scope: Should we capture all types of work (paid, unpaid family business, major volunteering) and caring under a single yes/no?
- Probes: Is the time threshold (≥ 10 hours/week) useful?
- Motivations: Do we need to differentiate motivations (supporting family vs. discretionary income), or is that out of scope?
- Data use: What is the minimal granularity needed to identify the “studying while working/caring” group?

9. Financial

Goal: Use globally clear terms and capture extent of financial stress.

- Terminology: Is “fees” clearer than “tuition” in your context? What other common costs should we name (software, technology, transport)?
- Timeframe: Should we ask “Have you ever...”, “in the past semester...”, or “how often...” to reflect ongoing vs. episodic concerns?
- Probes: Options like “never / sometimes / often / always” vs. specific frequencies (weekly, monthly).
- Wording: Does “Do you worry about...” or “Are you anxious about...” work better than “Do you have concerns about...”?

10. Health

Goal: Capture meaningful academic impact without over counting minor/acute illnesses.

- Construct focus: Would asking “How often do you miss lectures or deadlines because of health (mental or physical)?” better target impact?
- Probes: What response scale feels clear and progressive (e.g., never $\rightarrow$ several times a week)?
- Scope clarity: Should we explicitly include both mental and physical health in the wording?
- Benchmarking: Do we need a benchmark (e.g., “compared to your typical week”) or just focus on frequency?

11. Closing

- Final reflection: What question(s) felt most likely to misfire in your context? What change would fix it?
- Missed areas: Are there important dimensions of disadvantage we are not capturing?