

Research paper

Gendered and Cultural Bias in Elementary School Assessment: DIF as Evidence of Epistemic Inequality

Herwin Herwin ^{1*} , Shakila Che Dahalan ² , Muhammad Ilyas ³ , Ma'rufi Ma'rufi ³ 

¹ Universitas Negeri Yogyakarta, INDONESIA

² Universiti Pendidikan Sultan Idris, MALAYSIA

³ Universitas Cokroaminoto Palopo, INDONESIA

*Corresponding Author: herwin89@uny.ac.id

Citation: Herwin, H., Dahalan, S. C., Ilyas, M., & Ma'rufi, M. (2026). Gendered and cultural bias in elementary school assessment: DIF as evidence of epistemic inequality. *Feminist Encounters: A Journal of Critical Studies in Culture and Politics*, 10(1), 33. <https://doi.org/10.20897/femenc/18548>

Published: May 14, 2026

ABSTRACT

This study examines Differential Item Functioning (DIF) as a method for identifying gender and cultural bias in elementary school assessments. It employs an explanatory mixed-methods research design, with a strong emphasis on critical quantitative analysis supported by conceptual interpretation grounded in feminist theory. This approach was chosen to enable the empirical identification of bias in assessment while simultaneously providing a critical interpretation of the findings. Data were collected through documentation of assessment instruments used in schools and students' responses to test items, with participants drawn from diverse social and cultural backgrounds. Through DIF analysis, the study found that assessment items often advantage male students or those from local cultural contexts, thereby creating epistemic inequality in educational outcomes. These gender and cultural biases indicate that the knowledge recognized in education frequently fails to reflect the diversity of students' experiences. The findings underscore the need for more inclusive assessment practices that accommodate the diversity of students' experiences and challenge dominant norms in education. The study also highlights the importance of DIF analysis as a tool not only for improving measurement quality, but also for uncovering hidden power relations embedded in educational assessment practices, which often marginalize the knowledge of groups. These findings highlight the need to reconsider assessment practices as context-dependent and socially embedded, rather than purely neutral tools. By framing DIF as both a quantitative and critical approach, this study demonstrates its capacity to reveal concealed inequities in educational assessment and to guide the creation of more fair evaluation practices.

Keywords: differential item functioning (DIF), gender and cultural bias, epistemic inequality, inclusive educational assessment, feminist theory in education

Educational assessment, from a feminist lens, is not neutral but a site of knowledge production that mediates power and intersects with gender, culture, and social positioning. Epistemic injustice arises when assessment items privilege particular contexts and ways of knowing, marginalizing others (Buren et al., 2024; Rogers et al., 2024). Studies on epistemic justice emphasize that marginalized perspectives should be recognized as legitimate Knowledge claims within educational practice, demanding reflexivity, reciprocity, and relational accountability in assessment environments (Benschop, 2021; Francis, 2025). A critical approach to assessment must foreground intersectionality, power relations, and decolonial considerations to prevent epistemic exclusion and to

democratize what counts as valid know-how in learning. Where assessment reinscribes hegemonic epistemologies, it reproduces inequality; where it centers marginalized epistemologies, it expands educational justice and inclusion (Bedmar, 2022; Cook, 2025). Thus, reimagining assessment requires explicit attention to epistemic fairness, platforming diverse knowledges, and challenging the coloniality of knowledge embedded in conventional items (Amaral et al., 2025; Chung, 2021).

Assessment in elementary schools is often positioned as an objective and neutral technical instrument for measuring student learning outcomes (Walker et al., 2023). In formal educational practice, assessment is often understood as a value-free, standardized, and fair evaluation mechanism for all students (Laitusis & Karvonen, 2025). However, this view ignores the fact that assessment is never completely separated from the social, cultural, and ideological context in which it is designed and used (Czerwinski-Alley et al., 2024). Assessment questions are the product of a selection process of knowledge, language, and experience that reflects dominant values in society.

As pedagogical texts, assessment items carry implicit assumptions about gender roles, family norms, everyday activities, and cultural experiences considered “common” (Kuchynka et al., 2022). The representation of men and women in the context of the items, the division of domestic and public roles, and references to specific cultural practices indicate that assessments function as gendered cultural practices (Mamtani et al., 2022). Thus, assessments not only measure students' cognitive abilities but also affirm particular forms of knowledge and social experiences as legitimate standards (Balanta-Cobo et al., 2025; Chingarande et al., 2026; Ibrahim et al., 2024; Kuchynka et al., 2022).

In the context of multicultural primary education, not all students have equal access to the social and cultural contexts assumed in the questions (Prince, 2023). This is especially true in schools located in heterogeneous areas with diverse immigrant communities. Students from non-dominant cultural backgrounds, minority communities, or with different social experiences may face barriers in understanding the context of the questions, even if they have comparable academic abilities (Martínez et al., 2023; Scheffers, 2026). This situation places assessments at risk of reinforcing inequalities, not due to differences in learning ability, but due to differences in access to legitimate knowledge (Gibson, 2024; Martínez et al., 2023).

The concept of fairness in assessment is often reduced to uniform procedures and the same assessment standards for all students. However, uniformity is not always synonymous with fairness. Equal assessments for groups of students with different social and cultural backgrounds can result in hidden injustices (Deiri & Burkhard, 2025; Dogutas, 2025; Hauer et al., 2023; Nigar et al., 2025; Rezai et al., 2022). Therefore, fairness needs to be critically understood as substantive justice, namely the extent to which assessments provide equal opportunities for all students to demonstrate their abilities without being limited by gender and cultural biases.

Although critical and feminist scholarship has long highlighted gender and cultural biases in curricula and learning materials, assessment research remains dominated by technical approaches that rarely connect equity issues to power structures of knowledge (Abreu, 2026; Parisi, 2026). Many critiques of assessment bias are normative or descriptive in nature but are not supported by empirical evidence that systematically demonstrates how these biases operate within test items. This gap limits the power of critical research to influence assessment policy and practice.

Differential Item Functioning (DIF) offers an empirical approach to test whether a test item functions differently across groups of students with similar academic abilities (Herwin, 2022). When an item exhibits DIF based on gender or cultural background, the performance differences cannot be explained by ability alone but rather indicate structural biases in the item's construction (Ma et al., 2021; Ohiri et al., 2024; Teresi et al., 2021). Thus, DIF enables researchers to systematically and data-drivingly identify assessment inequities.

Within a critical epistemic framework, DIF findings not only point to technical measurement issues but also reflect epistemic inequalities in education. Items that favor certain groups indicate that the knowledge, experiences, and ways of thinking of dominant groups are legitimized as standards of truth (Acar et al., 2025; Chen et al., 2025; Halpern et al., 2025; Navarro-González et al., 2024). Conversely, the experiences of female students or students from non-dominant cultural backgrounds are at risk of being marginalized, thus making assessment a mechanism that reproduces knowledge inequities starting in elementary school.

Building on this framework, this study aims to examine gender and cultural bias in elementary school assessments through DIF analysis as empirical evidence of epistemic inequality. This study addresses the following research questions:

- 1) Do elementary school assessment items exhibit DIF based on gender and cultural background?
- 2) How do patterns of DIF reflect forms of epistemic inequality in educational assessment?
- 3) In what ways do gender and cultural biases intersect within assessment items?

By integrating psychometric approaches and critical perspectives, this study contributes to broadening the understanding of fairness in assessment, not merely as a technical issue, but as a matter of social justice and the politics of knowledge.

LITERATURE REVIEW

Feminist epistemology

Feminist epistemology critically interrogates the presumed neutrality and universality of knowledge, revealing how social context informs what is deemed "valid" or "testable" knowledge in education. It challenges the inherent biases in assessment practices that can silence diverse perspectives and uphold the interests of dominant cultural groups (Rizzolo et al., 2022; Wright, 2022). For instance, research demonstrates that assessments can perpetuate gender and racial biases. Specifically, the English Comprehension Test has been implicated in emphasizing language differences in ways that may unintentionally disadvantage certain groups (Arendse, 2021).

Furthermore, the alignment of epistemic beliefs with socio-economic backgrounds highlights how educational assessments might privilege certain cultural paradigms over others. Studies indicate that students' epistemic beliefs can be significantly shaped by their educational experiences, reinforcing systemic inequities and biases in assessment environments (Kinneer et al., 2023; Schiefer et al., 2022). Thus, assessments should be reconceptualized as epistemic artifacts that reflect and contribute to power relationships embedded within educational systems, allowing for greater representation and acknowledgment of marginalized voices (Sato, 2023). Incorporating feminist epistemology into educational assessment practices can promote more inclusive and equitable frameworks that validate diverse experiences and forms of knowledge.

Epistemic injustice

Miranda Fricker's concept of epistemic injustice sheds light on the systemic disadvantages faced by students whose lived experiences and knowledge are not acknowledged in assessments. This form of injustice does not stem from cognitive inability but rather from the failure of educational systems to recognize and value diverse knowledge bases (Hao et al., 2024). For example, educational assessments often reflect biases that favor certain cultural or social backgrounds, leading to an inequitable evaluation environment where some students are unable to demonstrate their true capabilities (Hauer et al., 2023; Johnson et al., 2022).

Research indicates that algorithmic biases in educational settings exacerbate these disparities, often perpetuating existing systemic inequalities and disadvantaging marginalized groups (Akgun & Greenhow, 2022; Nezami et al., 2024). Moreover, the reliance on traditional, biased assessment methods hinders fair opportunities for all students to showcase their understanding and skills (Turner et al., 2024). By failing to account for the diverse contexts from which students draw their knowledge, educational assessments can inadvertently contribute to institutionalized forms of epistemic injustice, ultimately depriving students of equitable and just evaluation (Nezami et al., 2024). Addressing epistemic injustice within education requires a reconceptualization of assessment practices that authentically reflect and include diverse perspectives and knowledge.

Cultural capital and educational assessments

Pierre Bourdieu's theory of cultural capital elucidates how educational systems contribute to social inequality by valuing specific cultural knowledge and practices that align with those of dominant groups. School assessments are often designed around the assumptions that favor students whose cultural backgrounds resonate with the dominant school's values, enabling them to navigate these assessments more effectively. For instance, children from advantaged socio-economic strata typically exhibit greater familiarity with the expected behaviors and knowledge in educational contexts, which facilitates their academic performance compared to their peers from less privileged backgrounds (Geven & Zwier, 2025).

When assessments do not recognize or incorporate the cultural capital of students from non-dominant backgrounds, these students encounter structural barriers that hinder their ability to express their knowledge (Farid et al., 2021). This inequity stems not from a lack of cognitive ability but from a disconnection between the lived experiences of marginalized groups and the cultural assumptions embedded in assessment questions (Ren & Liu, 2022). Consequently, the educational system inadvertently reinforces existing social hierarchies, further entrenching the disadvantages faced by these students (Ren, 2024). The implications of Bourdieu's cultural capital are critical for understanding how assessment practices perpetuate educational disparities, emphasizing the need for more inclusive approaches that acknowledge diverse cultural backgrounds.

Gendered knowledge in schooling

The concept of gendered knowledge in schooling highlights how curricula and assessments often reflect unequal representation of men's and women's roles, experiences, and achievements. Specifically, assessments can reflect biases through the choice of contexts, characters, and common activities depicted in questions, disproportionately favoring students who align with dominant cultural and gender norms (Fisher et al., 2025). For instance, when assessment items incorporate scenarios that predominantly reflect male experiences or

interests, this not only reinforces traditional gender roles but also limits the epistemic opportunities available to students who do not conform to these representations (Atias et al., 2025).

Research demonstrates that educational settings continuously reflect and reproduce societal gender norms, privileging a narrow set of experiences that some students may not connect with. This creates barriers not due to a lack of cognitive ability but because the knowledge deemed "valid" aligns more closely with certain gendered experiences, ultimately implementing an implicit exclusion of non-dominant gender identities from what is considered legitimate knowledge in assessments (Marabini, 2023; Stroupe et al., 2024). Furthermore, the consequent epistemic injustice leads to decreased engagement and performance among students who feel alienated by the curriculum and assessment design (Okoroji et al., 2023; Wee et al., 2023). Addressing gendered knowledge in assessments necessitates a critical examination of curricula to foster inclusivity and accurately represent diverse experiences, thereby mitigating entrenched biases and promoting equal epistemic opportunities for all students.

Assessment as social practice

Assessment, as a social and cultural practice, operates as a discursive site that constructs knowledge, privileges certain communities, and reproduces epistemic inequalities (power-relations embedded in contexts, language, and content) (Ruscoe et al., 2023; Wood & Hedges, 2025). Items function as cultural texts whose familiar or unfamiliar contexts shape who succeeds, thereby encoding normative biases and marginalizing alternative knowledges (Chan et al., 2025). Critical analyses of policy texts and curricula show how global discourses are locally genealogized, with policy borrowing needing careful contextualization to avoid reproducing inequities in assessment practices (Te Whāriki, EYFS; England) (Syed, 2024). CDA and CPA frameworks reveal how assessment discourse mediates power, legitimacy, and social justice in educational contexts, including language testing, multilingualism, and equity-oriented assessment design (Yu & Liu, 2025). Consequently, equitable assessment requires making biases visible, challenging deficit framings, and integrating culturally responsive, socially just approaches into assessment design and interpretation (CRA, HOTS, multimodality, and critical literacy perspectives) (Koh, 2023). While some sources converge on the central claim, others emphasize methodological tensions and contextual specificity, underscoring the need for nuanced, locally grounded critique in assessment discourse (Lavoie & Cavanagh, 2023).

Overall, these perspectives form an integrated analytical framework for understanding bias in educational assessment. Feminist epistemology challenges the assumption that knowledge is neutral, while the concept of epistemic injustice explains how certain groups are marginalized. Bourdieu's notion of cultural capital shows that assessment practices tend to privilege knowledge aligned with dominant cultural norms. Within this framework, DIF serves as an empirical tool to identify bias in test items, helping translate critical theoretical insights into measurable evidence of epistemic inequality in educational assessment.

METHODOLOGY

Research design and approach

This study employs an explanatory mixed methods design with a strong emphasis on critical quantitative analysis supported by conceptual interpretation grounded in feminist theory. This approach is chosen to enable the empirical identification of bias in assessment while simultaneously providing a critical interpretation of statistical findings. Rather than positioning quantitative methods as neutral instruments, the study deliberately situates psychometric analysis within a critical framework that interrogates institutionalized knowledge practices in elementary education.

Research context and participants

The research participants consist of elementary school students who completed written assessments in core subject areas within a socially and culturally diverse educational context. Beyond their numerical representation, these students come from varied social backgrounds, including differences in socio-economic status, cultural affiliations, and lived experiences that shape their engagement with educational content. The schools involved in this study are situated in contexts where cultural diversity and social stratification are present, reflecting broader structural inequalities within the education system.

This diversity is particularly relevant to the present study, as differences in students' cultural experiences and social positioning may influence how they interpret and respond to assessment items. Therefore, the inclusion of participants from varied backgrounds is not only a methodological consideration, but also a critical component for examining how assessment practices may reproduce or challenge existing inequalities.

Data collection techniques

Data were collected through documentation of assessment instruments used in schools and students' responses to test items. The instruments analyzed consisted of evaluation items employed in regular assessment practices, thereby reflecting authentic classroom assessment conditions. The use of naturalistic assessment data enables the study to capture bias operating implicitly within everyday assessment practices, rather than within instruments specifically designed for research purposes.

Differential item functioning analysis

Differential Item Functioning (DIF) analysis was employed to identify test items that function differently for groups of students with equivalent levels of academic ability. DIF is conceptualized as an indicator of unfairness when differences in the probability of answering an item correctly cannot be attributed to ability but are instead associated with membership in particular social groups, such as gender or cultural background. In this study, DIF analysis is conducted in a focused and concise manner to identify patterns of item bias without becoming entangled in excessive technical complexity, in line with the study's critical objectives

Critical methodological positioning

Methodologically, this study adopts a critical approach by rejecting the assumption that educational measurement is fully objective and instead using DIF as a tool to uncover hidden power relations and epistemic injustice in assessment. DIF is not treated as a neutral statistical technique, but as a means to identify how certain items advantage dominant groups by aligning with their knowledge and experiences, while marginalizing others with equivalent ability. By combining statistical analysis with theoretical reflection, the study shows that quantitative data can reveal underlying social values, norms, and inequalities in how knowledge is recognized and validated within the education system.

Ethical and political rationale for using DIF

The use of DIF in this study is not only methodological but also ethical and political, as it aims to uncover and challenge assessment biases that contribute to epistemic inequality. Identifying biased items is part of a broader commitment to educational justice, ensuring that assessments do not unfairly privilege certain groups while disadvantaging others. DIF also serves as a tool of accountability by providing empirical evidence to question claims of fairness and highlight how educational practices may marginalize certain forms of knowledge. In this way, DIF supports the development of more equitable and inclusive assessment systems while positioning quantitative analysis as a means of advocating for social justice in education.

RESULTS

Differential item functioning across gender and cultural groups

To provide a clearer overview of the identified patterns of DIF, [Table 1](#) summarizes the items exhibiting bias, the types of bias identified, and the groups advantaged or disadvantaged by each item.

Table 1

Summary of differential item functioning (DIF) across gender and cultural groups

Item Code	Content Context	Type of Bias	Advantaged Group	Disadvantaged Group	Type of DIF
Item 1	Economic activity (construction worker)	Gender	Male students	Female students	Uniform DIF
Item 2	Traditional market activity	Cultural	Local students	Immigrant students	Uniform DIF
Item 3	Agricultural environment	Gender & Cultural (Intersectional)	Local male students	Immigrant female students	Uniform DIF

As shown in [Table 1](#), the identified DIF patterns consistently indicate that certain assessment items advantage specific groups of students based on gender and cultural background, reinforcing patterns of epistemic inequality discussed in the following sections.

The analysis of DIF in elementary school assessments shows that some items are not neutral, as they advantage or disadvantage students based on gender and cultural background rather than purely academic ability. This indicates the presence of epistemic inequality, where access to recognized knowledge is uneven. Items related to socioeconomic contexts may favor certain genders, while those tied to dominant cultural experiences

tend to benefit students from those backgrounds, reflecting gendered and culturally biased knowledge. These patterns reveal that assessment instruments are shaped by dominant social norms, leading to epistemic injustice by marginalizing the knowledge of underrepresented groups. Therefore, DIF should not be seen merely as a technical issue, but as a critical lens to examine fairness in education and to question how knowledge is constructed and validated within assessment practices.

Patterns of gender bias in assessment items

In the analysis of DIF, several assessment items were found to exhibit gender bias, indicating unequal opportunities for male and female students to answer correctly. One particularly salient example is the following item:

Mr. Amir works as a construction labourer on a housing project in the city of Makassar. Every morning, he goes to work by motorcycle and returns home in the afternoon.

Mr. Amir's activity constitutes an economic activity in the form of ...

- A. production
- B. distribution
- C. consumption
- D. trade

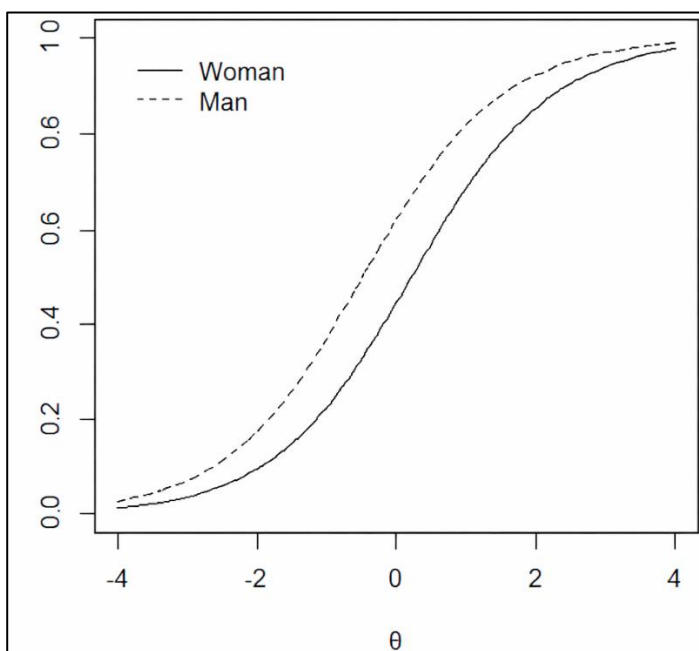
Empirically, this item demonstrates gender bias because it is more easily understood and mastered by male students, while female students appear to experience greater difficulty in forming the mental imagery required to grasp the context of the question. This bias is not merely related to knowledge of economic activities, but rather to the social experiences that shape ways of thinking and imagining particular situations.

It is important to note that the physical activity and out-of-home mobility mentioned in the item such as working on a housing project and riding a motorcycle represent contexts that are more familiar to many male students. They are more likely to possess mental representations of the kind of work performed by Mr. Amir due to social constructions that associate men with professions involving physical labour and mobility.

By contrast, for female students, even though they may understand the economic concepts being tested, they are often burdened by more complex social contexts. In many cultures, for example, women are more frequently associated with domestic work and roles that are more closely tied to the home. As a result, although their conceptual knowledge of production or economic activities may be equivalent to that of male students, they may find it more difficult to connect these concepts with experiences or mental images that are commonly encountered in their everyday lives. For clearer illustration, a plot of the DIF analysis results for this item is presented below.

Figure 1

An example of an item demonstrating gender-based inequity bias



The DIF plot for this item indicates a consistent difference in item functioning between female and male groups, suggesting the presence of uniform DIF. The item characteristic curves show that across nearly the entire range of latent ability (θ), male students have a higher probability of answering the item correctly than female students at equivalent ability levels. The leftward shift of the curve for the male group indicates that the item is relatively “easier” for them, whereas female students require a higher level of ability to achieve the same probability of success. These performance differences cannot be attributed solely to substantive ability differences but are likely influenced by item constructions that align more closely with the experiences, cultural representations, or knowledge practices of particular groups. Accordingly, this item can be interpreted as evidence of gender-based epistemic inequality in assessment, in which success probabilities are shaped by gender bias embedded in the design or contextual framing of the item, potentially leading to unfair evaluations of elementary students’ abilities.

This DIF illustrates that access to knowledge in assessment is not determined solely by conceptual understanding, but also by the social and cultural experiences that shape how students engage with test materials. Male students, who are often more familiar with the types of work depicted in the item, have a greater likelihood of answering correctly even when their academic ability is equivalent to that of female students.

From the perspective of *epistemic inequality*, this bias reveals how education systems often treat knowledge associated with male experiences as more legitimate and acceptable. The item not only tests understanding of economic concepts but also reflects the standardization of knowledge based on particular social and gendered experiences. This creates epistemic inequality, in which women’s experiences and knowledge are more frequently overlooked or deemed less relevant in the context of examinations and assessment.

This bias demonstrates how the knowledge considered valid and recognized within the education system is often rooted in dominant gender norms. In this case, an item that advantages male students does not merely measure economic knowledge objectively but also reinforces social norms that position men in the realm of physical and public work, while women are more often situated outside that context. This illustrates how educational knowledge systems can reinforce gender inequality and disadvantage groups whose experiences are not represented in the contexts embedded within assessment items.

Patterns of cultural bias in assessment items

In the analysis of DIF assessment items, several patterns of cultural bias were identified, reflecting inequalities in students’ understanding of and experience with different social contexts. One example of an item that illustrates cultural bias is the following:

Every morning, residents bring their agricultural produce to the traditional market using two-wheeled vehicles or carts. This activity demonstrates the relationship between economic activities and ...

- A. transportation
- B. communication
- C. education
- D. culture

This item exhibits cultural bias that may disadvantage students who come from outside the local community or those who lack experience with traditional markets or traditional means of transportation such as carts. The bias is linked to local cultural schemas that are more familiar to local students, who often have direct experience with traditional markets as part of their daily lives.

However, for migrant or newcomer students, who may be more accustomed to modern minimarkets or app-based transportation services, this item creates an unfair cognitive burden, even though they may possess an equivalent understanding of economic concepts. The cultural context embedded in the item is not symmetrical for all students. Those who are unfamiliar with traditional market life or cart-based transportation may conceptualize transportation differently, yet they are confronted with a cultural gap that hinders their ability to connect the tested context with their everyday experiences.

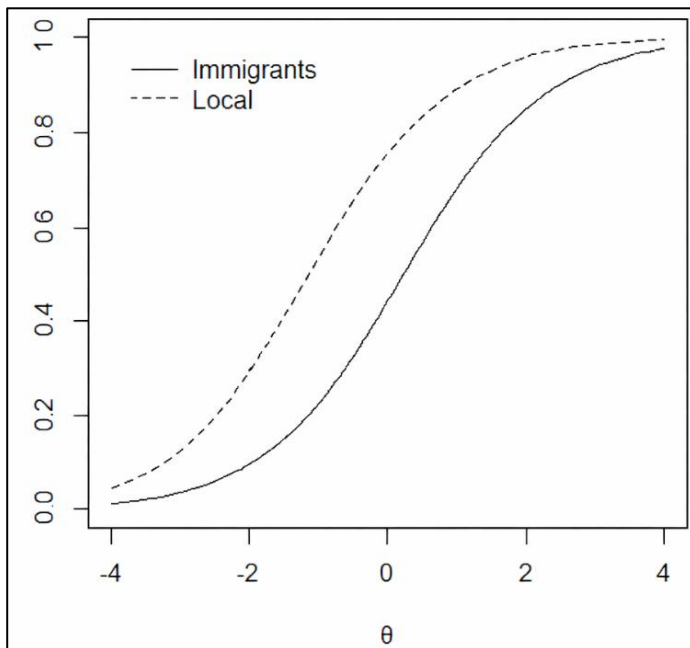
From the perspective of *epistemic inequality*, these findings reveal how local knowledge associated with traditional culture is often treated as universal within educational assessment. Assessment systems that employ specific cultural references demonstrate how access to knowledge recognized in examinations is frequently limited to the experiences and knowledge legitimized by dominant cultural groups.

This item not only assesses understanding of economic activities and transportation but also illustrates how certain forms of cultural knowledge are regarded as more legitimate and acceptable within the education system. Local students, who have direct experience with traditional markets and cart-based transportation, gain a cognitive advantage in answering the item, while migrant students or those unfamiliar with the cultural context experience greater difficulty in linking the concepts presented to their own experiences.

The inclusion of local cultural contexts in this item demonstrates how cultural gaps influence assessment processes, exacerbating epistemic inequality in the classroom. Knowledge considered legitimate within the education system is often tied to dominant cultural contexts, while other cultural experiences and forms of knowledge are frequently marginalized or ignored. This results in injustice in how knowledge is recognized and valued in education, benefiting those who belong to more dominant cultural contexts or are closer to socially accepted norms. A plot of the analysis results illustrating this disparity in opportunity indicating a situation of injustice is presented below.

Figure 2

An Example of an Item Demonstrating Culture-Based Inequity Bias



The DIF plot for this item reveals a consistent difference in item functioning between immigrant-background students and local students, indicating the presence of uniform DIF. The item characteristic curves show that across nearly the entire range of latent ability (θ), local students have a higher probability of answering the item correctly than immigrant students at equivalent ability levels. The leftward shift of the curve for the local student group indicates that the item is relatively easier for them, whereas immigrant students require a higher level of ability to attain the same probability of success. This pattern suggests that the observed performance differences do not fully reflect substantive ability differences but are likely influenced by cultural, linguistic, or knowledge representations embedded in the item that align more closely with the experiences of the local group. Accordingly, this item can be interpreted as evidence of culture-based epistemic inequality in elementary school assessment, in which the knowledge structures legitimized within the item tend to advantage dominant groups and may lead to unfair evaluations of immigrant students' abilities.

These DIF findings underscore the importance of more fully recognizing and valuing cultural diversity within education systems. Local cultural knowledge that may appear highly relevant and familiar to some students is not always applicable or easily accessible to students from different cultural backgrounds. Therefore, to ensure fairness in assessment, it is essential to adapt the contexts used in test items so that they are more inclusive and representative of the diversity of students' experiences.

Furthermore, items such as this illustrate how social and cultural norms shape the ways in which legitimate knowledge is defined in education. This cultural bias not only reflects injustice in how students' abilities are assessed but also reinforces the view that dominant cultural knowledge is the only form of knowledge worthy of recognition in the classroom.

Intersecting gender and cultural biases

In the analysis of DIF, it was found that several assessment items exhibit intersections of gender and cultural bias, affecting students in more complex ways than single-source bias based on only one factor. One example of an item that illustrates the intersection of gender and cultural bias is the following:

Rian often accompanies his father to the gardens and rice fields located in the hilly area around where they live. He knows that the area has fertile soil and many water streams.

The conditions of this area indicate the characteristics of a ... environment

- A. highland
- B. lowland
- C. coastal
- D. urban

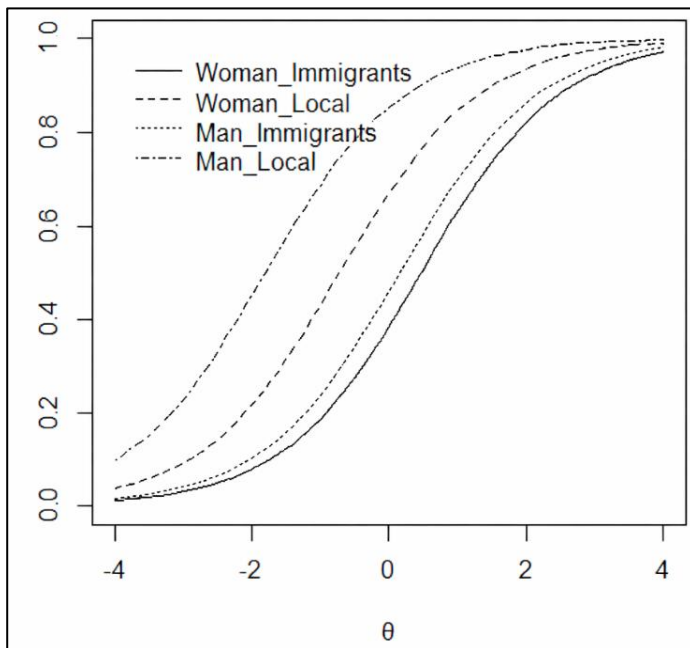
This item demonstrates bias that is not only related to local cultural context, but also to gender roles that are explicitly associated with masculine experiences in agricultural and farming settings. In this item, gardens, rice fields, and hilly landscapes are elements closely tied to local culture, in which many male students from the area are more familiar with agricultural life and have direct experience with such natural environments.

However, for migrant female students, the item contains two layers of unfamiliarity. On the one hand, they may not be accustomed to the local cultural context that depicts life in gardens and rice fields. On the other hand, they also encounter gender bias, as women's roles in many cultures are more often associated with domestic work or tasks within the home, rather than outdoor activities alongside their fathers working in fields or farms. This creates a *double disadvantage* for female students, who must navigate not only an unfamiliar cultural context but also gender bias that associates agricultural activities with masculinity.

For local male students, there is a cognitive advantage due to greater familiarity with the item's context, both culturally and through direct experience. As a result, they are in a more favorable position to answer the item correctly, even when their academic ability is equivalent to that of female students. The plot for this item is presented below to provide a more comprehensive illustration of the case.

Figure 3

An Example of an Item Demonstrating Inequity Bias Based on The Interaction Between Gender and Culture



The DIF plot for this item, which disaggregates groups by gender and cultural background, reveals a consistent and systematic pattern of DIF, indicating the presence of intersectional uniform DIF. The item characteristic curves show that local male students have the highest probability of answering the item correctly across nearly the entire range of latent ability (θ), followed by local female students, while both groups of immigrant students, particularly immigrant females, exhibit lower probabilities of success at equivalent ability levels. The progressively rightward shift of the curve for immigrant female students indicates that this group requires the highest level of latent ability to attain the same probability of success. This graded pattern suggests that performance differences do not merely reflect single-axis gender- or culture-based bias but rather arise from the interaction of both dimensions as embedded in the item construction. Accordingly, this item can be interpreted as evidence of intersectional epistemic inequality in elementary school assessment, in which the knowledge structures and contextual framing of the item simultaneously advantage dominant groups and marginalize students located at the intersection of gender and cultural identities, potentially leading to unfair evaluations of students' abilities.

When gender and cultural biases intersect in assessment items, they create deeper inequalities by privileging the knowledge and experiences of socially dominant groups. For example, items that reflect the experiences of local male students, such as familiarity with agriculture or outdoor activities, are often treated as more valid,

while the experiences of women or migrant students are less recognized. This results in epistemic injustice, where some students are disadvantaged not because of lower ability, but because their knowledge is not valued within the assessment. Therefore, these findings highlight the need for more inclusive assessment practices that recognize diverse student experiences across gender and cultural backgrounds, so that educational justice is achieved not only through equal standards but also through fair representation of knowledge.

Epistemic inequality through DIF findings

The findings of DIF indicate that assessment items in elementary schools do not merely measure cognitive ability but also reflect social structures that reinforce epistemic inequality. Some items provide unequal cognitive advantages based on gender and culture, thereby marginalizing forms of knowledge that do not align with dominant norms. For instance, items related to male-associated physical activities, such as working in gardens or construction projects, tend to benefit male students, while female students must exert additional effort to relate the concepts to their experiences. Similarly, items grounded in local cultural contexts, such as traditional markets, may disadvantage migrant students who are more familiar with modern or urban settings, creating cognitive gaps.

These inequalities become more complex when gender and cultural biases intersect, as seen in agricultural contexts that are more familiar to local male students and simultaneously reflect gendered roles. As a result, migrant female students may experience a double layer of disadvantage despite having equivalent abilities. Overall, the DIF findings highlight that assessments can reproduce social norms that disadvantage certain groups. Therefore, DIF serves as an important tool to uncover such inequalities and to encourage improvements in assessment practices, not only by revising biased items but also by rethinking how knowledge is defined and recognized to ensure greater inclusivity and fairness in education.

DISCUSSION

The phenomenon of DIF raises serious concerns regarding fairness in the implementation of educational assessments, thereby prompting questions about potential biases based on gender, culture, or other demographic factors. The presence of gender and cultural biases in teacher-made assessment items indicates that claims of neutrality in educational evaluation need to be critically examined (Navarro-González et al., 2024). The identification of DIF serves as an important tool for examining how test items may systematically disadvantage certain groups, such that measurements that appear to assess equivalent academic knowledge may in fact reflect differences in the social and cultural representations embedded within them.

The findings of this study can be more deeply understood through the lens of feminist epistemology, which challenges the assumption that knowledge represented in educational assessment is neutral and universally accessible. Feminist epistemology provides a critical reading of DIF as evidence of epistemic bias embedded in assessment, not mere measurement error (Brady, 2022; Egan, 2024). Across the referenced works, feminist frameworks emphasize that knowledge production in education is situated, gendered, and shaped by power relations, which can privilege male or dominant cultural ways of knowing in assessment design and interpretation (Worku, 2025). When DIF patterns repeatedly favor specific groups, this supports claims of deeper structures that reproduce inequality in schooling and credentialing, aligning with standpoint and intersectional critiques of “objectivity” in testing. Several sources argue for embedding feminist analyses into educational assessment policy and practice to counter bias, ensure inclusivity, and reframe assessment as a site of justice rather than neutrality. Where disagreements exist, they center the exact mechanisms of bias and the appropriate methodological responses (e.g., counterfactual analysis vs. participatory design), but converge on the view that DIF signals epistemic injustice requiring structural reform, not only calibration improvements (Finn & Brown, 2022).

The implication of gender bias is profound. Experiences reflecting traditionally masculine domains like mechanics or physical activities tend to be more accessible to male students, thus creating inequitable assessment conditions. This bias not only affects test scores but also the way knowledge itself is constructed and validated within the educational system. The findings align with a robust literature showing gender bias in assessment contexts that privilege masculine domains, and extend this to epistemic inequalities within broader assessment systems. Sometimes women are perceived as having a lower position in academic contexts (Moratti, 2021; Bharti, 2025). Prior work demonstrates male-favoring contexts (e.g., physical activities, male-dominated professions) can yield DIF in items even when ability is controlled, and similar patterns appear across textbooks and curricula where gendered representations shape learning experiences (Bearman & Ajjawi, 2025). Recent syntheses argue that bias in assessments arises not only from item functioning but from deeper sociocultural structures, including role expectations and normative gender identities within professional training (evidence from medicine, pharmacy, engineering) (Farrell, 2021; Winkel et al., 2021). Integrating DIF analysis with critical theory provides a more comprehensive account of how gender and cultural biases operate in assessment systems, distinguishing technical item bias from epistemic inequities embedded in curriculum, pedagogy, and institutional cultures

(Bochatay et al., 2022). Where prior work identified bias, these findings advance by linking DIF to a critical framework, illustrating how epistemic injustice is reproduced through assessment design and practice (Nedungadi et al., 2022).

Cultural bias presents a similar challenge in assessment fairness, which is particularly evident in standardized tests where local experiences and cultural contexts are foregrounded at the expense of wider applicability. Various studies demonstrate that assessments often do not uniformly translate across diverse cultural backgrounds, which contributes to epistemic injustice (Navarro-González et al., 2024). For example, local cultural scenarios are frequently utilized in item formulation, providing cognitive advantages to students familiar with these contexts while marginalizing those from different cultural backgrounds.

This issue is echoed in the findings by Navarro-González, which elaborates on the importance of measurement invariance in educational assessments and the challenges posed by cultural biases that disadvantage students from minority backgrounds (Navarro-González et al., 2024). Such findings highlight that knowledge regarded as "common" is often rooted in the experiences of dominant cultural groups, thus inequitable assessments risk perpetuating these hierarchies of knowledge validation.

The analysis of DIF shows that assessment items are not fully neutral, as they often reflect dominant gender and cultural perspectives that advantage certain groups. As a result, female and migrant students may be disadvantaged because their experiences do not align with the context of the items, with some facing multiple layers of bias. These findings highlight the importance of more fair and inclusive assessment practices.

The DIF findings align with a broader epistemic justice critique: assessment items often encode dominant social and cultural perspectives, privileging male or locally normative knowledge while marginalizing migrant or non-dominant voices (Rocha & Gottschalk, 2025; Vlase, 2024). This yields epistemic disadvantages not from ability gaps but from misalignment between items and diverse lived experiences, with intersectional effects (e.g., women who are migrants or from minority backgrounds may face compounded biases) (Holdsworth et al., 2023; Omran & Yousafzai, 2024). Consequently, fairness in assessment should entail inclusive representation of knowledge and recognition of diverse knowledges, rather than uniform questioning, to promote equitable assessment across identities and contexts (Balarin & Rodríguez, 2024; Cummings et al., 2023). Not all individuals experience bias uniformly; variability across intersectional positions underscores the complexity of bias dynamics within assessment practices (Byskov, 2021; Wee et al., 2023). These insights converge with calls for epistemic justice in education policy and practice, including gender-sensitive and intercultural approaches to assessment design and interpretation (Clavero & Galligan, 2021).

The findings of this study have important practical implications for improving assessment practices in education. First, assessment design should move beyond assumptions of neutrality by critically evaluating the cultural and gendered contexts embedded in test items. Item developers should ensure that assessment contexts are diverse, inclusive, and representative of students' varied experiences (Herwin et al., 2022). Second, the use of DIF analysis should be integrated into routine assessment evaluation processes as a diagnostic tool to identify and eliminate biased items before they are widely implemented. Third, teacher training and professional development programs should include critical awareness of gender and cultural bias in assessment, enabling educators to design more equitable evaluation instruments. Finally, adopting multiple and contextually varied assessment formats can help reduce the dominance of any single cultural or gendered perspective, thereby promoting a more inclusive and just education system.

CONCLUSION

This study demonstrates that elementary school assessment items may exhibit significant gender and cultural bias, as evidenced through Differential Item Functioning (DIF) analysis. The findings indicate that such biases are not merely technical issues, but reflect deeper epistemic inequalities in which certain forms of knowledge and experience are privileged over others. In particular, the results reveal that assessment items can advantage male and culturally dominant students, while disadvantaging female and non-dominant cultural groups, including in intersectional ways.

These findings underscore the importance of rethinking assessment practices as socially situated rather than neutral instruments. By positioning DIF as both a statistical and critical tool, this study highlights its potential to uncover hidden inequalities in educational assessment and to inform the development of more inclusive and equitable evaluation practices. Taken together, these findings underscore that the issue at stake is not merely measurement bias, but a broader concern of epistemic justice, where the recognition and validation of diverse forms of knowledge remain unevenly distributed within educational assessment practices.

However, this study has several limitations. First, the analysis is based on a limited set of assessment items and participants within a specific educational context, which may affect the generalizability of the findings.

Second, the study relies primarily on DIF analysis and does not incorporate complementary qualitative approaches that could provide deeper insights into students' experiences of assessment bias.

Future research is recommended to expand the scope of analysis by including larger and more diverse samples, as well as by integrating mixed or qualitative methods to further explore how students interpret and respond to assessment contexts. In addition, further studies could investigate strategies for designing culturally responsive and gender-inclusive assessment instruments, thereby contributing to more equitable education systems.

Acknowledgement

The authors would like to express their sincere gratitude to the participating schools, teachers, and students who contributed to this study. The authors also acknowledge the institutional support provided by Universitas Negeri Yogyakarta in facilitating this research.

Funding

This research was funded by Universitas Negeri Yogyakarta, Indonesia.

Ethical statement

This study was conducted in accordance with ethical research standards. Informed consent was obtained from all participants and their guardians prior to data collection. All data were collected, processed, and reported with strict adherence to confidentiality and anonymity principles.

Competing interests

The authors declare that there are no competing interests.

Author contributions

Herwin Herwin was responsible for conceptualization, methodology, formal analysis, writing of the original draft, and supervision. Shakila Che Dahalan contributed to validation, writing-review and editing, as well as the development of the theoretical framework. Muhammad Ilyas was responsible for data curation, investigation, and visualization. Ma'rufi Ma'rufi contributed to methodology support, data analysis assistance, and writing-review and editing. All authors have read and agreed to the published version of the manuscript.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. Due to ethical considerations, the data are not publicly available.

AI disclosure

The authors declare that artificial intelligence (AI) tools were used only for language refinement and editorial assistance. All conceptualization, analysis, interpretation, and writing of the manuscript were carried out by the authors. The authors take full responsibility for the content of this publication.

Biographical sketch

Herwin is a senior lecturer in the Department of Elementary Education at Universitas Negeri Yogyakarta, Indonesia. He holds a doctoral degree in Educational Research and Evaluation. His research focuses on educational assessment, psychometrics, and equity in education, with particular interest in epistemic justice and inclusive assessment practices.

Shakila Che Dahalan is a senior lecturer at Universiti Pendidikan Sultan Idris, Malaysia. She earned her PhD in Education with a specialization in curriculum and pedagogy. Her research interests include gender studies in education, curriculum development, and critical pedagogy in multicultural contexts.

Muhammad Ilyas is a professor in the field of education at Universitas Cokroaminoto Palopo, Indonesia, with expertise in educational measurement and evaluation. His research interests include educational assessment, data analysis in education, and classroom-based evaluation practices.

Ma'rufi is a lecturer at Universitas Cokroaminoto Palopo, Indonesia, with expertise in mathematics education. Her research focuses on mathematics education, assessment development, and quantitative research methods in education.

Disclaimer/Publisher's Note

The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and do not necessarily reflect the views of Lectito Publications and/or the editor(s). Lectito Publications and/or the editor(s) disclaim responsibility for any injury to persons or property resulting from any ideas, methods, instructions, or products referred to in the content.

REFERENCES

- Abreu, R. (2026). The complexity of giving voice: Intersectionality as method in the study of Instagram and marginalised identities. *Feminist Encounters: A Journal of Critical Studies in Culture and Politics*, 10(1), Article 9. <https://doi.org/10.20897/femenc/17897>
- Acar, E., Deiri, Y., & Yigit, F. (2025). A focused review of artificial intelligence in education: Evolution and challenges. *Journal of Interdisciplinary Research in Artificial Intelligence and Society*, 1(1), Article 3. <https://doi.org/10.20897/jirais/17640>
- Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2(3), 431–440. <https://doi.org/10.1007/s43681-021-00096-7>
- Amaral, I. G., Naves, F., & Rezende, A. F. (2025). Decolonizing knowledge, resisting coloniality of gender: Collective mobilizations at a Brazilian Public University. *Gender, Work & Organization*, 32(6), 2290–2307. <https://doi.org/10.1111/gwao.70009>
- Arendse, D. E. (2021). Addressing gender discrimination in cognitive assessment using the English Comprehension Test. *SA Journal of Industrial Psychology*, 47, 1–10. <https://doi.org/10.4102/sajip.v47i0.1776>
- Atias, O., Shavit, A., Kali, Y., & Baram-Tsabari, A. (2025). Deciphering the role of epistemic injustice in school-based citizen science: Sources, implications, and possible ways for mitigation. *Journal of Research in Science Teaching*, 62(5), 1350–1387. <https://doi.org/10.1002/tea.22006>
- Balanta-Cobo, S., Osorio-Mejía, A. M., Romero-Olmedo, G. A., & Martínez-Pérez, J. (2025). Measuring the gender gap in Unpaid Care Work in Colombia: A decomposition analysis. *Social Policy & Administration*, 59(6), 1055–1071. <https://doi.org/10.1111/spol.70016>
- Balarin, M., & Rodríguez, M. F. (2024). Shallow pedagogies as epistemic injustice: How uncritical forms of learning hinder education's contribution to just and sustainable development. *Global Social Challenges Journal*, 3, 49–67. <https://doi.org/10.1332/27523349Y2024D000000007>
- Bearman, M., & Ajjawi, R. (2025). Artificial intelligence and gender equity: An integrated approach for health professional education. *Medical Education*, 59(10), 1049–1057. <https://doi.org/10.1111/medu.15657>
- Bedmar, M. del C. D. (2022). Feminism, intersectionality, and gender category: Essential contributions for historical thinking development. *Frontiers in Education*, 7, 1–8. <https://doi.org/10.3389/feduc.2022.842580>
- Benschop, Y. (2021). Grand challenges, feminist answers. *Organization Theory*, 2(3), 1–19. <https://doi.org/10.1177/263178772111020323>
- Bharti, M. (2025). Engendering education and organisation: The women's reform movement in India. *Feminist Encounters: A Journal of Crifical Studies in Culture and Polifics*, 9(2), 1-10. <https://doi.org/10.20897/femenc/16788>
- Bochatay, N., Bajwa, N. M., Ju, M., Appelbaum, N. P., & van Schaik, S. M. (2022). Towards equitable learning environments for medical education: Bias and the intersection of social identities. *Medical Education*, 56(1), 82–90. <https://doi.org/10.1111/medu.14602>
- Brady, J. (2022). Exploring the role of black feminist thought in pre-service early childhood education: On the possibilities of embedded transformative change. *Contemporary Issues in Early Childhood*, 23(4), 392–407. <https://doi.org/10.1177/14639491221136584>
- Buren, V., Karam, C., & Afiouni, F. (2024). Editorial: Producing actionable knowledge about marginalized populations and communities: a challenge to editors and journals. *Equality, Diversity and Inclusion: An International Journal*, 43(2), 300–312. <https://doi.org/10.1108/EDI-03-2024-426>
- Byskov, M. F. (2021). What Makes Epistemic Injustice an “Injustice”? *Journal of Social Philosophy*, 52(1), 114–131. <https://doi.org/10.1111/josp.12348>
- Chan, J., Collins, K. A., Lee, R., Linton, J., Cherba, M., Christianson, T. D., Shawanda, A., Siden, E. G., & Wardman, M. (2025). A Scoping Review of Published Literature on the Linguistic Representation of Indigenous Peoples. *Journal of Language and Social Psychology*, 44(3–4), 441–480. <https://doi.org/10.1177/0261927X251318040>
- Chen, X., Aryadoust, V., & Zhang, W. (2025). A systematic review of differential item functioning in second language assessment. *Language Testing*, 42(2), 193–222. <https://doi.org/10.1177/02655322241290188>

- Chingarande, S. D., Muchabaiwa, W., & Muyambo, T. (2026). Good practices in gender mainstreaming in science at the women's university in Africa, Zimbabwe. *European Journal of STEM Education*, 11(1), Article 9. <https://doi.org/10.20897/ejsteme/17968>
- Chung, R. (2021). Structural health vulnerability: Health inequalities, structural and epistemic injustice. *Journal of Social Philosophy*, 52(2), 201–216. <https://doi.org/10.1111/josp.12393>
- Clavero, S., & Galligan, Y. (2021). Delivering gender justice in academia through gender equality plans? Normative and practical challenges. *Gender, Work & Organization*, 28(3), 1115–1132. <https://doi.org/10.1111/gwao.12658>
- Cook, E. A. (2025). Conflating the map with the territory: Challenges for evidence syntheses on homicide in a global context. *International Sociology*, 40(6), 944–963. <https://doi.org/10.1177/02685809251336694>
- Cummings, S., Dhewa, C., Kemboi, G., & Young, S. (2023). Doing epistemic justice in sustainable development: Applying the philosophical concept of epistemic injustice to the real world. *Sustainable Development*, 31(3), 1965–1977. <https://doi.org/10.1002/sd.2497>
- Czerwinski-Alley, N. C., Chithiramohan, T., Subramaniam, H., Beishon, L., & Mukaetova-Ladinska, E. B. (2024). The effect of translation and cultural adaptations on diagnostic accuracy and test performance in dementia cognitive screening tools: A systematic review. *Journal of Alzheimer's Disease Reports*, 8(1), 659–675. <https://doi.org/10.3233/ADR-230198>
- Deiri, Y., & Burkhard, T. (2025). The "then" and "now" politics of fear: A multilingual intimate duo-ethnography at the crossroads of language, religion, immigration, and education. *European Journal of Education & Language Review*, 1(1), Article 6. <https://doi.org/10.20897/ejelr/17648>
- Dogutas, A. (2025). A comparative analysis of immigrant children's educational policies: Türkiye and the United States. *European Journal of Education & Language Review*, 1(1), Article 2. <https://doi.org/10.20897/ejelr/17313>
- Egan, F. (2024). Teaching gender in and through uncertainty. *Journal of Sociology*, 60(1), 248–264. <https://doi.org/10.1177/14407833221135939>
- Farid, S., Abbasi, S. U. R. S., & Mahmood, Q. K. (2021). Modelling Bourdieusian social reproduction theory. *Social Indicators Research*, 157(1), 297–333. <https://doi.org/10.1007/s11205-021-02649-z>
- Farrell, S. (2021). A sociocultural learning framework for inclusive pedagogy in engineering. *Chemical Engineering Education*, 55(4), 192–204. <https://doi.org/10.18260/2-1-370.660-128660>
- Finn, G. M., & Brown, M. E. L. (2022). Ova-looking feminist theory: A call for consideration within health professions education and research. *Advances in Health Sciences Education*, 27(3), 893–913. <https://doi.org/10.1007/s10459-022-10108-8>
- Fisher, E., MacLennan, K., Mullally, S., & Rodgers, J. (2025). Neuro-normative epistemic injustice-consequences for the UK education crisis and school anxiety. *Neurodiversity*, 3, 1–13. <https://doi.org/10.1177/27546330251353565>
- Francis, L. (2025). Bidirectional intersectional supervision: Redefining power and equity for black clinicians. *Journal of Family Theory & Review*, 17(4), 883–894. <https://doi.org/10.1111/jftr.70008>
- Geven, S., & Zwier, D. (2025). Students' Interactional cultural capital and academic performance in test-and teacher-based assessments. *The British Journal of Sociology*, 76(3), 622–634. <https://doi.org/10.1111/1468-4446.13199>
- Gibson, M. L. (2024). Beyond virtue signaling: From talking about to enacting a pedagogical praxis of justice in urban classrooms. *Urban Education*, 59(8), 2550–2577. <https://doi.org/10.1177/00420859221107612>
- Halpern, B., Aydin, H., & Halpern, C. (2025). Seeing multilingual learners through media and AI: Pre-service teachers' perceptions in an ESOL course. *Journal of Interdisciplinary Research in Artificial Intelligence and Society*, 1(1), Article 4. <https://doi.org/10.20897/jirais/17647>
- Hao, J., von Davier, A. A., Yaneva, V., Lottridge, S., von Davier, M., & Harris, D. J. (2024). Transforming assessment: The impacts and implications of large language models and generative AI. *Educational Measurement: Issues and Practice*, 43(2), 16–29. <https://doi.org/10.1111/emip.12602>
- Hauer, K. E., Park, Y. S., Bullock, J. L., & Tekian, A. (2023). “My assessments are biased!” measurement and sociocultural approaches to achieve fairness in assessment in medical education. *Academic Medicine*, 98(8S), S16–S27. <https://doi.org/10.1097/ACM.0000000000005245>
- Herwin, H. (2022). DIF content of math test on learning assessment in elementary school. *AIP Conference Proceedings*, 2577(1). <https://doi.org/10.1063/5.0096007>
- Herwin, H., Pristiwaluyo, T., Ruslan, R., & Dahalan, S. C. (2022). Do Scoring techniques and number of choices affect the reliability of multiple-choice tests in elementary schools? *Cypriot Journal of Educational Sciences*, 17(4), 1258–1268. <https://doi.org/10.18844/cjes.v17i4.7149>

- Holdsworth, S., Turner, M., & Sandri, O. (2023). Gender bias in the Australian construction industry: Women's experience in trades and semi-skilled roles. *Social Sciences*, 12(11), 627. <https://doi.org/10.3390/socsci12110627>
- Ibrahim, M., Herwin, H., Retnawati, H., Firdaus, F. M., Umar, U., & Mufidah (2024). STEM learning for students' mathematical numeracy ability. *European Journal of STEM Education*, 9(1), 20. <https://doi.org/10.20897/ejsteme/15750>
- Johnson, M. S., Liu, X., & McCaffrey, D. F. (2022). Psychometric methods to evaluate measurement and algorithmic bias in automated scoring. *Journal of Educational Measurement*, 59(3), 338–361. <https://doi.org/10.1111/jedm.12335>
- Kinnear, B., Weber, D. E., Schumacher, D. J., Edje, L., Warm, E. J., & Anderson, H. L. (2023). Reconstructing neurath's ship: A case study in reevaluating equity in a program of assessment. *Academic Medicine*, 98(2), S50–S56. <https://doi.org/10.1097/ACM.00000000000005249>
- Koh, K. (2023). Designing Culturally Responsive Online assessments for equity-deserving students. *Cultural and Pedagogical Inquiry*, 15(1), 71–88. <https://doi.org/10.18733/cpi29712>
- Kuchynka, S. L., Eaton, A., & Rivera, L. M. (2022). Understanding and addressing gender-based inequities in STEM: Research synthesis and recommendations for U.S. K-12 education. *Social Issues and Policy Review*, 16(1), 252–288. <https://doi.org/10.1111/sipr.12087>
- Laitusis, C. C., & Karvonen, M. (2025). A case for reimagining universal design of assessment systems. *Educational Measurement: Issues and Practice*, 44(3), 5–14. <https://doi.org/10.1111/emip.12674>
- Lavoie, É., & Cavanagh, M. (2023). Writing instruction for social justice: an investigation into the components of a teacher preparation course. *Literacy*, 57(2), 132–148. <https://doi.org/10.1111/lit.12318>
- Ma, W., Terzi, R., & de la Torre, J. (2021). Detecting differential item functioning using multiple-group cognitive diagnosis models. *Applied Psychological Measurement*, 45(1), 37–53. <https://doi.org/10.1177/0146621620965745>
- Mamtani, M., Shofer, F., Scott, K., Kaminstein, D., Eriksen, W., Takacs, M., Hall, A. K., Weiss, A., Walter, L. A., Gallahue, F., Yarris, L., Abbuhl, S. B., & Aysola, J. (2022). Gender differences in emergency medicine attending physician comments to residents: A qualitative analysis. *JAMA Network Open*, 5(11), e2243134. <https://doi.org/10.1001/jamanetworkopen.2022.43134>
- Marabini, A. (2023). Paul Hirst, education and epistemic injustice. *Journal of Philosophy of Education*, 57(1), 77–90. <https://doi.org/10.1093/jopedu/qhac004>
- Martínez, P., Peoples, L., & Martin, J. (2023). Becoming culturally responsive: Equitable and inequitable translations of CRE theory into teaching practice. *The Urban Review*, 55(4), 476–504. <https://doi.org/10.1007/s11256-023-00658-5>
- Moratti, S. (2021). Contemporary fairy tales: Narrating women academics through metaphors. *Feminist Encounters: A Journal of Critical Studies in Culture and Politics*, 5(2), 1–13. <https://doi.org/10.20897/femenc/11157>
- Navarro-González, M. C., Padilla, J.-L., & Benítez, I. (2024). Analyzing measurement invariance for studying the gender gap in educational testing. *European Journal of Psychological Assessment*, 40(5), 412–426. <https://doi.org/10.1027/1015-5759/a000820>
- Nedungadi, S., Brown, C. E., & Paek, S. H. (2022). Differential item functioning analysis of the fundamental concepts for organic reaction mechanisms inventory. *Journal of Chemical Education*, 99(8), 2834–2842. <https://doi.org/10.1021/acs.jchemed.2c00242>
- Nezami, N., Haghghat, P., Gándara, D., & Anahideh, H. (2024). Assessing disparities in predictive modeling outcomes for college student success: The impact of imputation techniques on model performance and fairness. *Education Sciences*, 14(2), 136. <https://doi.org/10.3390/educsci14020136>
- Nigar, N., Dovchin, S., Yip, S. Y., & Sultana, S. (2025). Traversing self and other: The professional identity of immigrant English teachers. *Journal of Ethnic and Cultural Studies*, 12(5), 63–85. <https://doi.org/10.29333/ejecs/2295>
- Ohiri, S. C., Momoh, M., Christopher, O., Ikeanumba, I., & Benedict, C. (2024). Differential item functioning detection methods: An overview. *International Journal of Research Publication and Reviews*, 5(2), 1555–1564. <https://doi.org/10.55248/gengpi.5.0224.0505>
- Okoroji, C., Mackay, T., Robotham, D., Beckford, D., & Pinfeld, V. (2023). Epistemic injustice and mental health research: A pragmatic approach to working with lived experience expertise. *Frontiers in Psychiatry*, 14, 1–5. <https://doi.org/10.3389/fpsy.2023.1114725>
- Omran, W., & Yousafzai, S. (2024). Epistemic injustice and epistemic resistance: An intersectional study of women's entrepreneurship under occupation and patriarchy. *Entrepreneurship Theory and Practice*, 48(4), 981–1008. <https://doi.org/10.1177/10422587231208621>
- Parisi, M. (2026). "You can continue in Spanish—I understand": Intentionality and equity tensions in a translanguaging classroom. *American Journal of Qualitative Research*, 10(2), 1–28. <https://doi.org/10.29333/ajqr/18001>

- Prince, A. (2023). Culturally and socially responsible assessment: theory, research, and practice. *American Journal of Qualitative Research*, 7(1), 222–225. <https://doi.org/10.29333/ajqr/12871>
- Ren, H., & Liu, W. (2022). The effect and mechanism of cultural capital on Chinese residents' participation in physical activities. *Frontiers in Psychology*, 13, 1–7. <https://doi.org/10.3389/fpsyg.2022.848530>
- Ren, R. (2024). Searching for desirable bodies: How recruiters value physically exertive extracurricular activities for graduate hiring at elite professional firms in China. *The British Journal of Sociology*, 75(2), 187–200. <https://doi.org/10.1111/1468-4446.13074>
- Rezai, A., Namaziandost, E., Miri, M., & Kumar, T. (2022). Demographic biases and assessment fairness in classroom: insights from Iranian university teachers. *Language Testing in Asia*, 12(1), 8. <https://doi.org/10.1186/s40468-022-00157-6>
- Rizzolo, D., Sandifer, C., Kalabalik-Hoganson, J., & Lowy, N. (2022). A call to action for cultural humility in pharmacy education student assessments. *American Journal of Pharmaceutical Education*, 86(4), 8626. <https://doi.org/10.5688/ajpe8626>
- Rocha, M., & Gottschalk, C. (2025). Intergenerational formative epistemic injustice: A wittgensteinian critique of dominant world-pictures. *Educational Theory*, 75(6), 1060–1082. <https://doi.org/10.1111/edth.70052>
- Rogers, L. O., Moffitt, U., McLean, K. C., & Syed, M. (2024). Research as resistance: Naming and dismantling the master narrative of “good” science. *American Psychologist*, 79(4), 484–496. <https://doi.org/10.1037/amp0001246>
- Ruscoe, A., Barblett, L., & Barratt-Pugh, C. (2023). Discourses in power: Policy and curriculum demands in the first year of compulsory school. *Contemporary Issues in Early Childhood*, 24(3), 298–312. <https://doi.org/10.1177/1463949121990905>
- Sato, K. (2023). A non-ideal aim of redressing epistemic injustices in corruptive educational environments: Toward restorative epistemic justice. *Theory and Research in Education*, 21(3), 320–336. <https://doi.org/10.1177/14778785231207976>
- Schefers, S. E. (2026). Exploring intersectionality in identity research in multicultural education: Reflecting on the past to forge a more equitable future. *Asia Pacific Journal of Education and Society*, 14(1), Article 3. <https://doi.org/10.20897/apjes/17906>
- Schiefer, J., Edelsbrunner, P. A., Bernholt, A., Kampa, N., & Nehring, A. (2022). Epistemic beliefs in science—a systematic integration of evidence from multiple studies. *Educational Psychology Review*, 34(3), 1541–1575. <https://doi.org/10.1007/s10648-022-09661-w>
- Stroupe, D., Berk, L., & Kramer, A. (2024). Classroom communities for everyone. *Phi Delta Kappan*, 105(6), 32–37. <https://doi.org/10.1177/00317217241238105>
- Syed, H. (2024). Unravelling the deficit ideologies in English language education in Pakistan: A decolonial perspective. *TESOL Journal*, 15(4). <https://doi.org/10.1002/tesj.828>
- Teresi, J. A., Wang, C., Kleinman, M., Jones, R. N., & Weiss, D. J. (2021). Differential item functioning analyses of the patient-reported outcomes measurement information system (PROMIS®) measures: Methods, challenges, advances, and future directions. *Psychometrika*, 86(3), 674–711. <https://doi.org/10.1007/s11336-021-09775-0>
- Turner, L., Hashimoto, D. A., Vasisht, S., & Schaye, V. (2024). Demystifying AI: Current state and future role in medical education assessment. *Academic Medicine*, 99(1), S42–S47. <https://doi.org/10.1097/ACM.0000000000005598>
- Vlase, I. (2024). Unlocking the potential of the decolonial approach in migration studies. *Sociology Compass*, 18(9), 1–11. <https://doi.org/10.1111/soc4.70004>
- Walker, M. E., Olivera-Aguilar, M., Lehman, B., Laitusis, C., Guzman-Orth, D., & Gholson, M. (2023). Culturally responsive assessment: Provisional principles. *ETS Research Report Series*, 2023(1), 1–24. <https://doi.org/10.1002/ets2.12374>
- Wee, H. L., Karkkulainen, E. A., & Tateo, L. (2023). Experiences of epistemic injustice among minority language students aged 6–16 in the Nordics: A literature review. *Education Sciences*, 13(4), 367. <https://doi.org/10.3390/educsci13040367>
- Winkel, A. F., Telzak, B., Shaw, J., Hollond, C., Magro, J., Nicholson, J., & Quinn, G. (2021). The role of gender in careers in medicine: A systematic review and thematic synthesis of qualitative literature. *Journal of General Internal Medicine*, 36(8), 2392–2399. <https://doi.org/10.1007/s11606-021-06836-z>
- Wood, E., & Hedges, H. (2025). Curriculum in early childhood education: Global policy discourses and country-level responses in Aotearoa New Zealand and England. *The Curriculum Journal*, 36(1), 147–163. <https://doi.org/10.1002/curj.280>
- Worku, M. Y. (2025). Examining gender responsiveness in Ethiopian educational materials: Implications for sustainable development goals. *Sustainable Development*, 33(6), 8468–8478. <https://doi.org/10.1002/sd.70111>

- Wright, J. (2022). The deep roots of inequity: Coloniality, racial capitalism, educational leadership, and reform. *Educational Administration Quarterly*, 58(5), 693–717. <https://doi.org/10.1177/0013161X211029483>
- Yu, S., & Liu, S. (2025). Teaching writing in the Chinese primary and secondary school contexts: An analysis of the discourses of writing in L2 curriculum documents. *International Journal of Applied Linguistics*, 35(3), 1285–1297. <https://doi.org/10.1111/ijal.12699>