

BILC Shared Item Bank Project

Launched in 2020, the BILC Shared Item Bank (SIB) project represents a collaborative effort led by the Working Group (WG), comprising 14 language experts from Canada, Croatia, Estonia, Latvia, Romania, Slovakia, and Slovenia. This initiative, first conceived at the 2019 BILC Testing Workshop, aims to develop a bank of reading and listening items at STANAG 6001 levels 1, 2, and 3. These items can be utilized by NATO and partner countries as anchor items in their national reading and listening tests. Through this process BILC aims to:

- ensure uniform interpretation and application of the STANAG 6001 Language Proficiency scale.
- uphold the highest standards in item development, moderation, and analysis.
- strengthen linguistic interoperability among NATO and partner countries through standardization.

Development and Moderation Process

The project started with the creation of the reading item specifications in February 2020, which relied on the BILC Best Practices in STANAG 6001 Testing and the WG's collective expertise. The multiple-choice question format was selected as an objectively scored and the most commonly used testing method across NATO countries. By August 2020, a total of 47 items were submitted by the WG countries (12 at Level 1, 17 at Level 2, and 18 at Level 3). MS Teams platform was used for the storage and collaborative work on the items.

Through 14 online moderation sessions held between January 2021 and April 2022, the WG finalized 43 items. These items underwent a piloting process with teachers from Canada, Latvia, Slovakia, and Romania, resulting in the further refinement of the item pool.

Pre-Testing and Statistical Analysis

The WG compiled two pre-test versions, each containing 24 items, with six shared items across levels. Administered as paper-and-pencil tests to ensure uniform conditions, these tests aimed to assess item efficacy among a diverse volunteer group ranging in language ability from level 0+ to level 3(+). Approximately 1,000 candidates from 16 countries participated in the pre-testing phase.

For statistical analysis, WG employed various statistical tools like TiaPlus, jMetrik, snowIRT in jamovi, and the Dexter package in RStudio. These analyses helped evaluate items based on facility value, item-rest correlation, standard deviation, standard error, logit values, infit and outfit values for items and Cronbach's alpha for the whole population.

Outcomes and Impact

As a result of this rigorous process, the WG would like to share 14 high-quality anchor items (4 Level 1, 3 Level 2, and 7 Level 3) with NATO nations and partner countries. Additionally,

for educational and practice purposes, the WG has posted 6 other items (2 per level) on the NATO BILC website. They are available at the link below.

Conclusion

The SIB project marks a significant step in the continuous effort by the BILC community to harmonize language proficiency assessment across NATO and partner countries. It not only reinforces the standardized application of the STANAG 6001 scale but also promotes best practices in language testing.