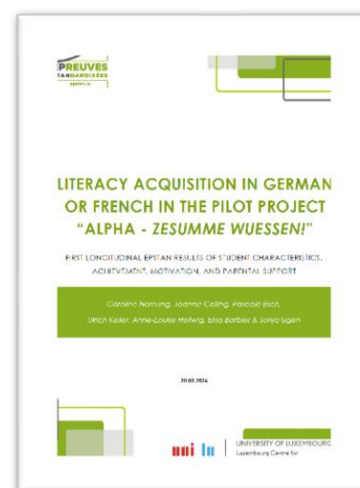


Scientific Review of the Longitudinal ALPHA Report

Independent Expert Opinions on

Hornung, C., Colling, J., Esch, P., Keller, U., Hellwig, A.-L., Barbier, E. & Ugen, S. (2026). *Literacy Acquisition in German or French in the Pilot Project “ALPHA - zesumme wuessen” – First longitudinal ÉpStan Results of Student Characteristics, Achievement, Motivation, and Parental Support*. Luxembourg Centre for Educational Testing (LUCET).

15 July 2026



Contents

Purpose and context	2
Executive summary	2
Independent expert opinions	3
Expertise 1 Prof. Dr. Andreas Hadjar	3
Expertise 2 Prof. Dr. Samuel Greiff	6
Expertise 3 Prof Dr. Jean-Marc Dewaele	10
Expertise 4 Prof. Dr. Marc Demeuse	11

Purpose and context

The Luxembourg Centre for Educational Testing (LUCET) has prepared a report on the ALPHA pilot project that has been under embargo due to diverging opinions with SCRIPT regarding methodological issues. It is important to note that key features of the study design regarding the number of students enrolled in the pilot project ALPHA as well as data availability, were determined by the implementation of the pilot itself and were therefore beyond the control of the authors of this report. These constraints necessarily influenced the methodological options available for the evaluation. Following questions raised regarding the methodological approach and interpretation of the longitudinal evaluation, LUCET invited four independent experts in educational evaluation and quantitative methodology to provide written scientific opinions on the study. This document includes the three expert opinions received to date.

The experts were asked to assess the appropriateness of the study design, the analytical approach, and the interpretation of the findings in light of the available data. The experts prepared their opinions independently. They received no financial compensation or other consideration for their contributions, and the opinions presented are their own.

For the sake of transparency, we disclose that two of the experts (Prof. Hadjar and Prof. Greiff) were previously affiliated with the University of Luxembourg, and two (Prof. Hadjar and Prof. Demeuse) currently serve on LUCET's Scientific Advisory Board. However, neither of these three experts has been involved in the present project, and there has been no substantial collaboration with the research team either during or following their affiliation with the University.

Executive summary

The expert opinions broadly converge in their assessment of the longitudinal evaluation of the ALPHA pilot project. They find the methodological approach and analytical procedures to be scientifically appropriate and consistent with established standards in educational research.

While noting the limitations described above regarding sample size, data availability, and study design, the experts consistently judge that these constraints are adequately acknowledged and transparently integrated into the analysis and interpretation of results. They further consider that the statistical methods employed are appropriate for the research questions addressed and that the conclusions drawn in the report are broadly supported by the empirical evidence.

In addition, the experts underline the value of the evaluation for both scientific understanding and policy discussion.

Independent expert opinions

Expertise 1 Prof. Dr. Andreas Hadjar

Prof. Dr. Andreas Hadjar
Université de Fribourg, Switzerland

Full Professor “Sociology, Social Policy, Social Research”

Scientific review

Hornung, C., Colling, J., Esch, P., Keller, U., Hellwig, A.-L., Barbier, E. & Ugen, S. (2026). Literacy Acquisition in German or French in the Pilot Project "ALPHA - zesumme wuessen" – First longitudinal ÉpStan Results of Student Characteristics, Achievement, Motivation, and Parental Support. Luxembourg Centre for Educational Testing (LUCET).

Background and summary of report:

This report centres on an evaluation study of the literacy pilot project "ALPHA - zesumme wuessen". This educational intervention was implemented in four primary schools in pre-primary education (Cycle 1.2) in the 2022/23 school year, allowing students an alphabetisation in terms of a literacy acquisition in a language fitting their home languages, rather than in German only. While the Francophone pupils would be assigned to a group acquiring literacy in French, the Germanophone pupils would be assigned to a group acquiring literacy in German. As the study authors mention in their report, this educational intervention introduced by the Luxembourgish Ministry is based on the scientifically backed "assumption that literacy acquisition is facilitated when students learn to read and write in a language linguistically closer to their home language(s)" (Hornung et al., p. 5; in more detail: p. 8f). The evaluation strategy of the related research project group centres on a comparative design, reassessing pupil's competencies in Cycle 3.1 and comparing the intervention groups with statistical control groups, employing a propensity-matching procedure to account for differences between the intervention and the control groups. The results do not show a considerable advantage regarding Mathematics as well as regarding reading and listening comprehension for the intervention groups who acquired literacy in either French or German (better fitting their home language) two years after the implementation of the intervention in Cycle 3.1. Slight differences were noted for Mathematics motivation, in favour of the German intervention group. The German intervention group also shows a profound benefit regarding subjective wellbeing, while the French intervention group appears to be disadvantaged in comparison to the other groups. The parents of the French intervention group perceive themselves as better able to support their children in (French) literacy acquisition than the parents of the French control group who would need to support the German literacy acquisition.

Evaluation of the study:

The study the report by Hornung et al. (2026) is based on appears to be of high quality. The comparative methodological design is convincing, as it even takes into account (socioeconomic and other) differences between the intervention groups and the control groups, as otherwise differences in the results may just be due to a selection bias. The same applies to the longitudinal design, reassessing the population of children who participated in the ALPHA study before in Cycle 3.1. Propensity-score matching is among the state-of-the-art statistical procedures to compare between different groups and assess a possible impact of interventions. Furthermore, the researchers also meaningfully addressed the fact that the competency measurement in Cycle 2.1 differed language-wise from the competency measurement in Cycle 3.1 in the French-language intervention group, making both measures not directly comparable to the results of the other four groups. The researchers accounted sufficiently to this problem, by a conceptual-equating procedure, developing the tests using the same *Plan d'Études* (MENFP, 2011) as a reference document and employing French translations including the same amount of items and texts as the German version.

The researchers carried out the study accompanied by a high level of reflection, including methodological reflection. Thus, they outline limitations of the study and indicate strong need for further research. The researchers and report authors – from my scientific perspective – rightly conclude that the study does not indicate a specific benefit of the intervention (the alphabetisation in a more fitting language) for the group of Francophone families for whom this intervention was designed. The Francophone children in the intervention group did not show an improved competency level than in the Francophone control group.

Being again an indication for their high scientific skills, the authors provide meaningful interpretations for these results in the final sections of their report. My own reading of these results, based on knowledge acquired during the study of international educational policy interventions to study inequalities, is that policymakers need to consider several side effects that may hinder the at first sight meaningful intervention to show strong positive effects in disadvantaged groups. As the state-of-research clearly indicates segregated settings as a driver of disadvantages and inequalities, the segregation of Francophone and Germanophone students may also come with the downsides of such segregation. With regard to the student population in Luxembourg, Francophone environments are dominated by disadvantaged groups characterised by low socio-economic status of the parents and immigrants (with Portuguese-speaking students as a particularly deprived group), while German and English-speaking environments mean a student composition that is much more advantaged. Certain home language-specific provisions need to be carried out in inclusive settings and deserve a high level of resources.

Fribourg, 28 May 2026

Andreas Hadjar

Prof. Dr. Andreas Hadjar
Full Professor in Sociology, Social Policy, Social Research

Expertise 2 Prof. Dr. Samuel Greiff

Prof. Dr. Samuel Greiff
Technical University of Munich, Germany

Full Professor "Educational Monitoring & Effectiveness"
Chairman of the Centre for International Student Assessment (ZIB)
Director PISA National Centre & National Project Manager

Munich, 08.06.2026

1 / 3

Independent Academic Review of the ALPHA Longitudinal Evaluation Report

To Whom it may concern,

I was asked to provide an independent academic assessment of the report Literacy Acquisition in German or French in the Pilot Project “ALPHA – zesumme wuessen” (2026). The present review considers the study’s scientific contribution, methodological approach, interpretation of findings, and overall value for the ongoing evaluation of the ALPHA pilot project.

Summary of the study

The report presents the first longitudinal findings from the „ALPHA – zesumme wuessen“ pilot project, which evaluates literacy acquisition in French as an alternative to the traditional German literacy pathway within Luxembourg’s multilingual education system. The study addresses a timely and relevant educational and policy question, namely whether alternative literacy pathways may support students from diverse linguistic backgrounds while maintaining comparable educational outcomes. The analyses follow the first cohort of ALPHA students from Cycle 2.1 (2023/24) to Cycle 3.1 (2025/26). The report examines academic achievement in mathematics, listening and reading comprehension, as well as students’ motivation, school wellbeing, and parental support.

Methodologically, the study combines comparisons with the national ÉpStan cohort and with matched reference groups constructed through propensity score matching. The longitudinal sample in ALPHA comprises 37 students in the ALPHA-French group and 46 students in the ALPHA-German group. The report transparently documents sample attrition and grade retention between measurement points and acknowledges the limitations associated with the relatively small sample sizes. Overall, the study provides an important first longitudinal perspective on the implementation and outcomes of the ALPHA pilot project and establishes a valuable foundation for future follow-up analyses.

Technical University of Munich
Centre for International Student
Assessment (ZIB e.V.).

PISA National Project Management
Arcisstraße 21
80333 Munich

Tel. +49 89 289 28272

pisa@tum.de
www.pisa.tum.de
www.zib.education

Bayerische Landesbank
IBAN-Nr.:
DE1070050000000024866
BIC: BYLADEMM
Steuer-Nr.: 143/241/80037
USt-IdNr.: DE811193231

Methodology of the study

The methodological approach in the report is appropriate and scientifically defensible given the constraints of evaluating a small-scale educational pilot under real-world conditions. In the absence of random assignment (which is neither feasible nor desirable in the context of a voluntary educational programme) the use of propensity score matching represents a widely accepted strategy for constructing comparable reference groups. At the same time, the well-known limitations of matching approaches should be acknowledged when interpreting the findings. In particular, matching can only account for observed characteristics and cannot fully exclude the possibility that unmeasured differences between groups may contribute to the observed outcomes that might increase further over time. This consideration is especially relevant in relatively small samples, where residual differences between groups may be difficult to detect and statistically adjust for. The report is appropriately cautious in this regard.

Similarly, the use of conceptually equated assessments across languages constitutes an important and necessary step towards comparability; however, as the authors recognize, conceptual equivalence cannot be assumed to eliminate all potential language-related differences in test functioning. Overall, the analytical strategy is well aligned with the aims of the study and provides a credible basis for examining the development of the ALPHA cohorts.

Limitations of the study

The report openly acknowledges the main limitations of the study, most notably the relatively small pilot cohorts, which were inherent to the design of the project. While the use of propensity score matching is appropriate in this context, the combination of small sample sizes and non-random assignment necessarily limits statistical precision and the extent to which findings can be generalized beyond the participating students and schools, both of which are small. Relatedly, the pilot currently involves only four schools, meaning that the reported findings should primarily be interpreted within the context of these specific implementations.

Given the limited sample size, statistical power to detect potentially meaningful differences is likely constrained. The absence of obvious (or statistically significant) group differences should not automatically be interpreted as evidence of equivalence. More generally, much of the evidence presented is descriptive in nature and is therefore particularly valuable for identifying promising patterns and generating hypotheses for future longitudinal analyses.

Overall assessment

Overall, the report provides a valuable contribution to the evaluation of the ALPHA pilot project. The study addresses an important educational question and makes appropriate use of the available longitudinal data. The authors are commendably explicit about the limitations of the design and generally reflect these in their interpretation of the findings.

At the same time, these limitations are substantial and constrain the conclusions that can be drawn, in particular from a policy perspective. Given the small sample sizes and the reliance on largely descriptive analyses, the study is better suited to documenting patterns within the

Technical University of Munich
Centre for International Student
Assessment (ZIB e.V.).

PISA National Project Management
Arcisstraße 21
80333 Munich

Tel. +49 89 289 28272

pisa@tum.de
www.pisa.tum.de
www.zib.education

Bayerische Landesbank
IBAN-Nr.:
DE1070050000000024866
BIC: BYLADEMM
Steuer-Nr.: 143/241/80037
USt-IdNr.: DE811193231

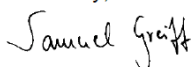
participating cohorts than to drawing strong inferences about performance differences, developmental trajectories, or programme effects as such. This is particularly relevant for questions regarding socio-economic inequalities, which motivate the project but cannot be examined in depth with the available data.

Recommendations

Overall, the report provides a strong basis for the ongoing evaluation of the ALPHA pilot project. The following recommendations are offered in a constructive spirit and are intended to further strengthen the clarity, interpretability, and scientific contribution of future reporting.

1. Further strengthen the distinction between descriptive findings and their interpretation: Given the exploratory nature of the study and the limited statistical power, a more explicit separation between empirical results and their broader interpretation would further enhance the transparency of the report. This is particularly relevant when discussing potential implications for educational inequalities and longer-term programme effects within the Luxembourgish educational system.
2. Differentiate more clearly between benchmarking and programme evaluation: The report serves two legitimate but distinct purposes: (a) positioning the ALPHA cohorts within the broader national educational context and (b) assessing potential programme-related differences through comparisons with matched reference groups. Presenting these two analytical perspectives more explicitly and separately would facilitate interpretation for both scientific and policy audiences.
3. Maintain a strong emphasis on the study's methodological limitations when drawing conclusions: The report appropriately discusses limitations related to sample size, attrition, educational pathways, and the non-randomized design. Given the importance of these constraints for the interpretation of the findings, it would be beneficial to ensure that they remain closely linked to the discussion of results and conclusions throughout the report.
4. Continue the longitudinal follow-up of the ALPHA cohorts: The current study provides an important first longitudinal assessment of the pilot project. Future follow-up analyses, based on additional cohorts and later educational stages, will be particularly valuable for determining whether the patterns observed at this stage remain stable over time and for strengthening the evidence base for future policy decisions.

Sincerely,



Prof. Dr. Samuel Greiff

*Full Professor "Educational Monitoring & Effectiveness"
Chairman of the Centre for International Student Assessment (ZIB)
Director PISA National Centre & National Project Manager*

Technical University of Munich
Centre for International Student
Assessment (ZIB e.V.).

PISA National Project Management
Arcisstraße 21
80333 Munich

Tel. +49 89 289 28272
pisa@tum.de
www.pisa.tum.de
www.zib.education

Bayerische Landesbank
IBAN-Nr.:
DE10700500000000024866
BIC: BYLADEMM
Steuer-Nr.: 143/241/80037
USt-IdNr.: DE811193231

Prof Dr. Jean-Marc Dewaele
University of London, UK

Emeritus Professor in Applied Linguistics and Multilingualism, Birkbeck, University of London
Honorary Professor, Institute of Education, University College London

Report on the LITERACY ACQUISITION IN GERMAN OR FRENCH IN THE PILOT
PROJECT "ALPHA - ZESUMME WUESSEN" 20.03.2026
FIRST LONGITUDINAL EPSTAN RESULTS OF STUDENT CHARACTERISTICS,
ACHIEVEMENT, MOTIVATION, AND PARENTAL SUPPORT
Luxembourg Centre for Educational Testing (LUCET), University of Luxembourg

London, 16/06/2026

TO WHOM IT CONCERNS

I have read this excellent report and I want to congratulate the authors on the quality of the analyses, the justification of the choice of statistical techniques -given the very small sample sizes- and the very careful and clear interpretation of the findings. The authors of this report did their utmost to match their very small experimental population with the reference groups.

There is no doubt in my mind that the "ALPHA - ZESUMME WUESSEN" pilot project did not have any significant impact on the performance of the 37 learners in the French groups and the 46 learners in the German group. There is absolutely no guarantee that a larger sample size would have yielded significant results. A larger sample size would simply have allowed more sophisticated statistical analyses that could have taken into account the multiple independent variables in the research design but I do not think that the results would have been any different. The assumption that literacy acquisition is facilitated when students learn to read and write in a language linguistically closer to their home language(s) may seem common-sensical but I would have been very surprised if that was indeed the case. I have found in my own research that linguistic distance may play a very minor role once learners have acquired sufficient meta-linguistic awareness about the different languages in their repertoire. I do not think these primary school children had any meta-linguistic awareness yet, which would have created affordances for them.

To conclude, this rigorous null-result study shows that that literacy acquisition is NOT facilitated when students learn to read and write in a language linguistically closer to their home language(s).



Jean-Marc Dewaele
Emeritus Professor in Applied Linguistics and Multilingualism, Birkbeck, University of London; Honorary Professor, Institute of Education, University College London.
Former president of the *European Second Language Association*, the *International Association of Multilingualism* and the *International Association for the Psychology of Language Learning*. General Editor of the *Journal of Multilingual and Multicultural Development* (Taylor & Francis).
Email: j.dewaele@bbk.ac.uk.

Expertise 4 Prof. Dr. Marc Demeuse

Prof. Dr. Marc Demeuse
University of Mons, Belgium

*Full Professor “Methodology and Training”
Head of “Institut d’administration scolaire”*

Université de Mons

Institut d'Administration Scolaire
Sciences de l'Enseignement
et de la Formation
Faculté de Psychologie et
des Sciences de l'Éducation

Prof. Dr. Marc DEMEUSE

Place du Parc, 20
B - 7000 Mons
tél +32(0)65 37 31 90
marc.demeuse@umons.ac.be
www.umons.ac.be/inas



Objet : En réponse à la demande d'évaluation externe indépendante du rapport longitudinal ALPHA (Prof. Dr. Sonja Ugen, 2026.05.10)
Homung, C., Colling, J., Esch, P., Keller, U., Hellwig, A.-L., Barbier, E. & Ugen, S. (2026).
L'alphabétisation en allemand ou en français dans le projet pilote « ALPHA - zesumme wuessen ». Premiers résultats Epstan longitudinaux sur les caractéristiques des élèves, les performances scolaires, la motivation et le soutien parental. Luxembourg Centre for Educational Testing (LUCET).

Le Luxembourg est un véritable laboratoire et cette étude met en évidence combien il est difficile de parvenir à un réel plurilinguisme à l'échelle d'un système éducatif où un nombre important de jeunes élèves est confronté, à l'école, à une autre langue que celle qu'ils parlent et, peut-être aussi, lisent et écrivent déjà à la maison. Le contexte de plurilinguisme, alors qu'il pourrait représenter un atout important, se heurte malheureusement aussi, comme le montrent les études internationales et nationales à d'importantes inégalités de performances au Luxembourg.

La volonté du projet ALPHA est de mettre en place des écoles publiques (n=4) dans lesquelles les jeunes élèves peuvent, dès le début de leur scolarité, apprendre à lire et à écrire soit en allemand, soit en français de manière à faire face aux difficultés rencontrées par les élèves qui ne pratiquent/maîtrisent pas la langue d'alphabétisation dans leur contexte familial. Naturellement, ce projet ne rencontre pas complètement toutes les situations possibles, notamment celle des élèves lusophones. L'hypothèse de base sur laquelle s'appuie ce projet est celle de la facilitation des apprentissages en mobilisant l'utilisation d'une langue proche de celle parlée à la maison plutôt que la langue officielle de scolarisation/alphabétisation lors du début des apprentissages scolaires. Le rapport du LUCET présente une évaluation du projet ALPHA en 3^e année de l'enseignement primaire, après une première évaluation en 1^e année, il y a deux ans, sur la même cohorte (groupe pilote).

Les élèves du groupe pilote sont progressivement insérés dans le monitoring national (EpStan), ce qui permet de comparer leurs caractéristiques (notamment performances académiques et motivation) à celles des autres élèves de manière à évaluer la mise en œuvre de ce projet (actuellement en C3.1, soit en 3^e année de l'enseignement primaire, après une première évaluation en C2.1 cf. le rapport publié en 2024). Compte-tenu de la taille de l'échantillon du groupe pilote, divisé en 2 sous-groupes (ALPHA-Français - n=37 - et ALPHA-Allemand - n=42 -, selon que les élèves apprennent à lire et à écrire en français ou en allemand) et des différences socio-économiques constatées entre le groupe pilote et la cohorte principale lors du premier rapport (niveau socio-économique plus faible dans le groupe pilote), la comparaison reste complexe à mener. Une méthode statistique baptisée « propensity score matching » a été mise en place de manière à paier chaque élève du groupe pilote et avec cinq élèves de la cohorte principale dès la première étude et une méthode alternative de sélection des élèves servant de point de comparaison est aussi présentée dans le rapport de manière à faire face aux problèmes d'attrition, sans que les conclusions ne soient réellement modifiées.

Les résultats observés en mathématiques, compréhension orale et compréhension écrite (en français ou en allemand, selon leur groupe d'appartenance) ne rencontrent malheureusement pas les attentes, soit une amélioration par rapport à leur groupe de référence. Des différences se marquent cependant en termes de motivation en mathématiques dans le groupe ALPHA-Allemand et de motivation pour l'apprentissage de leur langue d'alphabétisation et de soutien parental dans le groupe ALPHA-Français. Les auteurs font ici l'hypothèse que les parents se sentent plus compétents pour suivre leurs enfants dans leur propre langue. Au-delà des résultats, les auteurs proposent d'intéressantes hypothèses qu'il leur faudra éprouver par la suite, notamment quant aux caractéristiques des langues elles-mêmes (par exemple sur la transparence ou, au contraire, l'opacité de chacune des deux langues). D'autres hypothèses



résultent de la plus ou moins grande exposition à la langue d'alphabétisation dans les autres disciplines (notamment par comparaison aux Ecoles européennes publiques où les élèves peuvent choisir l'anglais, l'allemand ou le français comme langue principale d'enseignement).

De manière très transparente, les auteurs mettent bien en évidence certains problèmes méthodologiques liés au calibrage des épreuves en français ou à la taille et au pairage des échantillons, dans un dispositif longitudinal. Ils plaident pour un renforcement de la recherche avant généralisation de cette expérience menée à petite échelle, même s'ils se déclarent convaincus de l'intérêt de cette expérimentation dans le contexte luxembourgeois. Ils pointent notamment des paramètres qui ne semblent pas encore avoir été réellement pris en compte, comme l'exposition linguistique, la composition des classes et les approches pédagogiques mobilisées.

Le lecteur attentif aura sans doute aussi repéré d'autres aspects qui méritent d'être questionnés, comme les profils linguistiques dans les deux sous-groupes du groupe pilote : 52% d'élèves avec un profil « portugais » contre 27% de profils « français » et 16% de profils « luxembourgeois/allemand » dans le groupe ALPHA-Français alors qu'ils sont 52% de profils « luxembourgeois/allemands » contre 24% de « portugais » et 13% de « français » dans le groupe ALPHA-Allemand. Ces proportions sont néanmoins relativement bien respectées dans les groupes de références qui servent de comparaison. Le fait que les élèves de référence (cohorte principale) pour le groupe ALPHA-Français soient testés en allemand dans le système EpStan alors que le groupe pilote est testé en français ajoute encore aux difficultés.

Comme toute étude longitudinale, mais en particulier celles qui débutent avec de très faibles effectifs, le projet ALPHA n'est pas à l'abri de phénomènes d'attrition (y compris à cause de l'allongement de cycle octroyé à certains élèves, soit près d'un élève sur 5 dans le groupe pilote évalué deux ans plus tôt) et les comparaisons entre les groupes pilotes et la population constituent une véritable gageure. Les choix méthodologiques adoptés par les collègues du LUCET sont très certainement ingénieux, mais ils n'évitent pas les critiques qui peuvent être adressées à ce type de comparaisons statistiques et auxquelles les auteurs font face avec sérieux et modestie. Comme ils l'indiquent bien, la question de la généralisation de cette politique volontariste ne peut pas honnêtement être tranchée à ce stade. On ne peut donc que soutenir leur souhait de pouvoir poursuivre les travaux entamés, à la fois en améliorant les instruments de mesure (notamment en français), en étendant l'échantillon et en collectant les données qui leur permettent de mieux contrôler les éventuels effets. Un travail de recherche sur les approches pédagogiques mobilisées et sur la meilleure compréhension/modélisation de la maîtrise des langues en présence dans un contexte qui cumule à la fois différences socio-économiques et répartition linguistique ne peut qu'améliorer le projet ALPHA et son évaluation.



Marc DEMEUSE
Professeur ordinaire
Chef de service
Mons (BE), le 15 juillet 2026.