

Periodic Assessment of Research, Development, Artistic and Other Creative Activities VER 2026

Final Report of the Education Sub-panel

Overall, the sub-panel found a slight improvement in the quality of work since 2022. It is not clear to what extent this is substantive, or simply due to better choices of submissions and alignment with the research area. Too many pieces were still rated as 5 (below national recognition) due to irrelevance to education rather than quality necessarily.

Overall though, the quality of most work was still quite poor. Almost no studies had a research design. Most were small snapshots concerning relatively minor issue for education. There were no systematic reviews of prior evidence, and no robust evaluations of either policy or practice (randomised control trials or similar). There was very limited use of existing datasets, and nothing involving freely-available international datasets like PISA or TIMSS. There were no substantial sample sizes, no apparent consideration of data quality or missing data, and the impact of these on results. Much of the reporting was brief and formulaic, with background sections considerably longer than the methods, results and implications combined. There appears to be no habit of presenting a limitations section. The impression is of little passion for genuine discovery, and a limited set of skills/techniques for making genuine discoveries.

The five grades for outputs

We found no studies rated by the sub-panel as World-leading. This is a strong requirement for work in a field that is innovative, original, significant and robust, and that could take its place at the head of the best studies in the world. This might involve larger or longer-term

projects than were found here. None of the studies followed cases over extended time periods, or combined elements like process and impact evaluations. World-leading work would have required:

- Robust conceptual clarity, clear research questions, strong articulation of significance
- Appropriate, rigorous, and well executed methods (e.g., well designed statistical analysis, well-justified design)
- Clear contribution to the field based on warranted conclusions
- There would be no major threats to validity
- The work would be large-scale for its type, with complete and high quality data

We found only a couple of outputs deemed Internationally Excellent, largely for the reasons already given. We may have been a bit generous with these but wanted to encourage the research characteristics involved:

- Clear and relevant research aims
- Suitable analysis and interpretation
- Findings that could contribute meaningfully to the field

Crucial for all good studies, the reporting must be full, clear and honest. Currently, too many studies are brief, and/or apparently addressed to a minority audience. It does not matter how parochial a study is, it must be accessible to a wider audience

The best work that the sub-panel encountered was deemed Internationally Recognised. As above, this is nothing to do with whether the work itself was international or comparative. It is about whether the work would stand well with work going on in the same (or a similar) topic around the world. We took no account of where the report was published, or whether it had been peer-reviewed. However, it is worth noting that none of the work was in other

than local journals of unknown provenance, even where the topic was of international appeal. Perhaps researchers should be more ambitious here, if only to learn more about the “standards” expected elsewhere.

This Internationally Recognised work would be the norm (not the best) in the UK or elsewhere. It would be:

- Reasonably competent academic research
- Any contribution would be more modest rather than field-changing – unlikely to be cited much
- Suitable methods but perhaps limited in nature and design
- Limitations may be more pronounced (e.g. narrow samples, restricted analytical depth).

Most research was judged to be Nationally recognised. Again, this is perhaps generous since it was used to grade almost anything that was actually education research, as opposed to work below this standard which was mostly not research, and sometimes not at all relevant to education. Nationally Recognised usually denotes:

- Small attempts at empirical studies, of no great rigour or significance
- With research questions not well articulated
- Sometimes flawed or inappropriate analyses (tiny sub-groups, no comparators for comparative claims, significance tests with non-randomised cases...)

A high proportion of work submitted was graded as Below National Recognition. In practice this meant non-research or research so flawed or of so little consequence that it contributes nothing to the body of existing knowledge on any topic. This might involve:

- Serious methodological issues that compromise validity

- Poor alignment between aims, methods, and conclusions
- Unclear reporting or incoherence

Submissions included textbooks with not research basis, histories of institutions, and apparent research with no methods described at all.

The five grades for impact and environment

It was harder to judge the quality of both the social impact of studies and their creative environment. The criteria were not as easy to assess as our “sieve” for outputs. And the evidence presented was less objective, relying heavily on self-reports by institutions. There was little or no evidence of actual changes in society or classrooms directly attributable to the research (a rare phenomenon even for leading studies around the world). There is currently no clear link between the grades for outputs and the rest, but some sense of correlation between impact and environment. Nevertheless, the overall impression was that work on creating conditions for research and engaging with users was of a higher standard than the research itself. This could be due to the difficulty of judging self-reports, or it could be a sign of promise that future work will benefit from a sounder foundation. We hope it is the latter.

As with outputs neither impact case studies nor creative environment statements yielded a World Leading or Excellent grade.

Impact

True impact is intrinsically hard to produce and to demonstrate. It is largely not under the control of the researcher or institution, beyond making research as trustworthy and relevant as possible. For strong impact we expected the following generic factors:

- Well-established links to users and user bodies
- Robust underlying studies or reviews (weak studies should not have real-world impact)
- Independent examples where policy-makers, schools or other users cited the research as being influential in making their decisions.

Environment

For a strong environment we expected to see competitive, external grant-funding, throughput of PhDs, clear national and international links with other research organisations, and programmes to help build the capacity of research staff at all levels.

Suggestions for the future

Members of the sub-panel would be happy to offer help in any proposed capacity-building activities.¹

In reviews of literature and beyond, it might be useful for researchers to look beyond Slovakia and neighbouring countries – for ideas, substantive research and methods, even if only for critique.

Focus more on study design. Consider a wider range of approaches, including systematic reviews, secondary data analysis, web-crawling, and randomised control trials, for example.

¹Members of the Sub-panel are available to provide consultation in support of capacity-building activities. Applicants interested in this opportunity may contact the Ministry at ver@minedu.sk for contact details (*note by the Ministry*). / **Členovia komisie sú k dispozícii na poskytnutie konzultácie na podporu aktivít zameraných na budovanie kapacít. Žiadatelia, ktorí majú o túto možnosť záujem, môžu kontaktovať ministerstvo na adrese ver@minedu.sk pre kontaktné údaje (pozn. ministerstva).**

Sub-panel members

Chris Brown, University of Southampton

Antonio Francisco Canales Serrano, Universidad Complutense de Madrid

Pierre Dillenbourg, École Polytechnique Fédérale de Lausanne

Justin Dillon, University College London

Stephen Gorard, Durham University (Sub-panel Chair)

Dimitra Kokotsaki, Durham University

Zsuzsa Millei, Tampereen yliopisto

Martin Tomasik, Universität Zürich

Mathias Urban, Dublin City University

Andrew Wilkins, Goldsmiths, University of London

Nicola Yelland, The University of Melbourne