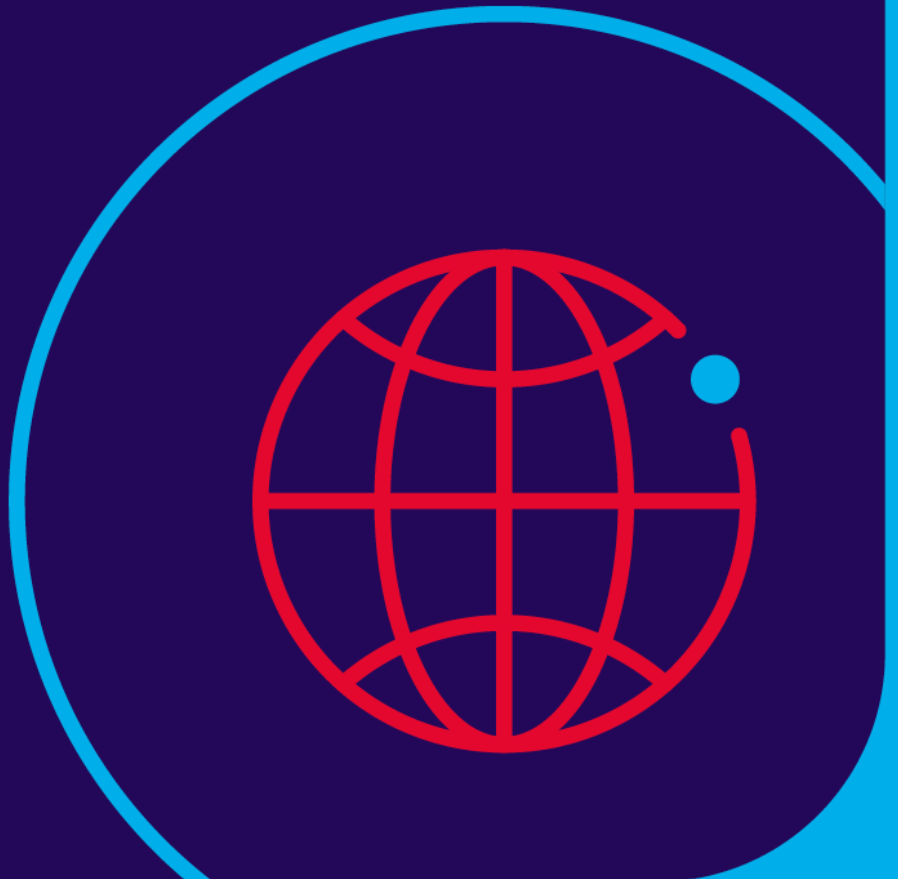


## Global Assessments

---

# Equality, diversity and inclusion (EDI) policy for test content creation and product quality review





# The equality, diversity and inclusion (EDI) policy for test content creation and product quality review

## Document details

|                                |  |
|--------------------------------|--|
| <b>Strategic Business Unit</b> | <b>English and Exams: Global Assessments</b> |
| <b>Document owner</b>          | Gemma Bellhouse                              |
| <b>Iteration author</b>        | Gemma Bellhouse                              |
| <b>Additional contributors</b> | Olena Rossi, Sheryl Cooke, Richard Spiby     |
| <b>Current version</b>         | 1.0  |
| <b>Issue date</b>              | 22 November 2022                             |

## Document details

| <b>Date</b> | <b>Version</b> | <b>Change details</b> |
|-------------|----------------|-----------------------|
|             | 0.1            |                       |
|             | 0.2            |                       |

## Access advice

This document has been prepared with referential links for the reader. The links may direct the reader to internal or external sites.



# Table of Contents

|   |           |
|---|-----------|
| <b>Table of contents</b> .....  | <b>3</b>  |
| <b>Introduction</b> .....   | <b>4</b>  |
| Background .....  | 4         |
| <b>The equality, diversity and inclusion (EDI) categories of test content creation</b> .....      | <b>6</b>  |
| 1. Group representation.....  | 6         |
| 2. Privilege balance .....  | 6         |
| 3. Sociocultural inclusion.....   | 7         |
| <b>The equality, diversity and inclusion (EDI) guidelines for image</b> .....                     | <b>8</b>  |
| 1. <b>Images</b> .....  | 8         |
| 1.1. Group representations.....   | 8         |
| 1.2. Privilege balance .....  | 8         |
| 2. <b>Audios and accents</b> .....  | 9         |
| 2.1. Background and rationale .....   | 9         |
| 2.2. Key principles underlying these guidelines.....  | 10        |
| 2.3. General recommendations: Group representation and accent variation .....                     | 10        |
| 2.2. General recommendations and guidelines .....   | 11        |
| <b>References</b> .....   | <b>12</b> |
| <b>The equality, diversity and inclusion (EDI) guidelines for short and long texts</b> .....      | <b>13</b> |
| 1. Privilege balance .....  | 13        |
| 2. Sociocultural inclusion.....   | 14        |
| <b>The equality, diversity and inclusion (EDI) checklist for item writing and reviewing</b> ..... | <b>15</b> |



# Introduction

## Background

The British Council’s work is based on building meaningful, enduring, and respectful relationships across different cultures. This cannot be accomplished without a strong commitment to equality, diversity and inclusion (EDI). EDI has long been a prioritised component of projects and processes. Activities to promote gender equality and equal opportunity in recruitment and promotion, as well as unconscious bias awareness, have increased since 2020.

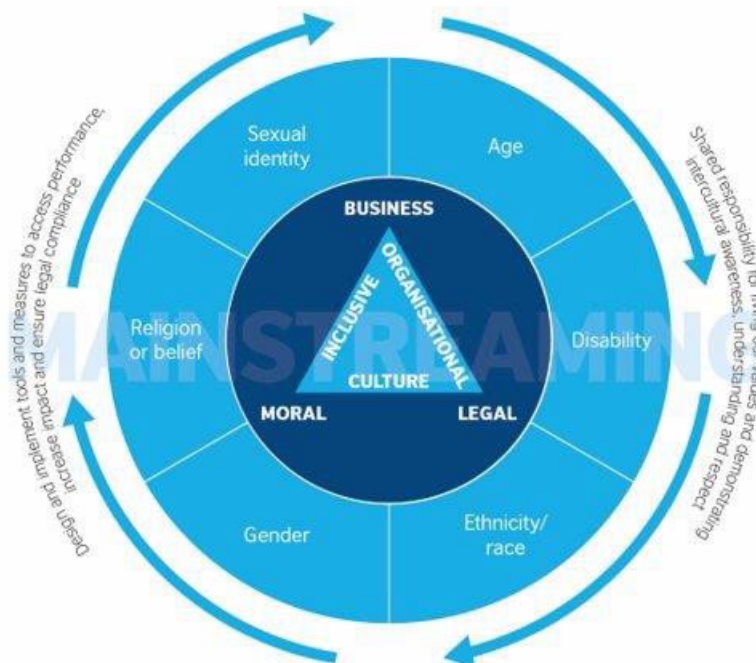


Figure 1: EDI Mainstreaming, Diversity Unit, British Council

The British Council refers to the idea of ‘mainstreaming’ EDI thinking into their processes (see diagram above). That is, that EDI is integral to each and every stage of every product and policy. Within the visual framework representation in the figure above, six areas of diversity go around the outer circle – Age, Disability, Ethnicity or race, Gender, Religion or belief, and Sexual identity. Recently a seventh focal category of ‘Socio-economic status’ was added.

These group categories are not exhaustive and should not be considered as comprehensive. The British Council aims to integrate these areas into the legal, business, and moral approaches of their products, with the hope of making everyone's experiences richer, and ultimately leading to more inclusive societies.



*Various forms of **sociocultural biases and formal and informal discrimination lead to social exclusion**, resulting in social vulnerability among individuals, households and disadvantaged groups and communities.*

***Stereotyping of any kind** hinders the cause of social integration by fostering misperceptions, bias and discrimination. Persons with disabilities, women, migrants, indigenous peoples and older persons all face misperceptions in their struggles for social acceptance and integration.*

### **UN Report on the World Social Situation 2003**

There is an opportunity for language test developers, test content creators, and item writers worldwide to build EDI stages into their processes, and to create test content using contemporary perspectives. English language tests are developed by the Global Assessments team at the British Council. Ultimately, creating test content that is inclusive and more diverse will also lead to test preparation and classroom activities that follow suit (in academic research, this impact is called ‘test washback’).

A new working group was set up in 2021 within the British Council by Judith Fairbairn with several aims. One of these aims was to review the Aptis test content with an ‘EDI lens’, in effect, to observe and record any presence of bias and stereotypes and to also observe any lack of diversity.

A professional review of test content produced by the Global Assessments team of the British Council, conducted in early 2021, indeed revealed the presence of test content that had the potential to uphold gender and cultural stereotypes and/or had a prevailing sense of Anglocentrism. Observations on these items were recorded and checked by Gemma Bellhouse, Test Product Quality Assurance Manager of Global Assessments, and Olena Rossi, PhD Language Testing and expert on item writing.

Potential stages for immediate action in which to implement an EDI review were identified, including:

1. test content creation
2. internal quality review
3. test content sign-off.

Informal general item writing and review guidelines were drafted and put into place in the following months by colleagues in Global Assessments. The implementation of these EDI guidelines as a formal policy began in November 2022. Furthermore, the mainstreaming of EDI into British Council processes continued with the dissemination of this document to multiple teams within Global Assessments including item writers, the test production team, quality assurance, as well as external consultants involved in test production. It is expected that this policy will immediately become a reference for all relevant and interested parties at the British Council who aspire to apply and uphold these EDI principles.



# The equality, diversity and inclusion (EDI) categories of test content creation

During the EDI Content Review in 2021, Gemma Bellhouse devised three categories: group representation, privilege balance, and sociocultural inclusion. This section will briefly define and give examples that fit within these categories.

## 1. Group representation

**Group representation refers to the general presence of minority and non-dominant groups.**

**The group categories adapted from the British Council Mainstreaming framework (Figure 1, above):**

1. **Different ages** (older, middle-aged, young adult, teenage, children; authority generation; vulnerable adults).
2. **Different ethnic and cultural groups** (majority and minority, including African people, West and East Asian people, Hispanic people, and Roma people).
3. **Different genders and sex** (men, women, transgender, intersex, other) and different marital status (single, married, civil partnership, divorced, separated, other), and also other matters related to family (Pregnancy, maternity, paternity and adoption).
4. **Different sexual orientations** (gay, lesbian, bisexual, heterosexual).
5. **Disabled people** (physical, sensory, learning, hidden, other) and different dependant responsibilities (childcare, eldercare, care for disabled and/or extended family).

## 2. Privilege balance

Privilege balance refers to class and racial privilege in authority roles and also the perspective and inclusion of gender roles, not limited to roles of authority in education, daily life and occupation. Privilege balance encompasses topics related to both people and context.

The British Council's Diversity Unit's Race Equality through Anti Racism guide says:

*Whiteness is increasingly being given attention, including through an examination of white privilege<sup>5</sup> and white fragility<sup>6</sup>. These concepts help highlight power and status, bringing attention to that which can be taken for granted and is often unspoken about. Racism is undeniably linked with other areas of unjustified discrimination. People are not simply identified by their*



*race/ethnicity. Self-identification and the way others see us is complex. We all have multiple identities, with different aspects that come to the fore at different times,*

### Section endnotes

<sup>5</sup> For more on white privilege see McIntosh, P. (2003). White privilege: Unpacking the invisible knapsack. In S. Plous (Ed.), *Understanding prejudice and discrimination* (pp. 191–196). McGraw-Hill.

<sup>6</sup> White fragility is where talking about racism triggers emotions or behaviours that close the conversation down. The phrase was coined by Robin DiAngelo in 2011

## 3. Sociocultural inclusion

Sociocultural inclusion refers to the representation and positive highlighting of community/society stereotypes and limiting the bias of Anglocentrism and middle-class norms. There is some overlap with the category of ‘privilege balance’ as the idea of privilege is also prevailing in this category. However, the idea of privilege is in reference to groups and their representation rather than individuals and their implied authority.

All test content, whenever possible, should reflect a conscious representation of minority groups and a positive sense of diverse communities from around the world.

Many of the sociocultural biases are unconscious, especially for content writers and reviewers who do not have a range of sociocultural experience. Some of these types of biases are highlighted in the eight sociocultural privileges as derived by Liu, Pickett and Ivey (2007):

1. The privileges of housing and neighbourhood: the middle-class privilege and expectation of safe and clean housing.
2. The privileges of economic liberty: the privilege to spend money freely and in ways that are self-satisfying and to feel relatively free from the problems of day-to-day needs.
3. The privileges of sociostructural support: the privilege of knowing that the government and power structures tend to benefit those in the middle class.
4. The privileges of power: the privilege of choice and the expectation of appropriate and respectful treatment by anyone.
5. Familiarity with middle-class behavioural norms: the privilege of being familiar with middle-class behaviours that allows one to negotiate the demands and expectations of middle-class culture.
6. The privileges of self-satisfaction: the privilege and expectation of being happy.
7. The privileges of leaving a heritage: the privilege of assuming that one’s family or children will be able to capitalise on the current successes.
8. The privileges of leisure: the privilege of having leisure time that is voluntary and does not jeopardize employment.

### Reference

Liu, W. M., Pickett Jr, T., & Ivey, A. E. (2007). White middle-class privilege: Social class bias and implications for training and practice. *Journal of multicultural counseling and development*, 35(4), 194-206.



# The equality, diversity and inclusion (EDI) guidelines for images and audios

## 1. Images

### 1.1. Group representation

Each individual version of a test, or each test form, or each test component (e.g. speaking, writing, etc) should have a minimum of 50 per cent of total image content that contains the presence of people from diverse groups. This is a general aim as a minimum that may be constrained by practicalities; ideally 100 per cent of the images in test content (that contain people) should have some representation of diverse groups.

Item writers can suggest certain details in pictures that will be shown with the tasks regarding ethnicity of person(s), gender, gender roles, and cultural representation. Please see the Section 3 of this document for the full list of groups that must be represented.

### 1.2. Privilege balance

Beyond the presence of diverse groups, representation must be handled carefully. This should be done by examining the authority that different genders and ethnic groups take in various roles and across contexts (home, school, and work). When possible, the authority should be given to the representative of a minority group and/or the female rather than the male in the task. When there are multiple instances of male and female representation within a task or test, the content should strive for a balance of authority and representation of both sexes (e.g. with two pictures within a task, a man can have the authority in one picture and a woman in the other). Additional equality grounds such as technical workers and physical labour workers and the geographical location of these roles should be considered.

**The following suggestions on authority and roles apply to all photos:**

#### *Home/Family Structure*

1. Men are taking care of children, rather than women as primary caregivers.
2. Men are doing domestic duties like cleaning and cooking.
3. Other types of family are represented other than the 'male-female-two children'.

#### *School/Work Roles*

1. Teachers as male in primary and secondary schools, and female at university level.
2. Women in leading positions and men asking questions.
3. Women in construction and other physical labour roles.





### *Sociocultural Inclusion*

1. Images should represent a range of cultural and community lifestyles and backgrounds.
2. Images should avoid representing Anglocentric, and privileged middle-class situations (a holiday on a yacht or all-inclusive resort; people holding glasses of wine looking at paintings in an art gallery).
3. Images should avoid representing the typical heterosexual, white western family tradition (e.g. a male and female parent, both of roughly the same age or the man is a bit older, with two children – girl and boy).
4. When images contain western contexts or landscapes, this should not be taken as the norm; when a different context is shown, it need not represent something 'exotic' or 'unusual'.
5. Work situations should be varied (e.g. not only an office situation or academic work). There should be manual, physical labour in multiple contexts and with a positive tone.
6. Locations should be representative of a range of geographical contexts.

## 2. Audios and accents

### 2.1. Background and rationale

The British Council's Assessment Research Group invited language testing expert, Luke Harding PhD, to consult with members of the test production and quality assurance teams of Global Assessments to discuss accent usage in language tests. On 21 June 2021, they held a workshop on this topic. Also present at the workshop was Sheryl Cooke, British Council, an expert in the listening component of language testing. This section and especially the tables and detailed guidance for items with audio content has been influenced by the workshop. The background, rationale, principles, and general guidelines were primarily authored by Sheryl Cooke of the Assessment Research Group.

**Two key elements of language testing** support the rationale for the development of the guidelines for accent use and representation in British Council-developed Aptis and other tests: **validation and washback**.

**What is the validation of tests and why is it important?** Validation refers to what McNamara terms as '...ensur[ing] the defensibility and fairness of interpretations based on test performance' (2006, p. 33). In other words, in order for a test to be valid, the decisions (interpretations) being made about someone's ability should be based on a chain of logical links that stretch from test-score to the ability being measured; the performance should be a reflection of that underlying ability. A key tenet of language test validation is ensuring that the Target Language Use (TLU) domain is reflected in the test, i.e. the real-world context in which the communicative ability will be used (Bachman and Palmer, 1996). To do so, we need to recognise the significant shift in the intended use of the English language from what Berns identified (amongst other assumptions) as 'everyone learning English does so in order to



interact with native speakers' (2006, p. 723). Indeed, most conversations in English over the last decade or more do not include a native speaker (NS), or first language (L1) user of English (Graddol, 2006; Kirkpatrick, 2007) and the function English performs in most contexts is to facilitate communication (Seidlhofer, 2009), i.e. to achieve mutual comprehensibility between speaker(s) and listener(s). This is also the case in higher education contexts in traditional English-speaking countries such as the UK and Australia (Cooke, 2020). The newest CEFR Companion volume has stated that the 'native speaker' is no longer the linguistic standard (Council of Europe, 2020). It would seem, then, that a focus on comprehensibility in English language testing is warranted.

**Tests have a significant impact on what teachers and learners focus on** in a language learning environment. Even in contexts in which curricula are changed, this washback results in teachers 'teaching to the test' and as such, is a key driver in the selection of a variety of English for study. It is a chicken-and-egg conundrum: one argument is to maintain a narrow range of standard (or prestige) varieties in language tests because this is what test-takers are familiar

with, yet if this is continued, exclusion will also be perpetuated.

**There are certainly challenges**, including the cost of using a range of accents in the Listening component, potential impact on reliability of introducing any variable (at least until test-takers orientate away from one taught variety to a comprehensible range), and potential resistance from test-takers, teachers, parents and other stakeholders to non-prestige varieties. Despite these hurdles, and as with other EDI considerations in this guide, underrepresentation has consequences on fairness and inclusion, and our position is one of inclusivity.

## 2.2. Key principles underlying these guidelines

1. The British Council is the leader in promoting EDI in language assessment; this includes actively promoting EDI with all stakeholders.
2. Validity includes the requirement that the test is representative of the TLU domain.
3. We recognise that prestige, native-speaker varieties are no longer benchmarks of communicative success.
4. While variation in English includes a host of linguistic and non-linguistic features, we recognise the significant role that accent plays as a marker of identity, socio-economic disparity and overall prestige, and the potential it has to be exclusive.
5. The British Council brand is recognised and trusted in English language teaching and testing; as such, we are able to drive the shift towards accent inclusivity.
6. We recognise the practical considerations inherent in producing Listening tests that are accent-inclusive but remain committed to leading on an incremental shift towards this goal.

## 2.3. General recommendations: Group representation and accent variation

The following are recommended percentages based on practical limitations of locating voice actors and are flexible as per the test's needs and any further constraints met by teams of test production.



| Recommended percentages and representations of accents in recordings |                                     |   |   |
|--|-------------------------------------|---|---|
| Level  | Group representation                | Countries/regions   | Percentage/number of tasks in a listening version |
| 1  | Great Britain                       | Scotland, Wales, Ireland, Caribbean, local and regional UK (e.g. East London, Northern England, South London) | 10-20%  |
| 2  | Outside UK but still L1 English     | Australia, Canada, New Zealand, America, South Africa   | 10-20%  |
| 3  | English as a lingua franca          | Singapore, Malaysia, Nigeria, Sudan, Ghana, India, Bangladesh, Pakistan                                       | 20-30%  |
| 4  | EFL (English as a Foreign Language) | Europe, Asia  | 20-30%  |
| 5  | Age representation                  | Younger and older speakers (recognisably)   | 10% younger than 18<br>10% older than 60          |

### 2.4. General recommendations and guidelines

1. Levels 3-5 are recommended for B1-C2 level tasks.
2. Within dialogues, the accents should be mixed (e.g. Scottish with African, Irish with American).
3. It is recommended to use professional voices who have not just been **trained** in accents, but professional actors who have **spent time in these regions and who are proficient speakers (C2 level) of the language** (e.g. Hindu for Indian, Italian for Italian). Voice actors tend to present inconsistencies and a potential lack of authenticity that can be quickly apparent to L1 speakers or people who are familiar with these languages.
4. Include a **broad range of accents** in Listening tests to ensure representation of the TLU domain:
  - Where the test is general, i.e. TLU domain is unspecified or ‘global’, ensure non-native varieties, regional varieties, and varieties associated with different ethnic and socio-economic groups in line with the approach to EDI described above.
  - Where the test is specific, i.e. for an intended TLU domain purpose (e.g. vocational training in Singapore), localise or ‘map’ the accents to this domain as closely as is practical.
5. Mitigate potential reliability issues by building up a **pool of voice-actors** representative of a (growing) range of accents to:
  - If possible, ensure ‘light’, comprehensible accents as pre-screened by a panel and, preferably, a range of listeners (potentially using a rating of



comprehensibility by British Council global colleagues).

- Ensure ‘authentic’ accents as self-identified by the speakers, avoiding accent mimicry; include mixed/hybrid accents (e.g. Japanese-American).
  - Include adequate samples of these voice-actors in trialling and pre-testing in order to build up a body of evidence to support reliability across a diverse test-taker candidature.
6. Make Listening tests fit-for-purpose by making them **more easily customisable**:
- Avoid asking item writers to indicate accent variety for two reasons:
    - A decision has been made not to include lexical and grammatic variation – as such, item writer Listening transcripts (i.e. written texts) can be repurposed by recording different accents where there is a demand, i.e. accent can be determined by the head of test development in order to ensure broad representation or where there is demand for a specific variety.
    - Mitigate bias and stereotyping, e.g. item writer preference for certain accents, or linking of particular accents with roles, content, names, etc.
  - Build up a tagged item bank, i.e. Listening items are tagged with accent information, so that test versions can be easily ‘mapped’ to accent specifications of the TLU domain.
7. Ensure clear **communication with all test stakeholders**:
- Include explanation of our approach in test literature.
  - Include a variety of accents in practice materials and make explicit to test-takers.
  - Integrate into the descriptors, e.g. ‘understands relevant varieties’.
  - Engage with test-score users to highlight the positive implications of accent variation.
  - Publish and present widely on the topic.
  - Include accent variety inclusion in a broader EDI checklist for a range of stakeholders.

## References

- Bachman, L.F. and Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Berns, M. (2006). Word Englishes and communicative competence. In B.B. Kachru, Y. Kachru, & C.L. Nelson (Eds.), *The handbook of world Englishes*, (pp 718–730).
- Cooke, S. (2020). Assessing real-world use of English as a Lingua Franca (ELF): A validity argument. *VNU Journal of Foreign Studies* 36(4) DOI: [10.25073/2525-2445/vnufs.4574](https://doi.org/10.25073/2525-2445/vnufs.4574).
- Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment. Companion volume*. Council of Europe.
- Graddol, D. (2006). *English Next*. London: British Council.



# The equality, diversity and inclusion (EDI) guidelines for short and long tests

For all components: Speaking, Listening, Reading, Writing and Grammar and Vocabulary.

## 1. Privilege balance

### 1.1. Some examples of gender stereotypes to avoid

- Clothing: women only wear dresses and all men wear suits. While these articles of clothing are still suitable, make a point to also include other articles of clothing.
- Punctuality: women are late getting ready, and men have to wait for them. If possible, avoid people being late or make it a neutral reason out of their control.
- Education: men must explain things to women and women don't always understand. Aim to write texts in which neither man nor woman is ignorant or ill-informed, and both are informed in different respects.
- Priorities: men forget holidays like birthdays and women get very upset about these types of scenarios. Try to avoid any scenarios in which people feel personally insulted or unappreciated. Use a healthy conflict resolution approach in which people feel heard, the conflict is easily resolved, and there is positive communication throughout.

### 1.2. Power Roles

- Women should be represented in power/authoritative roles, e.g. experts, university professors, company bosses, etc. If possible, either avoid mentioning the gender, e.g. 'The president was unhappy about the message' or have a balance of power, e.g. both the woman and man are a manager / lecturer / expert.
- Men should be given domestic roles: staying at home, cooking, cleaning, taking care of the children and generally supporting the family.

### 1.3. Relationships and interactions between men and women

- In a dialogue, the woman should not be the one who lacks knowledge, asks questions, is looking for information, or is persuaded by the man. Women should sound knowledgeable and assertive.
- Neither man nor woman should sound ignorant or be unnecessarily put in a negative light.
- Include friendships that are diverse in gender (men can be friends with women).

### 1.4. Names

- When multiple names are used in one text, aim to select female, male, and gender neutral. Use personalised names when possible and not 'he' and 'she'. Use 'he' and 'she' equally.



## 2. Sociocultural inclusion

### 2.1. Topics

- All texts should have inclusive scenarios, especially related to sociocultural topics (e.g. discussions on progress in EDI-related topics without a negative slant).
- Use a range of topics that are not UK-centred or spoken from the British, or American, or expat in Europe perspectives, e.g. holidays in France, going to places as a western tourist, hating learning French at school. Have sentences that represent life in different countries, people of different ethnic groups, and diversity in all senses.
- Use a range of topics that touch on global problems/perspectives (e.g. pollution in a city that is not in the UK). Use more culturally experienced voices within texts to avoid a bias towards a particular culture or societal group (e.g. travellers, university-educated people, those interested in the arts) as well as voices representing minority and indigenous groups, rather than always having the 'white gaze'.
- Business meetings can take place anywhere, not only English-speaking locations.
- Avoid gender stereotypes of sports and culture (e.g. for topics on football/science, a woman can be speaking, and for art/language, a man can be speaking).
- Avoid mentioning anything about doing well/badly in a test/exam, studying for an exam, or being nervous about an exam. This can trigger candidates who already have anxiety, and especially those who are given special accommodations to try to prevent anxiety.
- Avoid topics that might be upsetting, e.g. natural disasters, accidents, illnesses, crime (see PARSNIPS).

### 2.2. Relationships and interactions

- When two friends' names are given, use names from different nationalities to represent diverse connections.
- Use a variety of names from different cultures as well as a range of names from English speaking backgrounds. Be careful to not introduce gender stereotypes or cultural stereotypes with the names.
- When multiple voices exist in one text or task, use a diverse representation of cultural backgrounds, including different perspectives and opinions. Ask the item writers to draw on their experience from travel and living elsewhere and put themselves in someone else's shoes to give a variety of perspectives.



# The equality, diversity and inclusion (EDI) checklist for item writing and reviewing

An adaptation of the content creation checklist by the British Council Content Development Team.

| The Item Writer confirms that the content, characters, topics, images, and audios...   |          |       |
|--|----------|-------|
|  | Yes / No | Notes |
| 1. adhere to <b>PARSNIPS conventions</b> (politics, alcohol, religion, sex, narcotics, -isms and pork).  |          |       |
| 2. avoid <b>stereotypes</b> of gender, culture, race, religion, or ethnicity, or of any sociocultural and economic variety.  |          |       |
| 3. take Islamic, Christian, Jewish, and other religious sensitivities into account.  |          |       |
| 4. include references to <b>different or non-religious or philosophical beliefs</b> .  |          |       |
| 5. include an appreciation and representation of <b>gender equality</b> and a strong representation of <b>female roles</b> of authority.   |          |       |
| 6. include broader definition and/or representation of physical and invisible <b>disabilities</b> .  |          |       |
| 7. include references to <b>additional equality grounds</b> (such as full-time/part-time working, language, geographical location, other).   |          |       |
| 8. include references to <b>different marital statuses</b> (single, married, civil partnership, other) and also <b>pregnancy, maternity, paternity and adoption</b> (before/during/after). |          |       |
| 9. include broader definition and/or representation of <b>sexual orientation, gender identity, and sex characteristics</b> (gay, lesbian, bisexual, heterosexual, transgender, intersex).  |          |       |
| 10. include <b>scenarios</b> that are culturally inclusive <b>with community backgrounds</b> .   |          |       |