# Research Manifesto

This manifesto is the result of the symposium "Language Testing in Austria: Towards a Research Agenda" held in Klagenfurt on October 19th to 20th 2018. The working groups discussed desiderata for future research with reference to an adapted version of Weir's (2005) validity framework. This manifesto is an attempt to synthesise and summarise the suggestions made by the participants. The aim of the manifesto is to be a stimulant for future research. It does not aim to be all-encompassing or prescriptive in any sense.

This list provides a structured overview of the suggestions made.

1.  **Context validity**
    1.1. Research into genre-specific features of the text types (Textsorten) required in writing. Identifying such features is an important contribution to establishing contextual validity, because knowledge of these features is a prerequisite for defining criteria for correctness. This should be given high priority in L1 German and also in the modern foreign languages.
    1.2. Research into difficulty-generating features of items and tasks

2.  **Theory based validity**
    2.1. Cognitive strategy use (think-aloud, eye-tracking, mouse-click protocols)
    2.2. More detailed descriptions of the competence models underlying individual subtests of Matura (particularly Language in Use) and educational standards tests.
    2.3. Developing competence level descriptions for the classical languages in analogy to CEFR
    2.4. Developing competence level descriptions for L1 in analogy to CEFR

3.  **Scoring validity & Standardsetting**
    3.1. Research into difficulty-generating features of items
    3.2. Research into causes of discrepancies among expert judgements of item focus and difficulty
    3.3. Research into causes of discrepancies between estimated and empirical item difficulty

4.  **Rating validity & Benchmarking for writing and speaking**
    4.1. Rating scale validation
    4.2. Attempts to synthesize rating scales with aquisitional sequences identified in SLA
    4.3. Research into rater cognition
    4.4. Research into difficulty-generating features of tasks
    4.5. Research into rater reliability by multiple rating of a sample of matura performances

5.  **Consequential validity**
    5.1. Research into washback effects on teaching, learning, classroom testing, textbooks
    5.2. Use research designs which include a wide range of test stakeholders (learner, teachers, parents, government and official bodies, …).

6.  **Criterion-related validity**
    6.1. Research into language independent competences (in the receptive and the productive skills), e.g. correlation of English Writing with German Writing within BIST and within Matura.
    6.2. Research into convergent and discriminant validity of subtests within Educational Standards and of subtests within Matura (particularly Language in Use with aspects of Reading)
    6.3. Research into the predictive validity of BIST for Matura and Matura for success in tertiary-level studies.

**Further desiderata**
-   Promote assessment literacy
-   Promote a deeper understanding of test constructs in the teaching profession
-   Exploring the formative assessment potential of test systems
-   Cooperation in development and research at tertiary level
-   Development and research in the testing of young learners (Volksschule)
-   Research into the relationship among socio-economic variables and test performances
-   Research into factors complicating the implementation of testing systems
-   Create new and make existing corpora in all languages available for research purposes

ALPEN-ADRIA UNIVERSITÄT
KLAGENFURT | WIEN GRAZ