PISA READING COMPONENTS TRANSLATION AND ADAPTATION GUIDELINES

Document: PISA_ReadingComponents_TA_Guidelines.docx

John P. Sabatini
Kelly M. Bruce

Educational Testing Service
Princeton, NJ, USA

December 2011

TABLE OF CONTENTS

Part 1: Introduction

This document describes the procedures for *translating and adapting* the source version of English reading components items into the language(s) of assessment in your country.

The reading components measures are:

1. Word Meaning (Print Vocabulary)
2. Sentence Processing
3. Basic Passage Comprehension

1.1 Overview of the Measures

In previous cycles of PISA, students who fell into the bottom two levels of Reading Literacy (Levels 1 and 2) were defined by what they *could not* do with text. The reading components measures are designed to provide information about what these students *can do* with respect to the reading components, or the building blocks of reading literacy proficiency.

Word Meaning (Print Vocabulary) measures the extent to which students can recognize the printed forms of common objects. Sentence Processing measures the extent to which students can comprehend sentences of varying levels of complexity. Basic Passage Comprehension measures the extent to which students can comprehend literal meaning of connected text and make basic inferences about the text.

It is important to note that the task of translating and adapting the English source version is not meant to produce a literal, word-for-word translation into the language of assessment. Similarly, it is not meant to produce a version in which you have endeavored to preserve item-level difficulty. Rather, the task is to apply the following guidelines to produce item sets that maintain the critical features of the constructs across languages, as noted below.

Part 2: Translation & Adaptation of English Source Version

2.1 Word Meaning (Print Vocabulary)

Construct: Word Meaning (Print Vocabulary) is the ability to recognize the printed forms of objects and concepts given the particular orthographic and morphological structure of the language. In the PISA Reading Components assessment, this construct has been instantiated in a picture vocabulary task. Its purpose is to measure the extent to which students can recognize printed forms of common, everyday objects.

Task Description: The participant sees a black-and-white line drawing of a common object. Underneath the drawing are four words. The participant must circle the word that matches the drawing.

Word Meaning (Print Vocabulary) Item Type Example:

shoulder	shark	bird	radio
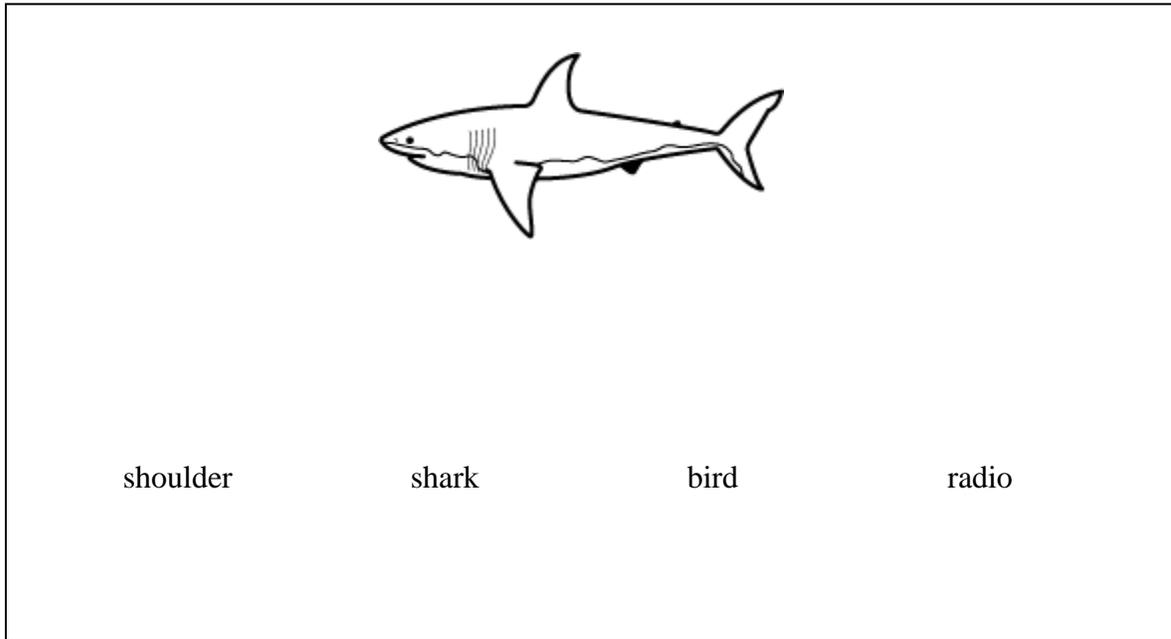
Terminology:

Every item contains four choices: the key and 3 distractors:
- **Key** – the correct choice, the word that matches the drawing
- **Distractor 1** – an incorrect choice, a word that is in the same semantic category as the key
- **Distractor 2a/2b** – an incorrect choice, a word that is based on the spelling pattern of the key
- **Distractor 3** – an incorrect choice, a word that is unrelated semantically to the key and is of intermediate length relative to the key and other distractors

Here is the previous example with the types added:

| shoulder | shark | bird | flower |
| Distractor 2a | Target | Distractor 1 | Distractor 3 |
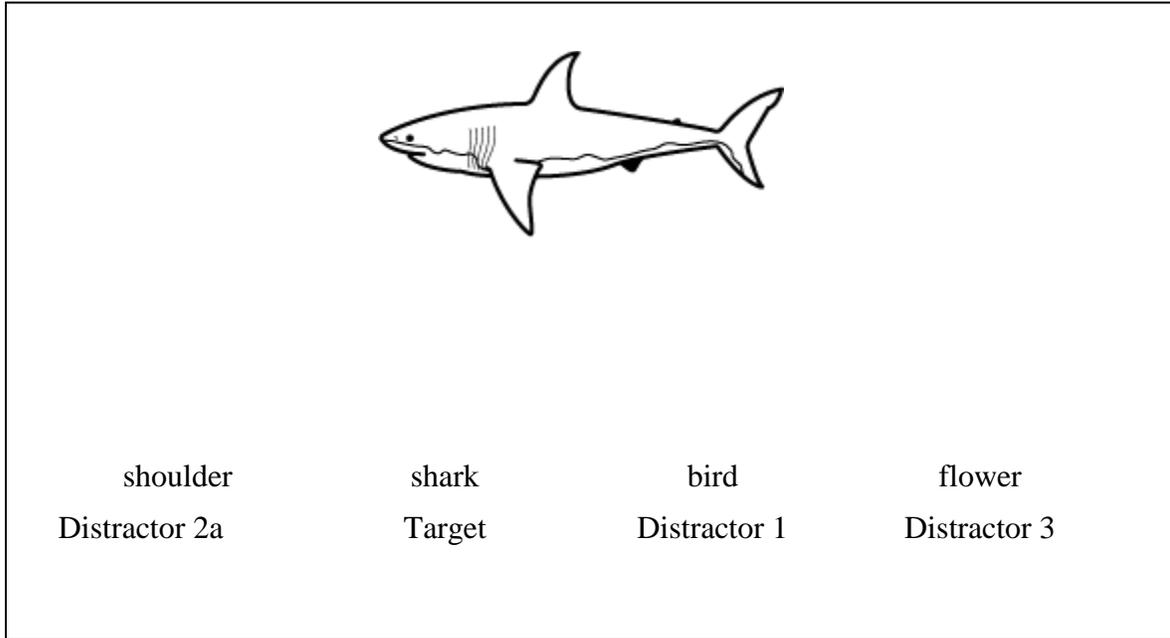
Array of Distractor Types:
The **Key** and **Distractor 1** are present in every item. **Distractors 2a**, **2b**, and **3** are distributed across the items. Please maintain the array (positioning) for the items in this set. See Table 1, below.

Translation & Adaptation Procedure:

Note: Target items were selected based on their commonality (i.e. everyday objects known across cultures) as well as their ability to be pictured (i.e. their concreteness, or imageability). All participating countries will use the same set of black-and-white line drawings to assess this component.

1. Translate the **Target** word into the language of assessment. Choose a translation that is the most common way of describing the object. For example, one of the images shows the picture of a baby. In English, one could translate this as "baby" (common, highly frequent) or "infant" (less common and frequent). Choose the more common, more frequent translation, i.e. in English, we choose "baby."

Similarly, choose a translation that describes the most generic description of the Target. For example, one of the images shows the picture of a bird. You should translate this as "bird" rather than a particular kind of bird, e.g. "robin," "crow," "sparrow," etc.

Translation may change the length of the word (letters and syllables) from the English source version. This is acceptable. If the translated **Target** is more than one word, then make sure at least one distractor is of the same word length as the target.

2. Translate the English version of **Distractor 1**. Again, as with the Target, if there is more than one translation for **Distractor 1**, choose the most common, most frequent word. Similarly, the translation of Distractor 1 may change the length of the word (letters and syllables) from the English. This is acceptable.

Note: if **Distractor 1** is translated as two (or more) words, *but the Target is translated as one word*, then:

Step 2a. Select another high frequency word within the same semantic category as the Target and document the fact that your Distractor 1 is different from the English source version because of this guideline.

Step 2b. If an alternate cannot be found, then generate one other distractor to take the place of 2a/2b/3 that is also two (or more) words. The remaining distractor should be one word and about the same length as the Target. Then document the differences from the English source versions because of this guideline.

3. At this point, you have the translations for the Target and Distractor 1 (both of which are part of each choice set for every item). Now, you will generate the remaining distractors (some combination of 2a/2b/3 depending on the array for that particular item) according to the following guidelines:

- Generate **Distractor 2a** by selecting a high-frequency word of roughly a length that is between the **Target** and **Distractor 1** AND has the same initial letter or cluster as your translation of the Target.

- Generate Distractor 2b by selecting a high-frequency word of roughly a length that is between the **Target** and **Distractor 1** AND has the same medial or final cluster as your translation of the Target.
  Note: If it is impossible to generate a word in your language with a spelling pattern that matches the Distractor type (2a vs. 2b) for that choice set, then try to generate a word of the opposite type of Distractor (2b for a 2a type; 2a for a 2b type).

- Generate **Distractor 3** by selecting a high-frequency, noun that in length (letters or syllables) falls between the shortest and longest length (in letters or syllables) of the **Target** and **Distractors 1** and **2a/2b**. Here are the steps to follow in generating **Distractor 3**.

Step 1. First, translate the English **Distractor 3**. If the translation is the appropriate length and does not contain the same initial or medial/final cluster as the

Target, then you are done. (Sharing the same final letter is to be avoided if possible, but it is acceptable.)

If not:

Step 2. Generate a high-frequency noun that falls between the shortest and longest length (in letters or syllables) between the target and **Distractors 1** and **2**. Make sure it is NOT the same semantic category as the target.

Below are examples of different item arrays and possible choice sets in English based on the sample "shark" item from above:

Example 1: Distractor 2a, Target, Distractor 1, Distractor 3

Distractor 2a	Target	Distractor 1	Distractor 3
same initial letter or cluster as target		same semantic category as target	unrelated noun of appropriate length (>4 letters, < 8 letters)
shoulder	shark	bird	teeth

Example 2: Distractor 2b, Target, Distractor 1, Distractor 3

Distractor 2b	Target	Distractor 1	Distractor 3
same medial or final letter or cluster as target		same semantic category as target	unrelated noun of appropriate length (>4 letters, < 6 letters)
g**ar**den	shark	bird	teeth

Example 3: Distractor 2b, Target, Distractor 1, Distractor 2a

Distractor 2b	Target	Distractor 1	Distractor 2a
same medial or final letter or cluster as target		same semantic category as target	same initial letter or cluster as target
g**ar**den	shark	bird	**sh**oulder

Other general guidelines:
- All distractors should be high-frequency, concrete nouns.
- If possible, use the Distractors only once.
- Where possible, avoid linguistic/morphological features unique in the language that would automatically cue a reader that a Distractor is implausible.
- Where possible, avoid making the **Target** word of a different overall length than all of the **Distractors**.
- Make sure there are a range of shorter and longer words in each item, OR that all words are of about the same length.

Note: By following these guidelines and procedures, only the **Target** and **Distractor 1** will be *translations* of the English source version. You will generate **Distractors 2a** and **2b** based on the translations of the Target and Distractor 1 and the guidelines provided above. **Distractor 3** may be a translation of the English source if it fits the guidelines described above; otherwise, you will generate it, too, based on the guidelines provided above.

2.2 Sentence Processing

Construct: Sentence Processing is the ability to comprehend written sentences of varying lengths. In the PISA Reading Components assessment, the construct has been instantiated in a sensibility judgment task. Its purpose is to measure the extent to which students can comprehend sentences of increasing lengths.

Task Description: Students see a set of sentences and decide if they make sense (YES) or do not make sense (NO) with respect to general knowledge about the real world, as in Example 1; or the internal logic of the sentence itself, as in Example 2.

Examples 1-2:

1. The sky is green.   YES   **NO**

2. If the house is taller than the tree, then the tree is shorter than the house.   **YES**   NO

Translation & Adaptation Procedure:

Note: The complexity of the target sentences in English was varied by adding modifiers, prepositional phrases, infinitives, and embedded clauses to the basic English grammatical structure of:

noun phrase + verb phrase + direct object

1. Translate the English source version sentence into the language of assessment.

    Step 1a: Adapt the grammatical structure (i.e. word order, representations of verb tense and plurals) as needed while **maintaining the part of the sentence that determines the sensibility judgment**.

    Note: Here are some guidelines as to how complexity varies across the set in English that may helpful to know as you translate:

    - Items 1-5 include items of "noun + verb", "noun + verb + object", and one modified subject or object.

- Items 6-8 differ from Items 1-5 in having one or two modifiers in the subject and object phrases.
- Items 9-10 differ from Items 6-8 in having a dependent or relative clause.
- Items 12-16 differ from items 9-10 in having a dependent or relative clause AND an additional prepositional phase, infinitive, or modifiers in the noun phrases.
- Items 17-20 differ from items 12-16 in having a wide mix of one or more relative and dependent clauses with embedded prepositional phrases, infinitives, modified noun phrases, and adverbials.

Step 2a: Adapt words used in the sentences to reflect common usage in your country as needed, i.e. "soccer" → "football."

2.3 Basic Passage Comprehension

Construct: Basic passage comprehension is the ability to understand the literal, "gist" meaning of connected text and to make low-level inferences across sentences in the text. In the PISA Reading Components assessment, the construct has been instantiated in an embedded cloze task. Its purpose is to measure the extent to which students understand the literal and inferential meaning of connected text.

Task Description: The participant sees a passage in which the sentences include an embedded cloze item (two word choices are given for a single blank). The participant reads the passage silently and circles the word that correctly completes each sentence.

Example 1, with correct responses in **bold**:

Michael was happy that he was getting a new computer. He had saved his files / **money** for several months to pay for it. He went to a store near his house to **buy** / sell the model he wanted.

Example 2, with correct responses in **bold**:

The company announced that it would close the factory.. Two hundred workers would be without a **job** / sky. But the company planned to build a new factory in the same **town** / cup. The workers were happy to hear the road / **news**.

Translation & Adaptation Procedure:

Note: The assessment contains different text types. They are texts that would be understood by most speakers of the target language if the texts were read aloud.

In Passage 1, the students are required to make low-level inferences across sentences in order to select the correct answer (see example 1, above). In Passages 2-4, the distractors are designed to be obviously wrong semantically (see example 2, above). The distractors are also mostly *high*

frequency obviously wrong choices to minimize the possibility that the respondent has never encountered one of the choices before.

1. Translate the English source passage, including distractors, into the language of assessment.

2. Adaptations to the structure of the sentences may be made to maintain the approximate level of difficulty of the passage, but the essential meaning of the passages must not be changed from the English originals.

3. Adaptations of individual words ("soccer" → "football") may be made to reflect common usage in your country.

4. Adaptations of proper nouns (places and people) may be adapted to reflect names used in your country.

5. The structure of the language of assessment may change the location of the cloze choice from the end of the sentence to another part of the sentence. This is acceptable. If possible, use the same word as in the English version. If not possible, choose another word from the same part of speech (noun, verb, adjective) as close to the end of the sentence as possible.

6. If a single word in the English version is a composite of two (or more) words in another language, but each word can be used individually, then write a distractor to one or another of the compound words.

7. If a pair of words in the English version is compounded into one word in another language, write a distractor for the one compounded word. The new distractor should not share the lexical meaning units (root, affix, word). For example, in this task (unlike the word meaning component), DO NOT create a key and distractor pair such as "telescope / telephone."

8. In general, both the key and distractor should be conjugated/declined so that the conjugated/inflected form grammatically fits the sentence. Please avoid conjugating/declining the key while leaving the distractor in its basic form.

9. If the English distractor when translated can be confused for a metaphor, idiom, or colloquial phrase or saying, then generate a distractor word in the same semantic category as the English word that does not make sense in the sentence.

10. If gender is encoded in the language, then include information in the choices that reflected in the choices. For example, articles (a, the) can be included in the choice as: das Boot / der Mann.