

**Innovative application of ICT in education (chair: Attila Pásztor)**

**T-13**

**AN INVESTIGATION OF THE VISUAL SUPPORT PROVIDED IN EFL ASSESSMENT TASKS**

**Gyula Tankó**

*Institute of Education, Eötvös Loránd University*

**Keywords:** pictures as visual support in assessment tasks; construct validity; item writing

For the justifiable assessment of foreign language ability, the ability to produce and understand discourse in meaningful communicative situations (Bachman & Palmer, 2010), the characteristics of the language assessment task must be such as to replicate as feasible the conditions of actual performance (Skehan, 1988) and require realistic discourse processing (Weir, 1990). Such a task characteristic is the addition of a still picture, which is intended to serve as visual support for the interpretation of the input. From a variety of graphics, pictures were found to have the most pronounced positive effect on reading comprehension (Guo et al., 2020). Furthermore, pictures are considered to have the potential to enhance listening comprehension if they define the context clearly and thereby help listeners activate task-relevant schema. However, if they are not adequate, pictures may become a source of difficulty and cause construct-irrelevant variance (Buck, 2001). To determine to what extent they helped create a real-life like context facilitating the test takers' meaningful use of language, this study investigated the pictures included in Reading (R), Use of English (UoE), and Listening (L) tasks as visual support intended to facilitate the processing of textual and aural input. Altogether 88 tasks (i.e., NR = 32; NUoE = 32; and NL = 24) were analysed, which were administered in the 2017–2020 examination periods as part of the high-stakes, advanced level EFL School-leaving examination in Hungary. A pilot sample of the pictures were first analysed by two coders with a coding scheme consisting of a codebook—which operationalised a taxonomy of visual-verbal relations (Tan et al., 2012)—and a recording form. Then the two coders independently coded the items in the remaining papers. Inter-coder reliability was substantial ( $\kappa = .782$ , 95% CI [.89, .99],  $p < .000$ ). The findings showed that in most tasks the pictures were used as illustrations, and a few of the pictures repeated the gist of the input, which set the context and could potentially help the activation of the schema necessary for comprehension. However, several pictures in the Reading and Use of English tasks were unconnected and most likely hindered discourse processing. Some pictures were also found to be dysfunctional because they were indiscernible, watermarked, and could become functional only after the input had been successfully processed. Some provided irrelevant and misleading details. Moreover, attempts made to illustrate more complex content failed. If they were used by the test takers, the dysfunctional pictures will have caused construct-irrelevant variance and therefore affected the interpretability and generalisability of the scores. Furthermore, they are likely to have also induced bias in the measurement process. Based on the findings, practical recommendations are formulated for item writers.

*The research was supported by OTKA K142536.*